

Pendugaan Angka Kematian Bayi dengan Menggunakan Model Poisson Bayes Berhirarki Dua-Level

NUSAR HAJARISMAN¹, KHAIRIL A.N.², KUSMAN SADIK³, I GUSTI P. PURNABA⁴

¹Program Studi Statistika, Universitas Islam Bandung, Jl Purnawarman 63, Bandung, Indonesia

^{2,3}Departemen Statistika, Institut Pertanian Bogor, Kampus IPB Darmaga, Bogor, Indonesia

⁴Departemen Matematika, Institut Pertanian Bogor, Kampus IPB Darmaga, Bogor, Indonesia
email: ¹ nusarhajarisman@yahoo.com, ^{2,3,4} khiril@yahoo.com

Abstract. Official institutions of national data providers such as the BPS-Statistics Indonesia is required to produce and present the statistical information, as necessary as a form of contributory BPS region in support of regional development policy and planning. There are survey conducted by BPS capability estimation techniques are still limited, due to the resulting estimators have not been able to directly assumed for small areas. In this article we propose the hierarchical Bayesian models, especially for count data which are Poisson distributed, in small area estimation problem. The model was developed by combining concept of generalized linear model and Fay-Herriot model. The results of the development of this model is implemented to estimate the infant mortality rate in Bojonegoro district, East Java Province.

Keywords: Fay-Herriot model, Poisson distribution, infant mortality rate.

Abstrak. Lembaga resmi penyedia data nasional seperti Badan Pusat Statistik (BPS) dituntut untuk memproduksi statistik dan menyajikannya sebagai informasi, sesuai kebutuhan daerah sebagai bentuk kontributif BPS dalam mendukung perencanaan dan kebijakan pembangunan daerah. Terdapat banyak kegiatan survey yang dilakukan oleh BPS dengan kemampuan teknik pendugaan yang masih terbatas, karena penduga yang dihasilkan belum mampu secara langsung menduga untuk area kecil. Oleh karena itu perlu dikembangkan suatu metode pendugaan area kecil, khususnya untuk data yang berbentuk cacahan, yang mampu memenuhi ketersediaan statistik area kecil. Model yang diusulkan dalam makalah ini adalah model Bayes berhirarki untuk data cacahan yang bersebaran Poisson untuk menyelesaikan masalah pendugaan area kecil dengan cara memadukan konsep pemodelan linear terampat dengan model Fay-Herriot. Kemudian, hasil-hasil dari pengembangan model ini diimplementasikan untuk menduga angka kematian bayi level kecamatan di Kabupaten Bojonegoro, Provnsi Jawa Timur.

Kata Kunci: model Fay-Herriot, sebaran Poisson, angka kematian bayi.

Pendahuluan

Masalah pendugaan area kecil (*small area estimation*, SAE) dalam dua dekade terakhir ini mulai banyak diperhatikan oleh banyak kalangan peneliti. Hal ini mencerminkan perlunya untuk menghasilkan suatu penduga untuk area dengan otoritas lokal atau domain kecil. Kesulitan dalam menghasilkan suatu penduga seperti itu, terutama untuk area-area kecil, adalah ukuran contoh yang dicapai melalui rancangan survey tertentu untuk keperluan nasional adalah terlalu kecil sebagai penduga langsung (*direct estimator*) dengan presisi yang dapat diterima. Oleh karena itu perlu suatu penduga taklangsung (*indirect estimator*) yang

'meminjam' kekuatan informasi dari area-area sekitarnya yang dihubungkan melalui suatu pendekatan model.

Salah satu metode pendugaan area kecil yang berkembang saat ini adalah metode Bayes, khususnya model Bayes berhirarki. Metode Bayes berhirarki ini dapat diterapkan pada model campuran linear terampat (*generalized linear mixed models*) yang digunakan untuk menangani data kategorik, seperti data biner dan data cacahan. Pada dasarnya, Rao (2003) serta Jiang dan Lahiri (2006) telah membahas cukup banyak tentang penerapan model linear terampat ini pada pendugaan area kecil ini. Walaupun penerapan dari

pendekatan Bayes pada masalah pendugaan area kecil sudah mulai banyak berkembang, akan tetapi masih sedikit apabila dibandingkan dengan pendekatan klasik utama. Masalah pendugaan area kecil untuk data kategorik ini lebih banyak dikembangkan untuk data biner, padahal dalam prakteknya sering kali ditemui data survey yang berbentuk cacahan (*count data*).

Menurut Torelli dan Trevisani (2007) sekarang ini pembahasan mengenai spesifikasi model nonlinear yang tepat pada saat penduga area kecil diperlukan untuk peubah kategorik relatif masih sedikit. Penggunaan pendekatan hirarki Bayes yang saat ini masih terbatas, padahal pendekatan ini mempunyai beberapa keuntungan, diantaranya yaitu: spesifikasi modelnya adalah langsung dan dapat digunakan untuk memodelkan berbagai sumber variasi yang berbeda, masalah inferensinya relatif lebih jelas dan komputasinya juga relatif lebih mudah dengan menggunakan teknik rantai Markov Monte Carlo (Markov *chain* Monte Carlo, MCMC).

Model alternatif untuk menduga angka kematian menurut peubah umur tertentu pada salah satu kategori penyakit kanker di Amerika Serikat juga telah dilakukan oleh Nandram, et al. (1999). Mereka menggunakan metode Bayes yang diterapkan pada empat jenis model yang berbeda. Masing-masing model itu mengasumsikan bahwa banyaknya yang mati pada area tertentu, dan kelompok umur tertentu adalah bersebaran Poisson. Hasilnya menunjukkan bahwa model yang mereka usulkan mampu menangkap efek area kecil dan efek regional dengan baik, dan juga mampu mendeteksi korelasi spasial sisaannya, sehingga memudahkan proses pendugaan parameternya. Nandram, et al. (2000) juga melakukan pemodelan yang hampir sama dengan apa yang telah dilakukan sebelumnya, akan tetapi di sini modelnya diterapkan untuk menduga angka kematian menurut peubah umur tertentu pada salah satu kategori penyakit *chronic obstructive pulmonary* di Amerika Serikat.

Dalam artikel ini akan ditunjukkan implementasi dari konsep pemodelan linear terampat pada masalah pendugaan area kecil. Model yang dikembangkan dalam penelitian ini adalah model regresi Poisson berhirarki dua-level dengan cara memadukan terminologi yang ada dalam model linear terampat dengan konsep metode Bayes berhirarki dua-level sedemikian rupa sehingga dapat diimplementasikan untuk menangani masalah pendugaan area kecil yang diwakili oleh model Fay-Herriot. Sebagaimana yang telah diketahui bahwa salah satu permasalahan dalam pemodelan Bayes berhirarki adalah pemilihan sebaran prior. Apabila sebaran prior ini diketahui maka inferensi dapat dengan mudah dilakukan dengan cara meminimumkan galat pos-

terior, menghitung daerah kepekatan sebaran posterior yang lebih tinggi dimensinya, atau mengintegrasikan parameter untuk menemukan sebaran prediktifnya.

Model Poisson Bayes berhirarki dua-level yang dikembangkan di sini menggunakan dua sebaran prior yang berbeda. *Pertama*, menggunakan sebaran prior gamma yang merupakan sebaran prior yang bersifat *conjugate* bagi sebaran Poisson. Dengan menggunakan sebaran prior yang bersifat *conjugate* ini akan memberikan kemudahan dalam penentuan sebaran posteriornya. *Kedua*, disini menggunakan sebaran prior yang bersifat *non conjugate* bagi sebaran Poisson, yaitu sebaran invers gamma. Penggunaan sebaran prior yang bersifat *non conjugate* ini tentu saja akan memberikan kesulitan dalam menemukan sebaran posterior dalam bentuk persamaan tertutup. Namun masalah ini dapat diatasi melalui penerapan rantai Markov Monte Carlo yang akan menghasilkan sederet peubah acak yang mendekati sebaran posteriornya. Merujuk pada teori yang dikembangkan oleh Gelman (2006) bahwa sebaran prior invers gamma tersebut diklasifikasikan sebagai sebaran prior yang bersifat *conditionally conjugate*. Hasil-hasil dari pengembangan model dalam penelitian ini akan diaplikasikan untuk menduga angka kematian bayi (AKB) level kecamatan pada Kabupaten Bojonegoro, Provinsi Jawa Timur.

Model Regresi Poisson Bayes Berhirarki Dua-Level

Model regresi Poisson berhirarki telah banyak digunakan untuk menganalisis berbagai jenis data yang berbentuk cacahan. Kebanyakan analisis yang dilakukan untuk pemetaan penyakit (*disease mapping*) dimulai dengan proses penarikan contoh Poisson. Clayton dan Klador (1987) menggambarkan pendekatan Bayes empirik yang memperhatikan kemiripan spasial antar angka kematian penyakit tertentu. Sementara itu Ghosh et al. (2009) membandingkan metode Bayes empirik dan Bayes berhirarki, yang diaplikasikan pada masalah kesehatan untuk subpopulasi yang bersifat minoritas. Sementara itu, Breslow dan Clayton (1993) menggunakan model campuran linear terampat untuk mempelajari masalah pemetaan penyakit ini. Sedangkan, Waller, et al. (1997) mengusulkan model Bayes berhirarki spatio-temporal untuk memodelkan angka kematian regional menurut ruang dan waktu termasuk didalamnya interaksi antara ruang dan waktu itu sendiri.

Saat ini sudah mulai banyak penelitian mengenai penerapan model Bayes berhirarki untuk menangani masalah pendugaan area kecil ini. Torabi dan Rao (2008) mengembangkan masalah

pendugaan area kecil dua-level melalui penduga generalized regression (GREG). Sementara itu, You dan Chapman (2006) membahas tentang pendugaan area kecil untuk level area. Mereka mengembangkan model untuk memperoleh ragam penarikan contoh melalui pendekatan model Bayes berhirarki dengan algoritma Gibbs penarikan contoh. Penggunaan model Bayes berhirarki juga dilakukan oleh Souza et al. (2009) dalam rangka memprediksi populasi area kecil, dimana model yang dikembangkan dilakukan melalui model pertumbuhan terstruktur secara spasial. Pengembangan model spasial yang dikombinasikan dengan model Bayes berhirarki untuk menangani masalah pendugaan area kecil juga dilakukan oleh You dan Zhou (2011). Hasil pengembangan model yang dilakukan oleh You dan Zhou (2011) ini diaplikasikan pada masalah data kesehatan.

Dalam makalah ini akan diusulkan pengembangan model regresi Poisson berhirarki yang pertama kali diusulkan oleh Christiansen dan Morris (1997), dimana model ini pada awalnya tidak dirancang untuk digunakan dalam masalah survey penarikan contoh. Sekali lagi, model ini dikembangkan dengan cara memadukan terminologi yang ada dalam model linear terampat dengan konsep metode Bayes, khususnya metode Bayes berhirarki, sedemikian rupa sehingga dapat diimplementasikan untuk menangani masalah pendugaan area kecil untuk data survey yang berbentuk data cacahan. Pengembangan model ini dimulai dengan mengasumsikan peubah acak yang diamati merupakan anggota dari keluarga eksponensial, sebagaimana yang muncul dalam konsep pemodelan linear terampat, bersyarat pada suatu parameter tertentu. Tujuan utama dari pengembangan model ini adalah membuat inferensi pada parameter tersebut yang juga dianggap sebagai peubah acak. Kemudian parameter tersebut dimodelkan dengan menggunakan model Fay-Herriot sebagai model dasar dalam konsep pendugaan area kecil. Selanjutnya, perpaduan dari kedua model tersebut akan distandarkan sedemikian rupa sehingga mewakili suatu model dalam kerangka kerja metode Bayes yang pada akhirnya akan terbentuk model Poisson Bayes berhirarki untuk menyelesaikan masalah dalam pendugaan area kecil.

Berikut ini pembahasan mengenai pengembangan model regresi Poisson Bayes berhirarki dua-level. Misalkan y_{ij} menyatakan banyaknya peristiwa 'sukses' atau dalam hal ini banyaknya kejadian yang mati pada unit pengamatan ke- j untuk area ke- i , n_i menyatakan populasi dalam area ke- i , serta θ_{ij} menyatakan angka kematian unit pengamatan ke- j pada area ke- i , dimana $\theta_{ij} = y_{ij}/n_{ij}$ (untuk $i = 1, 2, \dots, m$ dan $j = 1, 2, \dots, n_i$), serta m menunjukkan banyak area kecil yang diamati. Dalam hal ini y_{ij} adalah

peubah acak yang saling bebas dengan fungsi kepekatan peluang yang merupakan anggota dari keluarga eksponensial. Kemudian, fungsi kepekatan tersebut diparameterisasi terhadap parameter kanonik, α , dan parameter skala, β , dimana dan diasumsikan diketahui.

Parameter kanonik akan dimodelkan dengan menggunakan model Fay-Herriot. Untuk merumuskan model regresi Poisson Bayes berhirarki dua-level dilakukan dengan cara menentukan level 1 dari model deskriptif yang menyatakan sebaran dari vektor data yang diamati,

$$y = (y_{11}, \dots, y_{1n_1}, \dots, y_{m1}, \dots, y_{mn_m})^T$$

dengan syarat pada parameter individu $\{\theta_{ij}\}$: Pada level 2 terdapat dua sebaran prior yang dipertimbangkan, yaitu untuk menyatakan sebaran gamma untuk $\{\theta_{i\cdot}\}$ dengan syarat pada hyperparameter $\alpha = (\tau, \beta)$ dan level 2 untuk menyatakan sebaran invers-gamma untuk $\{\theta_{\cdot j}\}$ dengan syarat pada hyperparameter $\alpha = (\tau, \beta)$.

Pada dasarnya akan sangat sulit untuk menghitung besaran yang sedang dikaji dalam masalah parametrik yang bersifat nonlinear, sehingga perlu dilakukan penyederhanaan pendekatan masalah komputasi yang biasa digunakan, misalnya seperti di dalam metode rantai Markov Monte Carlo. Di sini akan dibahas mengenai metode Bayes dengan sebaran prior dua-tahap yang akan menghasilkan sebaran posterior bagi dua buah *hyperparameter*. Perlu diketahui bahwa metode yang saat ini berkembang biasanya tidak memperoleh sebaran posterior bersyarat dalam bentuk persamaan tertutup yang mengakibatkan contoh Gibbs agak sulit untuk digunakan (Gelfand dan Smith, 1990). Untuk mengatasi masalah tersebut kemudian digunakan algoritma Metropolis-Hasting. Namun perlu dicatat bahwa jika sebaran bersyarat posterior tidak baku berbentuk log konkaf, maka penarikan contoh Gibb dapat digunakan melalui algoritma Gilks-Wild (Nandram, 2000). Lebih jauh, Nandram (2000) menyatakan bahwa sebaran posterior bersama bagi parameter yang diamati akan bersifat *proper* untuk sembarang model.

Dalam penelitian ini, masalah komputasi dilakukan sebagaimana yang diusulkan oleh Ghosh et al. (1998) mengenai penerapan model linear terampat pada pendugaan area kecil. Proses komputasi dilakukan pada m buah area lokal atau m strata. Misalkan Y_{ij} menyatakan statistik cukup minimal (diskrit atau kontinu) yang berhubungan dengan unit ke- j dalam strata ke- i ($j = 1, \dots, n_i; i = 1, \dots, m$). Peubah acak diasumsikan sebagai peubah acak yang saling bebas yang merupakan anggota dari keluarga eksponensial. Dalam penelitian ini peubah acak respons yang diperhatikan adalah yang menyebar

Poisson $Y_{ij} \sim \text{Poisson}(\theta_{ij})$, dimana menurut McCullagh dan Nelder (1989) fungsi peluangnya dapat dituliskan dalam bentuk keluarga eksponensial sebagai berikut:

$$f(y_{ij}|\theta_{ij}, \phi_{ij}) = \exp\left[\frac{y_{ij} \log \theta_{ij} - \theta_{ij}}{a(\phi)} - \log y_{ij}!\right] \dots (1)$$

Dalam hal ini parameter alamiah $\theta_{ij} = \log \theta_{ij}$, $a(\phi) = \phi_{ij} = 1$, $\psi(\theta_{ij}) = \theta_{ij}$, serta $\rho(y_{ij}; \phi_{ij}) = \log y_{ij}!$.

Hal ini menunjukkan bahwa sebaran Poisson merupakan anggota dari keluarga eksponensial.

Untuk memudahkan mengkaitkan bentuk model keluarga eksponensial yang dituliskan dalam Persamaan (1) ke dalam masalah pemodelan Bayes berhirarki, maka fungsi kepekatan peluangnya dapat dituliskan lagi sebagai berikut:

$$f(y_{ij}|\theta_{ij}, \phi_{ij}) = \exp\left[\phi_{ij}^{-1}(y_{ij}\theta_{ij} - \psi(\theta_{ij})) + \rho(y_{ij}; \phi_{ij})\right] \dots (2)$$

dimana $j = 1, \dots, n_i; i = 1, \dots, m$. Fungsi kepekatan yang diberikan dalam (5) diparameterisasi terhadap parameter kanonik dan parameter skala. Dalam hal ini parameter skala diasumsikan diketahui nilainya.

Parameter alamiah terlebih dahulu dimodelkan sebagai:

$$h(\theta_{ij}) = \mathbf{x}_{ij}^T \boldsymbol{\beta} + u_i + \varepsilon_{ij} \quad (j = 1, \dots, n_i; i = 1, \dots, m)$$

... (3) dimana h merupakan fungsi naik; \mathbf{x}_{ik} adalah vektor rancangan berukuran $(p - 1)$, $\boldsymbol{\beta}$ adalah vektor koefisien regresi berukuran $(p \times 1)$, u_i merupakan efek acak, dan ε_{ik} adalah galat. Di sini diasumsikan bahwa u_i dan ε_{ij} adalah saling bebas dengan $u_i \sim N(0, \sigma_u^2)$ dan $\varepsilon_{ik} \sim N(0, \sigma^2)$.

Apabila diperhatikan lebih jauh, model yang diberikan dalam Persamaan (3) merupakan model Fay-Herriot yang dijadikan sebagai model dasar dalam pendugaan area kecil. Dengan demikian dapat dikatakan bahwa model linear terampat dapat dihubungkan ke masalah pendugaan area kecil melalui hubungan antara model dalam Persamaan (2) dan (3). Apabila melihat lebih jauh persamaan yang dinyatakan dalam (2) dan (3) tidak membentuk model Bayes berhirarki. Akan tetapi model tersebut akan dibakukan sedemikian rupa sehingga mewakili suatu model dalam kerangka kerja metode Bayes sebagaimana yang telah dilakukan oleh Ghosh et al. (1998).

Model pertama yang dipertimbangkan dalam penelitian ini adalah model Poisson Bayes berhirarki, dimana parameter merupakan suatu parameter yang berkenaan dengan angka kematian (*mortality rate*) yang diasumsikan mengikuti sebaran gamma. Perlu diketahui bahwa parameter yang menyebar gamma ini merupakan level pertama dari model Bayes berhirarki dua-level, sedangkan level kedua dari hirarki ini terletak pada parameter gamma b yang bersebaran

hyperprior, dan parameter gamma b yang menyebar *hyperprior*, dimana n dan r masing-masing menunjukkan parameter dari sebaran *hyperprior* tersebut.

Model kedua yang dipertimbangkan dalam penelitian ini adalah dengan menggunakan sebaran prior untuk parameter t berdasarkan pada sebaran invers gamma atau $\theta_{ij} \sim \text{Invers Gamma}(a, b)$, dengan mengambil nilai a dan b sama seperti pada model yang pertama, yaitu dengan mengambil nilai $a = b = 0.002$. Sedangkan prior untuk hyperparameter a dan b yang masing-masing juga mengikuti sebaran invers gamma.

Data yang digunakan dalam penelitian ini merupakan data sekunder, dimana sumber data yang digunakan adalah data hasil survey yang dilakukan oleh Badan Pusat Statistika (BPS), yaitu Survey Sosial Ekonomi Nasional (Susenas) 2010 dan Suvey Demografi dan Kesehatan Indonesia (SDKI) 2007. Adapun unit pengamatan yang dikaji dalam penelitian ini adalah level kecamatan yang ada di Kabupaten Bojonegoro Provinsi Jawa Timur.

Pokok-pokok atau komponen informasi yang dapat digali dari data Susenas 2007 adalah Keterangan Tempat; Keterangan Pokok Rumah Tangga; Keterangan Anggota Rumah Tangga; Keterangan Mortalitas Sejak Tahun 2004; Keterangan Perorangan Tentang Kesehatan Balita; Pendidikan; Ketenagakerjaan; Fertilitas dan KB; Keterangan Perumahan; Pengeluaran Rumah Tangga; Keterangan Sosial Ekonomi lainnya; serta Teknologi dan Informasi. Sedangkan data SDKI khusus dirancang untuk mengumpulkan berbagai informasi mengenai tingkat kelahiran, mortalitas, prevalensi keluarga berencana dan kesehatan khususnya kesehatan reproduksi.

Tujuan umum penyelenggaraan SDKI adalah dalam rangka mengumpulkan informasi mengenai kesehatan ibu dan anak serta informasi mengenai kesehatan reproduksi, prevalensi KB, pengetahuan tentang AIDS dan IMS serta prevalensi imunisasi. Sesuai dengan jenis data atau informasi yang dikumpulkan, kuesioner yang digunakan mencakup kuesioner untuk pengumpulan data rumah tangga dan kuesioner untuk pengumpulan data perorangan.

Tahapan Penelitian

Kegiatan utama dari penelitian ini mengimplementasikan model SAE melalui pendekatan Poisson Bayes berhirarki dua-level untuk menduga angka kematian bayi dan memprediksi jumlah kematian bayi untuk level kecamatan di Kabupaten Bojonegoro Provinsi Jawa Timur. Beberapa langkah yang akan dilalui pada tahap ini adalah eksplorasi data dan diagnostik model. Pertama, kegiatan utama dari eksplorasi data ini adalah melakukan uji kecocokan sebaran

Poisson dari peubah respons yang diamati, yaitu jumlah kematian bayi, serta pemilihan peubah tambahan (*auxilliary variable*) yang diperkirakan dapat digunakan untuk menduga angka kematian bayi. Pemilihan peubah tambahan ini akan menentukan kualitas pendugaan angka kematian bayi. Kedua, kegiatan utama dari diagnostik model ini adalah diagnostik kekonvergenan rantai Markov, serta diagnostik kecocokan model dengan cara menerapkan konsep pemodelan linear terampat dalam model Bayes berhirarki pada masalah pendugaan area kecil, termasuk didalamnya adalah melakukan analisis sisaan.

Hasil dan Analisis

Target utama dari aplikasi pemodelan Bayes berhirarki dua-level dalam menangani masalah pendugaan area kecil ini adalah untuk memprediksi banyaknya bayi yang mati dan menduga angka kematian bayi level kecamatan di Kabupaten Bojonegoro Provinsi Jawa Timur. Model yang dipertimbangkan adalah model Poisson Bayes berhirarki dua-level dengan menggunakan sebaran prior gamma dan invers gamma. Kerangka kerja dari model yang dikembangkan dalam penelitian ini merupakan perpaduan konsep model linear terampat dan model Bayes berhirarki untuk menangani masalah pendugaan area kecil, dimana metode komputasi yang digunakan adalah melalui metode rantai Markov Monte Carlo. Dalam penelitian ini menggunakan dua buah peubah tambahan (*auxiliary variables*) ke dalam model,

yaitu peubah PSLN peubah yang menyatakan persentase persalinan yang ditolong bukan oleh tenaga kesehatan, atau lebih jelasnya adalah persentase ibu bersalin di suatu wilayah dalam kurun waktu tertentu, pertolongan persalinan yang bukan oleh tenaga profesional: dokter spesialis kebidanan, dokter umum, bidan, pembantu bidan, pembantu bidan dan perawat bidan; serta peubah PEND, yaitu peubah yang menyatakan nilai tengah pendidikan terakhir yang ditempuh oleh ibu (dihitung dalam tahun).

Hal pertama yang dikaji dalam aplikasi ini adalah diagnostik kekonvergenan rantai Markov. Kemudian melakukan diagnostik kecocokan model dengan cara menerapkan konsep pemodelan linear terampat dalam model Bayes berhirarki pada masalah pendugaan area kecil, termasuk didalamnya adalah melakukan analisis sisaan; serta nilai prediksi untuk sebaran posterior, dimana ukuran statistik yang diamatinya adalah nilai tengah dan simpangan baku prediksi. Terakhir dilakukan kajian terhadap sifat-sifat dari penduga parameter model Poisson Bayes berhirarki, terutama yang berkaitan dengan ketidakbiasan dan akurasi dari penduga parameter yang diamati, termasuk didalamnya adalah galatbaku dari sebaran posteriornya.

Berdasarkan hasil diagnostik kekonvergenan untuk model Poisson Bayes berhirarki dengan menggunakan sebaran prior gamma belum sepenuhnya dikatakan konvergen, terutama yang berkaitan dengan parameter PEND (dilihat dari

Tabel 1
Diagnostik Kekonvergenan Rantai Markov melalui Statistik Uji Geweke
untuk Data Level Kecamatan di Kabupaten Bojonegoro

Parameter	Prior Gamma		Prior Invers Gamma	
	z	Pr > z	z	Pr > z
Konstanta	1.0445	0.2962	1.1690	0.2424
β_1	0.1772	0.8593	0.9580	0.3381
β_2	-2.0244	0.0429	-0.0081	0.9936

Tabel 2
Kriteria Informasi Devians dan Analisis Sisaan untuk Data Level Kecamatan
di Kabupaten Bojonegoro

Kriteria	Gamma	Invers Gamma
Nilai tengah devians	142.16	124.44
Devians yang dievaluasi pada nilai tengah posterior	135.87	115.74
Kriteria Informasi Devians (KID)	139.015	128.14
Sisaan Bayes	10.4753	6.0755
MSPE	14.1339	12.7704
MAPE	2.7820	2.6374

hasil statistik uji Geweke pada Tabel 1). Apabila belum mencapai kekonvergenan, biasanya periode *burn-in* perlu diperpanjang. Akan tetapi hal ini masih menyisakan porsi yang signifikan dari ruang yang dibangkitkan oleh sebaran posterior yang secara keseluruhan belum tereksplorasi. Sedangkan hasil diagnostik untuk setiap penduga parameter yang dihasilkan dari model dengan prior invers gamma sepenuhnya sudah mencapai kekonvergenan.

Terdapat tiga buah statistik yang digunakan untuk melakukan diagnostik kecocokan model, yaitu nilai tengah devians, devians yang dievaluasi pada nilai tengah posterior, serta kriteria informasi devians (KID). Sementara itu, untuk melakukan analisis sisaan juga digunakan tiga buah ukuran, yaitu sisasan Bayes, kuadrat tengah galat prediktif (*mean square predictive error*, MSPE), serta absolut tengah galat prediktif (*mean absolute predictive error*, MAPE). Berdasarkan hasil dari diagnostik kecocokan model dan analisis sisaan yang disajikan pada Tabel 2 tampak bahwa model Poisson Bayes berhirarki dengan sebaran prior invers gamma memberikan berbagai ukuran statistik relatif lebih kecil dibandingkan dengan model dengan sebaran prior gamma. Berdasarkan hasil dari diagnostik kecocokan model dan analisis sisaan ini dapat dikatakan bahwa model Poisson Bayes berhirarki dengan prior gamma lebih cocok diaplikasikan untuk menduga angka kematian bayi level kecamatan di Kabupaten Bojonegoro dibandingkan model dengan prior invers gamma. Walaupun demikian kesimpulan ini masih bersifat relatif karena dari berbagai ukuran statistik yang dikaji menunjukkan performa yang tidak jauh berbeda.

Tabel 3 berisi tentang ringkasan statistik

untuk sebaran posterior (nilai tengah dan simpangan baku) berdasarkan sebaran prior gamma dan invers gamma untuk data level kecamatan Kabupaten Bojonegoro. Diketahui bahwa tanda dari setiap penduga parameter yang diamati untuk kedua model tersebut adalah sama. Sebagai ilustrasi, misalnya penduga untuk parameter b_1 untuk model Poisson Bayes berhirarki dengan prior gamma dan invers gamma masing-masing adalah 0.3232 dan 1.6857, dimana keduanya bertanda positif. Parameter b_1 ini berkaitan dengan peubah tambahan PSLN. Dengan demikian dapat dikatakan bahwa semakin tinggi persentase ibu yang bersalin yang ditolong oleh bukan tenaga profesional di kecamatan tersebut, maka banyaknya bayi yang mati atau angka kematian bayinya pun semakin tinggi. Sementara itu tanda dari penduga parameter b_2 yang dihasilkan dari kedua model itu juga adalah sama, yaitu sebesar -0.5109 dan -1.0153, keduanya bertanda negatif. Parameter b_2 adalah parameter yang berhubungan dengan peubah PEND. Hasil ini dapat dimaknai bahwa semakin tinggi atau semakin lama pendidikan yang ditempuh oleh seorang ibu maka angka kematian bayinya pada kecamatan tersebut juga semakin rendah.

Walaupun nilai dari penduga parameter yang dihasilkan dari model Poisson Bayes berhirarki dengan prior invers gamma lebih besar daripada model dengan prior gamma, akan tetapi keduanya memberikan nilai simpangan baku yang tidak jauh berbeda. Hasil tersebut didukung juga dari perhitungan mengenai rasio antara galat baku Monte Carlo (GBMC) dan simpangan baku posterior (SBP) untuk rantai Markov yang diperoleh dari sebaran prior invers gamma untuk setiap parameter yang diamati relatif tidak jauh berbeda

Tabel 3 Ringkasan Statistik untuk Sebaran Posterior Berdasarkan Sebaran Prior Gamma dan Invers Gamma untuk Data Level Kecamatan Kabupaten Bojonegoro

Parameter	Prior Gamma		Prior Invers Gamma	
	Nilai tengah	Simpangan Baku	Nilai tengah	Simpangan Baku
Konstanta	0.0934	32.4533	0.5477	31.9449
β_1	0.3232	32.0055	1.6857	31.7652
β_2	-0.5109	32.2932	-1.0153	31.5746

Tabel 4 Galat baku Monte Carlo dan Simpangan Baku Posterior untuk Data Level Kecamatan Kabupaten Bojonegoro

Parameter	Sebaran Prior Gamma			Sebaran Prior Invers Gamma		
	GBMC	SBP	GBMC/SBP	GBMC	SBP	GBMC/SBP
Konstanta	1.2198	32.4533	0.0376	1.1130	31.9449	0.0348
β_1	1.0764	32.0055	0.0336	1.2095	31.7652	0.0381
β_2	1.1700	32.2932	0.0362	0.9876	31.5746	0.0313

dibandingkan dengan apa yang dihasilkan dari sebaran prior gamma, sebagaimana yang ditunjukkan pada Tabel 4.

Perlu diketahui bahwa simpangan baku posterior (SBP) dan galat baku Monte Carlo (GBMC) merupakan dua konsep yang sangat berbeda. Dalam hal ini, SBP menggambarkan ketidakpastian dalam parameter, sedangkan GBMC hanya menggambarkan ketidakpastian dalam penduga parameter sebagai hasil dari simulasi rantai Markov Monte Carlo. Lebih jauh, SBP merupakan fungsi dari ukuran contoh dalam suatu gugus data, sedangkan GBMC merupakan fungsi dari banyaknya iterasi dalam proses simulasi rantai Markov Monte Carlo. Dengan demikian, maka dapat dikatakan bahwa penduga parameter yang diperoleh melalui prior gamma mempunyai derajat ketakbiasan dan akurasi yang relatif sama dengan penduga parameter yang diperoleh dari prior invers gamma pada saat diaplikasikan untuk menduga angka kematian bayi level kecamatan di Kabupaten Bojonegoro Provinsi Jawa Timur.

Apabila diperhatikan lebih jauh hasil prediksi banyaknya bayi yang mati untuk level kecamatan di Kabupaten Bojonegoro Provinsi Jawa Timur ini, model Poisson Bayes berhirarki dengan prior invers gamma cenderung memberikan hasil prediksi yang tidak jauh berbeda dengan banyaknya bayi yang mati di masing-masing kecamatan tersebut. Demikian juga halnya hasil pendugaan angka kematian bayi level kecamatan yang dihasilkan model tersebut cenderung tidak jauh berbeda dengan angka kematian bayi yang dihitung secara langsung. Sementara itu, hasil prediksi dan dugaan angka kematian bayi level kecamatan yang diperoleh dari model Poisson Bayes berhirarki dengan prior gamma cenderung lebih kecil dibandingkan dengan data aktualnya.

Sebagai gambaran angka kematian bayi untuk Kabupaten Bojonegoro pada Tahun 2008 adalah sebesar 7.36, sebagaimana yang dilaporkan dalam Laporan Profil Kesehatan Provinsi Jawa Timur Tahun 2008 (Dinas Kesehatan Provinsi Jawa Timur, 2008). Sementara itu, secara agregat hasil pendugaan angka kematian bayi level kabupaten menurut model Poisson Bayes berhirarki dengan prior invers gamma adalah sebesar 7.67, sedangkan hasil pendugaan AKB menurut model dengan prior gamma adalah sebesar 5.32. Hal ini menunjukkan bahwa model Poisson Bayes berhirarki dengan prior invers gamma memberikan dugaan pada AKB yang tidak jauh berbeda dengan apa yang dilaporkan dalam profil kesehatan Provinsi Jawa Timur tersebut, sedangkan model dengan prior gamma justru memberikan dugaan AKB yang jauh lebih kecil. Namun demikian, satu hal yang perlu dicatat adalah bahwa AKB yang dimuat dalam Laporan Profil Kesehatan tersebut belum bisa menggambarkan AKB yang sebenarnya

di populasi.

Hal lain yang perlu diperhatikan adalah adanya beberapa hasil pendugaan AKB yang sepertinya tidak 'masuk akal'. Sebagai contoh, misalnya hasil pendugaan AKB pada Kecamatan Kadewan menurut model Poisson Bayes berhirarki dengan prior gamma adalah sebesar 44.42. Padahal AKB yang dihitung dari data aktual adalah sebesar 0, karena pada kecamatan tersebut dilaporkan tidak ada bayi yang meninggal. Sementara itu, hasil pendugaan AKB menurut model Poisson Bayes berhirarki dengan prior invers gamma hanya memberikan dugaan sebesar 5.76. Hasil lain yang juga cukup menonjol adalah terjadi Kecamatan Sekar, dimana AKB menurut data aktual adalah sebesar 50.00, sedangkan menurut model dengan prior gamma dan invers gamma masing-masing adalah sebesar 26.45 dan 50.80. Sekali lagi, telah ditunjukkan bahwa model Poisson Bayes berhirarki dengan prior gamma memberikan dugaan AKB yang relatif jauh berbeda dengan AKB yang dihitung dari data aktual, sementara itu model dengan prior invers gamma relatif memberikan hasil dugaan AKB yang tidak jauh berbeda dengan data aktual.

Dalam aplikasi ini telah ditunjukkan hasil-hasil dari model Poisson Bayes berhirarki dua-level dengan menggunakan dua buah sebaran prior yang berbeda, yaitu prior gamma dan prior invers gamma untuk menduga angka kematian bayi level kecamatan di Kabupaten Bojonegoro. Secara umum performa model dugaan yang ditunjukkan oleh kedua sebaran prior tersebut adalah hampir sama. Telah ditunjukkan bahwa ringkasan statistik sebaran posterior (yang ditunjukkan melalui nilai tengah dan simpangan bakunya) yang berasal dari sebaran prior gamma maupun invers gamma memberikan hasil yang tidak jauh berbeda.

Simpulan dan Saran

Hasil-hasil kajian yang telah dilakukan dalam penelitian ini menunjukkan bahwa performa model Poisson Bayes berhirarki dua-level ini dengan menggunakan dua buah sebaran prior (sebaran prior gamma dan invers gamma) yang berbeda adalah tidak jauh berbeda. Beberapa kriteria yang digunakan untuk membandingkan kedua model tersebut adalah derajat akurasi model, derajat kekonvergenan rantai Markov, serta derajat kecocokan model. Ketepatan penggunaan model ini sangat bergantung pada berbagai kondisi, seperti ukuran area, ketersediaan peubah penjelas yang baik di tingkat area, akurasi dari penduga ragam sampling. Pengembangan dari model HB ini menjadi dimungkinkan karena ketersediaan perangkat lunak yang menyediakan fasilitas untuk melakukan simulasi MCMC, sehingga proses pendugaan model menjadi relatif lebih mudah. Permasalahan yang masih relevan untuk dibahas

adalah untuk melakukan inferensi terutama yang berkaitan dengan pemilihan spesifikasi model yang 'terbaik'.

Dalam penelitian terdapat beberapa keterbatasan untuk dikaji lebih mendalam. Salah satu diantaranya adalah kemungkinan adanya data pencilan yang mungkin berpotensi sebagai data berpengaruh. Kemungkinan yang kedua adalah munculnya masalah heteroskedastisitas dalam data. Sebagaimana yang telah diketahui bahwa dalam pemodelan linear terampat dimana peubah respons yang diamatinya adalah berbentuk data cacahan yang bersebaran Poisson, seringkali dihadapkan pada suatu masalah yang disebut dengan masalah *overdispersion*. Perlu diketahui bahwa peubah acak respons yang diamati dalam penelitian ini adalah peubah acak yang saling bebas yang bersebaran Poisson. Dalam prakteknya asumsi bahwa peubah acak ini adalah saling bebas, terutama dalam analisis data kategorik, seringkali tidak terpenuhi. Kemungkinan terjadinya kedua masalah tersebut dalam penelitian ini belum dikaji lebih mendalam. Oleh karena itu, perlu kiranya dilakukan kajian yang mendalam tentang bagaimana mengatasi masalah data pencilan dapat dilakukan dengan cara membentuk suatu model Poisson Bayes berhirarki yang bersifat *robust*.

Daftar Pustaka

- Breslow, N.E., dan Clayton, D.G. (1993) "Approximate Inference in Generalized Linear Mixed Models." *Journal of the American Statistical Association*, 88, 9-25
- Christiansen, C. L. and Morris, C. N. (1997). "Hierarchical Poisson regression modeling." *Journal of the American Statistical Association*, 92, 618-632.
- Clayton, D. and Kaldor, J. (1987). "Empirical Bayes estimates of age-standardized relative risks for use in disease mapping." *Biometrics*, 43, 671-681.
- Datta, G.S., Rao, J.N.K., and Smith, D.D. (2005). "On Measuring The Variability of Small Area Estimators Under a Basic Area Level Model." *Biometrika*, 92, 183-196.
- Fay, R.E., dan Herriot, R. (1979). "Estimates of Income for Small Places: An Application of James-Stein Procedures to Census Data." *Journal of the American Statistical Association*, 74, 269-277.
- Gelfand, A.E., dan Smith, A.F.M. (1990). "Sampling-Based Approaches to Calculating Marginal Densities." *Journal of the American Statistical Association*, 85, 398-409.
- Gelman, A. (2006). "Prior distributions for variance parameters in hierarchical models." *Bayesian Analysis*, 1, 515-533.
- Ghosh, M., Natarajan, K., Stroud, T.W.F., and Carlin, B.P. (1998). "Generalized Linear Models for Small Area Estimation." *Journal of the American Statistical Association*, 93, 273-282.
- Jiang, J. dan Lahiri, P. (2001). "Empirical Best Prediction For Small Area Inference With Binary Data." *Ann. Inst. Statist. Math*, 53, 217-243.
- McCullah, P. dan Nelder, J.A. (1989) *Generalized Linear Models*. Second Edition. London: Chapman and Hall.
- Nandram, B. (2000). "Bayesian generalized linear models for inference about small area." In: Dey, D.K., Ghosh, S.K., and Mallick, B.K. (Eds.) *Generalized Linear Models: A Bayesian Perspective*. New York: Marcel Dekker, pp. 91-114.
- Nandram, B., Sendrask, J., and Pickle, L. (1999). "Bayesian Analysis of Mortality Rates For U.S. Health Service Areas." *Sankhya: The Indian Journal of Statistics*, 61, 146-165.
- Nandram, B., Sendrask, J., and Pickle, L.W. (2000). "Bayesian Analysis and Mapping of Mortality Rates for Chronic Obstructive Pulmonary Disease." *Journal of the American Statistical Association*, 95, 1110-1118.
- Rao, J.N.K. (2003) *Small Area Estimation*. New York: Wiley.
- Souza, D.F., Moura, F.A.S., dan Migon, H.S. (2009) "Small area population prediction via hierarchical models." *Survey Methodology*, 35, 203-214.
- Torabi, M. dan Rao, J.N.K. (2008) Small area estimation under a two-level model. *Survey Methodology*, 34, 11-17.
- Trevisani M, and Torelli, N. (2007). *Hierarchical Bayesian models for small area estimation with count data*. Working Paper: Dipartimento di Scienze Economiche e Statistiche, Università degli Studi di Trieste, Trieste, Italy.
- Waller, L., Carlin, A., Xia, H., and Gelfand, A. (1997). "Hierarchical spatio-temporal mapping of disease rates." *Journal of the American Statistical Association*, 92, 607-617.
- You, Y. dan Chapman, B. (2006) "Small Area Estimation Using Area Level Models and Estimated Penarikan contoh Variances." *Survey Methodology*, 32, 97-103.
- You, Y. dan Zhou, Q.M. (2011) "Hierarchical Bayes small area estimation under a spatial model with application to health survey data." *Survey Methodology*, 37, 25-37.