

Penerapan *Small Area Estimation* dengan Metode *Empirical Bayes* dalam Menduga Risiko Relatif Penyebaran TBC di Kabupaten Karawang Tahun 2021

Siti Nur Rohmah*, Nusar Hajarisman

Prodi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Islam Bandung, Indonesia.

*rsitinur02@gmail.com, nusarhajarisman@yahoo.com

Abstract. *Small Area Estimation* (SAE) is a statistical technique used to estimate subpopulation parameters (areas) with a small sample size. In *Small Area Estimation* there are methods for processing cacahan data, namely *Empirical Bayes* and Bayes hierarchical. This study used the Bayes empirical method with the Poisson-Gamma model in estimating the risk of TB disease in Karawang Regency. The purpose of this study is to apply the Bayes empirical method based on the Poisson-Gamma model to estimate the relative risk of TB disease in Karawang Regency and compare the results of direct estimators and bayes empirical estimators through the Mean Square Error (MSE) value. The results showed that by comparing the MSE values of the direct estimator of the standardized mortality ratio and the empirical estimator of Bayes, it was concluded that the Bayes empirical method did not provide better results than direct estimators. Although Bayes empirical estimators have a fairly good accuracy rate with an average MSE value of 0.0232, the MSE value is greater when compared to the direct estimator MSE average of 0.0059.

Keywords: *Empirical Bayes, Small Area Estimation, Poisson-Gamma Model, Relative Risk.*

Abstrak. *Small Area Estimation* (SAE) atau pendugaan area kecil merupakan teknik statistika yang digunakan untuk menduga parameter subpopulasi (area) dengan ukuran sampel kecil. Dalam *Small Area Estimation* terdapat metode untuk mengolah data cacahan, yaitu *Empirical Bayes* dan hierarchical Bayes. Penelitian ini menggunakan metode *Empirical Bayes* dengan model Poisson-Gamma dalam menduga risiko penyakit TBC di Kabupaten Karawang. Tujuan penelitian ini adalah menerapkan metode *Empirical Bayes* berbasis model Poisson-Gamma untuk menduga risiko relatif penyakit TBC di Kabupaten Karawang serta membandingkan hasil penduga langsung dan penduga *Empirical Bayes* melalui nilai Mean Square Error (MSE). Hasil penelitian menunjukkan dengan membandingkan nilai MSE penduga langsung standardized mortality ratio dan penduga *Empirical Bayes* disimpulkan bahwa metode *Empirical Bayes* tidak memberikan hasil yang lebih baik dibandingkan penduga langsung. Meskipun penduga *Empirical Bayes* memiliki tingkat keakuratan yang cukup baik dengan rata-rata nilai MSE sebesar 0.0232, tetapi nilai MSE tersebut lebih besar jika dibandingkan dengan rata-rata MSE penduga langsung sebesar 0.0059.

Kata Kunci: *Empirical Bayes, Small Area Estimation, Model Poisson-Gamma, Risiko Relatif.*

A. Pendahuluan

Dalam menduga subpopulasi, saat ini sering menggunakan *Small Area Estimation* (SAE) di berbagai bidang untuk memperoleh informasi di area dengan ukuran sampel kecil misal dalam lingkup kota/kabupaten, kecamatan, kelurahan/desa atau bahkan dalam lingkup rumah tangga. Informasi ini menjadi sangat penting sejalan dengan perkembangan era otonomi daerah di Indonesia karena dapat dijadikan bahan acuan dalam perencanaan, pengawasan dan kebijakan pemerintah setempat.

Secara umum terdapat tiga pendekatan untuk mendapatkan penduga parameter dalam SAE, yaitu penduga langsung (*direct estimation*), penduga tidak langsung (*indirect estimation*) dan penduga komposit (*composite estimation*) (Yensy, 2020). Untuk penduga langsung jika diterapkan pada data dengan ukuran sampel dan area kecil maka penduga langsung akan memiliki galat yang besar. Untuk mengatasi masalah tersebut, maka digunakan penduga tak langsung. Dalam mendapatkan penduga area kecil dengan penduga tak langsung, terdapat tiga metode yang umum digunakan yaitu *Empirical Best Linear Unbiased Prediction* (EBLUP) untuk data kontinu, *Empirical Bayes* (EB) dan *Hierarchical Bayes* (HB) untuk data biner atau cacahan (Rao, 2003).

Salah satu aplikasi SAE untuk data cacahan adalah dalam pemetaan penyakit. Dalam pemetaan penyakit, ukuran sampel kecil (jumlah kasus sakit) merupakan masalah yang sering dihadapi. Sehingga penduga langsung dalam menduga risiko relatif yaitu *Standardized Mortality Ratio* (SMR) menjadi tidak reliabel (Yensy, 2020). Penduga langsung SMR memiliki asumsi bahwa banyaknya pengamatan berdistribusi Poisson. Pada distribusi Poisson asumsi yang perlu dipenuhi yaitu variabel respon harus terjadi *equidispersi*, tetapi yang sering terjadi adalah variabel respon mengalami *overdispersi*. Masalah *overdispersi* dapat diatasi dengan menggunakan pendekatan *Poisson-Gamma* (Toyyibah, 2018). Oleh karena itu, metode alternatif dalam mengatasi masalah penduga langsung tersebut adalah dengan metode *Empirical Bayes* dengan model yang sering digunakan yaitu model *Poisson-Gamma* (Yensy, 2020).

SAE metode EB dengan model *Poisson-Gamma* telah banyak digunakan dalam menduga risiko relatif suatu penyakit. Nurul Astuty Yensy (2020) menggunakan metode EB dalam menduga risiko relatif penyakit kanker paru-paru. Kismiantini (2007) membahas penduga EB dengan variabel penyerta dan tanpa variabel penyerta. Penelitian itu memberikan hasil bahwa penduga dengan variabel memberikan ketelitian yang meningkat dibandingkan penduga tanpa variabel penyerta. Sedangkan Muslimatun Toyyibah, et al (2018) menggunakan EB model *Poisson-Gamma* dalam menduga risiko relatif penyakit TBC tanpa variabel penyerta.

Tuberculosis hingga saat ini masih menjadi masalah utama kesehatan masyarakat dan menjadi isu kesehatan global di semua negara. Berdasarkan data Badan Kesehatan Dunia atau *World Health Organization* (WHO) pada tahun 2020 terdapat 10 juta orang terinfeksi TBC dan 1,5 juta orang di dunia yang meninggal akibat Tuberculosis (TBC). Hal tersebut menjadikan TBC sebagai penyakit menular mematikan kedua setelah COVID-19.

Dalam SDGs tujuan nomor 3 terkait kesehatan menekankan tujuan untuk mampu menghentikan epidemi AIDS, TBC, malaria dan penyakit tropis lainnya. Kasus TBC sendiri ditargetkan akan berkurang sebesar 80% dan 90% untuk rasio kematian akibat TBC. Target tersebut diharapkan tercapai pada tahun 2030. Untuk mencapai target tersebut perlu adanya pemetaan penyakit pada tingkat area terkecil agar mampu menangani penyebaran penyakit menjadi lebih terarah.

Oleh karena itu, pada penelitian ini dilakukan pendugaan area kecil pada risiko relatif penyebaran TBC di Kabupaten Karawang dengan metode *Empirical Bayes* model *Poisson-Gamma*. Kabupaten Karawang dipilih karena menjadi salah satu kabupaten dengan kasus TBC terbanyak di Jawa Barat. Data yang digunakan berupa data Jumlah Penderita TBC sebagai variabel respon (Y), dan untuk variabel penyerta digunakan data Jumlah penderita HIV/AIDS (X1) dan Kepadatan penduduk (X2).

Berdasarkan latar belakang yang telah diuraikan, maka perumusan masalah dalam penelitian ini sebagai berikut: “Bagaimana perbandingan hasil penduga langsung dan penduga *Empirical Bayes* dalam menduga risiko relatif penyakit TBC di Kabupaten Karawang?”. Selanjutnya tujuan dalam penelitian ini adalah untuk mengetahui diantara penduga langsung

dan penduga *Empirical Bayes*, penduga apa yang lebih baik dalam menduga risiko relatif penyakit penyebara TBC di Kabupaten Karawang

B. Metodologi Penelitian

Small Area Estimation (SAE) atau pendugaan area kecil adalah teknik statistika yang bertujuan untuk menduga parameter pada area (domain) kecil, di mana ukuran sampel kecil atau bahkan tidak ada. Metode ini memanfaatkan data dari domain yang besar untuk menduga variabel yang diminati pada domain yang lebih kecil. SAE dipilih karena metode ini dapat meminimumkan *Mean Square Error* (MSE), sehingga data yang dihasilkan dapat lebih akurat (Ilham dan Hidayat, 2019).

Pendugaan parameter pada area kecil dapat dilakukan dengan pendugaan langsung (*direct estimation*) atau pendugaan tak langsung (*indirect estimation*). Pendugaan tak langsung dalam SAE dilakukan dengan cara memanfaatkan informasi dari variabel lain yang berhubungan dengan parameter yang diamati.

Terdapat dua komponen utama dalam mengembangkan model SAE yaitu model pengaruh tetap dan model pengaruh acak. Model pengaruh tetap (*fixed effect model*) yaitu asumsi keragaman pada area kecil bagi variabel responnya dapat diterangkan oleh hubungan keragaman pada informasi tambahan yang bersesuaian. Model pengaruh acak (*random effect*) yaitu asumsi keragaman spesifik area kecil tidak dapat diterangkan oleh informasi tambahan.

Pada SAE terdapat dua jenis model yaitu model level area dan model level unit (Rao, 2003). Model level area lebih sering digunakan dibandingkan level unit. Hal ini karena data untuk sampai ke level unit atau rumah tangga jarang tersedia.

Model level area menghubungkan penduga langsung pada area kecil dengan data pendukung dari domain lain untuk setiap area. Variabel area yang diamati dinotasikan dengan y_i untuk area ke- i dimana $i = 1, 2, \dots, m$ dengan m menyatakan total area. Dengan mengasumsikan y_i berhubungan dengan variabel tambahan $x_i = (x_{1i}, x_{2i}, \dots, x_{pi})^T$ dan parameter area kecil yang diamati adalah θ_i maka model linear diberikan sebagai berikut (Rao, 2003):

$$\theta_i = x_i^T \beta + b_i v_i \quad (1)$$

dimana:

$\beta = (\beta_1, \beta_2, \dots, \beta_p)^T$ merupakan vektor koefisien regresi berukuran $p \times 1$

b_i = konstanta positif yang diketahui

v_i = efek acak area yang diasumsikan independen, identik dan berdistribusi normal dengan $E(v_i) = 0$ dan $\text{var}(v_i) = \sigma_v^2 \geq 0$

Untuk menarik kesimpulan mengenai populasi, dimisalkan nilai penduga langsung $\hat{\theta}_i$ diketahui dan

$$\hat{\theta}_i = \theta_i + e_i \quad (2)$$

di mana galat sampling $e_i \sim N(0, \psi_i)$ dengan ψ_i diketahui.

Dengan menggabungkan Persamaan (2.5) dan (2.6) maka diperoleh model level area sebagai berikut:

$$\hat{\theta}_i = x_i^T \beta + b_i v_i + e_i \quad (3)$$

Salah satu penggunaan model level area adalah untuk pemetaan penyakit. Pemetaan tingkat kematian atau kejadian suatu penyakit pada area kecil merupakan alat yang banyak digunakan dalam penelitian kesehatan masyarakat. Data yang digunakan dalam pemetaan penyakit adalah data administratif mengenai jumlah kejadian dan variabel tambahan yang berkaitan, sehingga *sampling* tidak digunakan (Rao, 2003).

Pemetaan penyakit adalah metode yang digunakan oleh ahli epidemiologi, demografi medis, dan ahli biostatistik untuk memahami distribusi geografis suatu penyakit. Peta penyakit mungkin berguna bagi instansi pemerintah untuk mengalokasikan sumber daya atau mengidentifikasi bahaya yang terkait dengan penyakit.

Pada model pemetaan penyakit suatu wilayah dibagi kedalam m area yang tidak saling tumpang tindih. Penduga langsung θ_i yang sederhana untuk menduga risiko relatif (RR) pada penyebaran penyakit adalah SMR (*standardized mortality ratio*). SMR adalah indeks dari

insiden dan kematian akibat penyakit (Lahiri dan Maiti, 2002). SMR dapat diperoleh dengan

$$\hat{\theta}_i = \frac{y_i}{e_i} \quad (4)$$

Dan

$$e_i = n_i \left(\frac{\sum_{i=1}^m y_i}{\sum_{i=1}^m n_i} \right) \quad (5)$$

dimana

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

y_i = jumlah kasus yang teramati pada area ke- i

n_i = jumlah penduduk area ke- i

Dalam pemetaan penyakit diasumsikan bahwa banyaknya kasus yang teramati yaitu $y_i | \theta_i \sim \text{Poisson}(e_i \theta_i)$. Dibawah asumsi tersebut penduga bagi θ_i yaitu $\hat{\theta}_i = \frac{y_i}{e_i}$ merupakan penduga kemungkinan maksimum yang bersifat tak bias (Kismiantini, 2007). Karena penduga bagi θ_i bersifat tak bias, maka penduga MSE yaitu $MSE(\hat{\theta}_i)$ dapat ditentukan dengan sebagai berikut

$$MSE(\hat{\theta}_i) = \frac{y_i}{e_i^2} \quad (6)$$

Metode *Empirical Bayes*

Metode Bayes empirik atau *Empirical Bayes* (EB) merupakan salah satu metode yang cocok digunakan dalam menangani data biner dan data cacahan. Metode *Empirical Bayes* mengasumsikan bahwa prior tidak diketahui, selanjutnya data digunakan untuk memperoleh dugaan prior (Hadi dan Notodiputro, 2009). Kismiantini (2007) menjelaskan bahwa salah satu cara untuk mendapatkan distribusi prior adalah dengan konjugate. Jika distribusi prior menghasilkan distribusi posterior yang sama dengan dirinya, maka distribusi prior tersebut dikatakan konjugate. Semua distribusi yang termasuk dalam keluarga eksponensial memiliki konjugate, termasuk diantaranya kemungkinan Poisson memiliki prior konjugate distribusi Gamma.

Rao (2003) menyatakan bahwa metode EB dapat diringkas sebagai berikut:

1. Memperoleh fungsi densitas posterior dari parameter area kecil yang menjadi perhatian.
2. Menduga parameter model dari fungsi densitas marginal.
3. Memanfaatkan fungsi densitas posterior dugaan dalam membuat inferensi parameter area kecil yang menjadi perhatian.

Model Poisson-Gamma

Model Poisson adalah model standar untuk data cacahan. Model ini memiliki keterbatasan dalam rata-rata dan varians ketika digunakan untuk pendugaan parameter tunggal. Umumnya, data cacahan akan mengalami overdispersi. Maka, diperkenalkan model dua tahap untuk data cacahan, yang dikenal dengan model campuran Poisson-Gamma.

Wakefield dalam Kismiantini (2007) memperkenalkan model Poisson-Gamma dengan variabel penyerta. Dalam model Poisson-Gamma dilakukan dengan dua tahap. Tahap pertama $y_i | \theta_i \sim \text{Poisson}(e_i \mu_i \theta_i)$ dengan $\mu_i = \mu(x_i, \beta)$ menyatakan model regresi. Kemudian tahap kedua $\theta_i \sim \text{Gamma}(\alpha, \alpha)$ dengan rata-rata 1 dan varian $\frac{1}{\alpha}$. Fungsi densitas dari $y_i \sim \text{Poisson}(e_i \mu_i \theta_i)$ adalah :

$$f(y_i | \theta_i) = \frac{e^{-e_i \mu_i \theta_i} (e_i \mu_i \theta_i)^{y_i}}{y_i!}, \quad y_i = 0, 1, \dots \quad (7)$$

Fungsi densitas dari $\theta_i \sim \text{Gamma}(\alpha, \alpha)$ adalah

$$\pi(\theta_i) = \frac{\alpha^\alpha}{\Gamma(\alpha)} e^{-\alpha \theta_i} \theta_i^{\alpha-1}, \quad \theta_i \quad (8)$$

Distribusi marjinal $y_i | \beta, \alpha$ adalah binomial negatif dengan rata-rata dan varian:

$$E(Y_i | \beta, \alpha) = e_i \mu_i \quad (9)$$

$$\text{var}(Y_i | \beta, \alpha) = e_i \mu_i \left(1 + \frac{e_i \mu_i}{\alpha} \right) \quad (10)$$

Dugaan parameter prior yaitu $\hat{\beta}$ dan $\hat{\alpha}$ diperoleh dari distribusi marginal $y_i|\beta, \alpha \sim$ binomial negatif menggunakan pendugaan kemungkinan maksimum yang merupakan solusi dari metode regresi binomial negatif (Kismiantini, 2007).

Selanjutnya dengan teorema mendapatkan posterior yang didefinisikan sebagai $\theta_i|y_i, \beta, \alpha \sim \text{gamma}(y_i + \alpha, e_i\mu_i + \alpha)$ dan penduga EB bagi θ_i adalah

$$\hat{\theta}_i^{EB} = \hat{\theta}_i^{EB}(\hat{\beta}, \hat{\alpha}) = \hat{\gamma}_i \hat{\theta}_i + (1 - \hat{\gamma}_i) \hat{\mu}_i \quad (11)$$

dengan $\hat{\gamma}_i = \frac{e_i \hat{\mu}_i}{(\hat{\alpha} + e_i \hat{\mu}_i)}$, $\hat{\theta}_i$ adalah penduga langsung SMR dan $\hat{\mu}_i = \exp(x_i^T \hat{\beta})$ merupakan model regresi Binomial Negatif dengan memodelkan hubungan antara variabel y_i dengan variabel penyerta yang mempengaruhinya.

Mean square error (MSE) dapat digunakan untuk mengevaluasi kelayakan suatu model statistik. Hanike (2016) menggunakan ragam posterior untuk menduga $\text{MSE}(\hat{\theta}_i^{EB})$. Pendekatan ini dapat dilakukan dengan mensubstitusikan parameter yang telah diduga yakni $\hat{\beta}$ dan $\hat{\alpha}$ ke dalam persamaan sebagai berikut.

$$\text{MSE}(\hat{\theta}_i^{EB}) = \frac{(y_i + \alpha)}{(e_i \mu_i + \alpha)^2} \quad (12)$$

C. Hasil Penelitian dan Pembahasan

Penelitian ini menggunakan *Small Area Estimation* dengan metode *Empirical Bayes* model Poisson-Gamma. Sampel yang digunakan yaitu kecamatan di Kabupaten Karawang sebanyak 30 kecamatan. Data yang digunakan merupakan data sekunder yang diperoleh dari Badan Pusat Statistik (BPS).

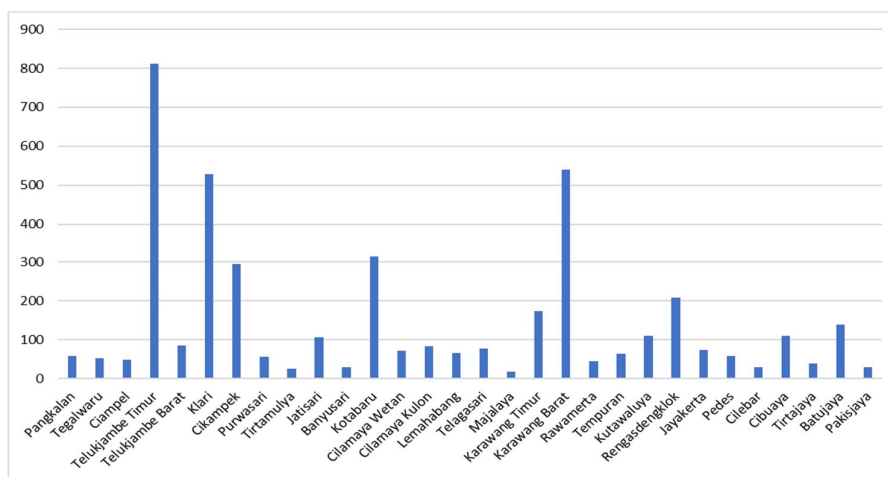
Deskripsi Data

Deskripsi mengenai kondisi penyakit TBC pada penelitian ini ditunjukkan sebagai berikut:

Tabel 1. Deskripsi Data

Statistik	TBC
Rata-rata	144.8667
Simpangan Baku	183.3791
Varians	33627.91
Minimum	18
Maksimum	812
Jumlah	4346

Sumber: Data Penelitian yang Sudah Diolah, 2022



Gambar 1. Grafik Jumlah Penderita TBC per Kecamatan

Pada Tabel 1 menampilkan deskripsi data TBC pada 30 kecamatan di Kabupaten Karawang. Berdasarkan deskripsi data dapat diketahui bahwa seluruh kecamatan terjangkit penyakit. Hal ditujukan dengan nilai minimum sebesar 18, yang artinya kecamatan memiliki penderita TBC paling sedikit berjumlah 18 orang. Kecamatan dengan penderita TBC tertinggi yaitu Kecamatan Telukjambe Timur dengan jumlah 812 orang, hal ini ditunjukkan pada Gambar 1. Berdasarkan nilai simpangan baku, variabel pengamatan TBC memiliki variansi yang lebih besar dibandingkan rata-rata banyaknya penderita TBC.

Analisis Overdispersi

Variabel respon yang diasumsikan memiliki sebaran Poisson sering melanggar asumsi equidispersi. Pengecekan ini dilakukan dengan menggunakan nilai Deviance dan Chi-square yang dibagi dengan derajat bebas. Jika Deviance/df dan Chi-square/df lebih besar dari 1, ini menunjukkan telah terjadi overdispersi. Pada Tabel 2 kedua kriteria diperoleh hasil Deviance/df dan Chi-square/df bernilai lebih dari 1. Oleh karena itu disimpulkan bahwa data Jumlah Penderita TBC di Kabupaten Karawang mengalami overdispersi.

Tabel 2. Hasil pengecekan disperse

Kriteria	df	Nilai	Nilai/df
<i>Deviance</i>	27	638.3113	23.6412
<i>Chi-square</i>	27	633.0986	23.4481

Pendugaan *Empirical Bayes* dengan Model Poisson-Gamma

Langkah awal untuk menentukan risiko relatif penyebaran TBC dengan metode *Empirical Bayes* adalah dengan menentukan penduga langsung terlebih dahulu. Penduga langsung yang digunakan adalah *standardized mortality ratio* (SMR) yang merupakan penduga sederhana untuk menduga risiko relatif (RR) dalam pemetaan penyakit. SMR diperoleh dengan

$$\hat{\theta}_i = \frac{y_i}{e_i}$$

dan

$$e_i = n_i \left(\frac{\sum_{i=1}^m y_i}{\sum_{i=1}^m n_i} \right)$$

dimana

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

y_i = jumlah kasus yang teramati pada area ke- i

n_i = jumlah penduduk area ke- i

Selanjutnya menghitung $\hat{\beta}$ dan $\hat{\alpha}$. Nilai dugaan β diperoleh dari regresi binomial negatif berupa koefisien regresi. Dan nilai dugaan α diperoleh dari $\frac{1}{\text{dispersi}}$.

Dengan menggunakan PROC GENMOD pada SAS, diperoleh parameter koefisien regresi sebagai berikut :

Tabel 3. pendugaan koefisien regresi binomial negatif

Parameter	df	Estimate	Standard Error	P-value
<i>Intercept</i>	1	-0.6604	0.1372	<.0001
X1	1	0.0231	0.0065	0.0004
X2	1	0.0001	0.0001	0.1087

Kemudian dilakukan uji signifikansi parameter model regresi Binomial Negatif dengan hipotesis sebagai berikut:

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0 ; j = 1, 2$$

Berdasarkan Tabel 3, dengan $\alpha = 5\%$ diperoleh $P\text{-value}$ untuk $x_1 = 0.0004 < \alpha$ sehingga keputusan tolak H_0 dan dapat disimpulkan bahwa nilai parameter β_1 signifikan. Sedangkan nilai $P\text{-value}$ untuk $x_2 = 0.1087 < \alpha$ sehingga keputusan terima H_0 dan dapat disimpulkan bahwa nilai parameter β_2 tidak signifikan. Oleh karena itu perlu dilakukan pemodelan ulang untuk persamaan regresi Binomial Negatif berdasarkan variabel yang signifikan.

Hasil analisis regresi Binomial Negatif berdasarkan variabel yang signifikan disajikan dalam tabel berikut ini:

Tabel 4. pendugaan parameter regresi Binomial Negatif berdasarkan variabel yang signifikan

Parameter	df	Estimate	Standard Error	P-value
Intercept	1	-0.4948	0.1005	<.0001
X1	1	0.0272	0.0064	<.0001
Dispersi	1	0.2145	0.0570	

Dari Tabel 4. diperoleh $\hat{\alpha} = \frac{1}{0.2145} = 4.662$ model regresi binomial negatif sebagai berikut:

$$\hat{\mu}_i = \exp(-0.4948 + 0.0272X1)$$

Setelah mendapatkan parameter $\hat{\beta}$ dan $\hat{\alpha}$ maka langkah selanjutnya adalah menentukan penduga *Empirical Bayes* $\hat{\theta}_i^{EB}$ dengan mensubstitusikan parameter $\hat{\beta}$ dan $\hat{\alpha}$ ke dalam persamaan persamaan berikut.

$$\hat{\theta}_i^{EB} = \hat{\theta}_i^{EB}(\hat{\beta}, \hat{\alpha}) = \hat{\gamma}_i \hat{\theta}_i + (1 - \hat{\gamma}_i) \hat{\mu}_i$$

dengan $\hat{\gamma}_i = \frac{e_i \hat{\mu}_i}{(\hat{\alpha} + e_i \hat{\mu}_i)}$, $\hat{\theta}_i = \frac{y_i}{e_i}$ adalah penduga langsung SMR dan $\hat{\mu}_i = \exp(x_i^T \hat{\beta})$ merupakan model regresi Binomial Negatif dengan memodelkan hubungan antara variabel y_i dengan variabel penyerta yang mempengaruhinya. Dan untuk mengevaluasi hasil dugaan, digunakan nilai mean square error (MSE) sebagai berikut

$$MSE(\hat{\theta}_i^{EB}) = \frac{(y_i + \alpha)}{(e_i \mu_i + \alpha)^2}$$

Hasil penduga *Empirical Bayes* beserta $MSE(\hat{\theta}_i^{EB})$ disajikan dalam Tabel 5 sebagai berikut.

Tabel 5. Hasil penduga dan MSE *Empirical Bayes*

No	Kecamatan	$\hat{\theta}_i$	$MSE(\hat{\theta}_i)$	$\hat{\theta}_i^{EB}$	$MSE(\hat{\theta}_i^{EB})$
1	Pangkalan	0.798	0.011	0.7464	0.0522
2	Tegalwaru	0.763	0.011	0.7220	0.0451
3	Ciampel	0.624	0.008	0.6268	0.0170
4	Telukjambe Timur	3.430	0.014	3.4192	0.0029
5	Telukjambe Barat	0.860	0.009	0.8122	0.0456
6	Klari	1.499	0.004	1.4703	0.0192
7	Cikampek	1.387	0.007	1.3795	0.0062
8	Purwasari	0.389	0.003	0.3899	0.0155
9	Tirtamulya	0.292	0.003	0.3072	0.0162
10	Jatisari	0.739	0.005	0.7394	0.0088
11	Banyusari	0.290	0.003	0.3034	0.0152
12	Kotabaru	1.285	0.005	1.2482	0.0271
13	Cilamaya Wetan	0.498	0.003	0.5014	0.0107
14	Cilamaya Kulon	0.707	0.006	0.6811	0.0311
15	Lemahabang	0.553	0.005	0.5391	0.0259

16	Telagasari	0.627	0.005	0.6093	0.0267
17	Majalaya	0.134	0.001	0.1555	0.0059
18	Karawang Timur	0.631	0.002	0.6243	0.0109
19	Karawang Barat	1.847	0.006	1.8276	0.0093
20	Rawamerta	0.457	0.005	0.4523	0.0247
21	Tempuran	0.553	0.005	0.5393	0.0263
22	Kutawaluya	1.018	0.009	0.9634	0.0450
23	Rengasdengklok	1.012	0.005	0.9901	0.0184
No	Kecamatan	$\hat{\theta}_i$	$MSE(\hat{\theta}_i)$	$\hat{\theta}_i^{EB}$	$MSE(\hat{\theta}_i^{EB})$
24	Jayakarta	0.617	0.005	0.6051	0.0220
25	Pedes	0.420	0.003	0.4209	0.0147
26	Cilebar	0.375	0.005	0.3795	0.0244
27	Cibuaya	1.170	0.012	1.0896	0.0603
28	Tirtajaya	0.305	0.002	0.3130	0.0142
29	Batujaya	0.975	0.007	0.9460	0.0248
30	Pakisjaya	0.400	0.006	0.4001	0.0295

Berdasarkan hasil yang dipaparkan, rata-rata dugaan risiko penyebaran TBC di Kabupaten Karawang dengan penduga langsung adalah sebesar 0.8136 dan dengan rata-rata MSE sebesar 0.0059. Sedangkan dengan penduga *Empirical Bayes* memberikan hasil rata-rata risiko relatif penyebaran TBC di Kabupaten Karawang sebesar 0.8067 dan rata-rata nilai MSE sebesar 0.0232.

D. Kesimpulan

Pendugaan area kecil dengan penduga langsung SMR dalam menduga risiko relatif penyebaran TBC di Kabupaten Karawang menghasilkan MSE yang relatif lebih kecil dari penduga *Empirical Bayes*. Berdasarkan hal tersebut menunjukkan bahwa penduga langsung memberikan hasil yang lebih baik dibandingkan penduga *Empirical Bayes*.

Acknowledge

Terima kasih kepada Bapak Dr. Nusar Hajarisman, S.Si., M.S. yang telah membimbing dalam penelitian ini.

Daftar Pustaka

- [1] Badan Pusat Statistik Kabupaten Karawang. (2022). Karawang dalam Angka 2022. Karawang: BPS Kabupaten Karawang.
- [2] Hanike, Yusrianti. (2016). Kajian Model Pendugaan Area Kecil Untuk Pendugaan Tingkat Pengangguran Menggunakan Pendekatan Bayes. Tesis tidak dipublikasikan. Program Pascasarjana, Program Studi Statistika, Institut Pertanian Bogor.
- [3] Kismiantini. (2007). Penduga Statistik Area Kecil Berbasis Model Poisson-Gamma. Tesis tidak dipublikasikan. Bogor: Program Pascasarjana, Program Studi Statistika, Institut Pertanian Bogor.
- [4] Rao JNK. 2003. *Small Area Estimation*. New York: John Wiley and Sons.
- [5] Toyyibah, Muslimatun., Komalasari, D., Fitriyani, N. (2018). *Small Area Estimation* Jumlah Penderita Penyakit TBC di Kabupaten Lombok Timur Menggunakan Metode *Empirical Bayes*. Eigen Mathematics Journal, 1(1), 35-38.
- [6] Yensy, Nurul Astuty. (2020). The *Small Area Estimation* by Using *Empirical Bayes* Method. Makalah dipresentasikan dalam International Conference on Educational Sciences and Teacher Profession, Atlantis Press.

- [7] Fitriani, Anisa Nur. (2021). *Selang Kepercayaan Koefisien Korelasi Berdasarkan Empirical Likelihood dan Penerapannya pada Data Rata-Rata Lama Sekolah dan Penduduk Miskin Kota/Kabupaten di Indonesia*, Jurnal Riset Statistika, 1(1), 51-56.