

PENDUGAAN ANGKA PUTUS SEKOLAH DI KABUPATEN SEMARANG DENGAN METODE PREDIKSI TAK BIAS LINIER TERBAIK EMPIRIK PADA MODEL PENDUGAAN AREA KECIL

Nandang Fahmi Jalaludin Malik¹, Abdul Hoyyi^{2*)}, Dwi Ispriyanti³

¹Mahasiswa Jurusan Statistika FSM UNDIP

^{2,3}Staf Pengajar Jurusan Statistika FSM UNDIP

ABSTRACT

Nowadays, small area information that has a small sample size is needed. A direct estimation in the small area will produce a large variance of values. In order of that, another alternative is needed that can be used is the indirect estimation. Small area estimation is an indirect estimation method that can be used to estimate parameters in a small area by utilizing information from outside the area, from the area itself, and from outside the survey. One of the methods that can be used is the empirical best linear unbiased prediction (EBLUP). EBLUP will be used to estimate the dropout rate for each village in the district of Semarang. Additional information used in this EBLUP method are the number of educational facilities, population, average expenditure per capita and distance from village to district. The results of EBLUP estimation showed that the lowest dropout rate village is Beji village and the highest is Pledokan village. Indirect estimation with EBLUP methods for the case of dropout rate in the district of Semarang has a coefficient variance 0,598% smaller than the coefficient variance that obtained from direct estimation.

Keywords: Dropout Rate, Small Area Estimation (SAE), Empirical Best Linear Unbiased Prediction (EBLUP)

1. PENDAHULUAN

1.1 Latar Belakang

Dalam perkembangan statistika banyak metode untuk mendapatkan data, diantaranya sensus dan survey. Sensus maupun survey berperan penting dalam proses pengambilan keputusan yang berbasis pada data. Survey sendiri sering dilakukan secara rutin baik di lembaga penelitian swasta maupun negeri. Tujuan utama dari survey adalah untuk mendapatkan informasi mengenai parameter populasi dengan mengefektifkan biaya yang tersedia. Selama ini survey rutin yang dilakukan oleh pemerintahan suatu negara hanya dirancang untuk memperoleh informasi data skala nasional. Persoalan muncul ketika dari survey seperti ini ingin diperoleh informasi untuk area yang lebih kecil, seperti informasi level provinsi, Kabupaten/kota bahkan mungkin level kecamatan dan desa/kelurahan.

Dalam konteks survey, penduga dikatakan langsung (*direct estimator*) apabila pendugaan terhadap parameter populasi pada wilayah hanya didasarkan terhadap data contoh yang diperoleh dari wilayah tersebut. Pendugaan langsung umumnya didasarkan pada teknik penarikan contohnya (*sampling technique*). Misalnya, *simple random sampling*, *stratified random sampling*, *cluster sampling*, dan sebagainya. Pendugaan secara langsung pada area kecil akan menghasilkan nilai varians yang besar jika contoh yang

diambil berasal dari survey yang dirancang untuk skala besar/nasional. Hal ini disebabkan oleh ukuran contoh yang diambil pada area tersebut kecil. Salah satu solusi yang digunakan adalah melakukan pendugaan tidak langsung dengan cara menambahkan variabel-variabel pendukung dalam menduga parameter. Variabel pendukung tersebut berupa informasi dari area lain yang serupa, survey terdahulu pada area yang sama, atau variabel lain yang berhubungan dengan variabel yang ingin diduga. Pendugaan tidak langsung tersebut dikenal sebagai pendugaan area kecil atau lebih dikenal dengan *Small Area Estimation* (SAE).

Berbagai metode pendugaan area kecil (*small area estimation*) telah dikembangkan khususnya menyangkut metode yang berbasis model (*model-based estimator*). Berkembangnya otonomi daerah di Indonesia semakin membutuhkan statistik area kecil. Setiap pemerintahan daerah memiliki wewenang lebih dalam memajukan daerahnya. Kebutuhan statistik area kecil pada level Kabupaten/kota, kecamatan ataupun desa/kelurahan sangat penting sebagai dasar pemerintah daerah untuk menyusun sistem perencanaan, pemantauan dan penilaian pembangunan daerah atau kebijakan penting lainnya.

Dari penjelasan di atas, maka penulis dalam hal ini akan mencoba mengaplikasikan metode *small area estimation* dalam menduga angka putus sekolah ditingkat Desa/Kelurahan yang ada di Kabupaten Semarang. Menurut Badan Pusat Statistik angka putus sekolah dihitung berdasarkan jumlah penduduk pada kelompok umur 7 sampai 12 tahun, 13 sampai 15 tahun dan 16 sampai 18 tahun yang tidak bersekolah lagi. Dalam mengaplikasikan metode tersebut, untuk menduga angka putus sekolah dengan berdasarkan kelompok umur diatas, penulis mengalami kesulitan dalam mendapatkan data yang akan menjadi variabel pendukungnya. Oleh karena itu, perhitungan angka putus sekolah pada tugas akhir ini tidak berdasarkan kelompok umur 7 sampai 12 tahun, 13 sampai 15 tahun dan 16 sampai 18 tahun, tetapi berdasarkan pada kelompok umur secara keseluruhan yaitu 7 sampai 18 tahun.

Parameter yang menjadi perhatian atau variabel responnya diambil dari hasil Survey Sosial Ekonomi Nasional (SUSENAS) tahun 2012, dan variabel pendukung diambil dari hasil sensus Potensi Desa (PODES) tahun 2012 yang dilaksanakan oleh Badan Pusat Statistik Provinsi Jawa Tengah. Permasalahan yang akan dibahas yaitu bagaimana mendapatkan hasil penduga tidak langsung dari angka putus sekolah dengan menggunakan salah satu metode dalam pendugaan area kecil yaitu prediksi takbias linier terbaik empirik atau biasa disebut dengan *Empirical Best Linear Unbiased Prediction* (EBLUP).

1.2 Tujuan Penelitian

Tujuan dari penulisan tugas akhir ini yaitu mendapatkan pendugaan angka putus sekolah disetiap desa/kelurahan yang ada di Kabupaten Semarang berdasarkan metode *Empirical Best Linear Unbiased Prediction* (EBLUP).

2. TINJAUAN PUSTAKA

2.1 Angka Putus Sekolah

Menurut Badan Pusat Statistik (BPS) Angka Putus Sekolah (APts) merupakan proporsi anak menurut kelompok usia sekolah yang sudah tidak bersekolah lagi atau yang tidak menamatkan suatu jenjang pendidikan tertentu. Adapun kelompok umur yang dimaksud adalah kelompok umur 7 sampai 12 tahun, 13 sampai 15 tahun dan 16 sampai 18 tahun. Hasil perhitungan angka putus sekolah ini digunakan untuk mengetahui banyaknya siswa putus sekolah di suatu jenjang pendidikan tertentu dan pada wilayah tertentu. Semakin tinggi angka putus sekolah berarti semakin banyak siswa yang putus sekolah di suatu jenjang pendidikan pada suatu wilayah.

Berdasarkan hasil Survey Ekonomi Nasional (SUSENAS) tahun 2011 diketahui bahwa sebagian besar anak berumur 7 sampai 17 tahun belum atau tidak sekolah lagi dikarenakan tidak ada biaya yaitu sebesar 56,95 persen, mengalami penurunan dibandingkan tahun 2010 sebesar 61,28 persen. Anak yang tidak bersekolah karena bekerja/mencari nafkah sebesar 3,39 persen, alasan menikah sebesar 2,37 persen, malu alasan ekonomi sebesar 0,95 persen, alasan sekolah jauh sebesar 0,81 persen, alasan tidak diterima sebesar 0,76 persen dan alasan menunggu pengumuman sebesar 0,52 persen. Masih tingginya anak berumur 7 sampai 17 tahun yang tidak bersekolah dengan alasan biaya, mencerminkan bahwa program pendidikan yang murah dan terjangkau masih belum dinikmati oleh masyarakat luas. Kondisi ini tidak sesuai dengan program pemerintah dalam penyediaan akses pendidikan yang dapat dinikmati oleh semua kalangan (BPS, 2011).

2.2 Profil Kabupaten Semarang

Luas wilayah Kabupaten Semarang 95.020,67 hektar atau sekitar 2,92 % dari luas provinsi Jawa Tengah, secara administratif terbagi menjadi 19 kecamatan terdiri dari 208 desa dan 27 kelurahan. Berdasarkan data sementara dari BPS, penduduk Kabupaten Semarang pada akhir 2011 berjumlah 938.802 jiwa dengan jumlah kepala keluarga sebanyak 284.018 KK. Berdasarkan pendidikan yang dimiliki penduduk, Kabupaten Semarang masih tergolong rendah karena hanya sekitar 2,36% yang memiliki ijazah DIV/S1 keatas, sementara persentase terbesar penduduk memiliki ijazah SD sederajat yaitu sebesar 29,36%, SMP sederajat sebesar 17,81% dan yang tidak memiliki ijazah SD sebesar 17,18%.

Desa di Kabupaten Semarang yang menjadi contoh dalam SUSENAS tahun 2011 berjumlah 72 desa dari 235 desa/kelurahan. Rata-rata setiap kecamatan diambil sampel 4 desa, kecuali kecamatan Ungaran timur sebanyak 7 Desa. Kecamatan Tenggaran dan Bringin sejumlah 5 desa. Kecamatan Banyubiru dan Pringapus sebanyak 3 Desa. Kecamatan Somowono dan Bancak sebanyak 2 Desa serta Kecamatan Pabelan yang hanya diambil 1 Desa (BPS, 2011).

2.3 Pendugaan Area Kecil

Pendugaan area kecil merupakan suatu metode yang digunakan untuk menduga parameter pada area kecil dengan memanfaatkan informasi dari luar area, dari dalam area itu sendiri, dan dari luar survey (Longford 2005).

Pendugaan parameter pada suatu *domain* dalam pendugaan area kecil dapat dilakukan dengan menggunakan pendugaan langsung dan pendugaan tidak langsung. Pendugaan langsung adalah pendugaan yang berbasis rancangan penarikan contoh (*design based*), sedangkan pendugaan tidak langsung terdiri dari pendugaan tidak langsung yang berdasar pada model implisit dan pendugaan tidak langsung yang berdasar pada model eksplisit. Kedua model tersebut menyediakan suatu *link* yang menghubungkan area-area kecil. Hal ini berarti bahwa dugaan tersebut mencakup data dari domain lain. Perbedaan dari kedua model itu adalah bahwa model implisit tidak mengikutkan pengaruh acak di dalam model, sedangkan dalam model eksplisit, terdapat pengaruh acak area yang menerangkan keragaman antar area. Menurut Rao (2003) penduga yang termasuk dalam eksplisit adalah *Empirical Best Linear Prediction* (EBLUP).

Model eksplisit disebut juga model area kecil diklasifikasi menjadi 2 jenis model (Rao, 2003), yaitu :

1. Model level area (*basic area level model*)

Merupakan model yang didasarkan pada ketersediaan data pendukung yang hanya ada untuk level area tertentu, misalkan $\mathbf{x}_i = (x_{1i}, x_{2i}, \dots, x_{pi})^T$, dan parameter yang akan diduga adalah θ_i yang diasumsikan mempunyai hubungan dengan data pendukung \mathbf{x}_i . Data pendukung tersebut digunakan untuk membangun model linear :

$$\theta_i = \mathbf{x}_i^T \boldsymbol{\beta} + b_i v_i \quad i = 1, 2, \dots, n \quad (2.1)$$

dengan $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^T$ merupakan $p \times 1$ vektor regresi dari parameter dan $v_i \stackrel{iid}{\sim} N(0, \sigma_v^2)$ sebagai pengaruh acak yang diasumsikan *independent and identically distributed* (*iid*) dan menyebar normal. Kesimpulan mengenai θ_i , dapat diketahui dengan mengasumsikan bahwa model penduga langsung y_i telah tersedia, yaitu :

$$y_i = \theta_i + e_i, \quad i = 1, 2, \dots, n \quad (2.2)$$

dan *error sampling* $e_i \stackrel{iid}{\sim} N(0, \sigma_e^2)$ dengan σ_e^2 diketahui. Kemudian kedua model tersebut digabung sehingga didapatkan model campuran :

$$y_i = \mathbf{x}_i^T \boldsymbol{\beta} + b_i v_i + e_i, \quad i = 1, 2, \dots, n \quad (2.3)$$

dengan b_i merupakan konstanta positif serta asumsi v_i dan e_i saling bebas.

Model linear campuran pada persamaan 2.3 dikenal sebagai model Fay-Heriot, dimana keragaman variabel respon di dalam area kecil diasumsikan dapat diterangkan oleh hubungan variabel respon dengan informasi tambahan (variabel pendukung) yang disebut sebagai model pengaruh tetap yaitu $\boldsymbol{\beta}$. Selain itu terdapat komponen keragaman spesifik area kecil yang tidak dapat diterangkan oleh informasi tambahan (variabel pendukung) dan disebut sebagai komponen pengaruh acak area

kecil yaitu v_i . Gabungan dari dua asumsi tersebut membentuk model pengaruh campuran atau model linier campuran.

2. Model level unit (*basic unit level model*)

Merupakan suatu model dimana data-data pendukung yang tersedia bersesuaian secara individu dengan data respon, misalkan $\mathbf{x}_{ij} = (x_{1ij}, x_{2ij}, \dots, x_{pij})^T$ sehingga didapatkan suatu model regresi :

$$y_{ij} = \mathbf{x}_{ij}^T \boldsymbol{\beta} + b_i v_i + e_{ij}, \quad j = 1, 2, \dots, n_i, \quad i = 1, 2, \dots, m$$

(2.4)

dengan asumsi $v_i \stackrel{iid}{\sim} N(0, \sigma_v^2)$ dan $e_{ij} \stackrel{iid}{\sim} N(0, \sigma_e^2)$ serta v_i saling bebas. Penelitian pada tugas akhir ini menggunakan model berbasis area karena data pendukungnya hanya ada pada level area tertentu yaitu pada level desa/kelurahan.

2.4 Metode *Empirical Best Linear Unbiased Prediction* (EBLUP)

Metode EBLUP untuk model linier campuran telah banyak digunakan dalam berbagai kasus. Estimator EBLUP dapat diperoleh dari mengestimasi nilai BLUP (*Best Linear Unbiased Predictor*) dengan menggantikan estimator yang cocok dengan varians parameternya. BLUP merupakan teknik penyelesaian model pengaruh campuran yang tidak tergantung pada kenormalan pengaruh acak (Rao, 2003).

Menurut Harville (1985) model linear campuran yang menjadi model dasar dalam pengembangan pendugaan area kecil adalah sebagai berikut :

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{v} + \mathbf{e} \quad (2.5)$$

dengan \mathbf{X} adalah matriks berukuran $n \times p$ dan \mathbf{Z} adalah matriks berukuran $n \times q$ serta \mathbf{Z} hanya mengandung intersep, sedangkan $\mathbf{v} \sim N(\mathbf{0}, \mathbf{G})$ dan $\mathbf{e} \sim N(\mathbf{0}, \mathbf{R})$. dimana $\mathbf{G} = \sigma_v^2 \mathbf{I}$ dan $\mathbf{R} = \sigma_e^2 \mathbf{I}$ serta $\text{Cov}(\mathbf{v}, \mathbf{e}) = \text{cov}(\mathbf{e}, \mathbf{v}) = \mathbf{0}$.

Dari persamaan model linear campuran persamaan 2.5 dapat diketahui bahwa sebaran bagi \mathbf{y} adalah menyebar normal dengan nilai tengah $\mathbf{X}\boldsymbol{\beta}$ dan ragam \mathbf{y} menurut Kurnia dan Notodiputro (2005) sebagai berikut:

$$\begin{aligned} \text{Var}(\mathbf{y}) &= \mathbf{V} = \text{var}(\mathbf{Z}\mathbf{v} + \mathbf{e}) = \mathbf{Z}\text{var}(\mathbf{v})\mathbf{Z}^T + \text{var}(\mathbf{e}) + \text{cov}(\mathbf{Z}\mathbf{v}, \mathbf{e}) + \text{cov}(\mathbf{e}, \mathbf{Z}\mathbf{v}) \\ &= \mathbf{Z}\mathbf{G}\mathbf{Z}^T + \mathbf{R} + \mathbf{Z}\text{cov}(\mathbf{v}, \mathbf{e}) + \text{cov}(\mathbf{e}, \mathbf{v})\mathbf{Z}^T \\ \text{Var}(\mathbf{y}) &= \mathbf{Z}\mathbf{G}\mathbf{Z}^T + \mathbf{R} \end{aligned} \quad (2.6)$$

Menurut Rao (2003) penduga BLUP berdasarkan model $y_i = \mathbf{x}_i^T \boldsymbol{\beta} + b_i v_i + e_i$ dimana y_i didapatkan dari pendugaan langsung, maka didapatkan model BLUP sebagai berikut :

$$\hat{\theta}_i^{BLUP} = \mathbf{x}_i^T \hat{\boldsymbol{\beta}} + v_i = \mathbf{x}_i^T \hat{\boldsymbol{\beta}} + \gamma(y_i - \mathbf{x}_i^T \hat{\boldsymbol{\beta}}) \quad (2.7)$$

dengan $\gamma = \left(\frac{\mathbf{G}}{\mathbf{G} + \mathbf{R}_i}\right)$, dan penduga bagi pengaruh acak (v_i) dari penduga BLUP menurut

McCulloch dan Searle (2001) diacu dalam Nissinen (2009) dihasilkan dari :

$$\hat{v}_i = E(v_i | y_i) = E(v_i) + \text{Cov}(v_i, y_i) [\text{var}(y_i)]^{-1} [y_i - E(y_i)] \quad (2.8)$$

dimana nilai harapan dan kovarian untuk menduga pengaruh acak diatas sebagai berikut :

- i. $E(\mathbf{v}) = \mathbf{0}$
- ii. $E(\mathbf{e}) = \mathbf{0}$

$$\text{iii. } E(y) = E(\mathbf{X}\beta + \mathbf{Z}v + e) = E(\mathbf{X}\beta) + E(\mathbf{Z}v) + E(e) = \mathbf{X}E(\beta) + \mathbf{Z}E(v) + E(e) \\ = \mathbf{X}\beta + \mathbf{Z}0 + 0 = \mathbf{X}\beta$$

$$\text{iv. } Cov(v, y) = \mathbf{GZ}^T$$

setelah didapatkan nilai harapan dan kovariannya, maka masukan hasil perhitungan tersebut ke persamaan berikut.

$$E(v|y) = E(v) + Cov(v, y)[var(y)]^{-1}[y - E(y)] = \mathbf{GZ}^T\mathbf{V}^{-1}[y - \mathbf{X}\beta]$$

maka penduga bagi pengaruh acak (v) adalah :

$$\hat{v} = \mathbf{GZ}^T\mathbf{V}^{-1}[y - \mathbf{X}\beta] \quad (2.9)$$

Dalam praktiknya β dan \mathbf{G} tidak diketahui sehingga perlu diduga terlebih dahulu dengan menggunakan metode *Maximum Likelihood* (ML). Sebelum menduga parameter \mathbf{G} harus diduga terlebih dahulu penduga β , menurut Kackar dan Harville diacu dalam Nissinen (2009) penduga β dihasilkan dari :

$$L(\beta, \mathbf{V}) = \frac{1}{(2\pi)^{\frac{1}{2}n}} |\mathbf{V}|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(y - \mathbf{X}\beta)^T \mathbf{V}^{-1}(y - \mathbf{X}\beta)\right)$$

dengan $c = \frac{1}{(2\pi)^{\frac{1}{2}n}}$ merupakan konstanta, maka fungsi *log-likelihood* nya:

$$l = \ln L(\beta, \mathbf{V}) = -\frac{1}{2} \ln |\mathbf{V}| - \frac{1}{2} (y - \mathbf{X}\beta)^T \mathbf{V}^{-1} (y - \mathbf{X}\beta)$$

$$\frac{dl}{d\beta} = \mathbf{X}^T \mathbf{V}^{-1} (y - \mathbf{X}\beta) = \mathbf{X}^T \mathbf{V}^{-1} y - (\mathbf{X}^T \mathbf{V}^{-1} \mathbf{X}) \beta$$

$$(\mathbf{X}^T \mathbf{V}^{-1} \mathbf{X}) \beta = \mathbf{X}^T \mathbf{V}^{-1} y$$

$$\beta = (\mathbf{X}^T \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{V}^{-1} y$$

(2.10)

dengan mensubstitusi β menjadi $\hat{\beta}$ maka penduga bagi β adalah : $\hat{\beta} = (\mathbf{X}^T \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{V}^{-1} y$ setelah penduga β didapatkan maka untuk mencari penduga \mathbf{G} harus dicari terlebih dahulu nilai dari σ_v^2 dengan melakukan iterasi sampai didapatkannya nilai yang konvergen (Rao, 2003), yaitu:

$$\sigma_v^{2(a+1)} = \sigma_v^{2(a)} + [\mathfrak{I}(\sigma_v^{2(a)})]^{-1} s(\hat{\beta}^{(a)}, \sigma_v^{2(a)})$$

dimana $\mathfrak{I}(\sigma_v^2) = \frac{1}{2} \sum_{i=1}^n \frac{b_i^4}{(\sigma_v^2 b_i^2 + R_i)^2}$ dan

$$s(\hat{\beta}^{(a)}, \sigma_v^{2(a)}) = -\frac{1}{2} \sum_{i=1}^n \frac{b_i^2}{\sigma_v^2 b_i^2 + R_i} + \frac{1}{2} \sum_{i=1}^n b_i^2 \frac{(y_i - x_i^T \hat{\beta})^2}{(\sigma_v^2 b_i^2 + R_i)^2}$$

dari hasil varians pengaruh acak (σ_v^2) tersebut, maka hasil tersebut dikalikan dengan matriks identitas untuk membentuk matriks \mathbf{G} . Setelah penduga bagi β dan \mathbf{G} didapatkan maka substitusi β oleh $\hat{\beta}$ serta \mathbf{G} oleh $\hat{\mathbf{G}}$ terhadap penduga BLUP ($\hat{\theta}_i^{BLUP}$), maka akan diperoleh penduga baru yang disebut *Empirical Best Linear Unbiased Predictor* (EBLUP).

$$\hat{\theta}_i^{EBLUP} = \hat{\theta}_i(y_i | \hat{\beta}, \hat{\mathbf{G}}) = x_i^T \hat{\beta} + \gamma(y_i - X_i^T \hat{\beta}) \quad (2.11)$$

dengan $\gamma = \left(\frac{\hat{\mathbf{G}}}{\hat{\mathbf{G}} + R_i}\right)$

3. DATA DAN METODE PENELITIAN

Data yang digunakan dalam penelitian ini merupakan data sekunder yaitu data SUSENAS 2012 dan PODES 2012 dengan informasi yang berbasis rumah tangga. Peubah respon yang menjadi perhatian dalam penelitian ini adalah angka putus sekolah pada beberapa desa di Kabupaten Semarang. Jumlah desa yang disurvei pada SUSENAS 2012 sebanyak 72 desa dari 235 desa dan 19 kecamatan di Kabupaten Semarang dengan jumlah kepala keluarga sebanyak 284.018 KK. Data Podes digunakan sebagai variabel pendukung dengan memilih beberapa variabel yang mempunyai korelasi dengan variabel angka putus sekolah, dan akhirnya diperoleh 4 (empat) variabel yang akan digunakan, yaitu :

X_1 = jumlah sarana pendidikan

X_5 = jumlah penduduk

X_6 = rata-rata jarak dari desa ke kabupaten (km),

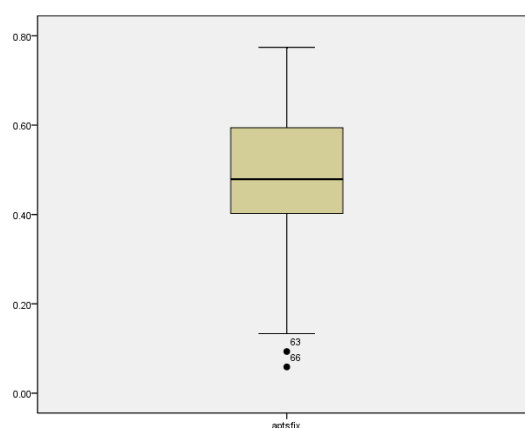
X_8 = rata-rata pengeluaran perkapita (unit)

Semua variabel pendukung di atas akan digunakan dalam pendugaan dengan metode EBLUP, penduga EBLUP ini dapat dihasilkan setelah dilakukan pendugaan terhadap komponen varians pengaruh acak (*efek random*) dengan menggunakan metode *Maximum Likelihood* (ML) dan pendugaan koefisien regresi β dengan metode *Weighted Least Square* (WLS).

Software yang digunakan dalam penelitian ini adalah *microsoft office excel 2010*, *microsoft office excel 2010* digunakan untuk mencari penduga langsung angka putus sekolah disetiap desadan *softwarestatistika* Minitab, SPSS untuk mencari statistik deskriptif dari data serta kenormalan data sedangkan *softwarestatistika* SAS digunakan untuk mencari pendugaan terhadap komponen varians pengaruh acak (*efek random*) dan penduga EBLUP.

Tabel 4.1 Nilai Statistik Angka Putus Sekolah Hasil Pendugaan Langsung

Statistik	Angka Putus Sekolah
Mean	0,478
Std. Deviasi	0,150
Koefisien Variansi	31,475
Minimum	0,059
Median	0,479
Maximum	0,774



Gambar 4.1 Diagram *Boxplot* Angka Putus Sekolah Hasil Pendugaan Langsung

4. HASIL DAN PEMBAHASAN

4.1 Pendugaan Langsung Angka Putus Sekolah (Apts)

Jumlah desa yang disurvei pada SUSENAS 2012 sebanyak 72 desa dengan rata-rata ukuran sampel yang diambil dari setiap desa/kelurahan sekitar 40 responden. Jumlah

sampel tersebut untuk masing-masing desa/kelurahan sangat kecil jika dibandingkan dengan jumlah penduduk di masing-masing desa/kelurahan. Pendugaan langsung angka putus sekolah (APts) untuk masing-masing desa diperoleh dengan membagi jumlah anak yang sudah tidak bersekolah lagi atau yang tidak menamatkan suatu jenjang pendidikan tertentu dengan jumlah murid pada tingkat pendidikan tertentu yang tersurvey dalam survey sosial ekonomi nasional (SUSENAS) tahun 2012. Hasil pendugaan langsung menunjukkan bahwa rata-rata angka putus sekolah desa-desa di Kabupaten Semarang sebesar 0,478.

Berdasarkan diagram *boxplot* yang diperlihatkan pada gambar 4.1 menunjukkan bahwa terdapat 2 (dua) desa/kelurahan yang menjadi pencilan, kedua desa/kelurahan tersebut adalah desa Ungaran dan desa Beji. Kedua desa tersebut memiliki angka putus sekolah yang lebih kecil dibandingkan yang lainnya. Salah satu alasan yang menjadikan desa ini memiliki angka putus sekolah yang lebih rendah dibandingkan dengan yang lainnya adalah letak desa tersebut yang berdekatan dengan wilayah perkotaan serta desa Ungaran terletak di kecamatan Ungaran Barat dan desa Beji terletak di kecamatan Ungaran timur yang merupakan ibukota Kabupaten Semarang, tepat berbatasan dengan Kota Semarang.

Pola angka putus sekolah untuk masing-masing desa di Kabupaten Semarang pada *boxplot* lebih lebar pada bagian atas. Hal ini menunjukkan bahwa persebaran angka putus sekolah setiap desa di Kabupaten Semarang lebih banyak berada diatas rata-rata nilai angka putus sekolah Kabupaten Semarang. Pola persebaran angka putus sekolah membentuk pola distribusi normal. Hal ini juga berdasarkan pada pengujian *Kolmogorov-smirnov* yang menunjukkan nilai Sig. (2-tailed) = 0,909 $> \alpha = 0,05$ ini berarti bahwa sudah cukup bukti bahwa data berdistribusi normal.

4.2 Pendugaan Tidak Langsung

4.2.1 Pemilihan Variabel Pendukung

Variabel awal yang dijadikan sebagai variabel pendukung ada sebanyak 8 variabel yang diambil dari data potensi desa (PODES) yang diambil pada tahun yang sama dengan data SUSENAS yaitu tahun 2012. Variabel pendukung tersebut akan diuji korelasinya terhadap variabel respon dengan menggunakan diagram pencar dan nilai korelasi *pearson*. Langkah pertama yang dilakukan untuk memilih variabel-variabel pendukung yang diasumsikan mempengaruhi angka putus sekolah dengan melihat diagram pencar dari masing-masing variabel pendukung dengan angka putus sekolah.

Berdasarkan diagram pencar variabel pendukung yang berkorelasi hanya 5 variabel yaitu jumlah sarana pendidikan, jumlah penduduk, jarak dari desa ke Kabupaten, jumlah gizi buruk dan rata-rata pengeluaran perkapita. Untuk lebih meyakinkan ada tidaknya korelasi dilakukan langkah selanjutnya yaitu uji formal dengan menghitung nilai korelasi *pearson*. Berdasarkan nilai korelasi *pearson* dari 8 variabel pendukung, 4 variabel yang memiliki korelasi dengan angka putus sekolah. Berdasarkan hasil yang ditunjukkan oleh diagram pencar dan nilai korelasi *Pearson* maka keempat variabel-variabel tersebut dapat

digunakan untuk menggambarkan angka putus sekolah pada beberapa desa di Kabupaten Semarang.

4.2.2 Pendugaan dengan Metode Empirical Best Linear Unbiased Prediction (EBLUP)

Sebelum melakukan pendugaan terhadap angka putus sekolah untuk masing-masing desa/kelurahan dengan metode EBLUP, terlebih dahulu dilakukan pendugaan terhadap koefisien regresi (β), pengaruh acak (v_i) dan varian dari pengaruh acak (σ_v^2). Hasil pendugaan tersebut kemudian digunakan untuk menduga angka putus sekolah desa/kelurahan di Kabupaten Semarang. Pendugaan β dan σ_v^2 dihasilkan dari pendugaan pada prosedur *proc mixed* dalam SAS. Nilai σ_v^2 adalah 0,01803 yang digunakan untuk membentuk matrik **G**. Pendugaan ini menggunakan metode *Maksimum Likelihood* (ML). Sedangkan nilai parameter β yang didapatkan adalah sebagai berikut.

Tabel 4.3 Nilai Dugaan Parameter Beta

X_i	Beta duga
X_0	0,5565
X_1	-0,00621
X_5	0,000003448
X_6	0,001857
X_8	-0,000000168

Tabel 4.4 Nilai Statistik Angka Putus Sekolah Hasil Pendugaan EBLUP

Statistik	Angka Putus Sekolah
Mean	0.478
Std. Deviasi	0.147
Koefisien Variansi	30.877
Minimum	0.067
Median	0.480
Maximum	0.768

Setelah Pendugaan β dan **G** dihasilkan maka langkah selanjutnya adalah membentuk model *Small Area Estimation* (SAE) menggunakan metode *Empirical Best Linear Unbiased Prediction* (EBLUP).

$$\hat{\theta}_i^{EBLUP} = 0,5565 + (-0,00621X_1) + (0,000003448X_5) + (0,001857X_6) + (-0,000000168X_8) + \{0,01803 / (0,01803 + R_i)\} \{y_i - (0,5565 + (-0,00621X_1) + (0,000003448X_5) + (0,001857X_6) + (-0,000000168X_8))\}$$

indek i melambangkan desa-desa yang digunakan untuk membangun model dan R_i merupakan *error sampling* yaitu sebesar 0,015538.

Setelah didapatkan model EBLUP, maka angka putus sekolah untuk masing-masing desa/kelurahan di Kabupaten Semarang dapat diduga. Rata-rata angka putus sekolah desa/kelurahan di Kabupaten Semarang tahun 2012 hasil pendugaan tidak langsung dengan metode EBLUP sebesar 0,478. Berdasarkan nilai koefisien variannya hasil dari pendugaan tidak langsung memiliki nilai koefisien variansi yang lebih kecil dibandingkan dengan penduga langsung yaitu sebesar 30,877%. Nilai duga angka putus

sekolah yang terbesar sebesar 0.768 yaitu di desa Pledokan. Sedangkan nilai duga angka putus sekolah terkecil ada di desa Beji sebesar 0.067.

Secara umum pendugaan angka putus sekolah pada area kecil dengan menggunakan metode EBLUP menghasilkan dugaan yang tidak berbeda jauh dengan penduga langsung. Selisih yang kecil ini disebabkan karena korelasi antara variabel pendukung dengan variabel respon yang rendah sehingga mempengaruhi dari nilai dugaan tidak langsung dan dapat dilihat juga dari nilai dugaan parameter β yang mendekati nol.

5. KESIMPULAN

Kesimpulan yang dapat diambil sebagai berikut :

1. Variabel pendukung yang diasumsikan mempengaruhi angka putus sekolah berdasarkan nilai korelasi *pearson* dan diagram pencar adalah jumlah sarana pendidikan, jumlah penduduk, rata-rata jarak dari desa ke kabupaten dan rata-rata pengeluaran perkapita.
2. Model *Small Area Estimation* (SAE) menggunakan metode *Empirical Best Linear Unbiased Prediction* (EBLUP) adalah

$$\hat{\theta}_i^{EBLUP} = 0,5565 + (-0,00621X_1) + (0,000003448X_5) + (0,001857X_6) + (-0,000000168X_8) + \{0,01803/(0,01803+R_i)\} \{y_i - (0,5565 + (-0,00621X_1) + (0,000003448X_5) + (0,001857X_6) + (-0,000000168X_8))\}$$

5.1 Saran

Adapun saran yang dapat disampaikan yaitu diperlukan kajian lebih lanjut dalam menyelesaikan masalah pendugaan pada area kecil dengan berbagai metode *Small Area Estimation* (SAE) yang ada. Supaya pendugaan area kecil bisa lebih berkembang.

Daftar Pustaka

- Kurnia, A. 2009. Prediksi terbaik empirik untuk model transformasi logaritma di dalam pendugaan area kecil dengan penerapan pada data susenas [disertasi]. Bogor: Bogor: Program Pascasarjana, Institut Pertanian Bogor.
- Kurnia, A., Notodiputro, K.A. 2005. *Pendekatan General Linear Mixed Model pada Small Area Estimation*. Forum Statistika dan Komputasi, Oktober 2005 p:12-16.
- Kurnia, A., Notodiputro, K.A. 2006. *EB-EBLUP MSE Estimator on Small Area Estimation with Application to BPS Data*. Paper presented International Conference on Mathematical Science. Bandung 19-21 Juni 2006.
- Kurnia, A., Notodiputro, K.A. 2007. *Generalized Additive Mixed Models for Small Area Estimation*. Proceeding at the 2nd International Conference on Mathematical Science 2007, 28-29 May 2007. University Teknologi Malaysia p:1-3.
- Nissinen, K. 2009. *Small Area Estimation with Linear Mixed Models from Unit-Level Panel and Rotating Panel Data*. University Printing House Jyvaskyla. University of Jyvaskyla Finland.
- Rao, J.N.K. 2003. *Small Area Estimation*. Jhon Willey and sons, Inc. New York.
- Sadik, K.A., Notodiputro, K.A. 2006. *Metode E-BLUP dalam Small Area Estimation untuk Model yang Mengandung Random Walk*. Forum Statistika dan Komputasi, Oktober 2006 p:37-41.