

SMALL AREA ESTIMATION UNTUK PENDUGAAN JUMLAH PENDUDUK MISKIN DI KOTA SEMARANG DENGAN PENDEKATAN KERNEL-BOOTSTRAP

¹Ujang Maulana, ²Moh Yamin Darsyah, ³Tiani Wahyu Utami

^{1,2,3}Program Studi Statistika FMIPA Universitas Muhammadiyah Semarang

Email: ujang.indigo@yahoo.com

ABSTRAK

Pendugaan area kecil dengan teknik pendugaan tak langsung memerlukan asumsi adanya hubungan linier antara rata-rata area kecil dengan variabel penyerta. Jika tidak ada hubungan linier antara rata-rata area kecil dan variabel penyerta maka tidak tepat 'meminjam kekuatan' dari area lain dengan menggunakan model linier dalam pendugaan tak langsung. Untuk mengatasi hal tersebut dikembangkan pendekatan nonparametrik. Salah satu pendekatan nonparametrik yang digunakan adalah pendekatan Kernel-Bootstrap. Pendugaan tak langsung dengan pendekatan SAE Kernel-Bootstrap digunakan untuk menduga angka jumlah penduduk miskin pada level kecamatan di Kota Semarang. Variabel yang digunakan dalam penelitian ini terdiri atas dua bagian, yakni Jumlah Penduduk Miskin (Y) sebagai variabel dependen, serta sebagai variabel penyerta: Penduduk Usia 65 Tahun keatas (X). Evaluasi hasil pendugaan dilakukan dengan melihat nilai MSE (Mean Square Error) penduga SAE Kernel-Bootstrap. Hasil pendugaan SAE Kernel-Bootstrap terbaik menggunakan replikasi $B=100$.

Kata Kunci : SAE, Kernel, Bootstrap, MSE

PENDAHULUAN

Otonomi daerah menyebabkan adanya pergeseran ketatanegaraan di Indonesia dari sentralisasi ke desentralisasi, dimana pemerintah daerah lebih leluasa dalam mengatur daerahnya sendiri. Sistem pemerintahan tersebut membuat statistik area kecil (*small area statistics*) pada saat ini sangat diminati dalam berbagai bidang dan sangat dibutuhkan untuk mendapatkan informasi-informasi pada area kecil, seperti pada lingkup kecamatan, ataupun desa/kelurahan. Sampel yang relatif sedikit menjadi kendala utama pada suatu penelitian, salah satu upaya yang dilakukan untuk mengatasi masalah tersebut dengan menambah jumlah sampel. Upaya lain yang bisa dilakukan adalah mengoptimalkan data yang tersedia

dengan metode *small area estimation* (SAE). Metode SAE merupakan suatu teknik statistika untuk menduga parameter-parameter subpopulasi dengan ukuran sampel kecil. Pendugaan langsung tidak mampu memberikan ketelitian yang cukup bila ukuran sampel dalam area kecil yang menjadi perhatian sedikit/ berukuran kecil, sehingga statistik yang dihasilkan akan memiliki varian yang besar atau bahkan pendugaan tidak dapat dilakukan karena tidak terwakili dalam survey [8].

Meminjam kekuatan dari area lain dengan menggunakan model linier dalam pendugaan tak langsung tidak tepat digunakan jika tidak terdapat hubungan linier antara rata-rata area kecil dan variabel penyerta. Meminjam kekuatan disini maksudnya meminjam informasi yang akan digunakan sebagai variabel

penyerta. Untuk mengatasi hal tersebut dikembangkan pendekatan nonparametrik. Salah satu pendekatan nonparametrik yang digunakan adalah pendekatan *Kernel-Based*.

Pendekatan dengan menggunakan fungsi kernel diusulkan karena fungsi ini didasarkan pada pendekatan penggunaan ketersediaan variabel-variabel umum antara sensus dan survei sehingga sesuai dengan pendugaan area kecil yang mengestimasi fungsi regresi berdasarkan informasi survei. Pendekatan *Kernel-Based* menawarkan teknik nonparametrik sebagai alternatif baru yang menjanjikan identifikasi fungsi regresi pada pendugaan area kecil. Metode ini lebih fleksibel dibanding dengan metode-metode sebelumnya yang menggabungkan pola-pola kovarian spasial untuk pendugaan area kecil. *Kernel-Based* juga memberikan prosedur yang fleksibel dalam pendugaan area kecil. Berdasarkan hal tersebut maka penulis tertarik untuk mempelajari *Kernel-Based* sebagai pendekatan nonparametrik dua tahap berdasarkan *Kernel Nadaraya-Watson* dalam metode SAE.

Beberapa penelitian yang berhubungan dengan *small area estimation* dengan pendekatan nonparametrik telah banyak dilakukan antara lain [1] melakukan pendugaan SAE dengan metode *kernel-bootstrap*, Darsyah dan Wasono (2013) dengan metode *nonparametric* dalam pendugaan IPM dan tingkat kemiskinan pada area kecil, [7] menggunakan *penalized spline*, [6] dengan pendekatan *two stage non-parametric*. Aplikasi Pendugaan pengeluaran perkapita pada level kecamatan di Kabupaten Sumenep dengan metode *nonparametric* [2].

Pada penelitian ini akan digunakan SAE berdasarkan pendekatan nonparametrik dengan fungsi *kernel-bootstrap* untuk mengestimasi jumlah penduduk miskin perkecamatan di Kota

Semarang. Metode SAE Kernel merupakan metode pendugaan parameter pada area kecil yang didasarkan pada fungsi Kernel dimana nilai yang diperoleh berdasar pada estimasi densitas kernel dari variabel yang diamati.

Ada dua konsep pokok yang digunakan untuk mengembangkan model pendugaan parameter small area, yaitu:

1. Model pengaruh tetap (fixed effect model) dimana asumsi bahwa keragaman di dalam small area variabel respon dapat diterangkan seluruhnya oleh hubungan keragaman yang bersesuaian pada informasi tambahan.
2. Model pengaruh acak (random effect) dimana asumsi keragaman spesifik small area tidak dapat diterangkan oleh informasi tambahan.

Gabungan antara kedua model tersebut membentuk model campuran (mixed model). Karena variabel respon diasumsikan berdistribusi normal maka pendugaan area kecil yang dikembangkan merupakan bentuk khusus dari *General Linear Mixed Model* (GLMM).

Model small area biasanya menggunakan model linier campuran dalam bentuk

$$y = X\beta + Zu + e \quad (1)$$

Dengan metode *Bootstrap* tidak perlu melakukan asumsi distribusi dan asumsi-asumsi awal untuk menduga bentuk distribusi dan pengujian-pengujian statistiknya [5].

Penjelasan bahwa untuk mengurangi bias yang relatif besar dari statistik area kecil dengan peubah penjelasnya dan untuk mendapatkan penduga MSE yang lebih baik dapat dirumuskan sebagai berikut [6]:

$$y_i = \mu_i + e_i \quad (2)$$

$$\mu_i = m(x_i) + u_i \quad (3)$$

dimana $i=1,2,\dots,n$ menunjukkan jumlah area kecil. Fungsi $m(\cdot)$ adalah

fungsi mulus (*smoothing function*) yang mendefinisikan hubungan antara x dan y . y_i adalah rata-rata area kecil yang tidak teramati, y_i adalah penduga langsung dari rata-rata area kecil, u_i galat peubah acak yang berdistribusi independen dan identik dengan $E(u_i) = 0$ dan $\text{var}(u_i) = \tau_u^2$, dan e_i berdistribusi independen dan identik dengan $E(e_i) = 0$ dan $\text{var}(e_i) = D_i$, dengan asumsi D_i diketahui. Persamaan 2.9 dan 2.10 disubstitusikan maka akan menghasilkan persamaan berikut:

$$y = m(x_i) + u + e \tag{4}$$

Penduga MSE dengan *bootstrap* diberikan oleh:

$$MSE^*(\hat{y}_i) = \frac{1}{J} \sum_{j=1}^B \left(\hat{y}_i^{*(j)} - y_i^{*(j)} \right)^2 \tag{5}$$

Cahyono (1998) menyatakan salah satu faktor yang mempengaruhi pendapatan adalah faktor umur. Umur produktif berkisar antara 15-64 tahun yang merupakan umur ideal bagi para pekerja.

METODE PENELITIAN

Sumber Data dan Variabel Penelitian

Data yang digunakan pada penelitian ini adalah data sekunder yang didapatkan dari Bappeda kota Semarang, dan Badan Pusat Statistik (BPS) kota Semarang. Data-data ini mencakup angka jumlah penduduk miskin dan variabel yang mempengaruhinya di Kota Semarang yang mencakup 16 kecamatan. Variabel yang digunakan dalam penelitian ini adalah 2 variabel yang terdiri dari 1 variabel dependen dan 1 variabel penyerta dengan rincian sebagai berikut:

1. Variabel dependen (Y) yaitu angka jumlah penduduk miskin tiap kecamatan di Kota Semarang.

2. Variabel penyerta (X) yaitu penduduk usia 65 tahun keatas per kecamatan di Kota Semarang. Berdasarkan variabel-variabel yang ada.

Metode Analisis

Adapun langkah analisis yang dilakukan dalam penelitian ini adalah:

1. Dengan menggunakan data variabel prediktor (x_i) dan variabel respon (y_i), hitung

$$\hat{m}_h(x) = \frac{1}{m} \sum_{i=1}^m w_{hi}(x) y_i$$

2. Hitung

$$\hat{\tau}_u^2 = \max\left\{0, \frac{1}{m-1} \sum_{i=1}^m W_{hi}(x) [y_i - \hat{m}_h(x)]^2 - 1\right\}$$

3. Substitusikan

$$\hat{y}_i = \hat{x}_i y_i + (1 - \hat{x}_i) \hat{m}_h(x_i) \quad \text{dengan}$$

$$\hat{x}_i = \frac{\hat{\tau}_u^2}{\hat{\tau}_u^2 + D_i}$$

4. Menghitung $MSE(\hat{y}_i)$ dilakukan dengan pendekatan *bootstrap*

$$MSE^*(\hat{y}_i) = \frac{1}{J} \sum_{j=1}^B \left(\hat{y}_i^{*(j)} - y_i^{*(j)} \right)^2$$

Dalam penelitian ini dengan pendekatan *bootstrap*, adapun langkah-langkahnya adalah sebagai berikut:

- a. Sampel pada data x didefinisikan sebagai data sampel berukuran n yang terdiri dari $x_i = x_1, x_2, x_3, \dots, x_n$ dengan x_i sebagai vektor data pengamatan.
- b. Sampel pada data x diambil secara acak dengan pengembalian sebanyak n kali. Kemudian diperoleh data sampel baru yang didefinisikan sebagai x^* . Sampel data x^* terdiri dari anggota data asli, akan tetapi mungkin sebagian data tidak akan muncul, atau muncul hanya sekali atau dua kali

semuanya tergantung kepada randomisasinya.

- c. Langkah ke 2 dilakukan beberapa kali sebanyak B, sehingga mendapatkan himpunan data bootstrap $(x^{*1}, x^{*2}, \dots, x^{*B})$ dimana setiap sampel bootstrap merupakan sampel acak yang saling independen
- d. Replikasi bootstrap $(x^{*1}, x^{*2}, \dots, x^{*B})$ didapatkan dengan cara menghitung nilai $s(x)$ pada masing-masing sampel bootstrap. Nilai $s(x)$ merupakan suatu nilai taksiran statistik dari masing-masing sampel bootstrap. Rataan dari suatu sampel bootstrap didefinisikan sebagai $\bar{s}(x)$, dimana $\bar{s}(x)$ adalah rata-ran dari $(x^{*1}, x^{*2}, \dots, x^{*B})$.

5. Mendapatkan hasil pendugaan dan mendeskripsikannya.

Software yang digunakan dalam penelitian ini adalah *Minitab 14*, *R Program*, dan *Microsoft Excel 2007*.

HASIL PENELITIAN

Pendugaan SAE dengan pendekatan kernel dengan variabel yang digunakan untuk menduga jumlah penduduk miskin di Kota Semarang adalah penduduk usia 65 tahun keatas. Maka langkah selanjutnya adalah menduga $\hat{m}_h(x_i)$ dengan menggunakan pendugaan kernel Nadaraya-Watson sebagai kernel dua tahap yaitu $\hat{m}_h(x_i) = \frac{\sum_i K_h(x-x_i)y_i}{\sum_i K_h(x-x_i)}$, dimana

$K_h(\cdot)$ adalah fungsi kernel dengan bandwidth h dan $\hat{m}_h(x) = \frac{1}{h} \sum_{i=1}^n K(\frac{x}{h})$.

Fungsi kernel yang dipakai dalam penelitian ini adalah fungsi epanechnikov. Penduga kernel di atas linier terhadap y_i , dan dapat ditulis

$$\hat{m}_h(x_i) = \frac{1}{m} \sum_{i=1}^m w_{hi}(x) y_i, \text{ dimana}$$

$$w_{hi}(x) = \frac{K_h(x-x_i)}{\frac{1}{m} \sum_i K_h(x-x_i)}. \quad \text{Hasil}$$

pendugaan $\hat{m}_h(x_i)$ dapat dilihat pada Tabel 1.

Tabel 1. Hasil Pendugaan Kernel Penduduk Miskin (x 10.000 jiwa)

No.	Kecamatan	$\hat{m}_h(x_i)$
1	Semarang Tengah	2,331315
2	Semarang Utara	2,331366
3	Semarang Timur	2,331274
4	Gayamsari	2,331273
5	Genuk	2,331273
6	Pedurungan	2,331275
7	Semarang Selatan	2,331277
8	Candisari	2,331273
9	Gajahmungkur	2,331310
10	Tembalang	2,331308
11	Banyumanik	2,331337
12	Gunungpati	2,331281
13	Semarang Barat	2,331351
14	Ngaliyan	2,331283
15	Mijen	2,331342
16	Tugu	2,331421

Setelah diketahui pendugaan kernel untuk setiap area, maka dilakukan pendugaan pembobot pengaruh acak untuk setiap area $\hat{\chi}_i = \frac{\hat{\tau}_u^2}{\hat{\tau}_u^2 + D_i}$, dimana

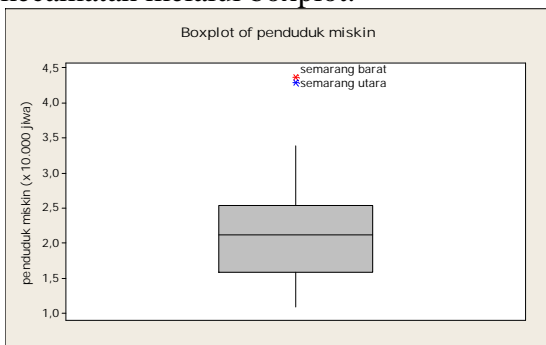
$\hat{\tau}_u^2$ merupakan penduga varian antar area dan D_i merupakan varian tiap area. Maka penduga untuk rata-ran area kecil $\hat{m}_i = \hat{\chi}_i y_i + (1 - \hat{\chi}_i) \hat{m}_h(x_i)$. Hasil pendugaan area secara lengkap dapat dilihat pada Tabel 2.

Tabel 2. Perbandingan Penduduk Miskin dan Hasil Pendugaan Penduduk Miskin SAE-Kernel (x 10.000 jiwa)

No	Kecamatan	Data asli	SAE-kernel
1	Semarang Tengah	1,6613	1,7140
2	Semarang Utara	4,2907	3,9849
3	Semarang Timur	2,1587	2,1592
4	Gayamsari	2,2202	2,2203
5	Genuk	2,4541	2,4539
6	Pedurungan	2,5695	2,5682
7	Semarang Selatan	2,0403	2,0430
8	Candisari	2,4430	2,4428
9	Gajahmungkur	1,5561	1,6111
10	Tembalang	3,3901	3,3188
11	Banyumanik	1,5079	1,6022
12	Gunungpati	1,9872	1,9933
13	Semarang Barat	4,3637	4,0901
14	Ngaliyan	2,0834	2,0887
15	Mijen	1,4783	1,5823
16	Tugu	1,0933	1,3733

Berdasarkan Tabel 2 menunjukkan bahwa terdapat perbedaan yang cukup mencolok, dimana pada jumlah penduduk miskin paling sedikit berada di Kecamatan Tugu sebesar 10.933 jiwa, sedangkan hasil pendugaan penduduk miskin SAE-Kernel pada kecamatan yang sama sebesar 13.733 jiwa. Penduduk miskin terbanyak berada pada Kecamatan Semarang Barat sebesar 43.637 jiwa, sedangkan untuk pendugaan SAE-Kernel pada kecamatan yang sama sebesar 40.901 jiwa.

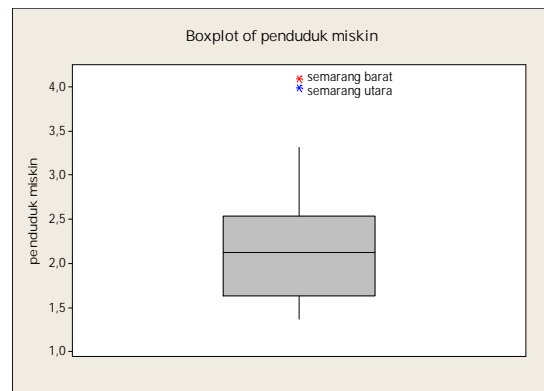
Selanjutnya kita akan membandingkan antara jumlah penduduk miskin data asli dengan pendugaan SAE-Kernel salah satunya dengan melihat koefisien keragaman penduduk miskin antar kecamatan melalui boxplot.



Gambar 1. Boxplot Penduduk Miskin

Gambar 1 menunjukkan bahwa terdapat dua kecamatan yang menjadi pencilan yaitu Kecamatan Semarang Barat dan Kecamatan Semarang Utara. Kedua kecamatan tersebut memiliki penduduk miskin yang besar dengan selisih cukup jauh dengan penduduk miskin kecamatan lain di Kota Semarang. Hal ini dapat dipahami karena mayoritas pensiunan bertempat tinggal di kedua kecamatan tersebut, sehingga menyebabkan kedua kecamatan tersebut mempunyai penduduk tidak produktif yang cukup banyak bila dibandingkan kecamatan lain sehingga secara tidak langsung mempengaruhi banyaknya penduduk miskin di kedua wilayah tersebut.

Pola penduduk miskin pada tiap kecamatan di Kota Semarang pada boxplot lebih lebar pada bagian bawah. Hal ini menunjukkan bahwa persebaran penduduk miskin setiap kecamatan di Kota Semarang lebih banyak berada di bawah rata-rata banyaknya penduduk miskin.



Gambar 2. Boxplot Penduduk Miskin Pendugaan SAE-Kernel

Gambar 2 pola penduduk miskin di setiap kecamatan di Kota Semarang pada boxplot hampirimbang antara lebar bagian atas dan lebar bagian bawah. Hal ini menunjukkan bahwa persebaran penduduk miskin di Kota Semarang berimbang. Jadi hampir setengah dari total kecamatan di Kota Semarang yang berada di bawah rata-rata penduduk

miskin, hal ini menjelaskan bahwa tingkat kesejahteraan masyarakat di Kota Semarang belum merata.

Ada beberapa kecamatan yang memiliki pencilan tinggi salah satunya ialah Kecamatan Semarang Barat dimana sebagai tempat bermukim para pensiunan. Kecamatan Semarang Barat memiliki penduduk usia 65 tahun paling tinggi di Kota Semarang, hal ini yang menjadi faktor utama yang mengakibatkan tingginya penduduk miskin, didukung dengan tidak sehatnya lingkungan, dan rendahnya tingkat pendidikan.

Setelah dilakukan pendugaan terhadap penduduk miskin di Kota Semarang dengan metode SAE-Kernel bootstrap, maka langkah berikutnya adalah menduga nilai MSE. Dalam pendugaan SAE-Kernel, dilakukan koreksi terhadap nilai MSE dengan menggunakan metode resampling *Bootstrap*.

Berikut adalah ringkasan hasil nilai rata-rata MSE dari masing-masing replikasi bootstrap.

Tabel 3. Statistik Replikasi Bootstrap

Replikasi bootstrap	Rata-rata MSE
50	0,188
100	0,175
150	0,186
200	0,184

Berdasarkan Tabel 3 menunjukkan bahwa nilai rata-rata MSE terkecil adalah pada replikasi bootstrap $B=100$ yaitu sebesar 0,175. Hasil tersebut juga menjelaskan bahwa pendugaan jumlah penduduk miskin dengan *Small Area Estimation* Kernel-Bootstrap dengan replikasi bootstrap $B=100$ merupakan metode terbaik.

KESIMPULAN

Hasil pendugaan rata-rata jumlah penduduk miskin di Kota Semarang dengan pendekatan SAE-Kernel sebesar 23.278 jiwa. Model *Small Area Estimation* dengan pendekatan kernel-bootstrap dengan replikasi $B=100$ untuk menduga jumlah penduduk miskin pada level kecamatan di Kota Semarang dapat menghasilkan dugaan yang presisi, dapat ditunjukkan dari nilai MSE yang dihasilkan sebesar 0,175.

DAFTAR PUSTAKA

- [1] Darsyah, M. Y, Rumiati, AT, Otok, B.W. 2012. *Small Area Estimation terhadap Pengeluaran Per Kapita di Kabupaten Sumenep dengan Pendekatan Kernel-Bootstrap*. Prosiding Seminar MIPA. UNESA.
- [2] Darsyah, M. Y. 2013. *Small Area Estimation terhadap Pengeluaran Per Kapita di Kabupaten Sumenep Dengan Pendekatan Nonparametrik*. Jurnal Statistika Volume 1 Nomor 2. Universitas Muhammadiyah Semarang.
- [3] Darsyah, M.Y. dan Wasono, R., 2013. *Pendugaan Tingkat Kemiskinan di Kabupaten Sumenep dengan Pendekatan SAE*. Prosiding Seminar Nasional Statistika. UII. Yogyakarta.
- [4] Darsyah, M.Y. dan Wasono, R., 2013. *Pendugaan IPM pada Area Kecil di Kota Semarang dengan Pendekatan Nonparametrik*. Prosiding Seminar Nasional Statistika. Universitas Diponegoro.
- [5] Kurnia A, dan Notodiputro KA. (2006). *Penerapan Metode Bootstrap dalam Pendugaan Area Kecil*. Forum Statistika dan Komputasi, April 2006, Vol. 11, hal. 12-16.

- [6] Mukhopadhyay P, dan Maiti T. (2004). *Two Stage Non-Parametric Approach for Small Area Estimation*. Proceedings of ASA Section on Survey Research Methods, hal. 4058-4065.
- [7] Opsomer et al. (2004). *Nonparametric Small Area Estimation Using Penalized Spline Regression*. Proceedings of ASA Section on Survey Research Methods, hal.1-8.
- [8] Prasad, N.G.N. dan Rao, J.N.K. (1990). *The Estimation of The Mean Squared Error of The Small Area Estimators*. Journal of American Statistical Association, 85, hal.163-171.