

ESTIMASI PENGELUARAN PER KAPITA DI KABUPATEN REMBANG DENGAN PENDEKATAN SAE-NONPARAMETRIK

Iswahyudi Joko Suprayitno¹, Moh Yamin Darsyah², Budiharto³

¹ *Program Studi Pendidikan Matematika, UNIMUS*

² *Program Studi Statistika, UNIMUS*

³ *FKIP Universitas Terbuka*

matematikawan.mr.joe@gmail.com

Abstrak

Small Area Estimation (SAE) merupakan suatu teknik statistika untuk menduga parameter-parameter subpopulasi yang ukuran sampel nya kecil. Teknik pendugaan ini “*borrowing information*” memanfaatkan data dari domain besar (seperti data sensus data susenas) untuk menduga variabel yang menjadi perhatian pada domain yang lebih kecil yang selanjutnya dikenal pendugaan tidak langsung. Pendugaan langsung tidak mampu memberikan ketelitian yang cukup bila ukuran sampel dalam area kecil, sehingga statistik yang dihasilkan akan memiliki varian yang besar atau bahkan menghasilkan pendugaan yang bias. Penelitian *SAE* diaplikasikan untuk memetakan Pengeluaran Per kapita di Kabupaten Rembang dimana variabel kepadatan penduduk berperan signifikan dalam mempengaruhi pengeluaran per kapita daerah. Kecamatan Rembang merupakan kecamatan dengan pengeluaran per kapita tertinggi di Kabupaten Rembang sedangkan kecamatan yang memiliki pengeluaran per kapita terendah di Kecamatan Sumber.

Kata kunci : *Small Area Estimation, Nonparametrik, Pengeluaran Per kapita*

Small Area Estimation (SAE)

I. Pendahuluan

Suatu area disebut area kecil apabila sampel yang diambil pada area tersebut tidak mencukupi untuk melakukan pendugaan langsung dengan hasil dugaan yang akurat (Rao,2003). Dewasa ini pendugaan area kecil menjadi sangat penting dalam analisis data survei karena adanya peningkatan permintaan untuk menghasilkan dugaan parameter yang cukup akurat dengan ukuran sampel kecil.

merupakan suatu teknik statistika untuk menduga parameter-parameter subpopulasi yang ukuram sampel nya kecil. Teknik pendugaan ini memanfaatkan data dari domain besar seperti data sensus dan data susenas untuk menduga variabel yang menjadi perhatian pada domain yang lebih kecil. Pendugaan sederhana pada area kecil yang didasarkan pada penerapan model desain penarikan contoh (*design-based*) disebut sebagai pendugaan langsung (*direct*

estimation) dimana pendugaan langsung tidak mampu memberikan ketelitian yang cukup bila ukuran sampel dalam area itu kecil sehingga statistik yang dihasilkan akan memiliki varian yang besar atau bahkan pendugaan tidak dapat dilakukan karena berakibat estimasi yang bias (Rao, 2003).

Sebagai alternatif teknik pendugaan untuk meningkatkan efektivitas ukuran sampel dan menurunkan eror maka dikembangkan teknik pendugaan tak langsung (*indirect estimation*) untuk melakukan pendugaan pada area kecil dengan ketelitian yang cukup. Teknik pendugaan ini dilakukan melalui suatu model yang menghubungkan area terkait melalui penggunaan informasi tambahan atau variabel penyerta (*model based*) inilah yang selanjutnya dikenal dengan konsep *SAE*.

Terdapat dua masalah pokok dalam pendugaan area kecil. Masalah pertama adalah bagaimana menghasilkan suatu dugaan yang akurat dengan ukuran sampel kecil pada suatu domain atau area kecil. Masalah kedua yaitu bagaimana menduga *Mean Square Error (MSE)*. Solusi untuk masalah tersebut adalah dengan "meminjam informasi" baik dari dalam area, luar area, maupun luar survei.

Berbagai penelitian yang berkaitan dengan pendugaan area kecil dengan pendekatan nonparametrik antara lain Darsyah (2013) menggunakan *small area estimation* dengan metode *kernel-bootstrap*, Anwar (2008) melakukan *small area estimation* dengan metode *kernel learning*, Indahwati, Sadik dan Nurmasari (2008) melakukan *small area estimation* dengan pendekatan pemulusan kernel, Opsmer (2005) melakukan *small area estimation* menggunakan *penalized spline*, Mukhopadhyay dan Maiti (2004) melakukan *small area estimation* dengan pendekatan nonparametrik.

Penelitian *SAE* diaplikasikan untuk mengestimasi pengeluaran per kapita di Kabupaten Rembang dengan pendekatan nonparametrik dengan menggunakan fungsi *kernel-bootstrap*. Kabupaten Rembang merupakan salah satu daerah tertinggal di Propinsi Jawa Tengah maka dibutuhkan kajian strategis dan mendalam untuk mengetahui informasi yang akurat sehingga pembangunan daerah tepat sasaran.

Dalam kebanyakan aplikasi pendugaan area kecil, digunakan asumsi model linier campuran dan pendugaannya sensitif terhadap asumsi ini. Jika asumsi kelinieran antara rata-rata area kecil dan

variabel penyerta tidak terpenuhi, maka “meminjam kekuatan” dari area lain dengan menggunakan model linier tidak tepat.

$$y_i = \theta_i + \epsilon_i \quad (1)$$

$$\theta_i = m(x_i) + u_i \quad (2)$$

Dimana $i = 1, 2, \dots, m$ menyatakan banyaknya area kecil. Fungsi $m(\cdot)$ adalah fungsi mulus (smoothing function) yang mendefinisikan relasi antara x dan y . θ_i adalah rata-rata area kecil yang tidak teramati, y_i adalah pendugaan langsung dari rata-rata area kecil yang tersampel, u_i galat peubah

$$y = m(x_i) + u + \epsilon \quad (3)$$

Untuk menduga $m(x_i)$, Mukhopadhyay dan Maiti (2004)

$$\hat{m}_h(x_i) = \frac{\sum_i K_h(x-x_i)y_i}{\sum_i K_h(x-x_i)} \quad (4)$$

Dimana $K_h(\cdot)$ adalah fungsi kernel dengan *bandwidth* h dan $K_h(u) = \frac{1}{h}K(u/h)$ dengan $K_h(\cdot)$ memenuhi:

- i. $K(\cdot)$ simetri
- ii. $K(\cdot)$ terbatas dan kontinu pada daerah hasil x

$$\hat{m}_h(x_i) = \frac{1}{m} \sum_{i=1}^m W_{hi}(x)y_i \quad (5)$$

$$\text{Dimana } W_{hi}(x) = \frac{K_h(x-x_i)}{1/m \sum_i K_h(x-x_i)}$$

Berdasarkan definisi di atas, penduga terbaik dari rata-rata area kecil θ_i adalah

Mukhopadhyay dan Maiti (2004) menggunakan model

acak yang berdistribusi independen dan identik dengan $E(u_i) = 0$ dan $var(u_i) = \sigma_u^2$, dan ϵ_i berdistribusi independen dan identik dengan $E(\epsilon_i) = 0$ dan $var(\epsilon_i) = D_i$, dengan asumsi D_i diketahui. Substitusi persamaan (1) dan (2) akan menghasilkan persamaan berikut:

menggunakan pendugaan kernel Nadaraya-Watson

$$\text{iii. } \int K(a)da = 1$$

Fungsi kernel yang sering dipakai adalah fungsi normal dan besarnya *bandwidth* dipilih $h \propto n^{1/5}$ (silverman, 1986). Penduga di atas linier terhadap y_i , dan dapat ditulis sebagai

$$\hat{\theta}_i = \hat{\gamma}_i y_i + (1 - \hat{\gamma}_i) \hat{m}_h(x_i) \quad (6)$$

Dimana $\hat{\gamma}_i = \frac{\hat{\sigma}_u^2}{\hat{\sigma}_u^2 + D_i}$ dan $\hat{\sigma}_u^2$ merupakan penduga dari σ_u^2 .

$$\hat{\sigma}_u^2 = \max \left\{ 0, \frac{1}{m-1} \sum_{i=1}^m W_{hi}(x) [y_i - \hat{m}_h(x_i)]^2 - D \right\} \quad (7)$$

Untuk *MSE* pendugaan area kecil berikut:

$$MSE(\hat{\theta}_i) = \frac{D \hat{\sigma}_u^2}{\hat{\sigma}_u^2 + D} + (1 - \gamma)^2 mse[\hat{m}_h(x_i)] + 2D^2 (\hat{\sigma}_u^2 + D)^{-3} mse(\hat{\sigma}_u^2) \quad (8)$$

Namun demikian pendugaan *MSE* diatas mempunyai kelemahan karena ada informasi yang terputus dan tidak ada rumus

$$mse^*(\hat{\theta}_i) = \frac{1}{J} \sum_{j=1}^B (\hat{\theta}_i^{*(j)} - \theta_i^{*(j)})^2 \quad (9)$$

dimana J adalah banyaknya populasi bootstrap, $\hat{\theta}_i^{*(j)}$ adalah penduga rataan area kecil $ke-i$ dari populasi bootstrap $ke-j$, dan

jadinya, maka untuk pendugaan *MSE* bisa dilakukan dengan pendekatan bootstrap berikut persamaannya

$\theta_i^{*(j)}$ adalah nilai sebenarnya rataan area kecil $ke-i$ dari populasi bootstrap $ke-j$.

II. Metode

Selanjutnya akan dijelaskan mengenai Kajian Model *SAE*, Sumber Data dan Variabel Penelitian, Metodologi Penelitian sebagai berikut.

Model area kecil biasanya menggunakan model linier campuran dalam bentuk $y = X\beta + Zu + e$, dimana X adalah matriks peubah penyerta, Z adalah vektor acak yang biasa dikenal sebagai pengaruh area kecil, dan e adalah vektor dari galat sampel (Rao, 2003). Penduga rataan area kecil di duga dengan melalui estimor *best linier unbiased precictor (BLUP)* $\hat{\theta}_i = \gamma_i \hat{\theta}_i + (1 - \gamma_i) z_i^T \tilde{\beta}$. Mukhopadhyay dan Maiti (2004) menggunakan pendugaan

kernel Nadaraya-Watson $\hat{m}_h(x_i) = \frac{\sum_i K_h(x-x_i) y_i}{\sum_i K_h(x-x_i)}$, dimana estimor pendugaan untuk area kecil diberikan $\hat{\theta}_i = \hat{\gamma}_i y_i + (1 - \hat{\gamma}_i) \hat{m}_h(x_i)$, dengan $\hat{\gamma}_i = \frac{\hat{\sigma}_u^2}{\hat{\sigma}_u^2 + D_i}$. Dimana nilai $\hat{\gamma}_i$ merupakan estimator pembobot yang didapat dari *best linier unbiased precictor (BLUP)* yang merupakan bentuk parametrik. Jadi ada kemungkinan untuk penduga area kecil diatas merupakan penduga semiparametrik karena terdapat 2 (dua) pendekatan yaitu pendekatan nonparamtrik dan pendekatan parametrik.

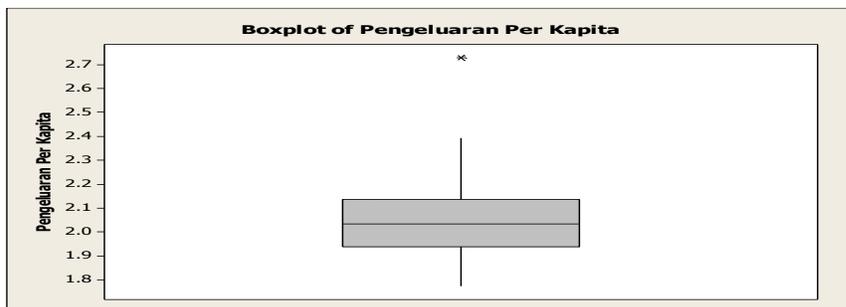
Sumber data yang digunakan dalam penelitian ini untuk pengeluaran per kapita

berasal dari data Survei Sosial Ekonomi Nasional (SUSENAS) BPS Tahun 2013 dan Rembang Dalam Angka 2014. Dimana variabel respon yang menjadi perhatian dalam penelitian ini adalah pengeluaran per kapita pada masing-masing kecamatan di

Kabupaten Rembang. Variabel penyerta yang digunakan yaitu Jumlah prosentase keluarga nelayan, Jumlah prosentase keluarga pengguna Jamkesmas, kepadatan penduduk.

III. Hasil Penelitian

a. Sebaran Data

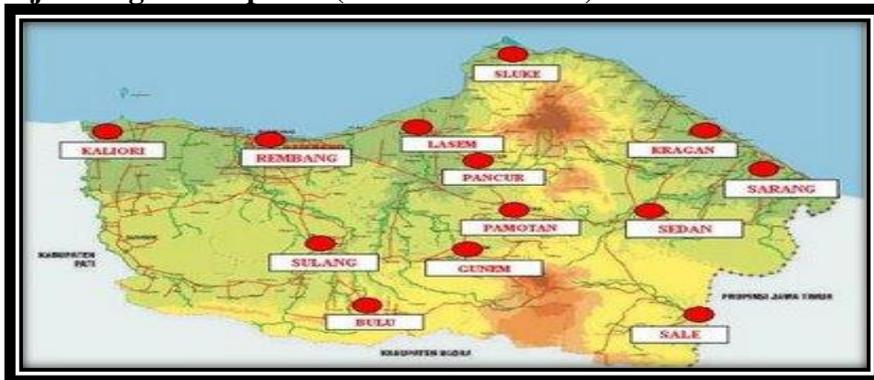


Gambar 1. *Boxplot* Pendugaan Pengeluaran Per Kapita dengan *SAE*

Gambar 1 menunjukkan adanya pencilan pendugaan yaitu Kecamatan Kota Rembang. Lebar bagian atas dan bagian bawah sama hal ini mengindikasikan bahwa

keragaman pengeluaran per kapita pada kecamatan di Kabupaten Rembang tidak terlalu mencolok.

b. Uji Keragaman Spasial (*Heterokedastisitas*)



Gambar 2. Peta Kabupaten Rembang

Pengujian keragaman spasial menggunakan uji *Breusch-Pagan (BP)* menghasilkan nilai BP sebesar 7,76 dengan nilai-p (0,059) yang kurang dari taraf nyata 10%, sehingga diperoleh keputusan tolak H_0 yang berarti bahwa terdapat keragaman spasial pada data pengeluaran per kapita pada tiap kecamatan di Kabupaten Rembang

tahun 2014. Adanya keragaman spasial pada pengeluaran per kapita tersebut menunjukkan bahwa setiap kecamatan di Kabupaten Rembang memiliki karakteristik tersendiri, sehingga diperlukan pendekatan lokal untuk memodelkan dan untuk mengatasi keragaman yang terjadi pada pengeluaran per kapita.

c. Koefisien Parameter

Tabel 1. Koefisien Parameter

Variabel	Koefisien Parameter		
	Minimum	Median	Maksimum
Intersep	1,727	2,468	11,240
X ₁	-0,894	0,021	0,993
X ₂	-0,033	-0,027	0,409
X ₃	0,0006	0,003	0,004
SSE		43,976	
R ²		70,79%	

Nilai R² yang diperoleh dari model sebesar 70,79%. Hal ini berarti keragaman pengeluaran per kapita disebabkan oleh persentase rumah tangga nelayan, persentase jumlah rumah tangga

pengguna jamkesmas dan kepadatan penduduk sebesar 70,79%, sedangkan 20,21 % sisanya disebabkan oleh adanya faktor lainnya yang turut mempengaruhi besarnya pengeluaran per kapita.

d. Pengujian Model

Goodness of fit atau pengujian kesesuaian untuk model dilakukan untuk mengetahui faktor lokasi yang berpengaruh

terhadap tingkat pengeluaran per kapita di Kabupaten Rembang.

Tabel 2. Uji Kesesuaian Model

	SSE	Df	F _{hitung}	P _{value}
Model	43,977	9,023	2,775	0,032

Berdasarkan tabel 2 di atas didapatkan nilai p_{value} (0,042) yang berarti nilai p_{value} kurang dari taraf nyata 5% (0,032<0,05). Hal ini

berarti tolak H₀ karena nilai p_{value} lebih kecil dari taraf nyata 5%, yang artinya ada faktor pengaruh geografis pada model.

Tabel 3. Signifikansi Parameter

No	Kecamatan	Variabel
1	Bulu	X _{2,3}
2	Gunem	-
3	Kaliori	X _{1,2,3}
4	Kragan	X _{1,2,3}
5	Lasem	X _{1,2,3}
6	Pamotan	X _{2,3}
7	Pancur	X _{2,3}
8	Rembang	X _{1,2,3}
9	Sale	X ₃
10	Sarang	X _{1,3}
11	Sedan	-
12	Sluke	-
13	Sulang	X _{2,3}
14	Sumber	-

Berdasarkan tabel 3 diperoleh hasil bahwa ada 10 kecamatan yang dipengaruhi oleh variabel kepadatan penduduk (X₃), dan ada 4 kecamatan yang tidak berpengaruh pada ketiga variabel yang digunakan dalam

penelitian ini. Hal ini diduga karena ada variabel lain yang lebih signifikan yang mempengaruhi pengeluaran per kapita di Kabupaten Rembang.

IV. Kesimpulan

Berikut kesimpulan yang diperoleh dari hasil penelitian yang telah dilakukan antara lain:

- a) Aplikasi Small Area Estimation dapat di kombain dengan Regresi Spasial.
- b) Variabel yang paling berpengaruh terhadap pengeluaran per kapita di Kabupaten Rembang secara keseluruhan yaitu kepadatan penduduk walaupun variabel persentase keluarga nelayan cukup besar.
- c) faktor lokasi/area antar kecamatan berpengaruh terhadap tingkat pengeluaran per kapita di Kabupaten Rembang.
- d) Nilai R^2 yang diperoleh dari model sebesar 70,79%. Hal ini berarti keragaman persentase jumlah keluarga miskin disebabkan oleh persentase rumah tangga nelayan, persentase jumlah rumah tangga pengguna jamkesmas dan kepadatan penduduk sebesar 70,79%, sedangkan 20,21 % sisanya disebabkan oleh adanya faktor lainnya yang turut mempengaruhi pengeluaran per kapita.

V. Daftar Pustaka

[BPS]. **Badan Pusat Statistik.** (2014). <http://www.bps.go.id/glossary/2014>.

Darsyah, M.Y, Rumiati, A.T, Otok, B.W.

(2012). *Small Area Estimation terhadap Pengeluaran Per Kapita di Kabupaten Sumenep dengan pendekatan Kernel-Bootstrap.* Prosiding Seminar Nasional MIPA UNESA, Surabaya.

Darsyah, M.Y dan Wasono, R. (2013).

Pendugaan Tingkat Pengeluaran per kapita di Kabupaten Sumenep dengan pendekatan SAE. Prosiding Seminar Nasional Statistika UII, Yogyakarta.

Darsyah, M.Y dan Wasono, R. (2013).

Pendugaan IPM pada Area Kecil di Kota Semarang dengan Pendekatan Nonparametrik. Prosiding Seminar Nasional Statistika UNDIP, Semarang.

Darsyah, M.Y. (2013).

Small Area Estimation terhadap Pengeluaran Per Kapita di Kabupaten Sumenep dengan Pendekatan Nonparametrik. Jurnal Statistika. Vol 1 Nomer 2.

Eubank, R. L. (1988).

Spline Smoothing and Nonparametric Regression. New York: Marcel Dekker.

Fay, R.E. dan Herriot, R.A. (1979).

Estimates Income for Small Places: An Application of James-Stein Procedures to Census Data. *Journal of American Statistical Association*, 74, pp.269-277.

Hardle, W. (1994). *Applied Nonparametric Regression*. New York: Cambridge University Press.

Hastie, T. dan Tibshirani, R.J. (1990). *Generalized Additive Models*. New York: Chapman and Hall.

Mukhopadhyay P, Maiti T. (2004). *Two Stage Non-Parametric Approach for Small Area Estimation*. Proceedings of

ASA Section on Survey Research Methods: 4058-4065.

Opsomer et al. (2004). *Nonparametric Small Area Estimation Using Penalized Spline Regression*. Proceedings of ASA Section on Survey Research Methods: 1-8.

Rao JNK. (2003). *Small Area Estimation*. New Jersey: John Wiley & Sons, Inc.