SMALL AREA ESTIMATION TERHADAP PENGELUARAN PER KAPITA DI KABUPATEN SUMENEP DENGAN PENDEKATAN NONPARAMETRIK

Moh Yamin Darsyah
Prodi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Muhammadiyah Semarang
Alamat e-mail : mydarsyah@yahoo.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. Adapun pendugaan langsungnya 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. Pendugaan tak langsung dengan pendekatan SAE Nonparametrik di gunakan untuk menduga pengeluaran per kapita pada level kecamatan di Kabupaten Sumenep. Evaluasi hasil pendugaan dilakukan dengan membandingkan nilai RRMSE (*Relative Root Mean Square Error*) penduga langsung dengan nilai RRMSE (*Relative Root Mean Square Error*) penduga tidak langsung SAE Nonparametrik, hasil pendugaan SAE Kernel-Bootstrap memberikan hasil dugaan dengan presisi yang lebih teliti.

Kata kunci : *SAE, Nonparametrik, RRMSE*

PENDAHULUAN

Statistik area kecil (*small area statistic*) dewasa ini sangat diminati dalam berbagai bidang. Pendugaan pada area kecil (*small area estimation*) sangat dibutuhkan untuk mendapatkan informasi – informasi pada area kecil misalnya pada lingkup kota/kabupaten, kecamatan, kelurahan/ desa. Informasi tersebut menjadi sangat penting sejalan dengan berkembangnya era otonomi daerah di Indonesia karena bisa digunakan sebagai acuan untuk menyusun perencanaan, pemantauan serta kebijakan pemerintah sehingga pembangunan di daerah terencana dan berbasis informasi.

Small Area Estimation merupakan suatu teknik statistika untuk menduga parameter- parameter subpopulasi yang ukuran 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 [9].

Sebagai alternatif teknik pendugaan untuk meningkatkan efektivitas ukuran sampel dan menurunkan *error* maka dikembangkan teknik pendugaan tak langsung (*inderct 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 *Small Area Estimation*.

Secara statistik metode dengan memanfaatkan informasi tambahan akan mempunyai sifat "meminjam kekuatan" dari hubungan antara rataan area kecil dan informasi tambahan tersebut.Semua teknik pendugaan tak langsung mempunyai asumsi adanya hubungan linier antara rataan area kecil dengan variabel penyerta yang digunakan sebagai informasi tambahan dalam pendugaan tersebut. Berbagai teknik pendugaan area kecil yang sering digunakan seperti *Empirical Bayes*, *Hirarical Bayes*, EBLUP, Pendekatan sintetik dan komposit semuanya menggunakan prosedur parametrik.

Jika tidak ada hubungan linier antara rataan 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.

Pendekatan dengan menggunakan fungsi kernel diusulkan karena fungsi ini didasarkan pada pendekatan penggunaan ketersediaan variabel- variabel umum antara data sensus dan survei sehingga sesuai dengan pendugaan area kecil yang mengestimasi fungsi regresi berdasarkan informasi tambahan. Pendekatan kernel menawarkan teknik nonparametrik sebagai alternatif baru yang menjanjikan identifikasi fungsi regresi pada pendugaan area kecil. Metode ini lebih fleksibel dibanding dengan metode parametrik yang menggabungkan pola-pola kovarian spasial untuk pendugaan area kecil.

Berbagai penelitian yang berkaitan dengan pendugaan area kecil dengan pendekatan nonparametrik antara lain *small area estimation* dengan metode *kernel learning* [5], *small area estimation* dengan pendekatan pemulusan kernel [4], *small area estimation* menggunakan *penalized spline* [7], *small area estimation* dengan pendekatan nonparametrik [6].

Berdasarkan Master Plan Madura 2008 pertumbuhan dan perkembangan pulau Madura relatif lambat hal ini bisa dilihat dari rendahnya pendapatan per kapita penduduk yang masih dibawah rata- rata pendapatan per kapita Provinsi Jawa Timur, salah satu kabupaten yang terletak di Pulau Madura yaitu Kabupaten Sumenep yang memiliki populasi 1.041.915 jiwa dengan luas 2000 kilometer persegi dengan yang terbagi dua bagian yaitu wilayah daratan dan wilayah kepulauan, tentu saja hal ini sangat menarik untuk menjadi objek penelitian. [3] meduga area kecil untuk pengeluaran per kapita pada level kecamatan di Kabupaten Sumenep dengan pendekatan *empirical bayes* tetapi disini penulis menduga pengeluaran per kapita dengan pendekatan nonparametrik.

Evaluasi hasil Pendugaan dilakukan dengan membandingkan nilai RRMSE pengeluaran per kapita yang sudah diketahui melalui pendugaan langsung dengan nilai RRMSE pengeluaran per kapita melalui pendugaan tak langsung dalam hal ini menggunakan metode nonparametrik.

Menurut Badan Pusat Statistika, pengeluaran rata- rata per kapita menunjukkan besarnya pengeluaran setiap anggota rumah tangga dalam kurun waktu sebulan, sedangkan definisi rumah tangga adalah sekelompok orang yang mendiami sebagian atau seluruh bangunan fisik dan biasanya tinggal bersama serta makan dari satu dapur .

Suatu area disebut area kecil apabila sampel yang diambil pada area tersebut tidak mencukupi untuk melakukan pendugaan langsung dengan hasil dugaan yang akurat [9]. 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.

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, maupu lauar survei.

Pendugaan parameter pada suatu area kecil dapat dilakukan dengan pendugaan langsung maupun pendugaan tidak langsung. Pendugaan langsung merupakan pendugaan berdasarkan data sampel atau design sampling, dimana hasil pendugaan langsung pada suatu area kecil merupakan pendugaan tak bias meskipun memiliki varian yang besar dikarenakan dugaannya diperoleh dari ukuran sampel kecil. Pendugaan tak langsung merupakan pendugaan dengan cara memanfaatkan informasi variabel lain yang berhubungan dengan parameter yang diamati, dimana variabel penyerta terpilih diharapkan bisa memberikan pengaruh pada parameter yang diamati.

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 [9]. Dalam kebanyakan aplikasi pendugaan area kecil, digunakan asumsi model linier campuran dan pendugaannya sensitif terhadap asumsi ini. Jika asumsi kelinieran antara rataan area kecil dan variabel penyerta tidak terpenuhi, maka "meminjam kekuatan" dari area lain dengan menggunakan model linier tidak tepat. [6] menggunakan model

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

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

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

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

Untuk menduga $m(x_i)$, [6] menggunakan pendugaan kernel Nadaraya-Watson

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

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

i. $K(.)$ simetri
ii. $K(.)$ terbatas dan kontinu pada daerah hasil x
iii. $\int K(a)\partial a = 1$

Fungsi kernel yang sering dipakai adalah fungsi normal dan besarnya bandwidht dipilih $h \infty n^{1/5}$ [10]. Penduga di atas linier terhadap y_i, dan dapat ditulis sebagai

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

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 rataan area kecil θ_i adalah

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

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.

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

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

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

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

METODOLOGI PENELITIAN

Sumber data dan variabel penelitian

Sumber data yang digunakan dalam penelitian ini untuk pengeluaran per kapita berasal dari data Survei Sosial Ekonomi Nasional (SUSENAS) BPS Tahun 2009 dan variabel penyerta berasal dari Sumenep Dalam Angka 2010. Dimana variabel respon yang menjadi perhatian dalam penelitian ini adalah pengeluaran per kapita pada masing – masing kecamatan di Kabupaten Sumenep. Variabel penyerta yang digunakan yaitu kepadatan penduduk.

Metode Analisis

Tahapan-tahapan analisis yang dilakukan pada penelitian ini dijelaskan pertujuan penelitian ini dijelaskan sebagai berikut.

a. Menduga pengeluaran per kapita rumah tangga per kecamatan di Kabupaten Sumenep dengan pendekatan SAE .

Memilih variabel bantu x_i yang diasumsikan mempengaruhi dan mampu menggambarkan pengeluaran per kapita, salah satu nya dengan melihat nilai *Korelasi Pearson* .

$$\rho = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{(n-1)s_x s_y} \quad (2.17)$$

Selanjutnya berikut langkah- langkah algoritma SAE- Pendekatan kernel

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 $\hat{\sigma}_u^2 =$ (2.18) $\max\left\{0, \frac{1}{m-1}\sum_{i=1}^{m} W_{hi}(x)[y_i - \hat{m}(x_i)]^2 - 1\right\}$
3. Subtitusikan $\hat{\theta}_i = \hat{\gamma}_i y_i + (1-\hat{\gamma}_i)\hat{m}(x_i)$ dengan $\hat{\gamma}_i = \frac{\hat{\sigma}_u^2}{\hat{\sigma}_u^2 + D_i}$
4. Menghitung $MSE(\hat{\theta}_i)$ akan dilakukan dengan *bootstrap*

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

b. Menduga pengeluaran per kapita rumah tangga per kecamatan di Kabupaten Sumenep dengan menggunakan metode pendugaan langsung.

1. Menghitung pengeluaran per kapita rumah tangga untuk masing-masing kecamatan secara langsung (*direct estimation*), dengan cara:

   $y_i = \frac{p_i}{q_i}$

   Dimana:

   y_i = pengeluaran per kapita kecamatan ke-i

   p_i = total pengeluaran rumah tangga sebulan di kecamatan ke-i

   q_i = banyaknya anggota rumah tangga di kecamatan ke-i

   i = 1, 2, ..., 27

2. Menghitung nilai MSE dari hasil estimasi pengeluaran per kapita rumah tangga dengan metode *direct estimation*

$$MSE = \frac{\sum_i (y_i - \hat{y})2}{(n-1)}$$

c. Membandingkan nilai RRMSE pendugaan kernel – bootstrap dengan nilai pendugaan langsung, dimana perhitungan RRMSE sebagai berikut:

$$\mathrm{RRMSE}(\hat{\theta}_i) = \frac{\sqrt{\mathrm{MSE}(\hat{\theta}_i)}}{\hat{\theta}_i} \times 100\%$$

HASIL PENELITIAN

Model area kecil biasanya menggunakan model linier campuran yang memuat *fixed* dan random *effect*, dimana rataan area kecil dapat dikombinasikan secara linier antara *fixed* dan random *effect* melalui BLUP. Estimator BLUP merupakan semacam parameter yang dapat meminimumkan MSE dan tidak tergantung pada normality dari random effect tetapi tergantung pada varian dan kovarian dari random *effect* yang dapat di estimasi melalui metode moment atau metode *restricted maximum likelihood* (REML). Secara umum 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, u dan e adalah independen dengan mean 0 dan diberikan kovarian matriks G dan R yang tergantung dari parameter varian δ = $(\boldsymbol{\delta_{1i}}, \dots, \boldsymbol{\delta_{pi}})^{\boldsymbol{T}}$. Selanjutnya kita asumsikan Var = Var (δ) = R+ ZGZ^T, μ = $1^T\beta$+$m^T v$.

Berikut estimator BLUP dari μ

$$\tilde{\mu}^H = t(\delta, y) = 1^T + m^T\tilde{v} = 1^T\tilde{\beta} + m^T GZ^T V^{-1}(y - X\tilde{\beta}) \qquad (10)$$

Dimana

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

Selanjutnya diperoleh best linier unbiased estimator untuk β

$$\tilde{v} = \tilde{v}(\delta) = GZ^T V^{-1}(y - X\tilde{\beta})$$

Basic area level model yaitu model yang didasarkan pada ketersediaan data pendukung yang hanya ada untuk level area tertentu, misalkan **zi** = $(\boldsymbol{z_{1i}}, \dots, \boldsymbol{z_{pi}})^{\boldsymbol{T}}$ dengan **zi** adalah suatu vektor, i adalah banyaknya area dan p adalah banyaknya peubah penyerta, dan parameter yang akan diduga θ_i, diasumsikan mempunyai hubungan dengan **zi**. Data pendukung tersebut digunakan untuk membangun model:

$$\hat{\theta}_i = z_i^T\beta + b_i u_i + e_i,\ i = 1, \dots, m \qquad (11)$$

dengan β merupakan vektor koefisien regresi untuk data pendukung **zi** dan $u_i \sim N(0, \boldsymbol{\sigma}^2_u)$, sebagai pengaruh acak yang diasumsikan normal dan $e_i \sim N(0, \psi_i)$. Persamaan (4.1) yang merupakan model yang dibentuk dari model linier campuran. Selanjutnya

$$y_i = \hat{\theta}_i,\ X_i = z_i^T,\ Z_i = b_i$$

$$V_i = u_i,\ e_i = e_i,\ \beta = (\beta_1, \dots, \beta_p)^T$$

Sehingga

$$V_i = \psi_i + \sigma_v^2 b_i^2$$

[8] menduga rataan area kecil dari model (1) dengan estimator BLUP

$$\tilde{\theta}_i^H = z_i^T\tilde{\beta} + \gamma_i(\hat{\theta}_i - z_i^T\tilde{\beta}) \qquad (12)$$

$$= \gamma_i\hat{\theta}_i + (1-\gamma_i)z_i^T\tilde{\beta} \qquad (13)$$

dimana

$$\gamma_i = \sigma_v^2 b_i^2/(\psi_i + \sigma_v^2 b_i^2)$$

$$\tilde{\beta} = \tilde{\beta}(\sigma_v^2) = [\textstyle\sum_{i=1}^m z_i z_i^T/(\psi_i + \sigma_v^2 b_i^2)]^{-1}[\sum_{i=1}^m z_i\hat{\theta}_i/(\psi_i + \sigma_v^2 b_i^2)]$$

Estimator BLUP $\tilde{\theta}_i^H$ dapat diartikan sebagai rata- rata terboboti pada pendugaan tak langsung sedangkan $z_i^T\tilde{\beta}$ merupakan estimator sintetik. Model varian $\sigma_v^2 b_i^2$ relatif terhadap total varian $\psi_i + \sigma_v^2 b_i^2$. Jika model $\psi_i + \sigma_v^2 b_i^2$ relatif kecil maka nilai $\gamma_i$akan relatif kecil juga. Untuk itu akan dilakukan pendugaan ragam antar area σ_v^2, rumus dasar estimator σ_v^2 :

$$E\left[\frac{\sum_i(\hat{\theta}_i - z_i^T\tilde{\beta})^2}{(\psi_i + \sigma_v^2 b_i^2)}\right] = E[h(\sigma_v^2)] = m - 1 \qquad (14)$$

Dimana

$$\tilde{\beta} = \tilde{\beta}(\sigma_v^2)$$
$$h(\sigma_v^2) = m - 1$$

Dalam aplikasi [3] menyarankan untuk melakukan iterasi 0 - 10. Berikut

$\sigma_v^{2(0)} = 0$

$\sigma_v^{2(a+1)} = \sigma_v^{2(a)} + \frac{1}{h_*^{'}\left(\sigma_v^{2(a)}\right)}\left[m - 1 - h\left(\sigma_v^{2(a)}\right)\right]$

Dimana

$$h_*^{'}(\sigma_v^2) = -\sum_i b_i^2 \left(\hat{\theta}_i - z_i^T\hat{\beta}\right)^2 / (\psi_i + \sigma_v^2 b_i^2)^2$$

Karena persamaan dirasa sangat sulit karena harus melakukan iterasi 0 sampai 10 maka diberikan simpel estimator

$\tilde{\sigma}_{vs}^2$ = maks(0, $\tilde{\sigma}_{vs}^2$),

$\tilde{\sigma}_{vs}^2 = \frac{1}{m-1}\left[\sum_i (b_i^{-1}\hat{\theta}_i - z_i^T\hat{\beta}_{WLS})^2 - \sum_i \frac{\psi_i}{b_i^2}\left(1 - \tilde{h}_{ii}\right)\right]$,

dimana

$$\hat{\beta}_{WLS} = \left(\sum_i \tilde{z}_i \tilde{z}_i^T\right)^{-1}\left(\sum_i \tilde{z}_i\, \hat{\theta}_i b_i\right)$$

[6] 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 estimtor pendugaan untuk area kecil diberikan $\hat{\theta}_i = \hat{\gamma}_i y_i + (1 - \hat{\gamma}_i)\hat{m}_h(x_i)$, secara keseluruhan rataan area kecil antara estimator SAE-Pendekatan kernel dengan estimator BLUP sama, dimana pada pendekatan kernel lebih fleksibel karena tidak memerlukan teknik penggabungan kovarian. Untuk estimator pembobot area dengan pendekatan kernel $\hat{\gamma}_i = \frac{\hat{\sigma}_u^2}{\hat{\sigma}_u^2 + 1}$ diperoleh dari estimator BLUP sehingga model $\hat{\theta}_i = \hat{\gamma}_i y_i + (1 - \hat{\gamma}_i)\hat{m}_h(x_i)$ yang terbentuk adalah semiparamtrik.

Selanjutnya dilakukan pendugaan baik pendugaan langsung maupun tidak langsung, berikut disajikan dalam Tabel 1.

Tabel 1. Hasil Pendugaan Pengeluaran per kapita (x Rp.100.000,00)

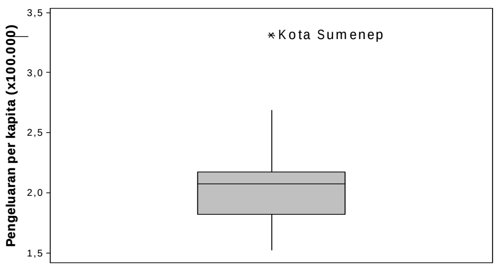
No	Kecamatan	Penduga Langsung	SAE Kernel
1	Pragaan	1.95459	2.00322
2	Bluto	1.52799	1.77645
3	Saronggi	1.86102	1.95357
4	Talango	1.88380	1.96573
5	Kalianget	1.62802	1.88380
6	Kota Sumenep	3.31510	2.72683
7	Lenteng	2.24386	2.15731
8	Ganding	2.13588	2.09995
9	Guluk Guluk	1.91476	1.98243
10	Pasongsongan	1.69374	1.86496
11	Ambunten	2.01365	2.03510
12	Rubaru	2.12178	2.09263
13	Dasuk	2.07330	2.06690
14	Manding	2.68897	2.39429
15	Batuputih	2.13224	2.09833
16	Gapura	2.15088	2.10828
17	Batang Batang	1.78621	1.91443
18	Dungkek	1.69120	1.86396
19	Gayam	2.10407	2.08352
20	Sapeken	2.30692	2.19142
21	Arjasa	2.20532	2.13744
22	Giligenteng	1.89215	1.93908
23	Batuan	1.89881	1.97059
24	Nonggunong	2.15956	2.17543
25	Ra'As	1.81602	1.93132
26	Kangayan	2.31962	2.19785
27	Masalembu	2.03310	2.03583

Tabel 2. Nilai Statistik Pengeluaran per kapita (x Rp.100.000,00)

Statistik	Langsung	SAE-Kernel
Mean	2.0683	2.0591
Standar Deviasi	0.3892	0.1891
Koef Varians	18.82	9.18
Minimum	1.5280	1.7765
Maksimum	3.3151	2.7268
Jangkauan	1.7871	0.9504

Eksplorasi data dilakukan terhadap pendugaan pengeluaran per kapita dengan pendekatan SAE- Kernel pada level kecamatan di Kabupaten Sumenep. Rata-rata pengeluaran per kapita rumah tangga di Kabupaten Sumenep pada tahun 2009 sebesar Rp 205.910,00. Pengeluaran per kapita antar kecamatan di Kecamatan Sumenep tidak terlalu beragam yang ditunjukkan oleh nilai koefisien varians sebesar 9,18% dan standar deviasi sebesar 0,1891. Kecamatan Bluto memiliki pengeluaran per kapita paling kecil sebesar Rp 177.650,00 dan Kecamatan Kota Sumenep memiliki pengeluaran per kapita tertinggi sebesar Rp 272.6800,00. Kecamatan Bluto memiliki pengeluaran per kapita terendah karena disebabkan mayoritas masyarakat bekerja di sektor pertanian dan termasuk daerah yang mempunyai kepadatan penduduk tinggi. Kecamatan Kota Sumenep memiliki pengeluaran per kapita tertinggi di sebabkan karena kota sumenep merupakan pusat pemerintahan sekaligus pusat perekonomian di Kabupaten Sumenep sehingga ini menjadi daya magnet bagi masyarakat sekitar untuk bermigrasi di kecamatan kota sumenep hal ini bisa ditunjukan dengan kepadatan penduduk paling tinggi di Kabupaten Sumenep.

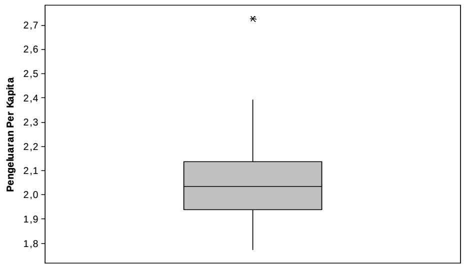
Selanjutnya membandingkan antara pengeluaran per kapita pendugaan langsung dengan pendugaan SAE-Kernel salah satunya dengan melihat koef keragaman pengeluaran per kapita antar kecamatan melalui boxplot.



Gambar 1. *Boxplot* Pendugaan Langsung Pengeluaran Per Kapita

Gambar 1. menunjukkan bahwa ada satu kecamatan yang menjadi pencilan yakni Kecamatan Kota Sumenep. Kecamatan Kota Sumenep memiliki pengeluaran per kapita yang paling besar dengan selisih cukup jauh dengan pengeluaran perkapita kecamatan lain di Kabupaten Sumenep. Hal ini dapat dipahami karena Kecamatan Kota Sumenep merupakan ibukota kabupaten sehingga pusat kegiatan pemerintahan dan perekonomian berada di kecamatan tersebut.

Pola pengeluaran per kapita di setiap kecamatan di Kabupaten Sumenep pada *boxplot* lebih lebar pada bagian bawah. Hal ini menunjukkan bahwa persebaran pengeluaran per kapita setiap kecamatan di Kabupaten Sumenep lebih banyak berada di bawah rata-rata nilai pengeluaran per kapita.



Gambar 2. *Boxplot* Pendugaan Pengeluaran Per Kapita Dengan Pendekatan SAE-Kernel

Gambar 2 menunjukan adanya pencilan pendugaan yaitu kecamatan kota Sumenep. Lebar bagian atas dan bagian bawah sama hal ini mengindikasikan bahwa keragaman pengeluaran per kapita pada kecamatan di Kabupaten Sumenep tidak terlalu mencolok.

Untuk mengetahui kebaikan masing-masing model pendugaan akan dilakukan dengan membandingkan RRMSE pendugaan langsung dengan RRMSE pendugaan SAE Kernel-Bootstrap. Dimana nantinya RRMSE terkecil dari model pendugaan merupakan penduga terbaik untuk area kecil. Untuk replikasi bootstrap diberikan B=100 yang tersaji dalam Tabel 3.

Tabel 3. Nilai Statistik Pendugaan RRMSE

Statistik	Pendugaan Langsung	SAE Bootstrap
Mean	0.77	0.51
Standar Deviasi	0.99	0.05
Minimum	3.21	0.41
Maksimum	11.27	0.60
Jangkauan	8.06	0.19

Tabel 4. Hasil Pendugaan RRMSE

No	Kecamatan	Penduga Langsung	SAE Bootstrap
1	Pragaan	8.112191	0.58
2	Bluto	9.198523	0.60
3	Saronggi	10.21500	0.52
4	Talango	9.148489	0.48
5	Kalianget	9.388531	0.59
6	Kota Sumenep	6.855927	0.41
7	Lenteng	9.047110	0.53
8	Ganding	9.606263	0.53
9	Guluk Guluk	9.015315	0.51
10	Pasongsongan	10.32860	0.56
11	Ambunten	8.534539	0.52
12	Rubaru	10.04373	0.47
13	Dasuk	11.26767	0.51
14	Manding	9.803575	0.42
15	Batuputih	9.322334	0.48
16	Gapura	8.779233	0.49
17	Batang Batang	9.318675	0.54
18	Dungkek	10.68886	0.56
19	Gayam	8.935445	0.43
20	Sapeken	9.569151	0.46
21	Arjasa	7.772604	0.48
22	Giligenteng	4.296472	0.55
23	Batuan	9.987213	0.59
24	Nonggunong	3.457565	0.46
25	Ra'As	4.987932	0.54
26	Kangayan	3.875792	0.51
27	Masalembu	3.210147	0.54

Dari pendugaan langsung diketahui kecamatan yang memiliki RRMSE tertinggi yaitu kecamatan dasuk. Sedangkan pendugaan RRMSE dengan pendekatan SAE kernel-bootstrap memiliki nilai RRMSE rekatif kecil dan variansi antar kecamatan kecil. Untuk pendugaan RRMSE dengan pendekatan SAE kernel-bootstrap akan mengalami perubahan tergantung banyaknya replikasi B yang dilakukan.

Setelah dilakukan pendugaan RRMSE baik pendugaan langsung maupun pendugaan SAE kernel-bootstrap, diketahui dari masing- masing pendugaan nilai RRMSE terkecil diduga dengan pendekatan SAE kernel-bootstrap.

KESIMPULAN

Pendugaan tak langsung dengan pendekatan SAE Kernel-Bootstrap dapat memberikan hasil pendugaan yang lebih presisi dan akurat dibanding pendugaan langsung yang bisa dilihat dari nilai RRMSE yang dihasilkan masing- masing pendugaan. Disamping itu SAE sangat efektif untuk ukuran sampel yang kecil dimana dapat menurunkan *error* dan variannya.

Selanjutnya untuk kedepan bisa digunakan kernel multivariabel agar lebih bisa menggambarkan parameter yang diamati. Dan bisa juga dibandingkan dengan metode parametrik.

DAFTAR PUSTAKA

[1] [BPS]. Badan Pusat Statistik. 2012.http://www.bps.go.id/glossary/2012. [12 November 2012]

[2] Fauzi, H.(2011). *Small Area* Estimation Terhadap Pengeluaran Per Kapita Kabupaten Sumenep dengan Pendekatan Empirical- Bayes. *Makalah Kolokium ITS*, Surabaya (Tidak Dipublikasikan)

[3] 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, hal. 269-277.

[4] Indahwati, Sadik K, Nurmasari R.(2008). Pendekatan Pemulusan Kernel Pada Pendugaan Area Kecil. *Prosiding Seminar Nasional Matematika*. Universitas Negeri Yogyakarta.

[5] Kurnia, A. (2008). Modifikasi General Regression dan Pendekatan Nonparametrik Pada Pendugaan Area Kecil. *Makalah Kolokium*. (Tidak Dipubilkasikan). Institut Pertanian Bogor. Bogor.

[6] Mukhopadhayay P, Maiti P.(2004). *Two stage Non Parametric Approachfor Small Area Estimation.* Proceedings of ASA Section on

Survey Research Metodhs: 3447-3452.
[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, N.J.K.(1990). The Estimation MSE of The Small Area Estimation. *Journal of ASA*,85,pp.163-171.
[9] Rao, N.J.K.(2003). *Small Area Estimation*. New Jersey: John Wiley&Sons, Inc.
[10] Silverman,B.W.(1986). *Density Estimation for Statistics and Data Analysis*. London: Chapman and Hall.