ESTIMASI FUNGSI DENSITAS MULTIVARIAT SECARA SEMIPARAMETRIK

Mochamad Sonhaji*

ABSTRAK

ESTIMASI FUNGSI DENSITAS MULTIVARIAT SECARA SEMIPARAMETRIK. Tujuan dari paper ini adalah untuk membangun suatu kelas dari estimasi-estimasi semiparametrik dan menemukan estimator yang cocok untuk kelas ini. Gagasannya adalah dengan mengalikan estimasi fungsi densitas parametrik dengan fungsi koreksi yang ditawarkan. Kami membahasnya dalam multivariat. Estimator yang baru haruslah mudah digunakan, bila dihadapi problem dalam penggunaan estimasi secara parametrik ataupun nonparametrik. Disajikan juga perbandingan terhadap estimator kernel dengan beberapa contoh.

Kata kunci : Estimator Hjort-Glad, Estimator Lokal Likelihood, Estimator Hjort-Jones, fungsi kernel, *bandwidth*, MISE, AMISE.

ABSTRACT

SEMIPARAMETIC ESTIMATION ON MULTIVARIAT DENSITY FUNCTIONS. The objectives of this paper are to develop a class of semiparametric estimations and to find an appropriate estimator for this class. The idea is to multiply an initial parametric density estimate with a proposal correction function. We consider this study in the multidimensional case. A new estimator should be particularly useful when either parametric or nonparametric methods have problems. The comparisons with the traditional kernel estimator are presented by some practical examples.

PENDAHULUAN

Dimisalkan $\mathbf{X}_1$, $\mathbf{X}_2,\ldots,\mathbf{X}_n$ adalah observasi-observasi berdimensi d yang independen dari fungsi densitas $f(\mathbf{x})$ yang tidak diketahui. Dari observasi-observasi itu, kami mencoba untuk mengestimasi $f(\mathbf{x})$. Ada dua kelas pendekatan dalam mengestimasi densitas $f(\mathbf{x})$ yaitu pendekatan secara *parametrik* dan pendekatan secara *nonparametrik*. Dalam pendekatan parametrik, $f(\mathbf{x})$ diasumsikan sebagai anggota dari

* Metodologi Statistik, Badan Pusat Statistik (BPS)

kelas parametrik berdimensi terbatas $\{g(\mathbf{x}, \theta)|\ \theta \in \Theta\}$, di mana Θ adalah ruang parametrik dari θ (mungkin berupa vektor). Masalah estimasi $f(\mathbf{x})$ dalam kelas ini adalah estimasi θ. Dengan menggunakan esimator θ yaitu $\hat{\theta}$, ditentukan bahwa estimatornya $\hat{f}(\mathbf{x}) = g(\mathbf{x}, \hat{\theta})$.

Seandainya $f(\mathbf{x})$ tidak mungkin untuk diasumsikan sebagai anggota dari kelas parametrik, maka penting untuk mengadakan pendekatan dengan cara lain yaitu pendekatan nonparametrik. Ada banyak metode pendekatan secara nonparametrik. Supaya terdapat hubungan dengan makalah ini, kami sajikan pendekatan secara nonparametrik dengan menggunakan fungsi kernel. Fungsi kernel itu biasanya berupa fungsi densitas yang simetris. Keuntungan dalam menggunakan fungsi kernel adalah terdapatnya kemudahan dalam penghitungan sifat dasar estimator seperti bias dan variansi, serta didapatkan jawaban andal dan sederhana untuk pengembangan aplikasi yang lebih luas. Estimator yang terkenal dengan menggunakan fungsi kernel adalah estimator densitas kernel tradisional, ref. Wand dan Jones (1995), sebagai berikut:

$$\tilde{f}(\mathbf{x}, \mathbf{H}) = \frac{1}{n}\sum_{i=1}^{n} K_{\mathbf{H}}(\mathbf{x} - \mathbf{X}_{\mathbf{i}}), \tag{1}$$

di mana $\mathbf{H}$ adalah matrik orde d x d definit positif simetris yang dinamakan *matrik bandwith* atau disingkat *bandwith*, and K adalah fungsi densitas bervariat d yang simetris yang dinamakan *fungsi kernel*, dengan ketentuan

$K_{\mathbf{H}}(\mathbf{x}) = |\mathbf{H}|^{-1/2} K(\mathbf{H}^{-1/2}\mathbf{x})$.

Sifat dasar $\tilde{f}(\mathbf{x})$ adalah untuk $n \to \infty$, $\mathbf{H} \to \mathbf{0}$, $n^{-1}|\mathbf{H}|^{-1/2} \to 0$

$$\text{Bias } \tilde{f}(\mathbf{x}) = \frac{1}{2}\sum_{j=1}^{d} h_j^2 \sigma(K_j)^2\, \mathrm{H}_{\mathbf{x}}[f(\mathbf{x})]_{jj} + O\left(\sum_{j=1}^{d} h_j^4\right)$$

$$\text{Var } \tilde{f}(\mathbf{x}) = \frac{R(K) f(\mathbf{x})}{n|\mathbf{H}|^{1/2}} - \frac{f(\mathbf{x})^2}{n} + O\left(\frac{\sum_{j=1}^{d} h_j^2}{n|\mathbf{H}|^{1/2}}\right)$$

di mana

$R(K) = \int K(\mathbf{z})^2 d\mathbf{z}$ $\quad \sigma(K_j)^2 = \int z_j^2 K(\mathbf{z}) d\mathbf{z}$, dan $\mathrm{H}_{\mathbf{x}}[f(\mathbf{x})]_{jj}$ adalah elemen (j,j) dari $\mathrm{H}_{\mathbf{x}}[f(\mathbf{x})]$. Notasi dari $\mathrm{H}_{\mathbf{x}}[f(\mathbf{x})]$ adalah matrik Hessian dari $f(\mathbf{x})$.

Karena pendekatan nonparametrik ini memberikan jawaban yang fleksibel demikian pula kemudahan aplikasi dari pendekatan parametrik, akhir-akhir ini sering didiskusikan pendekatan baru yaitu *semiparametrik*. Contohnya, Hjort and Glad (1995) dan Hjort and Jones (1996), yang akan dibahas lebih lanjut di sesi berikutnya. Tujuan menggunakan semiparametrik adalah untuk menemukan metode estimasi yang lebih baik dibanding metode parametrik maupun nonparametrik, terutama jika mengalami kesulitan bila menggunakan kedua metode tersebut. Dalam paper ini kami juga tawarkan suatu kelas estimasi densitas parametrik dan mencoba menemukan estimator yang terbaik dalam kelas ini. Pemunculan fungsi kernel akan memerlukan spesifikasi lebih lanjut tentang bandwith **H**, akan tetapi tidak kami bahas didalamnya. Secara tradisional, *Mean Integrated Square Error* (MISE) ataupun *Asymptotic MISE* (AMISE) sering digunakan sebagai alat evaluasi dari suatu estimator. Kami akan menggunakan AMISE untuk berdiskusi lebih lanjut tentang penemuan sebuah estimator baru.

ESTIMATOR DENSITAS SEMIPARAMETRIK

Estimasi fungsi densitas untuk observasi-observasi merupakan persoalan dasar dalam statistik, akan tetapi, suatu hal yang sulit. Keuntungan menggunakan metode parametrik adalah memungkinkan dalam mempelajari observasi-observasi itu dengan mudah. Akan tetapi jika observasi-observasi itu jauh dari model parametrik, maka didapati bahaya jika menggunakan metode parametrik dalam keakuratan densitas estimatornya. Dalam hal ini pendekatan secara nonparametrik perlu untuk dipikirkan. Dari kefleksibelan pendekatan nonparametrik, dan penyelesaian yang sederhana dari pendekatan parametrik, para ilmuan statistik berusaha untuk membuat metode gabungan nonparametrik dan parametrik, yaitu *metode semiparametrik*. Metode ini berkembang pesat saat ini.

Dalam paper ini kami menfokuskan estimasi densitas secara semiparametrik menggunakan metode Hjort dan Glad (1995) dan Hjort dan Jones (1996). Metode semiparametrik ini didesain lebih baik dibandingkan dengan estimator kernel (1) dalam lingkup nonparametrik dengan pemberian estimasi awal secara parametrik, seperti fungsi normal, yang hasilnya tidak kehilangan ketelitian terlalu banyak pada saat dihadapkan masalah fungsi densitas sebenarnya jauh dari kelas parametrik. Gagasannya adalah dengan mengalikan estimasi densitas parametrik dengan *fungsi koreksi*. Prosedur teknisnya adalah sebagai berikut.

Dimisalkan $g(\mathbf{x},\hat{\theta})$ adalah fungsi densitas parametrik yang merupakan estimasi kasar dari $f(\mathbf{x})$**.** Penjelasan khas gagasan ini adalah

$$f(\mathbf{x}) = g(\mathbf{x},\hat{\theta})\frac{f(\mathbf{x})}{g(\mathbf{x},\hat{\theta})} \equiv g(\mathbf{x},\hat{\theta})\xi(\mathbf{x}).$$

Jadi masalah pengestimasian fungsi densitas $f(\mathbf{x})$ secara semiparametrik adalah untuk menemukan estimator fungsi koreksi $\xi(\mathbf{x})$, yaitu

$$\hat{\xi}(\mathbf{x}) = \left(\widehat{f}(\mathbf{x}) / g(\mathbf{x},\hat{\theta}) \right).$$

Di bawah ini adalah contoh-contoh versi dari fungsi koreksi.

Contoh 1. **Estimator Hjort-Jones**

Dimisalkan $\mathbf{x}$ adalah sebuah titik. Konsep fundamental dari estimator Hjort-Jones yaitu menentukan optimum ξ sebagai estimatornya dengan meminimumkan fungsi tertimbang

$$q(\mathbf{x}, \xi) = \int K_{\mathbf{H}}(\mathbf{t} - \mathbf{x})\{f(\mathbf{t}) - g(\mathbf{t},\hat{\theta})\,\xi\}^2 d\mathbf{t}$$

Penimbang $K_{\mathbf{H}}$ $(\mathbf{t} - \mathbf{x})$ didesain secara lokal terhadap $\mathbf{t}$ yang dekat dengan $\mathbf{x}$. Pada prakteknya dengan meminimumkan

$$q_n(\mathbf{x}, \xi) = \xi^2 \int K_{\mathbf{H}}(\mathbf{t} - \mathbf{x})g(\mathbf{t},\hat{\theta})^2 d\mathbf{t} - \frac{2\xi}{n}\sum_{i=1}^{n} K_{\mathbf{H}}(\mathbf{X}_i - \mathbf{x})g(\mathbf{X}_i,\hat{\theta})$$

Nilai optimum ξ adalah

$$\hat{\xi} = \hat{\xi}_{HJ}(\mathbf{x}) = \frac{n^{-1}\sum_{i=1}^{n} K_{\mathbf{H}}(\mathbf{X_i} - \mathbf{x})g(\mathbf{X_i},\hat{\theta})}{\int K_{\mathbf{H}}(\mathbf{t} - \mathbf{x})g(\mathbf{t},\hat{\theta})^2 d\mathbf{t}}$$

Jadi didapatkan estimator $f(\mathbf{x})$ sebagaimana di bawah ini.

$$\hat{f}_{HJ}(\mathbf{x}) = g(\mathbf{x},\hat{\theta})\hat{\xi}_{HJ}(\mathbf{x}) = g(\mathbf{x},\hat{\theta})\frac{n^{-1}\sum_{i=1}^{n} K_{\mathbf{H}}(\mathbf{X_i} - \mathbf{x})g(\mathbf{X_i},\hat{\theta})}{\int K_{\mathbf{H}}(\mathbf{t} - \mathbf{x})g(\mathbf{t},\hat{\theta})^2 d\mathbf{t}}.$$

Ini adalah versi multivariat dari estimator yang didiskusikan dalam Hjort dan Jones (1996, Sesi 5.6).

Contoh 2. **Estimator Lokal Likelihood**

Gagasan utama dalam mendapatkan fungsi koreksi untuk contoh ini adalah dengan memaksimumkan fungsi *local log-likelihood*

$$L(\mathbf{x},\xi) = \int K_{\mathbf{H}}(\mathbf{t} - \mathbf{x})[\, f(\mathbf{t}) \log\{g(\mathbf{t},\hat{\theta})\,\xi\} - g(\mathbf{t},\hat{\theta})\,\xi\,]d\mathbf{t}.$$

Sebagai pengganti F yang tidak diketahui, kami menggunakan distribusi empiris F_n sebagaimana contoh 1, sehingga $L(\mathbf{x}, \xi)$ bisa di gantikan dengan

$$L_n(\mathbf{x},\xi) = \tilde{f}(\mathbf{x})\log\xi - \xi\int K_{\mathbf{H}}(\mathbf{t} - \mathbf{x})g(\mathbf{t},\hat{\theta})d\mathbf{t}.$$

Tanpa mengalami kesulitan, estimator Lokal Likelihood diperoleh dalam bentuk

$$\hat{f}_{LL}(\mathbf{x}) = g(\mathbf{x},\hat{\theta})\hat{\xi}_{LL}(\mathbf{x}) = g(\mathbf{x},\hat{\theta})\frac{\hat{f}(\mathbf{x})}{\int K_{\mathbf{H}}(\mathbf{t}-\mathbf{x})g(\mathbf{t},\hat{\theta})d\mathbf{t}},$$

di mana

$$\hat{\xi}_{LL}(\mathbf{x}) = \arg\max_{\xi} L_n(\mathbf{x},\xi) = \frac{\hat{f}(\mathbf{x})}{\int K_{\mathbf{H}}(\mathbf{t}-\mathbf{x})g(\mathbf{t},\hat{\theta})d\mathbf{t}},$$

ref. Hjort and Jones (1996, Sesi 5.5).

Contoh 3. **Estimator Hjort-Glad**

Hjort dan Glad menawarkan estimator :

$$\hat{f}_{HG}(\mathbf{x}) = g(\mathbf{x},\hat{\theta})\hat{\xi}_{HG}(\mathbf{x}) = g(\mathbf{x},\hat{\theta})\,\frac{1}{n}\sum_{i=1}^{n}\frac{K_{\mathbf{H}}(\mathbf{X}_i-\mathbf{x})}{g(\mathbf{X_i},\hat{\theta})},$$

di mana

$$\hat{\xi}_{HG}(\mathbf{x}) = \left(\widehat{f(\mathbf{x})/g(\mathbf{x},\hat{\theta})}\right)$$

$$= \frac{1}{n}\sum_{i=1}^{n}\frac{K_{\mathbf{H}}(\mathbf{X}_i-\mathbf{x})}{g(\mathbf{X_i},\hat{\theta})}.$$

Ini adalah estimator yang alami untuk $\hat{\xi}(\mathbf{x})$, ref. Hjort and Glad (1995, Sesi 7).

Mereka berargumentasi dalam papernya bahwa estimator mereka bekerja baik menggunakan 15 *test densitas* list dalam paper Marron dan Wand (1992) dengan start normal dalam kasus univariat atau d = 1. Estimator ini memenangkan 12 dari 15 test dibanding dengan metode kernel yang biasanya.

Perbedaan dari estimator-estimator itu terletak kepada fungsi koreksinya. Yang menarik, fungsi-fungsi koreksi itu seolah-olah mengikuti suatu ketentuan khusus. Ini memberikan motivasi kepada kami untuk menawarkan suatu bentuk umum dari estimator-estimator di atas.

Dimisalkan $g(\mathbf{x},\hat{\theta})$ adalah estimasi densitas parametrik. Ini adalah estimasi kasar dari $f(\mathbf{x})$. Dan $\mathbf{x}$ ditentukan sebagai titik tetap. Kami menawarkan fungsi

$$Q(\mathbf{x},\xi\,|\,\alpha) = \int K_{\mathbf{H}}(\mathbf{t}-\mathbf{x})\frac{\{f(\mathbf{t})-g(\mathbf{t},\hat{\theta})\xi\}^2}{g(\mathbf{t},\hat{\theta})^{\alpha}}\,d\mathbf{t}$$

untuk mendapatkan estimator dari ξ. Catatan bahwa $g(\mathbf{t},\hat{\theta})^{-\alpha}$ adalah penimbang lain pada fungsi $q(\mathbf{x},\xi)$ dalam contoh 1, di mana $\alpha \geq 0$. Secara praktek, estimator dari ξ dapat diperoleh dengan mengoptimumkan

$$Q_n(\mathbf{x},\xi\,|\,\alpha) = \xi^2\int K_{\mathbf{H}}(\mathbf{t}-\mathbf{x})g(\mathbf{t},\hat{\theta})^{2-\alpha}\,d\mathbf{t} - \frac{2\xi}{n}\sum_{i=1}^{n}\int K_{\mathbf{H}}(\mathbf{X_i}-\mathbf{x})g(\mathbf{X_i},\hat{\theta})^{1-\alpha}.$$

Akhirnya kami menemukan estimator baru:

$$\hat{f}_{\alpha}(\mathbf{x}) = g(\mathbf{x},\hat{\theta})\hat{\xi}_{\alpha}(\mathbf{x}) = g(\mathbf{x},\hat{\theta})\frac{n^{-1}\sum_{i=1}^{n}K_{\mathbf{H}}(\mathbf{X}_i-\mathbf{x})g(\mathbf{X}_i,\hat{\theta})^{1-\alpha}}{\int K_{\mathbf{H}}(\mathbf{t}-\mathbf{x})g(\mathbf{t},\hat{\theta})^{2-\alpha}\,d\mathbf{t}} \qquad (2)$$

di mana

$$\hat{\xi}_{\alpha}(\mathbf{x}) = \arg\min_{\xi} Q_n(\mathbf{x},\xi\,|\,\alpha)$$

$$= \frac{n^{-1}\sum_{i=1}^{n}K_{\mathbf{H}}(\mathbf{X}_i-\mathbf{x})\,g(\mathbf{X}_i,\hat{\theta})^{1-\alpha}}{\int K_{\mathbf{H}}(\mathbf{t}-\mathbf{x})g(\mathbf{t},\hat{\theta})^{2-\alpha}\,d\mathbf{t}}.$$

Kami menegaskan (2) merupakan suatu kelas umum untuk estimasi semiparametrik. Keuntungan dalam penggunaan fungsi $Q_n(\mathbf{x},\xi\,|\,\alpha)$ adalah untuk membuat suatu kelas yang mencakup ketiga estimator di atas, dengan kata lain

$$\hat{f}_0(\mathbf{x}) = \hat{f}_{HJ}(\mathbf{x}),\ \hat{f}_1(\mathbf{x}) = \hat{f}_{LL}(\mathbf{x}),\ \hat{f}_3(\mathbf{x}) = \hat{f}_{HG}(\mathbf{x}).$$

SIFAT ESTIMATOR

Dalam sesi ini kami mempelajari bias dan variansi dengan menentukan beberapa asumsi pada g, f, $\mathbf{H}$, dan K sebagai berikut :

(A1) $\frac{\partial^2}{\partial x_i \partial x_j} f(\mathbf{x})$ dan $\frac{\partial^2}{\partial x_i \partial x_j} g(\mathbf{x},\theta)$ adalah kontinyu untuk i, $j = 1,2,\ldots,d$

(A2) $\mathbf{H} = \mathrm{diag}(h_1^2,\ldots, h_d^2)$ sehingga $n^{-1}|\mathbf{H}|^{-1/2} = (nh_1\ldots h_d)^{-1}$ dan semua elemen $\mathbf{H}$ mendekati nol jika $n \to \infty$.

(A3) K merupakan fungsi kernel variat-d yang *bounded* dan *compact*, dengan memenuhi

$$K(\mathbf{x}) = \prod_{i=1}^{d} K(x_j),$$

$$\int K(\mathbf{z})d\mathbf{z} = 1, \int \mathbf{z}K(\mathbf{z})d\mathbf{z} = \mathbf{0}$$

dan

$$\int \mathbf{z}\mathbf{z}^T K(\mathbf{z})d\mathbf{z} = \mathrm{diag}\,(\sigma(K_1)^2,\ldots,\sigma(K_d)^2).$$

Tentu saja penghitungan untuk bias dan variansi estimator (2) cukup sulit, dikarenakan terdapat multidimensi yang ganda yaitu parameter dan observasi itu sendiri. Sebagai referensi yang umum, diketahui bahwa ada satu fungsi $T(\theta)$ sehingga $\hat{\theta} = T(F_n)$ dan $\theta = T(F)$ dengan fungsi *influence*

$$I(\mathbf{x}) = \lim_{\varepsilon\to 0}\{T((1-\varepsilon)F + \varepsilon\delta_{\mathbf{x}}) - T(F)\}/\varepsilon$$

di mana $\delta_{\mathbf{x}}$ adalah massa unit pada titik $\mathbf{x}$. Dalam kondisi regular, didapat

$$\hat{\theta} = \theta_0 + \frac{1}{n}\sum_{i=1}^{n} I(\mathbf{X}_i) + \frac{b}{n} + \varepsilon_n,$$

di mana $\gamma_n = Op(n^{-1})$ dengan rata-rata $O(n^{-2})$ dan b/n adalah bias dari $\hat{\theta}$. Disini $I_I = I(\mathbf{X}_I)$ mempunyai rata-rata nol, ref. Lehmann (1983). Pada hal maximum likelihood, diketahui bahwa

$$I(\mathbf{x}) = \Im^{-1}\frac{\partial \log g(\mathbf{x},\theta_{\mathbf{0}})}{\partial\theta\partial\theta^{T}} \quad \text{di mana} \quad \Im = \mathrm{E}_0\left[-\frac{\partial^2 \log g(\mathbf{x},\theta_{\mathbf{0}})}{\partial\theta\partial\theta^{T}}\right].$$

Lemma 1. *Dimisalkan* $g_0(\mathbf{x}) = g(\mathbf{x}, \theta_0)$ *adalah estimasi pendekatan parametrik yang sudah ditetapkan dan terbaik untuk* $f(\mathbf{x})$. *Berdasarkan asumsi* (A1)-(A3), *estimator*

$$f_\alpha(\mathbf{x})^* = g_0(\mathbf{x})\frac{n^{-1}\sum_{i=1}^{n} K_{\mathbf{H}}(\mathbf{X_i}-\mathbf{x})g_0(\mathbf{X_i})^{1-\alpha}}{\int K_{\mathbf{H}}(\mathbf{t}-\mathbf{x})g_0(\mathbf{t})^{2-\alpha}\,d\mathbf{t}}$$

mempunyai sifat

$$\text{Bias } f_\alpha(\mathbf{x})^* = \frac{1}{2}\sum_{j=1}^{d} h_j^2\sigma(K_j)^2 B_{\alpha,j}(\mathbf{x}) + O\left(\sum_{j=1}^{d} h_j^4\right), \tag{3}$$

$$\text{Var } f_\alpha(\mathbf{x})^* = \frac{R(K)f(\mathbf{x})}{n\prod_{j=1}^{d} h_j} - \frac{f(\mathbf{x})^2}{n} + O\left(\frac{\sum_{j=1}^{d} h_j^2}{n\prod_{j=1}^{d} h_j}\right), \tag{4}$$

di mana $B_{\alpha,j}(\mathbf{x}) = \dfrac{\mathrm{H}_{\mathbf{x}}[g_0(\mathbf{x})^{1-\alpha} f(\mathbf{x})]_{jj}\, g_0(\mathbf{x}) - \mathrm{H}_{\mathbf{x}}[g_0(\mathbf{x})^{2-\alpha}]_{jj}\, f(\mathbf{x})}{g_0(\mathbf{x})^{2-\alpha}}.$

Lemma di atas menerangkan bahwa parameter yang digunakan sudah ditentukan dengan kata lain *fixed start*, dan ini bisa diaplikasikan untuk penghitungan bias dan variansi jika menggunakan start parametrik. Dalam kasus $g_0(\mathbf{x})$ berbentuk distribusi *uniform*, akan memberikan bentuk estimator seperti estimator kernel (1).

Teorema 2. *Berdasarkan asumsi* (A1)-(A3), *estimator semiparametrik* (2) *mempunyai sifat*

$$\text{Bias } \hat{f}(\mathbf{x}) = \frac{1}{2}\sum_{j=1}^{d} h_j^2\sigma(K_j)^2 B_{\alpha,j}(\mathbf{x}) + O\left(\sum_{j=1}^{d}\left(h_j^4 + \frac{h_j^2}{n}\right) + n^{-2}\right), \tag{5}$$

$$\text{Var } \hat{f}(\mathbf{x}) = \frac{R(K)f(\mathbf{x})}{n\prod_{j=1}^{d} h_j} - \frac{f(\mathbf{x})^2}{n} + O\left(\frac{\sum_{j=1}^{d} h_j^2}{n\prod_{j=1}^{d} h_j} + n^{-2}\right). \tag{6}$$

Pembuktian untuk Lemma 1 dan Teorema 2 tidak disertakan dalam paper ini karena sangat panjang. Teorema 2 menerangkan bahwa variansi tidak tergantung pada α. Fakta ini akan memberikan diskusi menarik dalam pengembangan suatu estimator baru dalam kelas ini.

Remark 3. *Jika* $f(\mathbf{x}) = \{g(\mathbf{x}, \theta) \mid \theta \in \Theta\}$, *atau* $f(\mathbf{x}) = g_0(\mathbf{x})$ *maka bagian dari bias* $\hat{f}_\alpha(\mathbf{x})$ *yaitu* $O\left(\sum_{j=1}^{d} h_j^2\right)$ *menjadi nol.*

EVALUASI

Untuk mengevaluasi estimator, kami menggunakan MISE. MISE adalah ukuran *smoothing* yang merupakan nilai harapan dari integral kwadrat dari selisih estimator dan targetnya. Dalam hal ini,

MISE dari $\hat{f}_\alpha$ adalah

$$\mathrm{MISE}\left\{\hat{f}_\alpha\right\} = \mathrm{E}\int\left(\hat{f}_\alpha(\mathbf{x}) - f(\mathbf{x})\right)^2 d\mathbf{x}$$

MISE ini mempunyai dua komponen yaitu integral dari variansi dan bias kwadrat, yaitu

$$\begin{aligned}\mathrm{MISE}\left\{\hat{f}_\alpha\right\} &= \int\left\{\mathrm{Bias}\ \hat{f}_\alpha(\mathbf{x})\right\}^2 d\mathbf{x} + \int \mathrm{Var}\ \hat{f}_\alpha(\mathbf{x})\,d\mathbf{x} \\ &= \mathrm{AMISE}\left\{\hat{f}_\alpha\right\} + o\left(\prod_{j=1}^{d}\frac{1}{nh_j} + h_j^4\right)\end{aligned}$$

di mana

$$\mathrm{AMISE}\left\{\hat{f}_\alpha\right\} = \frac{R(K)}{n\prod_{j=1}^{d} h_j} + \int\left\{\frac{1}{2}\sum_{j=1}^{d} h_j^2 \sigma(K_j)^2 B_{\alpha,j}(\mathbf{x})\right\}^2 d\mathbf{x}.$$

Dalam paper ini, kami akan memakai nilai AMISE untuk menemukan suatu estimator baru yang bisa bersaing dengan estimator lain dalam kelas ini. Dari hasil nilai AMISE, diketahui bahwa variansi tidak tergantung pada α , sehingga kami tertarik kepada integral dari bias kwadrat. Target kami adalah untuk menyelidiki α . Apakah mungkin bisa menemukan optimum dari α pada nilai AMISE sehingga diperoleh suatu estimator terbaik dalam kelas (2).

Pertama perhatikan bagian dari bias yaitu $B_{\alpha,j}(\mathbf{x})$. Kami dapatkan

$$B_{\alpha,j}(\mathbf{x}) = B_j^{(1)}(\mathbf{x}) - \alpha B_j^{(2)}(\mathbf{x}) \tag{7}$$

di mana

$$B_j^{(1)}(\mathbf{x}) = \frac{2}{g_0(\mathbf{x})}\frac{\partial g_0(\mathbf{x})}{\partial x_j}\frac{\partial f(\mathbf{x})}{\partial x_j} + \frac{\partial^2 f(\mathbf{x})}{\partial x_j^2} - \frac{f(\mathbf{x})}{g_0(\mathbf{x})}\frac{\partial^2 g_0(\mathbf{x})}{\partial x_j^2} - \frac{2f(\mathbf{x})}{g_0(\mathbf{x})^2}\left(\frac{\partial g_0(\mathbf{x})}{\partial x_j}\right)^2$$

$$B_j^{(2)}(\mathbf{x}) = \frac{2}{g_0(\mathbf{x})}\frac{\partial g_0(\mathbf{x})}{\partial x_j}\frac{\partial f(\mathbf{x})}{\partial x_j} - \frac{2f(\mathbf{x})}{g_0(\mathbf{x})^2}\left(\frac{\partial g_0(\mathbf{x})}{\partial x_j}\right)^2$$

Maka (5) bisa ditulis seperti berikut:

$$\text{Bias } \hat{f}_\alpha(\mathbf{x}) = \sum_{j=1}^{d}\frac{h_j^2\sigma(K_j)^2}{2}\left\{B_j^{(1)}(\mathbf{x}) - \alpha B_j^{(2)}(\mathbf{x})\right\} + O\left(\sum_{j=1}^{d}\left(h_j^4 + \frac{h_j^2}{n}\right) + n^{-2}\right). \tag{8}$$

Dari hitungan di atas, adalah sangat sulit untuk mendapatkan nilai optimum α karena masih bercampur dengan elemen *bandwidth* **H**. Oleh sebab itu perlu ditentukan semua elemennya sama yaitu $h_i = h$ untuk $i = 1, 2, \ldots, d$. Dari persamaan (6) dan (8), AMISEnya bisa ditulis sebagai

$$\text{AMISE}\left\{\hat{f}_\alpha\right\} = h^4\left\{C^{(1)}\alpha^2 - 2C^{(2)}\alpha + C^{(3)}\right\} + \frac{R(K)}{nh^d}, \tag{9}$$

di mana

$$C^{(1)} = \frac{1}{4}\sum_{i=1}^{d}\sum_{j=1}^{d}\sigma(K_i)^2\sigma(K_j)^2\int B_i^{(2)}(\mathbf{x})B_j^{(2)}(\mathbf{x})d\mathbf{x},$$

$$C^{(2)} = \frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\sigma(K_i)^2\sigma(K_j)^2\int B_i^{(1)}(\mathbf{x})B_j^{(2)}(\mathbf{x})d\mathbf{x},$$

$$C^{(3)} = \frac{1}{4}\sum_{i=1}^{d}\sum_{j=1}^{d}\sigma(K_i)^2\sigma(K_j)^2\int B_i^{(1)}(\mathbf{x})B_j^{(1)}(\mathbf{x})d\mathbf{x},$$

di mana $B_i^{(1)}$ dan $B_i^{(2)}$ seperti yang dijelaskan dalam (7).

Kita menekankan alasan mengapa menggunakan parameter α, karena bentuk dari bias (8) mengarahkan bias kwadratnya pada bentuk $C^{(1)}\alpha^{(2)}$-$2C^{(2)}\alpha+C^{(3)}$, yang mengartikan bahwa ada α_0 yang optimum, sehingga

adalah estimator ter baik dalam kelas ini. Akhirnya $\hat{f}_{\alpha_0}$ didapatkan nilai optimumnya jika f $\alpha_0 = C^{(2)} / C^{(1)}$.

PERBANDINGAN SECARA ASIMTOT

Tujuan sesi ini adalah untuk membandingkan secara asimtot antara nilai AMISE estimator baru (2) dengan estimator kernel tradisional (1). Untuk itu kami gunakan kondisi d = 2 dan $\sigma(K_i) = \sigma(K)$ untuk i = 1, 2. Ini mengartikan bahwa perbandingan ini tidak dipengaruhi oleh fungsi kernel yang digunakan. Dari (9), AMISE dari estimator baru menjadi

$$\text{AMISE}\left\{\hat{f}_\alpha\right\} = h^4\sigma(K)^4\left\{C^{(1)}\alpha^2 - 2C^{(2)}\alpha + C^{(3)}\right\} + \frac{R(K)}{nh^2} \qquad (10)$$

di mana

$$C^{(1)} = \frac{1}{4}\int\sum_{i=1}^{2}\sum_{j=1}^{2}B_i^{(2)}(\mathbf{x})B_j^{(2)}(\mathbf{x})d\mathbf{x},$$

$$C^{(2)} = \frac{1}{2}\int\sum_{i=1}^{2}\sum_{j=1}^{2}B_i^{(1)}(\mathbf{x})B_j^{(2)}(\mathbf{x})d\mathbf{x},$$

$$C^{(3)} = \frac{1}{4}\int \sum_{i=1}^{2}\sum_{j=1}^{2} B_i^{(1)}(\mathbf{x})B_j^{(1)}(\mathbf{x})d\mathbf{x}.$$

Sementara itu AMISE dari estimator kernel tradisional (1) dengan kondisi yang sama, mempunyai bentuk

$$\mathrm{AMISE}\left\{\tilde{f}\right\} = h^4\sigma(K)^4 C + \frac{R(K)}{nh^2} \tag{11}$$

di mana

$$C = \frac{1}{4}\int \sum_{i=1}^{2}\sum_{j=1}^{2} \frac{\partial^2 f(\mathbf{x})}{\partial x_i^2}\frac{\partial^2 f(\mathbf{x})}{\partial x_j^2} d\mathbf{x}.$$

Untuk perbandingan AMISE, adalah cukup untuk membandingkan bagian $C^{(1)}\alpha^2$-$2C^{(2)}\alpha + C^{(3)}$ dalam persamaan (10) dengan C pada persamaan (11) Estimator yang terbaik adalah estimator yang mempunyai AMISE paling kecil dibanding dengan yang lain. Jika diberikan contoh kongkrit dari fungsi $f(\mathbf{x})$ and $g(\mathbf{x})$, maka keduanya bisa dihitung dan dibandingkan. Perbandingan disini ditujukan untuk menentukan estimator mana yang lebih cocok untuk dipakai. Dalam paper ini akan dibuktikan bahwa estimator baru kami lebih baik dibanding dengan metode kernel tradisional, dengan menggunakan fungsi normal sebagai fungsi parametrik awalnya, dan $f(\mathbf{x})$ termasuk *three normal mixtures*, dengan kata lain,

$$f(\mathbf{x}) = p_1\phi_{\Sigma_1}(\mathbf{x}-\mu_1) + p_2\phi_{\Sigma_2}(\mathbf{x}-\mu_2) + p_3\phi_{\Sigma_3}(\mathbf{x}-\mu_3)$$

di mana $p_1 \geq 0$, $p_1 + p_2 + p_3 = 1$, $\mu_I = (\mu_{I,1}, \mu_{I,2})$, $\Sigma_i = \sigma_I \mathbf{I}$. Dengan rata-rata dan variansi dari $f(\mathbf{x})$:

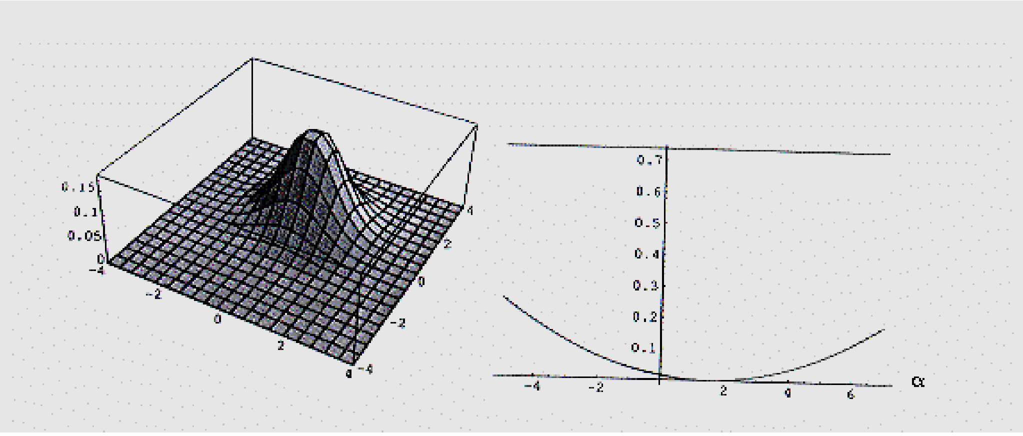
$\mu_0 = p_1\mu_1 + p_2\mu_2 + p_3\mu_3,$

$$\Sigma_0 = \sum_{i=1}^{3} p_1\left\{\Sigma_i + \mu_i\mu_i^T\right\} - \left\{\sum_{i=1}^{3} p_i\mu_i\right\}\left\{\sum_{i=1}^{3} p_i\mu_i\right\}^T$$

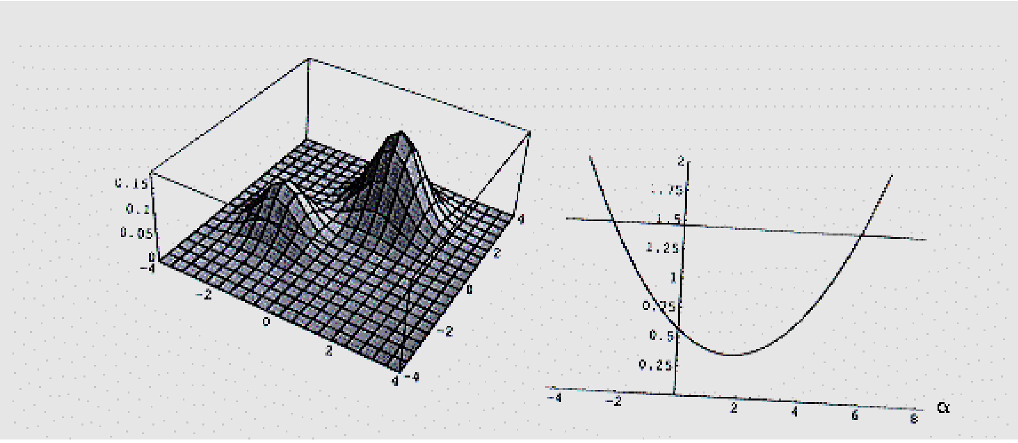
yang merupakan parameter terbaiknya. Disini digunakan fungsi parametrik awal yaitu sebuah fungsi normal. Sehingga parametrik awalnya adalah $g_0(\mathbf{x}) = g(\mathbf{x}, \mu_0, \mu_0, \Sigma_0)$, sebab rata-rata dan variansi distribusi normal tersebut membuat jarak pendek pada densitas sesungguhnya $f(\mathbf{x})$ untuk *Kulback-Leibler distance*. Pada kondisi demikian bisa didapatkan nilai integral dari bias kwadrat untuk membandingkannya.

Beberapa contoh perbandingan nilai AMISE estimator baru dengan estimator kernel.

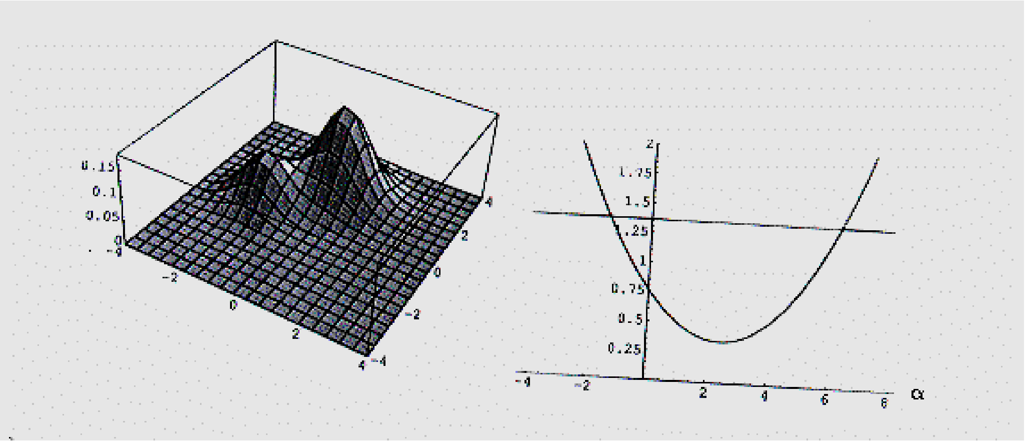
Grafik kiri: pemisalan fungsi densitas sebenarnya dalam *three normal mixtures.*
Grafik kanan: perbandingan graph C dengan $C^{(1)}a^{2}-2C^{(2)}a+C^{(3)}$.



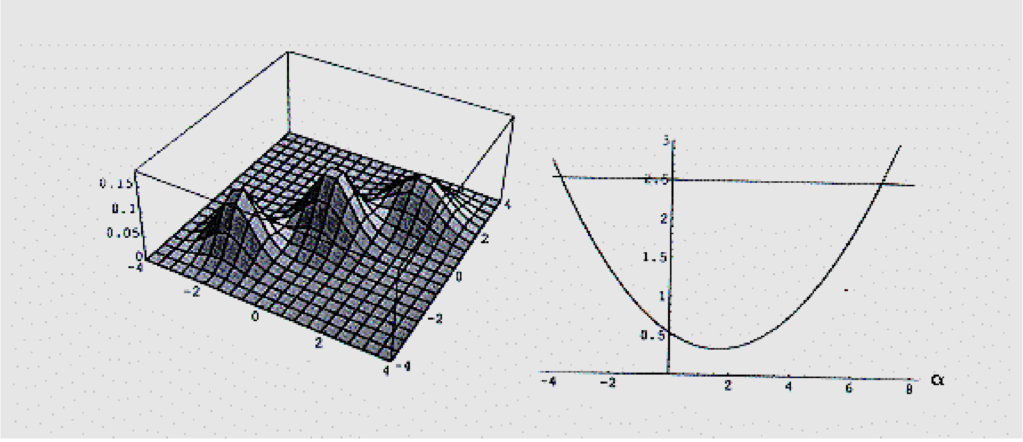
Graph 1: $p_1 = 0$, $p_2 = 0$, $\mu_{3,1} = 0$, $\mu_{3,2} = 1$, $\sigma_3 = 0{,}9$; $\alpha_0 = 1{,}5$



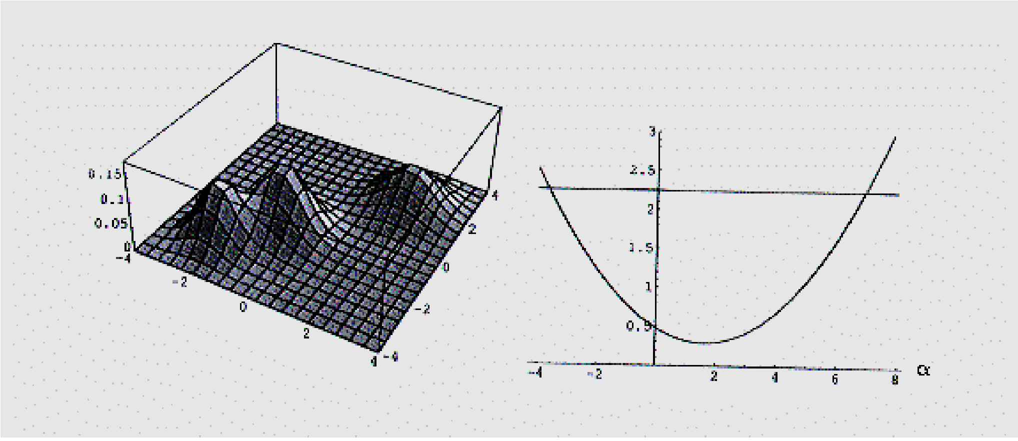
Graph 2: $p_1 = 0$, $p_2 = 0{,}3$, $\mu_{2,1} = -1{,}9$, $\mu_{2,2} = -0{,}6$, $\mu_{3,1} = 0{,}6$, $\mu_{3,2} = 1{,}9$, $\sigma_2 = 0{,}7$, $\sigma_3 = 0{,}8$; $\alpha_0 = 1{,}74172$



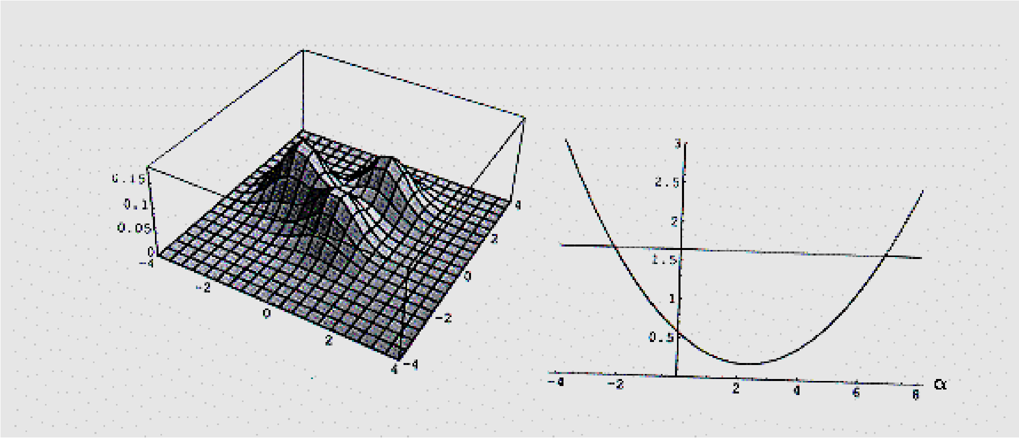
Graph 3: $p_1 = 0$, $p_2 = 0{,}3$, $\mu_{2,1} = -1{,}85$, $\mu_{2,2} = 0{,}1$, $\mu_{3,1} = -0{,}1$, $\mu_{3,2} = 1{,}85$, $\sigma_2 = 0{,}7$, $\sigma_3 = 0{,}8$; $\alpha_0 = 2{,}43908$



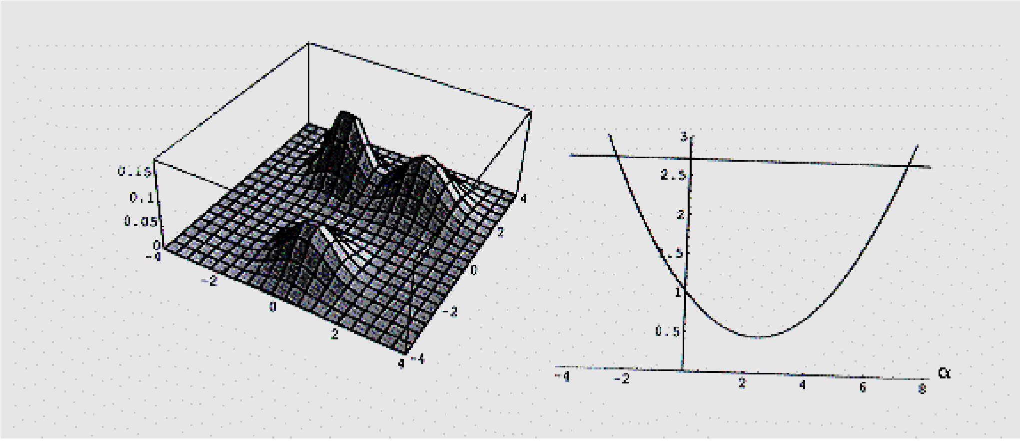
Graph 4: $p_1 = 0{,}33$, $p_2 = 0{,}4$, $\mu_{1,1} = 2$, $\mu_{1,2} = 2$, $\mu_{2,1} = 0$, $\mu_{2,2} = 0$, $\mu_{3,1} = -2$, $\mu_{3,2} = -2$, $\sigma_1 = 0{,}8$, $\sigma_2 = 0{,}7$, $\sigma_3 = 0{,}6$; $\alpha_0 = 1{,}60167$



Graph 5: $p_1 = 0{,}33$, $p_2 = 0{,}4$, $\mu_{1,1} = 2{,}1$, $\mu_{1,2} = 2{,}1$, $\mu_{2,1} = -0{,}8$, $\mu_{2,2} = -0{,}8$, $\mu_{3,1} = -2{,}2$, $\mu_{3,2} = -2{,}2$, $\sigma_1 = 0{,}8$, $\sigma_2 = 0{,}7$, $\sigma_3 = 0{,}6$; $\alpha_0 = 1{,}62995$



Graph 6: $p_1 = 0{,}393$, $p_2 = 0{,}31$, $\mu_{1,1} = -1{,}85$, $\mu_{1,2} = 0{,}5$, $\mu_{2,1} = 0{,}5$, $\mu_{2,2} = 1{,}85$, $\mu_{3,1} = 0$, $\mu_{3,2} = -0{,}5$, $\sigma_1 = 0{,}7$, $\sigma_2 = 0{,}7$, $\sigma_3 = 0{,}7$; $\alpha_0 = 2{,}29352$



Graph 7: $p_1 = 0{,}33$, $p_2 = 0{,}4$, $\mu_{1,1} = -1{,}3$, $\mu_{1,2} = 2$, $\mu_{2,1} = 1{,}5$, $\mu_{2,2} = 2$, $\mu_{3,1} = 0$, $\mu_{3,2} = -2{,}3$, $\sigma_1 = 0{,}6$, $\sigma_2 = 0{,}8$, $\sigma_3 = 0{,}7$; $\alpha_0 = 2{,}39882$

KESIMPULAN

Metode kernel tradisional bekerja lebih baik dibanding dengan metode parametrik, akan tetapi dari bukti-bukti simulasi kami, estimator baru kami lebih baik dari metode kernel. Lebih dari itu estimator tersebut merupakan suatu kelas estimasi semiparametrik yang mencakup estimator Hjort-Jones, estimator Lokal Likelihood dan estimator Hjort-Glad. Kesimpulan akhir, kami tegaskan bahwa ada estimator yang terbaik dalam kelas ini.

UCAPAN TERIMA KASIH

Kami sangat berterima kasih kepada Prof. Kanto Naito, Shimane Univ. dan

Prof. Maretsugu Yamasaki, Shimane Univ., yang mendorong kami dalam penelitian ini. Demikian pula kepada Sri Budianti M.S. (Deputi Perencanaan dan Analisa Statistik-BPS) dan Herminto Widodo, B. Sc., (staf Metodologi Statistik BPS) yang telah membantu meninjau paper kami.

DAFTAR PUSTAKA

1. HJORT, N. L., GLAD, I. K., "Nonparametric Density Estimation with a Parametric Start", Ann. Statist. **23** (1995) 882-904

2. HJORT, N. L., JONES, M. C., "Locally Parametric Density Estimation", Ann. Statist., **24** (1996) 1619-1647

3. LEHMANN, E.L, *Theory of Point Estimation*, John Wiley & Sons, (1983)

4. MARRON, J. S. AND WAND, M. P., "Exact Mean Integrated Squared Error", Ann.. Statist., **20** (1992) 712-736

5. WAND, M. P., JONES, M. C., *Kernel Smoothing*, Publ. By Chapman & Hall, London, UK, (1995)

DISKUSI

MUSTANGIMAH

1. Dalam situasi *sampling* yang bagaimana, metode semiparametrik dapat diterapkan?
2. Dari evaluasi yang Anda lakukan, apakah sisaan (*error*) yang diperoleh memenuhi azas acak (*random*)?

MOCHAMAD SONHAJI

1. Pada dasarnya semiparametrik *approach* memberikan solusi yang lebih tepat terhadap semua data tanpa batasan kondisi. Jikalau data tersebut mempunyai densitas yang cenderung kepada parametrik model maupun nonparametrik model, adalah tidak masalah. Sebab semiparametrik *approach* memberikan jaminan kearah itu. Lebih dari itu solusi dengan semiparametrik memberikan penawaran *estimator* yang lebih baik dibanding dengan parametrik *approach* ataupun semiparametrik *approach*.
2. Kami selalu memakai *random* data, sehingga estimator yang ditawarkan tidak tergantung pada data itu sendiri.

ARKO

1. Sudahkah Anda membandingkan *estimator* anda dengan *estimator* lain menggunakan data lapangan? Benarkah *estimator* anda terbaik berdasarkan perbandingan tersebut?
2 Apa dasar pemilihan nilai α untuk *estimator* anda? Bagaimana bila $\alpha > 2$ atau $\alpha > 0$? Berapakah nilai α yang terbaik?

MOCHAMAD SONHAJI

1. Kami membatasi *estimator* yang ditawarkan menggunakan Fungsi Kernel. Tentunya ada beberapa motode semiparametrik yang lain. Kami lebih memfokuskan kepada *estimator* secara semiparametrik yaitu *Estimator Hjort Jones*, *Estimator Local Likehood*, dan *Estimator Hjort-Glad*. Berdasarkan

temuan kami, ditemukan suatu khas semiparametrik yang mencakup 3 *estimato*r tersebut Kemudian kami mengklaim bahwa ada *estimator* yang terbaik dalam kelas ini . Selanjutnya *estimator* kami, dibandingkan dengan tradisional *Kernel estimator* adalah jauh lebih baik, dengan catatan disini simulasinya menggunakan *true density* yang tergolong dalam *three normal mixtures*. Data yang digunakan adalah *random*, bisa digunakan data lapangan.

2. α tersebut untuk membentuk suatu kelas dalam semiparametrik model. Nilai α sebenarnya bisa nilai *real* R. Seperti yang telah kami jelaskan dalam simulasi paper kami, bahwa ada α_o (α optimum) yang membentuk $\hat{f}_{\alpha o}(\mathbf{x})$ yaitu *estimator* yang terbaik dalam kelas semiparametrik yang kami temukan.

DAFTAR RIWAYAT HIDUP

1. Nama : MOCHAMAD SONHAJI
2. Tempat/Tanggal Lahir : Surabaya, 27 Juni 1973
3. Instansi : Badan Pusat Statistik (BPS)
4. Pekerjaan / Jabatan : Staf Desain Statistik Biro Metodologi
5. Riwayat Pendidikan : (setelah SMA sampai sekarang)
   - Shimane Univ. (major Mathematics) 1993-1997 (S1)
   - Shimane Univ. (major Mathematics) 1997-1999 (S2)
   - Bahasa Jepang dan Bahasa Inggris
6. Pengalaman Kerja : –
7. Organisasi Professional :
   - Shimane Muslim Association (SMA) : 1995-1997 IMAM SMA
   - Shimane Muslim Association (SMA) : 1997-1998 President SMA
   - Islamic Guidance Society (IGS) : 1997-1999 Editor jurnal Islam Al-Irshaad
   - Muslim Student Association -Japan (MSA-J) : 1997-1998