

Estimasi Parameter Model *Mixture Of Mixture* Untuk Pengeluaran Rumah Tangga Pada Data Susenas Kota Semarang

Zaenal Abidin

Jurusan Matematika FMIPA Universitas Negeri Semarang

Abstrak

Badan Pusat Statistik (BPS) memiliki tanggung jawab untuk menyediakan data yang dibutuhkan dalam perencanaan pembangunan. Beberapa data sosial yang dihasilkan BPS diperoleh melalui Survei Sosial Ekonomi Nasional (Susenass). Susenas didesain untuk mendapatkan data sosial penduduk dalam lingkup yang luas. Tujuan dari penelitian ini adalah untuk memperoleh suatu fungsi distribusi *mixture of mixture* yang dibentuk dari fungsi densitas rumah tangga dalam bentuk *mixture*.

Dalam penelitian ini, data yang akan digunakan adalah data persentase pengeluaran rata-rata per kapita sebulan menurut jenis pengeluaran dan golongan pengeluaran per kapita sebulan dari data Susenas Kota Semarang Juli 2007 kota Semarang, untuk daerah perkotaan saja. Sebelum dilakukan analisis, dilakukan eksplorasi data dengan mengelompokkan data ke dalam tiap jenis pengeluaran rumah tangga. Tahap berikutnya adalah menentukan distribusi data dari jenis pengeluaran rumah tangga. Kemudian menentukan nilai parameter distribusi prior setiap jenis pengeluaran rumah tangga. Distribusi prior yang digunakan adalah prior konjugate (sekawan). Selanjutnya menentukan masing-masing fungsi posterior untuk masing-masing fungsi sebaran pada fungsi *mixture* wilayah dengan Bayesian – MCMC dengan menggunakan paket program WinBugs 1.4. Terakhir dilakukan estimasi fungsi sebaran *mixture of mixture* untuk tiap jenis pengeluaran rumah tangga. Dalam penelitian ini model *mixture of mixture* yang dihasilkan adalah: $z(x|\theta, \pi) = 0.5011f_1(x|\theta_1) + 0.4989f_2(x|\theta_2)$.

Kata kunci: MCMC, *Mixture of Mixture*, Susenas.

A. Pendahuluan

Badan Pusat Statistik (BPS) memiliki tanggung jawab untuk menyediakan data yang dibutuhkan dalam perencanaan pembangunan. Beberapa data sosial yang dihasilkan BPS diperoleh melalui Survei Sosial Ekonomi Nasional (Susenass). Melalui Susenas, kemajuan pembangunan dapat dievaluasi secara berkelanjutan karena ketersediaan data setiap tahunnya. Susenas didesain untuk mendapatkan data sosial penduduk dalam lingkup yang luas. Data yang disajikan dalam Susenas meliputi tingkat kabupaten/kota. Estimasi populasi juga dilakukan sampai tingkat tersebut. Dasar pengambilan sampel Susenas adalah pengeluaran penduduk dari suatu rumah tangga. Pengeluaran ini digunakan sebagai kerangka untuk menentukan sampel dengan *linear systemic*

sampling. Dengan mempertimbangkan metode pengambilan sampel tersebut, maka akan diteliti besarnya peran serta masing-masing jenis pengeluaran rumah tangga dalam memberikan peran terhadap sosial ekonomi masyarakat kota Semarang dengan menggunakan model *mixture*. Selanjutnya penggabungan fungsi distribusi *mixture* yang terbentuk antar jenis pengeluaran rumah tangga digunakan distribusi *mixture of mixture*.

Permasalahan yang akan dibahas adalah bagaimana fungsi distribusi data Susenas kota Semarang dengan pendekatan *mixture* menggunakan analisis Bayesian dan Markov Chain Monte Carlo (MCMC). Tujuan dari penelitian ini adalah untuk memperoleh suatu fungsi distribusi *mixture of mixture* yang dibentuk dari fungsi densitas rumah tangga dalam bentuk *mixture*.

B. Kajian Pustaka

1. Fungsi Kepadatan Peluang

Fungsi kepadatan peluang (*probability density function*) merupakan suatu konsep dasar dari ilmu statistika, karena melalui fungsi kepadatan peluang dapat diperoleh gambaran tentang distribusi dari suatu variabel acak tertentu. Jika terdapat data pengamatan yang berasal dari suatu fungsi kepadatan peluang yang tidak diketahui, maka untuk mengetahui distribusi data tersebut diperlukan estimasi fungsi kepadatan peluang berdasarkan data pengamatan. Pengertian fungsi kepadatan peluang secara umum dapat dijelaskan sebagai berikut.

Definisi 1.

Fungsi kepadatan peluang $f(x)$ merupakan gambar dan distribusi probabilitas dari peubah acak tertentu, misalkan: X . Berdasarkan definisi tersebut, sifat-sifat dari suatu fungsi kepadatan peluang $f(x)$ adalah:

1. $0 \leq f(x) \leq 1$
2. $\int_{-\infty}^{\infty} f(x) dx = 1$ (1)
3. $P(a < X < b) = \int_a^b f(x) dx$

(Walpole, 1989)

2. Goodness of Fit Test

Uji keselarasan (*goodness of fit test*) biasanya melibatkan pengujian sebuah sampel random dari beberapa distribusi yang tidak diketahui, sehingga pengujian hipotesis nol dari fungsi distribusi yang tidak diketahui tersebut sesuai dengan distribusi yang ditentukan (Conover, 1999).

Uji keselarasan ini bisa menjadi alat yang bermanfaat untuk mengevaluasi sampai seberapa jauh suatu model mampu menggambarkan keadaan yang nyata. Uji keselarasan yang digunakan dalam penelitian ini adalah uji Kolmogorov-Smirnov (K-S). Uji K-S

dikembangkan oleh Andrei Nikolaevich Kolmogorov dan Nikolai Vasilyevich Smirnov, matematikawan asal Rusia pada tahun 1930-an. Uji K-S secara luas digunakan untuk uji non parametrik (Biswaset *all*, 2008).

Penggunaan uji K-S ini perlu memperhatikan dua buah fungsi distribusi, yaitu: distribusi kumulatif yang dihipotesiskan $F_0(x)$ dan distribusi kumulatif yang teramati $S(x)$. Untuk suatu nilai x , $F(x)$ adalah peluang bahwa nilai variabel acak X kurang dari atau sama dengan x , dengan kata lain $F(x) = P(X \leq x)$. Untuk uji status ini menetapkan bahwa $F(x) < F_0(x)$, statistikujinya adalah:

$$D^+ = \sup_x [F_0(x) - S(x)] \quad (2)$$

Dalam bentuk grafik, D adalah jarak vertikal terjauh antara $F_0(x)$ dan $S(x)$. Jika data sampel telah ditarik dari suatu distribusi yang dihipotesiskan, maka ketidaksesuaian antara $F_0(x)$ dan $S(x)$ untuk nilai-nilai x yang teramatitidakbolehterlalubesar. Untuk menentukan apakah dalam situasi tertentu D cukup besar untuk menolak H_0 , maka nilai D hasil perhitungan dibandingkan dengan nilai-nilai yang diberikan pada tabel kuantil-kuantil statistik uji Kolmogorov.

3. Distribusi Mixture

Kemampuan analisis *mixture* sebagai metode pada data dengan sifat campurannya menunjukkan keunggulan metode analisis ini jika dibandingkan dengan metode statistik yang lain. Beberapa karakteristik data dari beberapa kelompok data yang merupakan satu kesatuan dapat dianalisis dan ditunjukkan gejala nya melalui model yang terbentuk pada sebuah *mixture* (Brahmana, 2003).

Masing-masing komponen *mixture* dalam penelitian ini adalah kelompok data yang dibedakan berdasarkan jenis pengeluaran, mempunyai distribusi tertentu dalam satu kelas distribusi yang sama, sehingga fungsi densitas *mixture*-nya adalah:

$$f(x|\theta, k) = k_1 g_1(x|\theta_1) +$$

$$k_2 g_2(x|\theta_2) + \dots + k_m g_m(x|\theta_m) \quad (3)$$

dimana $f(x|\theta)$ menotasikan fungsi densitas *mixture* dari setiap kelompok data, $g_i(x|\theta_i)$ menotasikan fungsi densitas dari komponen *mixture* yang mempunyai parameter model θ_i dengan θ_i sebuah vektor. Parameter k_i menotasikan parameter proporsi dari komponen *mixture* sehingga $\sum_{i=1}^m k_i = 1$. (Iriawan, 2001).

4. Distribusi *Mixture of Mixture*

Jenis pengeluaran rumah tangga pada data survei sosio-ekonomi nasional (susenas) kota Semarang terdiri dari dua yaitu pengeluaran berupa makanan dan bukan makanan dan untuk menghitung densitas dari setiap jenis pengeluaran rumah tangga digunakan pendekatan *mixture*. Kemudian untuk menghitung densitas pengeluaran rumah tangga keseluruhan digunakan metode *mixture of mixture*.

Pada distribusi ini, data yang diteliti tersusun dari beberapa grup yang dalam hal ini disebut dengan kelompok. Setiap kelompok terdiri dari beberapa komponen dengan proporsi yang bervariasi untuk setiap komponennya, sehingga densitas *mixture of mixture* adalah:

$$p(x|\theta, \pi) = \pi_1 f_1(x|\theta_1) + \pi_2 f_2(x|\theta_2) + \dots + \pi_n f_n(x|\theta_n) \quad (4)$$

(Setiabudi, 2003).

5. Metode *Markov Chain Monte Carlo* (MCMC)

Markov Chain Monte Carlo (MCMC) adalah teknik sampling. Pada setiap langkah algoritma, *state* yang diusulkan dibandingkan dengan *state* saat ini, dan yang lainnya ditolak. *State* yang diterima menjadi *current state*, yang digunakan untuk langkah berikutnya. Distribusi yang diinginkan di dekat dengan menggunakan rantai Markov yang dibentuk oleh sampel diterima (McDuff, 2010)

Iriawan (dalam Setiabudi, 2003) mendefinisikan rantai Markov pada *state space E* sebagai suatu deret variabel random $\{\Theta^{(t)}\}_{t \geq 0}$ dimana nilai untuk masing-masing variabel random tersebut berada dalam *state space E*. Penjelasan lebih lengkap rantai Markov diberikan pada definisi sebagai berikut.

Definisi 2.

Misalkan $\{\Theta^{(t)}\}_{t \geq 0}$ merupakan deret dari suatu variabel random dengan $\Theta^{(t)} \in E$, maka $\Theta^{(t)}$ dikatakan sebagai rantai Markov jika:

$$P(\Theta^{(t+1)} = \theta^{(t+1)} | \Theta^{(0)} = \theta^{(0)}, \Theta^{(1)} = \theta^{(1)}, \dots, \Theta^{(t)} = \theta^{(t)}) \quad (5)$$

$= P(\Theta^{(t+1)} = \theta^{(t+1)} | \Theta^{(t)} = \theta^{(t)})$ untuk semua nilai $t \geq 0$ dan $\theta^{(k)} \in E$, dimana $k = 0, 1, \dots, (t+1)$.

Berdasarkan Definisi

2, dijelaskan bahwa pada saat $t \geq 0$, distribusi dari *state* berikutnya $\Theta^{(t+1)}$, dengan syarat diberikan *state* tertentu $\Theta^{(t)}$, dan *state* sebelumnya $\Theta^{(0)}, \Theta^{(1)}, \dots, \Theta^{(t-1)}$, hanya akan bergantung pada *state* tertentu pada saat t , yaitu $\Theta^{(t)}$ baik untuk komponen diskrit maupun kontinu.

Pada analisis Bayesian, penggunaan MCMC dapat mempermudah analisis, sehingga keputusan yang diambil dari hasil analisis akan dapat dilakukan dengan tepat dan cepat. Menurut Iriawan (dalam Setiabudi, 2003) ada dua kemudahan yang dapat diperoleh dari penggunaan metode MCMC pada analisis Bayesian. Pertama, metode MCMC dapat menyederhanakan bentuk integral yang kompleks dengan dimensi yang besar menjadi bentuk integral yang lebih sederhana dengan satu dimensi. Kedua, dengan menggunakan metode MCMC, estimasi densitas data dapat diketahui dengan cara membangkitkan suatu rantai Markov yang berurutan sebanyak N , $(\Theta^{(T+t)}, 0 \leq t \leq N)$ dimana T adalah suatu bilangan yang cukup besar.

6. Survei Sosial Ekonomi Nasional (Susenas)

Susenas didesain untuk mendapatkan data sosial penduduk dalam lingkup yang luas. Data yang dikumpulkan pada survei ini adalah meliputi data pendidikan, kesehatan/gizi, perumahan/lingkungan, kriminalitas, sosial budaya, pendapatan dan pengeluaran rumahtangga, *travel*, dan kesejahteraan rumah tangga.

Semula kegiatan Susenas disetiap tahunnya hanya mencakup sebagian bidang saja dan diulang setiap tiga tahun. Pada tahun 1992, sistem pengumpulan data diperbarui. Data yang digunakan untuk menyusun indikator kesejahteraan penduduk di modul (informasi yang dikumpulkan setiap 3 tahun) diikutsertakan dalam kor (informasi yang dikumpulkan setiap tahun), sehingga dari kegiatan Susenas akan tersedia data untuk memantau taraf kesejahteraan penduduk, merumuskan program pemerintah dan menganalisis dampak dari program peningkatan kesejahteraan penduduk.

Penjelasan tentang pengumpulan informasi dari modul yang dilakukan setiap tiga tahun. Pada tahun pertama, pendataan difokuskan pada pendapatan dan pengeluaran rumah tangga. Pada tahun kedua, pendataan dilakukan pada kesejahteraan sosial budaya, perjalanan dan kriminalitas. Sedangkan pada tahun ketiga, pendataan dilakukan pada bidang kesehatan, gizi, pendidikan, dan perumahan. Informasi dalam modul lebih lengkap dan luas dibandingkan dengan data dalam kor. Pada kor sendiri, data dikumpulkan dalam rangka mendapatkan informasi penting untuk mengantisipasi beberapa perubahan yang mungkin terjadi setiap tahun. Informasi kor juga dapat digunakan melakukan perencanaan jangka pendek

Susenas sangat potensial untuk menggambarkan tentang kesejahteraan rakyat. Data Susenas diharapkan dapat digunakan untuk menentukan kebijakan dan keputusan diberbagai bidang. Beberapa hal yang penting yang harus

dilaksanakan di masa mendatang adalah perumusan masalah, pemantauan, dan evaluasi untuk menemukan solusi. Karena kemajuan teknologi dan komputer membuat analisis pekerjaan menjadi lebih mudah dan secara tidak langsung data Susenas menjadi lebih tersosialisasikan. (BPS, 2001)

C. Metode Penelitian

Dalam penelitian ini, data yang digunakan adalah data persentase pengeluaran rata-rata per kapita sebulan menurut jenis pengeluaran dan golongan pengeluaran per kapita sebulan dari data Susenas Kor Juli 2007 kota Semarang, untuk daerah perkotaan saja. Variabel dalam penelitian ini adalah variabel-variabel yang berhubungan langsung dalam penentuan sampel dan estimasi parameter yaitu variabel pengeluaran rumah tangga.

Tahapan analisis dalam penelitian ini adalah: eksplorasi data dilakukan dengan pengelompokan data ke dalam tiap jenis pengeluaran rumah tangga. Tahap berikutnya adalah menentukan distribusi data dari jenis pengeluaran rumah tangga. Kemudian menentukan nilai parameter distribusi prior setiap jenis pengeluaran rumah tangga. Distribusi prior yang digunakan adalah prior konjugate (sekawan). Selanjutnya menentukan masing-masing fungsi posterior untuk masing-masing fungsi sebaran pada fungsi *mixture* wilayah dengan Bayesian – MCMC dengan menggunakan paket program WinBugs 1.4. Terakhir dilakukan estimasi fungsi sebaran *mixture of mixture* untuk tiap jenis pengeluaran rumah tangga.

D. Hasil dan Pembahasan

Masing-masing data yang digunakan dalam penelitian ini, dikelompokkan berdasarkan jenis pengeluaran penduduk yaitu:

- a. Pengeluaran kelompok makanan meliputi:

1. Padi-padian
 2. Umbi-umbian
 3. Ikan
 4. Daging
 5. Telur dan susu
 6. Sayur-sayuran
 7. Kacang-kacangan
 8. Buah-buahan
 9. Minyak dan lemak
 10. Bahan minuman
 11. Bumbu-bumbuan
 12. Konsumsi lainnya
 13. Makanan dan minuman
 14. Minuman alkohol
 15. Tembakau dan sirih
- b. Pengeluaran kelompok bukan makanan meliputi:
1. Perumahan
 2. Aneka barang dan jasa
 3. Biaya pendidikan
 4. Biaya kesehatan
 5. Pakaian dan alas kaki
 6. Barang tahan lama
 7. Pajak dan asuransi
 8. Keperluan pesta

Selanjutnya identifikasi bentuk distribusi dari masing-masing kelompok data tersebut, menggunakan uji Kolmogorov-Smirnov. Bentuk distribusi masing-masing kelompok data pada kolom (3) dan nilai masing-masing parameternya pada kolom (4) dan (5) dapat dilihat pada Tabel 1.

Analisis yang diterapkan pada penelitian ini yaitu estimasi distribusi parameter dari data Susenas kota Semarang. Tipe prior yang digunakan adalah prior konjugate (sekawan) dan distribusi prior dari masing-masing kelompok data dapat dilihat pada Tabel 1 kolom (6) dan (7).

Dengan melalui proses iterasi 5000 kali, diperoleh bobot masing-masing kelompok data dan fungsi densitas *mixture* untuk kelompok makanan, yaitu:

$$\begin{aligned}
 f_1(x|\theta, k) = & 0,06828g_{11}(x|\theta_{11}) \\
 & + 0,06708g_{12}(x|\theta_{12}) \\
 & + 0,06502g_{13}(x|\theta_{13}) \\
 & + 0,06694g_{14}(x|\theta_{14}) \\
 & + 0,0672g_{15}(x|\theta_{15}) \\
 & + 0,06765g_{16}(x|\theta_{16}) \\
 & + 0,06552g_{17}(x|\theta_{17}) \\
 & + 0,06594g_{18}(x|\theta_{18}) \\
 & + 0,06585g_{19}(x|\theta_{19}) \\
 & + 0,06778g_{110}(x|\theta_{110}) \\
 & + 0,06697g_{111}(x|\theta_{111}) \\
 & + 0,06628g_{112}(x|\theta_{112}) \\
 & + 0,06499g_{113}(x|\theta_{113}) \\
 & + 0,06762g_{114}(x|\theta_{114}) \\
 & + 0,06688g_{115}(x|\theta_{115})
 \end{aligned}$$

dengan:

g_{11} : Jenis pengeluaran padi-padian, berdistribusi normal dengan mean 9,807143 dan standar deviasi 7,191522.

g_{12} : Jenis pengeluaran umbi-umbian, berdistribusi normal dengan mean 0,441429 dan standar deviasi 0,29852.

g_{13} : Jenis pengeluaran ikan, berdistribusi normal dengan mean 2,23 dan standar deviasi 0,815904.

g_{14} : Jenis pengeluaran daging, berdistribusi normal dengan mean 2,021429 dan standar deviasi 0,224457.

g_{15} : Jenis pengeluaran telur dan susu, berdistribusi normal dengan mean 3,807143 dan standar deviasi 0,571539.

g_{16} : Jenis pengeluaran sayur-sayuran, berdistribusi normal dengan mean 3,635714 dan standar deviasi 1,784964.

g_{17} : Jenis pengeluaran kacang-kacangan, berdistribusi normal dengan mean 2,398571 dan standar deviasi 1,324895.

g_{18} : Jenis pengeluaran buah-buahan, berdistribusi normal dengan mean 1,777143 dan standar deviasi 0,313566.

g_{19} : Jenis pengeluaran minyak dan lemak, berdistribusi normal dengan

- mean 1,942857 dan standar deviasi 1,06042.
- g_{110} : Jenis pengeluaran bahan minuman, berdistribusi normal dengan mean 1,792857 dan standar deviasi 0,555659.
- g_{111} : Jenis pengeluaran bumbu-bumbuan, berdistribusi normal dengan mean 1,077143 dan standar deviasi 0,346781.
- g_{112} : Jenis pengeluaran konsumsi lainnya, berdistribusi normal dengan mean 1,962857 dan standar deviasi 0,876617.
- g_{113} : Jenis pengeluaran makanan dan minuman, berdistribusi normal dengan mean 8,824286 dan standar deviasi 1,923702.
- g_{114} : Jenis pengeluaran minuman alkohol, berdistribusi normal dengan mean 0,015714 dan standar deviasi 0,029358.
- g_{115} : Jenis pengeluaran tembakau dan sirih, berdistribusi normal dengan mean 4,062857 dan standar deviasi 1,480942.

Tabel 1. Bentuk distribusi dan nilai parameter masing-masing kelompok data

No	Jenis Pengeluaran	Distribusi	Nilai Parameter		Distribusi Parameter	
			μ	σ	μ	σ
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Kelompok Makanan						
1.	Padi-padian	Normal	9,807143	7,191522	Normal	Gamma
2.	Umbi-umbian	Normal	0,441429	0,29852	Normal	Gamma
3.	Ikan	Normal	2,23	0,815904	Normal	Gamma
4.	Daging	Normal	2,021429	0,224457	Normal	Gamma
5.	Telur dan susu	Normal	3,807143	0,571539	Normal	Gamma
6.	Sayur-sayuran	Normal	3,635714	1,784964	Normal	Gamma
7.	Kacang-kacangan	Normal	2,398571	1,324895	Normal	Gamma
8.	Buah-buahan	Normal	1,777143	0,313566	Normal	Gamma
9.	Minyak dan lemak	Normal	1,942857	1,06042	Normal	Gamma
10.	Bahan minuman	Normal	1,792857	0,555659	Normal	Gamma
11.	Bumbu-bumbuan	Normal	1,077143	0,346781	Normal	Gamma
12.	Konsumsi lainnya	Normal	1,962857	0,876617	Normal	Gamma
13.	Makanan dan minuman	Normal	8,824286	1,923702	Normal	Gamma
14.	Minuman alkohol	Normal	0,015714	0,029358	Normal	Gamma
15.	Tembakau dan sirih	Normal	4,062857	1,480942	Normal	Gamma
Kelompok bukan makanan						
16.	Perumahan	Normal	27,48143	8,161931	Normal	Gamma
17.	Aneka barang dan jasa	Normal	10,88429	1,414777	Normal	Gamma
18.	Biaya pendidikan	Normal	6,028571	0,891374	Normal	Gamma
19.	Biaya kesehatan	Normal	2,787143	1,860078	Normal	Gamma
20.	Pakaian dan alas kaki	Normal	2,191429	0,237236	Normal	Gamma
21.	Barang tahan lama	Normal	1,667143	1,449721	Normal	Gamma
22.	Pajak dan asuransi	Normal	1,792857	0,848267	Normal	Gamma
23.	Keperluan pesta	Normal	1,371429	0,862234	Normal	Gamma

Begitu pula dengan estimasi fungsi densitas *mixture* untuk kelompok data bukan makanan, yaitu:

$$f_2(x|\theta, k) = 0,1265g_{21}(x|\theta_{21}) + 0,1249g_{22}(x|\theta_{22}) + 0,1239g_{23}(x|\theta_{23}) + 0,1258g_{24}(x|\theta_{24}) + 0,1223g_{25}(x|\theta_{25}) + 0,1251g_{26}(x|\theta_{26}) + 0,1265g_{27}(x|\theta_{27}) + 0,125g_{28}(x|\theta_{28})$$

dengan:

- g_{21} : Jenis pengeluaran perumahan, berdistribusi normal dengan mean 27,48143 dan standar deviasi 8,161931.
- g_{22} : Jenis pengeluaran aneka barang dan jasa, berdistribusi normal dengan mean 10,88429 dan standar deviasi 1,414777.
- g_{23} : Jenis pengeluaran biaya pendidikan, berdistribusi normal dengan mean 6,028571 dan standar deviasi 0,891374.
- g_{24} : Jenis pengeluaran biaya kesehatan, berdistribusi normal dengan mean 2,787143 dan standar deviasi 1,860078.
- g_{25} : Jenis pengeluaran pakaian dan alas kaki, berdistribusi normal dengan mean 2,191429 dan standar deviasi 0,237236.
- g_{26} : Jenis pengeluaran barang tahan lama, berdistribusi normal dengan mean 1,667143 dan standar deviasi 1,449721.
- g_{27} : Jenis pengeluaran pajak dan asuransi, berdistribusi normal dengan mean 1,792857 dan standar deviasi 0,848267.
- g_{28} : Jenis pengeluaran keperluan pesta, berdistribusi normal dengan mean 1,371429 dan standar deviasi 0,862234.

Gabungan dari fungsi distribusi *mixture* yang terbentuk antara jenis pengeluaran kelompok makanan dan bukan makanan adalah berupa fungsi distribusi *mixture of mixture*. Berdasarkan proses iterasi 5000 kali diperoleh bobot

untuk fungsi distribusi *mixture* jenis pengeluaran kelompok makanan mencapai 0,5011, sedangkan untuk jenis pengeluaran kelompok bukan makanan mencapai 0,4989, sehingga fungsi *mixture of mixture* yang terbentuk adalah:

$$z(x|\theta, \pi) = 0,5011f_1(x|\theta_1) + 0,4989f_2(x|\theta_2)$$

dengan:

- f_1 : Jenis pengeluaran rumah tangga untuk kelompok makanan.
- f_2 : Jenis pengeluaran rumah tangga untuk kelompok bukan makanan.

Persamaan z yang diperoleh merupakan persamaan *mixture of mixture* yang memberikan makna bahwa jenis pengeluaran kelompok makanan f_1 memberikan peran 0,5011 terhadap sosial ekonomi masyarakat kota Semarang, dimana f_1 memiliki distribusi *mixture* yang didalamnya mengandung 15 macam pengeluaran rumah tangga. Kelimabelas macam pengeluaran tersebut berdistribusi normal dengan parameter mean dan standar deviasi yang berbeda-beda. Sedangkan jenis pengeluaran kelompok bukan makanan f_2 memberikan peran 0,4989 terhadap sosial ekonomi masyarakat kota Semarang, dimana f_2 memiliki distribusi *mixture* yang didalamnya mengandung 8 macam pengeluaran rumah tangga. Kedelapan macam pengeluaran tersebut berdistribusi normal dengan parameter mean dan standar deviasi yang berbeda-beda. Dengan demikian pendekatan *mixture* sebagai suatu metode mampu mencampurkan beberapa kelompok data yang memiliki distribusi yang berbeda, kemudian beberapa kelompok *mixture* tersebut dicampur kembali dengan pendekatan *mixture of mixture* dengan tanpa meninggalkan sifat-sifat yang ada pada masing-masing kelompok data yang dalam hal ini adalah jenis pengeluaran.

E. Penutup

Berdasarkan hasil analisis dan pembahasan di atas, dapat ditarik

beberapa kesimpulan sebagai berikut. Bentuk model *mixture of mixture* suatu data sangat dipengaruhi oleh klasifikasi datanya sebagai penentu banyaknya grup data sebagai komponen *mixture*. Semakin banyak grup data yang dibentuk akan membuat semakin banyak komponen distribusi dalam *mixture*-nya dan jugasemakin dapat mempresentasikan datanya sehingga menjadi lebih teliti. Dalam penelitian ini fungsi *mixture of mixture* yang terbentuk adalah:

$$\begin{aligned} z(x|\theta, \pi) \\ &= 0,5011f_1(x|\theta_1) \\ &+ 0,4989f_2(x|\theta_2) \end{aligned}$$

Sebagaimana diketahui bahwa dalam Susenas terdapat cukup banyak variabel yang dicakup, sedangkan penelitian ini hanya mencakup satu variabel saja, yaitu jenis pengeluaran penduduk suatu rumah tangga. Oleh karena itu, ada baiknya apabila dalam penelitian ini dilanjutkan dengan melibatkan lebih banyak variabel yang ada di Susenas sehingga bentuk model *mixture* yang diperoleh akan lebih menggambarkan suatu gambaranyanglebih lengkap.

DAFTAR PUSTAKA

- Biswas, S., Ahmad, S., Molla, M.K.I., Hirose, K. & Nasser, M. 2008. *Kolmogorov-Smirnov Test in Text-Dependent Automatic Speaker Identification*. Engineering Letter. Vol.16. No.4.20 November 2008, tersedia di http://www.engineeringletters.com/issues_v16/issue_4/EL_16_4_01.pdf, diakses pada tanggal 18 Juli 2008.
- BPS. 2001. *National Socio – Economic Survey*. Jakarta: Workmanual.
- Brahmana, E.T. 2003. *Estimasi Densitas Produk Domestik Regional Bruto (PDRB) Propinsi Jawa Timur dengan Menggunakan Pendekatan Mixture dari Beberapa Mixture*. Tesis. Surabaya: Program Pascasarjana ITS.
- Conover, W. J.1999. *Practical Nonparametric Statistical, 3rd Edition*, pp.428-433 (6.1), New York:John Wiley & Sons, Inc.
- Iriawan, N.2003. *Workshop Pemodelan Data dengan Markov Chain Monte Carlo (MCMC) Menggunakan WinBUGS 1.4*. Modul. Surabaya: ITS.
- McDuff, D.2010. *A Human-Markov Chain Monte Carlo Method For Investigating Facial Expression Categorization*. Proceeding International Conference on Cognitive Modeling. Philadelphia. Agustus 5-8, 2010.
- Setiabudi, B. 2003. *Estimasi Parameter Populasi Dengan Fungsi Mixture of Mixture Pada Data Susenas Kabupaten Sidoarjo*. Tesis. Surabaya: Program Pascasarjana ITS.
- Walpole, R. E. 1989. *Ilmu Peluang dan Statistika untuk Insinyur dan Ilmuwan*. Edisi Ke 4. Bandung: ITB-Press.