PENGELOMPOKAN NEGARA BERDASARKAN KASUS *STUNTING* DENGAN MODEL *FINITE MIXTURE* NORMAL MENGGUNAKAN PENDEKATAN *BAYESIAN*

(State Management Based on Stunting Case with Mixed to Normal Model Using Bayesian Approach)

Adella Okky Herashanti[1)], Irwan Susanto[2)], dan Isnandar Slamet[3)]
[1,2,3)]Program Studi Statistika, Universitas Sebelas Maret, Surakarta
e-mail: adellaokky@student.uns.ac.id

Abstract. Stunting (short stature) is a chronic condition characterized by stunted growth due to malnutrition over a long period of time. Stunting needs special attention because it can hamper children's physical and mental development and is associated with an increased risk of illness and death. The high incidence of stunting in the world is still a health problem that is still being faced by WHO [7]. Therefore, it is necessary to group countries based on stunting cases in order to facilitate appropriate policy making in overcoming and preventing stunting. In this study, 151 countries were grouped based on stunting cases with a normal finite mixture model. The model is estimated using Bayesian estimation through Birth-Death Markov Chain Monte Carlo (BD-MCMC) method. The results of the study show that there are five groups of countries based on stunting cases, where the first group consists of 30 countries and the second group consists of 32 countries which are groups with low stunting rates. While the third group has 33 members, the fourth group has 34 countries and the fifth group has 22 countries which is a group with a high stunting rate.

Keywords: BD-MCMC, *Finite Mixture,* Pengelompokan, *Stunting*.

1. Pendahuluan

Stunting (tubuh pendek) merupakan suatu kondisi kronis yang digambarkan oleh pertumbuhan yang terhambat akibat malnutrisi dalam jangka waktu yang panjang. Dengan kata lain *stunting* merupakan suatu kondisi terlambatnya pertumbuhan anak yang ditandai dengan tinggi badan anak lebih pendek dibandingkan dengan tinggi badan anak-anak lain di usia yang sama [6]. *Stunting* perlu menjadi perhatian khusus karena bisa membuat perkembangan motorik dan mental anak terhambat serta berkaitan dengan risiko kesakitan dan kematian yang meningkat [3]

Berdasarkan WHO [7], pada tahun 2018 terdapat sebanyak 21,9% atau sekitar 149 juta anak dibawah umur 5 tahun di dunia mengalami *stunting*. Angka tersebut berarti 1 dari 4 anak dibawah umur 5 tahun mengalami *stunting.* Kejadian *stunting* pada tahun 2018 paling banyak terjadi di Asia yaitu sekitar 55% dari total kejadian *stunting* di dunia dan diikuti oleh Afrika sebanyak 39%.

Oleh karena itu diperlukan pengelompokan negara berdasarkan kasus *stunting* supaya diketahui negara mana saja yang memiliki kasus *stunting* yang sama guna mempermudah pengambilan kebijakan yang tepat dalam mengatasi dan mencegah *stunting*. Pada penelitian ini digunakan metode analisis klaster dengan model *finite mixture*. Parameter dari model *finite mixture* akan diestimasi menggunakan pendekatan *Bayesian* melalui *Birth-Death Markov Chain Monte Carlo* (BD-MCMC) dengan penentuan komponen yang tepat menggunakan besarnya probabilitas marginal *posterior*.

Penelitian mengenai pengelompokan kasus *stunting* pernah dilakukan oleh Wulandari dan Kurniawan [8] dalam jurnal publikasinya yang berjudul *Pengelompokan Kabupaten/Kota di Jawa Timur Berdasarkan Kasus Stunting Balita Menggunakan Algoritma Fuzzy Particle Swarm Optimization-Fuzzy C-Means* dan hasilnya didapatkan dua kelompok kasus *stunting*.

2. Metodologi

Penelitian ini menggunakan data kasus *stunting* per negara tahun 2019 terdiri dari 151 negara. Data yang digunakan merupakan data sekunder yang diperoleh dari *website kaggle*. Analisis klaster pada penelitian ini dilakukan dengan menggunakan *model based clustering*. Estimasi parameter dilakukan dengan menggunakan pendekatan *Bayesian* melalui algoritma *Birth-Death Markov Chain Monte Carlo* (BD-MCMC).

Langkah-langkah pemodelan *finite mixture* yang dilakukan adalah sebagai berikut :

1. Identifikasi pola distribusi data kasus *stunting* apakah memenuhi pola *univariat multimodal*. Identifikasi dilakukan dengan menggunakan uji *Goodness of Fit*. Jika memenuhi, maka terdapat indikasi adanya model *finite mixture*.
2. Estimasi model *finite mixture* Normal dengan pendekatan *Bayesian* melalui metode *Birth-Death Markov Chain Monte Carlo* (BD-MCMC).
3. Menentukan jumlah komponen yang tepat dengan menggunakan probabilitas *marginal posterior*, yaitu dipilih probabilitas *marginal posterior* terbesar. Kemudian menginterpretasikan hasil pengelompokan.

Teori pada penelitian ini antara lain model *finite mixture*, estimasi *Bayesian*, *Birth-Death Markov Chain Monte Carlo* (BD-MCMC) dan uji *goodness of fit*.

2.1 Model *Finite Mixture*

Model *finite mixture* didasarkan pada representasi fungsi distribusi probabilitas (kumulatif) dan berlaku pada distribusi diskrit maupun kontinu. Suatu vektor variabel random $x = [x_1, x_2, \ldots, x_n]^T$ yang bertipe diskrit atau kontinu dapat dikatakan berdistribusi *finite mixture* jika memiliki fungsi kepadatan probabilitas $f(x_i)$ yang didefinisikan sebagai

$$f(x_i) = w_1 f_1(x_i) + w_2 f_2(x_i) + \cdots + w_K f_K(x_i)$$

dengan:

$f_K(x_i)$: fungsi kepadatan probabilitas *mixture* untuk $k = 1,2,\ldots,K$;

w_K : parameter *weight*

$w = [w_1, w_2, \ldots, w_K]^T$: vektor parameter *weight*

Nilai-nilai w harus memenuhi:

$$0 \leq w_k \leq 1$$

dan

$$\sum_{k=1}^{K} \quad w_k = 1$$

Densitas *mixture* dapat ditulis dengan mengasumsikan semua komponen distribusi *finite mixture* berasal dari distribusi probabilitas yang memiliki vektor θ, yaitu

$$f(x_i; \Psi) = w_1 f_1(x_i; \theta_1) + w_2 f_2(x_i; \theta_2) + \cdots + w_K f_K(x_i; \theta_K)$$

$$\sum_{k=1}^{K} \quad w_k f_k(x_i; \theta_k)$$

Dengan $\Psi = [w, \theta]^T$dan $\theta = [\theta_1, \theta_2, \ldots, \theta_K]^T$ [1]. Model statistika yang mengimplementasikan konsep distribusi *mixture* pada pemodelannya disebut sebagai model *finite mixture* [5].

2.2 Estimasi *Bayesian*

Dalam estimasi *Bayesian*, diperoleh data sampel dari populasi serta memperhitungkan distribusi awalnya yang disebut dengan distribusi *prior*. Selanjutnya, dari distribusi *prior* tersebut dapat ditentukan distribusi *posterior*-nya sehingga akan didapatkan estimator *Bayesian*. Data yang diobservasi dalam estimasi *Bayesian* disebut x dan parameter data θ. Teorema bayes digunakan untuk menentukan distribusi θ dengan syarat x, yaitu :

$$p(x) = \frac{l(\theta)p(\theta)}{p(x)} \qquad (1)$$

Persaman (1) dapat dituliskan sebagai berikut:

$$p(x) \propto l(\theta)p(\theta) \qquad (2)$$

Dengan kata lain, *posterior* didapatkan dari *likelihood* yang berisi informasi yang tersedia dari sampel dikalikan dengan *prior* parameter yang berisi informasi dari data sebelumnya. Persamaan (2) memperlihatkan bahwa distribusi *posterior* akan proposional saat *likelihood* dikalikan dengan distribusi *prior*-nya, sehingga jika diterapkan pada model *mixture*, maka θ dinyatakan sebagai vektor parameter yang memuat semua parameter model *mixture* [2].
Diasumsikan data $x_1, \dots, x_n$ adalah observasi yang independen dari model *mixture* normal dengan k komponen yang tidak diketahui:

$$p(w_i, \theta_i) = w_1 f(x; \theta_1) + \cdots + w_k f(x; \theta_k) \tag{3}$$

Dengan θ_k merupakan parameter model komponen ke-k, yaitu (μ_k, σ_k^2). Pemodelan *mixture* Normal (3) juga dapat dituliskan sebagai berikut:

$$p(w_i, \mu_i, \sigma_i^2) = w_1 f(x; \mu_1, \sigma_1^2) + \cdots + w_k f(x; \mu_k, \sigma_k^2)$$

Dalam membuat dugaan distribusi parameter, akan lebih mudah dengan mengasumsikan setiap observasi x_j bangkit dari komponen *mixture* z_j yang spesifik tetapi tidak diketahui. Dimana $z_1, z_2, \dots, z_n$ yang diasumsikan sebagai realisasi dari variabel random diskrit yang terdistribusi independen dan identik $Z_1, Z_2, \dots, Z_n$ dengan distribusi *prior*-nya sebagai berikut

$$Pr(k, w) = w_i \qquad (j = 1, \dots, n; i = 1, \dots, k)$$

Kemudian, dapat didefinisikan distribusi *prior* bersama untuk parameter k, w, μ, σ^2, akan diasumsikan bahwa distribusi *prior* bersama bisa difaktorkan sebagai:

$$f(z, k) \propto f(k, w) f(k) f(\mu|k) f(\sigma^2|k)$$

Menurut Stephens [4], diasumsikan distribusi *prior truncated* Poisson pada k komponen:

$$p(k) \propto \frac{\lambda^2}{k!} \qquad (k = 1, \dots., k_{max})$$

Dengan bersyarat pada k, definisi distribusi *prior* untuk parameter yang tersisa adalah sebagai berikut

$$w \sim D(\gamma)$$

$$\mu_i \sim N(\xi, \kappa^{-1}) \qquad (i = 1, \dots, k)$$

$$\sigma_i^2 \sim \Gamma(\alpha, \beta) \qquad (i = 1, \dots, k)$$

Dimana $D(\gamma)$ dinotasikan sebagai distribusi Dirichlet dengan parameter γ.
Likelihood untuk parameter k, w, μ, σ^2 adalah

$$L(k, w, \mu, \sigma^2) = \prod_{j=1}^{n} [w_1 f(x; \mu_1, \sigma_1^2) + \cdots + w_k f(x; \mu_k, \sigma_k^2)]$$

sehingga diperoleh distribusi *posterior*-nya sebagai berikut

$$p(x^n) \propto L(k, w, \mu, \sigma^2)\, p(k, w, \mu, \sigma^2)$$

2.3 *Birth-Death Markov Chain Monte Carlo* (BD-MCMC)

Penggunaan metode *Markov Chain Monte* Carlo (MCMC) dalam estimasi *Bayesian* bisa mempermudah analisis untuk menemukan distribusi *posterior* yang mana membutuhkan proses integrasi yang rumit dan cukup lama. Metode tersebut melibatkan pembangunan dari suatu rantai Markov $\{\Theta^{(t)}\}$ dengan distribusi *posterior* $p(\theta|x^n)$ dari parameter θ [4]**.** Pada umumnya, estimasi *bayesian mixture* menggunakan algoritma MCMC terbatas hanya pada kasus dimensi vektor parameter banyaknya model tetap. Akibatnya, terdapat pengembangan metode MCMC yang bisa menyelesaikan masalah tersebut, yaitu metode BD-MCMC. Menurut Stephens [4], metode BD-MCMC bisa digunakan untuk melakukan pemodelan *mixture* dengan banyak komponen yang tidak diketahui.

Dalam metode BD-MCMC, ukuran k untuk *finite mixture* berubah sedemikian sehingga kelahiran dan kematian dari komponen *mixture* terjadi dalam waktu kontinu dengan distribusi bersama stasioner *posterior* dari parameter *mixture*. Proses kelahiran dalam komponen *mixture* terjadi saat *constant rate* yang sama dengan parameter θ di distribusi *prior*.

Didefinisikan *birth* and *death* pada Ω adalah sebagai berikut

Births: jika pada waktu t dan proses terjadi pada $y = \{(w_1, \theta_1), \dots, (w_k, \theta_k)\} \epsilon\ \Omega_k$ dan kelahiran (*birth*) dikatakan terjadi pada $(w, \theta)\ \epsilon\ [0,1] \times \theta$, maka proses lompat ke

$$y \cup (w, \theta) \coloneqq \{(w_1(1 - w_1), \theta_1), \dots, (w_k(1 - w_k), \theta_k), (w, \theta)\}\ \epsilon\ \Omega_{k+1}$$

Deaths: jika pada waktu t dan proses terjadi pada $y = \{(w_1, \theta_1), \dots, (w_k, \theta_k)\} \epsilon\ \Omega_k$ dan kematian (*death*) dikatakan terjadi pada $(w_i, \theta_i)\ \epsilon\ y$, maka proses lompat ke

$$y \backslash (w_i, \theta_i) \coloneqq \left\{\left(\frac{w_1}{(1 - w_i)}, \theta_1\right), \dots, \left(\frac{w_{i-1}}{(1 - w_i)}, \theta_{i-1}\right), \left(\frac{w_{i+1}}{(1 - w_i)}, \theta_{i+1}\right), \dots, \left(\frac{w_k}{(1 - w_i)}, \theta_1\right)\right\} \epsilon\ \Omega_k$$

Proses kelahiran meningkatkan jumlah komponen satu persatu, sedangkan proses kematian menurunkan jumlah komponen satu persatu. Baik proses kelahiran maupun proses kematian merupakan proses Poisson yang independen.

Stephens [4] mendefinisikan algoritma dari proses *birth-death* adalah sebagai berikut: Dimulai dengan inisiasi model $y = \{(w_1, \mu_1, \sigma_1^2), \dots, (w_k, \mu_k, \sigma_k^2)\} \epsilon\ \Omega_k$, dimana Ω_k dinotasikan sebagai ruang parameter dari model *mixture* dengan k komponen, lalu iterasikan langkah-langkah berikut:

1. *Birth rate* $\beta(y) = \lambda_b$
2. Hitung *death rate* untuk setiap komponen. *Death rate* untuk komponen ke-i yaitu

$$\delta_i(y) = \lambda_b \frac{L(y \backslash (w_i, \mu_i, \sigma_i^2))}{L(y)} \frac{p(k - 1|\mu_i, \sigma_i^2)}{kp(k|\mu_i, \sigma_i^2)}, \quad i = 1, \dots, k$$

3. Hitung jumlah total *death rate* $\delta(y) = \sum_i \delta_i(y)$

4. Simulasikan waktu untuk lompatan selanjutnya dari suatu distribusi eksponensial dengan *mean* $1/(\beta(y)+\ \delta(y))$
5. Simulasikan tipe lompatan: *birth* atau *death* dengan probabilitasnya masing-masing
$$Pr\ Pr\ (birth)\ = \frac{\beta(y)}{\beta(y)+\ \delta(y)} \qquad Pr\ Pr\ (death)\ = \frac{\delta(y)}{\beta(y)+\ \delta(y)}$$
6. Sesuaikan y untuk merefleksikan *birth* atau *death* (mengubah komponen *mixture*)
7. Kembali ke langkah 2

Setelah melakukan algoritma proses *birth-death*, maka dilanjutkan proses membentuk rantai Markov. Menurut Stephens [4] dalam pembentukan rantai Markov diperlukan distribusi penuh bersyarat posterior tiap parameter, distribusinya sebagai berikut
$$p(z_j|k,w_i,\mu_i,\sigma_i^2) \propto w_i\ N(\mu_i,\sigma_i^2)$$
$$p(k,\mu_i,\sigma_i^2,k)\ \ \sim D(\gamma+n_1,\dots,\gamma+n_k)$$
$$p(k,w_i,\sigma_i^2,k)\ \ \sim N\big((n_i\sigma_i^{-2}+\kappa)^{-1}(n_i\sigma_i^{-2}\underline{x}_i+\kappa\xi),(n_i\sigma_i^{-2}+\kappa)^{-1}\big)$$
$$p(k,w_i,\mu_i,k)\ \ \sim\Gamma\left(\alpha+\frac{1}{2}n_i,\beta+\frac{1}{2}\sum_{j:z_z=i}(x_j-\mu_i)^2\right)$$
Untuk $i=1,\dots,k$ dan $j=1,\dots,n$, dimana n_i adalah jumlah observasi yang ditempatkan pada kelas i $(n_i=\#\{j:z_z=i\})$ dan $\underline{x}_i$ adalah rata-rata observasi yang ditempatkan di kelas i.

Algoritma pembentukan rantai Markov yang didefinisikan oleh Stephens [4] adalah sebagai berikut:
1. Menjalankan proses *birth-death* pada waktu t_0
2. Membangkitkan $(z^n)^{(t+1)}$ dari $p(z|k^{(t+1)},w^{(t)},\mu^{(t)},\sigma^{2(t)},x^n)$
3. Membangkitkan $w^{(t+1)}$ dari $p(w|k^{(t+1)},\mu^{(t)},\sigma^{2(t)},x^n,z^n)$
4. Membangkitkan $\mu^{(t+1)}$ dari $p\big(k^{(t+1)},w^{(t+1)},\sigma^{2(t)},x^n,z^n\big)$
5. Membangkitkan $(\sigma^2)^{(t+1)}$ dari $p(\sigma^2|k^{(t+1)},w^{(t+1)},\mu^{(t+1)},x^n,z^n)$

Komponen untuk *finite mixture* menggunakan metode BD-MCMC dapat dipilih dengan melihat probabilitas marginal *posterior*. Berdasarkan Stephens [4], probabilitas marginal *posterior* dapat dihitung dengan rumus berikut:
$$Pr\ Pr\ (x^n)\ = \frac{1}{N}\ \#\{t:k^{(t)}=i\}$$
$$\approx\ \frac{1}{N}\#\big\{t:k^{(t)}=i\big\} \qquad (N\ besar)$$
Besar komponen dipilih berdasarkan probabilitas marginal *posterior* terbesar.

2.4 Uji *Goodness of Fit*

Identifikasi data untuk mengetahui apakah berdistribusi *univariat multimodal* bisa dilakukan dengan menggunakan uji *Goodness of Fit.* Langkah-langkah melakukan uji *Goodness of Fit* adalah sebagai berikut:

1. Menentukan hipotesis
   Hipotesis dalam uji *Goodness of Fit* adalah:
   H_0 : Distribusi data sesuai dengan distribusi yang diuji
   H_1 : Distribusi data tidak sesuai dengan distribusi yang diuji
2. Menentukan taraf signifikansi α
3. Menghitung statistik uji
   Statistik uji yang digunakan adalah sebagai berikut:
   $$A = -n - \frac{1}{n}\sum_{i=1}^{n} \quad [2i - 1]\left[ln\ ln\left(F(x_i)\right) + ln\ (1 - F(x_{n+1-i}))\right]$$
   Dimana:
   A : statistik uji untuk metode Anderson-Darling
   n : ukuran sampel
   $F(x_i)$: nilai fungsi distribusi kumulatif dari distribusi yang diuji
   i : sampel ke- i dari data yang telah diurutkan

   Modifikasi dari metode Anderson-Darling menggunakan rumus berikut :
   $$A^* = A\left(1 + \frac{0{,}75}{n} + \frac{2{,}25}{n^2}\right)$$
   dan
   $$c_\alpha = a_\alpha\left(1 - \frac{b_\alpha}{n} + \frac{d_\alpha}{n^2}\right)$$
   Dengan a_α, b_α, d_α ditunjukkan dalam tabel nilai kritis Anderson-Darling
4. Menentukan daerah kritis
   Jika $A^* > c_\alpha$ dengan $p_{value} < \alpha$ maka H_0 akan ditolak
5. Menarik kesimpulan.

3. Hasil dan Pembahasan

3.1 Identifikasi Pola data

Identifikasi pola data dilakukan untuk mengetahui apakah data berdistribusi *univariat multimodal* dengan menggunakan uji *Goodness of Fit.* Salah satu bentuk dari uji *Goodness of Fit* yang digunakan pada penelitian ini adalah uji *Anderson-Darling*. Berikut merupakan hasil pemeriksaan pada data *stunting* tahun 2019.

Hipotesis yang digunakan adalah:

H_0 : Distribusi data sesuai dengan distribusi yang diuji

H_1 : Distribusi data tidak sesuai dengan distribusi yang diuji

Daerah kritis:

H_0 ditolak jika $p - value < \alpha$ = 0,05

Statistik Uji:

Tabel 1. Statistik Uji

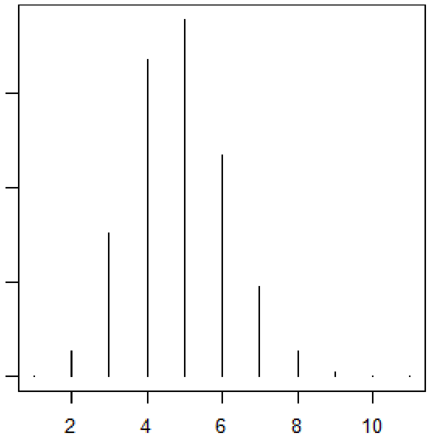
Distribusi	*p-value*
Normal	0,007

Berdasarkan statistik uji di atas, H_0 ditolak karena besar *p-value* distribusi yang diuji bernilai lebih kecil dari $\alpha = 0{,}05$, sehingga dapat dikatakan bahwa distribusi data *stunting* tidak sesuai dengan distribusi yang diuji.

Jadi, dapat disimpulkan bahwa data kasus *stunting* cocok untuk dimodelkan dengan model *finite mixture*, karena data tersebut tidak menunjukkan pola distribusi *unimodal* melainkan *multimodal*.

3.2 Pemodelan

Data *stunting* yang telah dicek dan menunjukan pola distribusi *multimodal*, kemudian dicari komponen *mixture*-nya untuk membentuk model *mixture* Normal. Diperoleh komponen distribusi *mixture* untuk data *stunting* yaitu sebanyak 5 komponen. Hal tersebut sesuai Gambar 1 yang menunjukkan bahwa probabilitas *marginal posterior* untuk komponen 5 lebih tinggi dibandingkan dengan komponen lainnya.



Gambar 1. Probabilitas *marginal posterior*

Estimasi model *finite mixture* yang diperoleh adalah

$$f(x_i) = \hat{w}_1 N(\hat{\mu}_1, \hat{\sigma}_1) + \hat{w}_2 N(\hat{\mu}_2, \hat{\sigma}_2) + \hat{w}_3 N(\hat{\mu}_3, \hat{\sigma}_3) + \hat{w}_4 N(\hat{\mu}_4, \hat{\sigma}_4) + \hat{w}_5 N(\hat{\mu}_5, \hat{\sigma}_5) \quad (4)$$

Dalam model (4) terdapat parameter untuk tiap komponen dan diperoleh hasil estimasi parameter yang ditampilkan dalam tabel 2 berikut:

Tabel 2. Estimasi Parameter *Finite Mixture* Normal

Parameter	Komponen 1	Komponen 2	Komponen 3	Komponen 4	Komponen 5
Pi	0,2	0,21	0,22	0,22	0,15
Mu	5,55	16,97	26,69	36,02	49,23
Sig	33,01	52,46	61,34	62,01	53,32

Dari tabel 2, terlihat bahwa besar parameter bobot (*weight*) dari tiap komponennya yaitu $\hat{w}_1 = 0{,}2$, $\hat{w}_2 = 0{,}21$, $\hat{w}_3 = 0{,}22$, $\hat{w}_4 = 0{,}22$, dan $\hat{w}_5 = 0{,}15$. Tabel tersebut juga menunjukkan parameter-parameter distribusi Normal yaitu $\hat{\mu}_1 = 5{,}55$ dan $\hat{\sigma}_1 = 5{,}75$ untuk komponen pertama, $\hat{\mu}_2 = 16{,}97$ dan $\hat{\sigma}_2 = 7{,}24$ untuk komponen kedua, $\hat{\mu}_3 = 26{,}69$ dan $\hat{\sigma}_3 = 7{,}83$ untuk komponen ketiga, $\hat{\mu}_4 = 36{,}02$ dan $\hat{\sigma}_4 = 7{,}87$ untuk komponen keempat, dan $\hat{\mu}_5 = 49{,}23$ dan $\hat{\sigma}_5 = 7{,}30$ untuk komponen kelima.

Berdasarkan model *finite mixture* yang didapat, maka dapat diketahui deskripsi dari masing-masing klaster atau kelompok negara. Klaster pertama terdiri atas 20% dari populasi atau terdapat 30 negara yang rata-rata kasus *stunting*-nya sebesar 5,55%. Begitupula untuk kluster kedua yang terdiri atas 21% atau 32 negara dengan rata-rata kasus *stunting*-nya sebesar 16,97 %, klaster ketiga terdiri atas 22% dari populasi atau terdapat 33 negara yang rata-rata kasus *stunting*-nya sebesar 26,69 %, klaster keempat terdiri atas 22% dari populasi atau terdapat 34 negara yang rata-rata kasus *stunting*-nya sebesar 36,02 %, dan klaster terakhir yaitu klaster kelima terdiri atas 20% dari populasi atau terdapat 22 negara yang rata-rata kasus *stunting*-nya sebesar 53,32 %.

Klaster pertama dan kedua didominasi sebagian negara Asia dan beberapa negara Amerika atau Eropa. Klaster pertama dan kedua bisa dikatakan merupakan klaster yang memiliki kasus *stunting* yang rendah karena negara-negara yang termasuk dalam klaster tersebut merupakan negara maju. Sedangkan klaster ketiga, keempat, dan kelima merupakan klaster yang didominasi oleh negara yang berada di Afrika dan sebagian negara yang berada di Asia. Ketiga klaster tersebut termasuk klaster dengan kasus *stunting* yang tinggi karena kebanyakan negara yang masuk kedalam klaster tersebut merupakan negara miskin dan negara berkembang.

4. Kesimpulan

Data kasus *stunting* per negara di dunia tahun 2019 menunjukkan pola *multimodal* sehingga data tersebut cocok untuk dimodelkan dengan model *finite mixture*. Dengan pendekatan Bayesian melalui BD-MCMC diperoleh hasil yaitu model *finite mixture* Normal dengan lima komponen. Kelima komponen tersebut menggambarkan lima klaster dimana klaster pertama terdiri dari 30 negara, klaster kedua terdiri dari 32 negara, klaster ketiga terdiri dari 33 negara, klaster keempat terdiri dari 34 negara, dan klaster kelima terdiri dari 22 negara. Klaster pertama dan kedua didominasi sebagian negara Asia dan beberapa negara Amerika atau Eropa. Sedangkan klaster ketiga, keempat, dan kelima merupakan klaster yang didominasi oleh negara yang berada di Afrika dan sebagian negara yang berada di Australia.

Daftar Pustaka

[1] Fruhwirth-Schnatter, S., (2006), *Finite Mixture and Markov Switching Models*. New York: Springer.

[2] Iriawan, N., 2000, *Computationally Intensive Approaches to Inference in NeoNormal Linier Models*, Australia.

[3] Purwandini, K., & Kartasurya, M. I., (2013), Pengaruh Pemberian Mikronutrient Sprinkle Terhadap Perkembangan Motorik Anak Stunting Usia 12-36 Bulan. *Journal of Nutrition College*, **2(1)**, 50-59.

[4] Stephens, M., (2000), Bayesian analysis of mixture models with an unknown number of components-an alternative to reversible jump methods. *Annals of statistics*, 40-74.

[5] Susanto, I. dan Handajani, S.S., (2020), Pemodelan Distribusi Pendapatan Rumah Tangga Per Kapita di Indonesia dengan Model Finite Mixture. *Media Statistika*, **13(1)**, 13-24.

[6] Welasasih, B.D. dan Wirjatmadi, R.B., (2012), Beberapa Faktor Yang Berhubungan Dengan Status Gizi Balita Stunting. *The Indonesian Journal of Public Health*, **3(8)**, 99-104.

[7] World Health Organization., (2019), *Levels and trends in child malnutrition: key findings of the 2019 edition* (No. WHO/NMH/NHD/19.20). World Health Organization.

[8] Wulandari, S., & Kurniawan, R., (2019), Pengelompokan Kabupaten/Kota di Jawa Timur Berdasarkan Kasus Stunting Balita Menggunakan Algoritme Fuzzy Particle Swarm Optimization-Fuzzy C-Means. *Jurnal Statistika Universitas Muhammadiyah Semarang*, ***7*(1).**