

---

**Meningkatkan Ketahanan Wilayah Melalui Estimasi *Underreported* Data  
Kejahatan Dengan Pendekatan Bayes  
(Studi Di Kota Ambon, Provinsi Maluku)**

***Herlin Venny Johannes***

Departemen Statistik, Fakultas Matematika dan Ilmu Pengetahuan Alam,  
Universitas Padjajaran  
Email: [herlin.jo86@gmail.com](mailto:herlin.jo86@gmail.com)

***Septiadi Padmadisastra***

Departemen Statistik, Fakultas Matematika dan Ilmu Pengetahuan Alam,  
Universitas Padjajaran  
Email: [s\\_padmadisastra@yahoo.com](mailto:s_padmadisastra@yahoo.com)

***Bertho Tantular***

Departemen Statistik, Fakultas Matematika dan Ilmu Pengetahuan Alam,  
Universitas Padjajaran  
Email: [bertho@unpad.ac.id](mailto:bertho@unpad.ac.id)

***ABSTRACT***

*This paper presented a study for the number of crime that run into underreporting counts. The purpose of the analysis was to estimated parameter of the model which was the actual number of crime. The model was a mixture of the poisson and the binomial distributions developed by Winkelmann (1996). The parameters of the model were estimated by Bayesian approach and Markov Chain Monte Carlo simulation using Gibbs sampling algorithm. Determination the convergence of the algorithm using trace plot, autocorrelation plot and ergodic mean plot. In the end, estimator of the parameters of the underreported counts model were the simulation sample mean that calculated from the simulation sample of iteration after burn in period until the last iteration.*

***Keywords: Underreporting counts, Bayesian, MCMC, Gibbs sampling***

***ABSTRAK***

Penelitian ini mengkaji permodelan data tingkat kejahatan yang mengalami *underreporting counts*. Tujuan analisis ini adalah untuk menaksir parameter model yaitu banyaknya jumlah tindak kejahatan yang sebenarnya. Model yang digunakan adalah hasil penggabungan antara distribusi poisson dan distribusi binomial yang dikembangkan oleh Winkelmann (1996). Penaksiran parameter model dilakukan melalui pendekatan Bayes dan simulasi *Markov Chain Monte Carlo* menggunakan algoritma *Gibbs sampling*. Penentuan konvergensi algoritma dilakukan melalui *trace plot*, *autocorrelation plot*, dan *ergodic mean plot*. Taksiran parameter model diperoleh dari rata-rata nilai sampel hasil simulasi yang dihitung dari iterasi setelah *burn in period* sampai dengan iterasi yang terakhir.

***Kata kunci: Underreporting counts, Bayesian, MCMC, Gibbs sampling***

## PENGANTAR

Kebutuhan fundamental setiap manusia terdiri dari kebutuhan biologis seperti makan, minum serta tidur dan kebutuhan sosial, seperti status sosial, peranan sosial, aktualisasi diri dan rasa aman. Saat ini dapat dikatakan bahwa rasa aman merupakan salah satu kebutuhan dasar manusia dalam menjalankan aktivitas sehari-harinya. Menurut Abraham Maslow dalam teori hierarki, kebutuhan manusia meliputi rasa aman berada pada tingkatan kedua dibawah kebutuhan dasar manusia seperti sandang, pangan dan papan. Hal ini menunjukkan bahwa rasa aman merupakan kebutuhan manusia yang penting (Maslow, 1943).

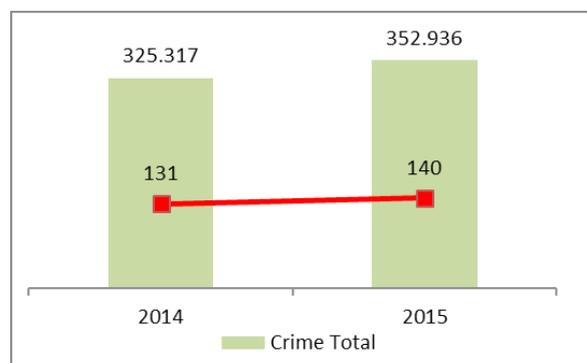
Rasa aman (*security*) merupakan salah satu hak asasi yang harus diperoleh atau dinikmati setiap orang. Hal ini tertuang dalam UUD Republik Indonesia 1945 Pasal 28G ayat 1 yang menyebutkan: “Setiap orang berhak atas perlindungan diri pribadi, keluarga, kehormatan, martabat, dan harta benda yang di bawah kekuasaannya, serta berhak atas rasa aman dan perlindungan dari ancaman ketakutan untuk berbuat atau tidak berbuat sesuatu yang merupakan hak asasi”.

Rasa aman merupakan variabel yang sangat luas karena mencakup berbagai aspek dan dimensi, mulai dari dimensi politik, hukum, pertahanan, keamanan, sosial dan ekonomi. Sejalan dengan itu, statistik dan indikator yang biasa digunakan untuk mengukur rasa aman masyarakat merupakan indikator negatif, misalnya jumlah angka kejahatan (*crime total*), jumlah orang yang berisiko terkena tindak kejahatan (*crime rate*) setiap 100.000 penduduk. Semakin tinggi angka kriminalitas menunjukkan semakin banyak tindak kejahatan yang terjadi di masyarakat yang merupakan indikasi bahwa masyarakat merasa semakin tidak aman.

Kriminalitas merupakan ancaman terbesar bagi kemanusiaan yang secara langsung dapat melemahkan ketahanan wilayah. Ketahanan wilayah adalah kondisi dinamik suatu wilayah yang meliputi seluruh aspek kehidupan masyarakat yang terintegrasi, berisi keuletan dan ketangguhan yang mengandung kemampuan mengembangkan kekuatan wilayah dalam menghadapi dan mengatasi segala tantangan, ancaman, hambatan serta gangguan baik yang datang dari dalam dan dari luar, yang langsung maupun tidak langsung (Lemhannas, 2000).

Selama periode Tahun 2014–2015, jumlah kejadian kejahatan atau tindak kriminalitas di Indonesia mengalami peningkatan. Seperti yang disajikan pada gambar 1, data di Biro Pembinaan dan Operasional, Mabes Polri memperlihatkan memperlihatkan jumlah kejadian kejahatan (*crime total*) pada tahun 2014 sebanyak 325.317 kasus meningkat pada tahun 2015 menjadi 352.936 kasus. Sejalan dengan itu, jumlah orang yang berisiko terkena tindak kejahatan (*crime rate*) setiap 100.000 penduduk diperkirakan sebanyak 131 orang pada tahun 2014 meningkat menjadi 140 orang pada tahun 2015.

Gambar 1  
Jumlah Kejahatan (*Crime Total*) dan Tingkat Resiko Terkena Kejahatan (*Crime Rate*) Indonesia, Tahun 2014-2015



Sumber: BPS, 2016

Dari segi jumlah kejahatan untuk level provinsi/polda, selama tahun 2015 Polda Metro Jaya mencatat jumlah kejahatan terbanyak (44.461 kasus), disusul oleh Polda Jawa Timur (35.437 kasus), dan Sumatera Utara (35.248 kasus). Sedangkan Polda Maluku, Papua Barat, dan Maluku Utara jumlah kejadian kejahatan berturut-turut sebanyak 1.843; 1.356; dan 814, merupakan tiga Polda dengan jumlah kejahatan paling sedikit. Jumlah kejahatan bisa sangat dipengaruhi dengan banyaknya jumlah penduduk di suatu wilayah.

Hal ini terlihat dari meskipun Jawa Timur dari sisi jumlah kejadian (*crime total*) termasuk kategori tiga terbesar, tetapi dari sisi tingkat resiko terkena kejahatan (*crime rate*) termasuk kelompok lima terendah. Terkait *crime rate*, Polda Sulawesi Utara mencatat tingkat resiko terkena kejahatan tertinggi yakni 328 (setiap 100.000 penduduk diperkirakan sebanyak 328 orang yang berisiko terkena tindak kejahatan), disusul oleh Polda Sulawesi Tengah (317) dan Sumatera Barat (317). Peningkatan yang terjadi baik dari segi frekuensi maupun intensitas kejahatan mencerminkan tingkat keamanan masyarakat yang rendah.

Kriminalitas yang terjadi di negara berkembang seperti Indonesia seringkali diakibatkan oleh kepadatan penduduk akibat dari persebaran penduduk yang tidak merata antar daerah karena tingginya arus urbanisasi, pembangunan yang tidak merata sehingga terjadi ketimpangan pembangunan antar daerah, tingkat pendidikan yang rendah, ketimpangan pendapatan, banyaknya pengangguran dan lain sebagainya. Baik secara nasional maupun daerah, kriminalitas yang terjadi di berbagai wilayah Indonesia disebabkan oleh hal yang sama, begitu juga dengan tingkat kriminalitas yang terjadi di Maluku.

Jumlah kejahatan yang dilaporkan kepada Polda Maluku ada sejumlah 1.843 kasus kejahatan atau kriminalitas selama periode tahun 2015. Jika dilihat dari tahun 2014 dan tahun 2015, jumlah kriminalitas yang terjadi di Maluku mengalami peningkatan (BPS, 2016). Dari sisi jumlah kejadian kejahatan (*crime total*), Kota Ambon merupakan daerah di Maluku dengan jumlah kejahatan paling banyak. Pada tahun 2014, tercatat ada 1003 kasus yang di laporkan ke Polres Ambon, namun mengalami penurunan di tahun 2015 menjadi 919 kasus dan terus mengalami penurunan di tahun 2016 menjadi 898 kasus. Dari sisi tingkat resiko terkena kejahatan (*crime rate*), pada tahun 2016 Kota Ambon mencatat tingkat resiko terkena kejahatan tertinggi yakni 223 (setiap 100.000 penduduk diperkirakan sebanyak 223 orang yang berisiko terkena tindak kejahatan), disusul oleh Kabupaten Maluku Tenggara (218) dan Kabupaten Maluku Tenggara Barat (121). Peningkatan yang terjadi baik dari segi frekuensi maupun intensitas kejahatan mencerminkan tingkat keamanan masyarakat yang rendah.

Namun data kriminalitas ini hanya mencakup seluruh peristiwa atau kejadian kriminalitas yang dilaporkan oleh masyarakat atau aksi kriminalitas yang pelakunya tertangkap tangan oleh kepolisian. Mengingat masih tingginya keengganan masyarakat untuk melapor, diduga data yang dihasilkan berada dalam keadaan "*underreported*". Artinya, kejadian kriminalitas yang sesungguhnya diduga lebih besar dari yang dilaporkan. Dengan kata lain, angka gelap (*dark number*) kejahatan masih relatif besar.

Selama ini, permodelan dari sebuah data cacahan sering tidak memperhitungkan adanya kemungkinan *underreported count*

seperti regresi poisson dan regresi binomial. Model untuk *underreported* data pertama kali diperkenalkan oleh Moran's (1952). Kemudian, beberapa peneliti mengembangkan model poisson untuk *underreported count* yang dikembangkan oleh Winkelmann (1996), model regresi binomial negatif yang dikembangkan oleh Mukhopadhyay dan *generalized poisson-poisson mixture model* yang dapat digunakan untuk *misreporting* yang dikembangkan oleh Pararai (2010). Padmadisastra dan Supriadi (2014) meneliti penyusunan peta penyakit dengan memperhitungkan adanya kemungkinan *underreported*. Dalam penelitiannya, penaksir untuk tingkat kesakitan dikoreksi dengan memperhatikan keadaan *underreported* dalam data.

Penelitian tentang tindak kejahatan yang dilakukan selama ini tidak memperhitungkan adanya kemungkinan *underreported*. Penelitian ini diusulkan untuk melakukan penaksiran angka kriminalitas yang sebenarnya dengan memperhitungkan keadaan tersebut sehingga diharapkan akan lebih meningkatkan akurasi hasil estimasi tindak kejahatan. Penelitian ini dilakukan dengan melibatkan kovariat ke dalam model regresi yang dapat dimanfaatkan untuk melakukan prediksi mengenai efek kovariat terhadap estimasi tindak kejahatan.

### **Underreporting Data Kejahatan**

Data mengenai banyaknya kejadian kejahatan di suatu wilayah yang diperoleh dengan cara mencacah banyaknya (*count*) kejadian merupakan bilangan bulat *non-negatif*. Data cacahan seperti ini sering mengalami kerusakan karena berbagai alasan. Kerusakan yang dimaksud adalah bahwa data yang dilaporkan tidak sesuai (*misreporting*) dengan keadaan sebenarnya.

*Misreporting count* ini dapat terjadi dalam setiap sistem pelaporan. Li et al (dalam Pararai, 2010) menyatakan bahwa dalam sebuah model, *misreporting* dapat terjadi sebagai  $y_i, i = 1, 2, \dots, n$  berbeda dari nilai sebenarnya  $y_i^*, i = 1, 2, \dots, n$ .

Dengan demikian *misreporting count* terbagi menjadi dua, yaitu *underreporting count* dan *overreporting count*. *Underreporting* adalah sebuah masalah yang dapat terjadi saat pengumpulan data, ketika pencacahan sebuah kejadian untuk beberapa alasan menjadi tidak lengkap (Neubauer, 2011). *Underreported count* terjadi jika banyak kejadian yang diamati, dilaporkan lebih kecil dari banyak kejadian yang sebenarnya ( $y_i < y_i^*$ ). Sebaliknya *overreported count* terjadi jika kejadian yang sedang diamati dilaporkan melebihi dari banyaknya kejadian yang sebenarnya ( $y_i > y_i^*$ ).

Penyebab *underreporting count* akan berbeda dalam berbagai bidang aplikasi. Sebagai contoh, Morocoima (dalam Scollnik, 2006) menyatakan bahwa tidak semua kecelakaan lalu lintas dilaporkan kepada pihak kepolisian. Ada banyak alasan yang dinyatakan oleh peneliti tersebut mengenai mengapa kecelakaan lalu lintas mungkin tidak dilaporkan. Salah satunya dapat disebabkan seorang individu yang mengalami kecelakaan tidak tahu kepada siapa kecelakaan harus dilaporkan. Akibatnya, banyak kecelakaan lalu lintas yang tercatat oleh pihak kepolisian akan berada dalam keadaan *underreported*. Pada penelitian lain mengenai *underreported data* yang dilakukan oleh Rian (2014) menyatakan bahwa *underreported count* pada banyaknya produk yang terjual terjadi sebagai keterlambatan input data yang terjadi secara berulang.

Selanjutnya dalam penelitian ini ada sejumlah korban tindak kejahatan yang tidak

melaporkan kejadian kejahatan yang dialami ke pihak yang berwajib (polisi) sehingga datanya tidak tercatat. Berdasarkan *Bureau of Justice Statistics, U.S. Department of Justice* (2012) dalam laporannya “*Victimizations Not Reported to the Police*” ada beberapa alasan mengapa korban kejahatan tidak melapor ke polisi, yaitu: takut akan pembalasan dari pelaku, tidak ingin mendatangkan kesulitan bagi pelaku, korban merasa tidak cukup penting untuk melapor, polisi tidak mau atau tidak bisa membantu, penyelesaian masalah dengan cara lain atau masalah pribadi dan alasan lainnya.

Sejalan dengan hal tersebut, alasan *underreporting* data kejahatan di negara-negara berkembang dapat disebabkan oleh beberapa alasan, yaitu sistem pelaporan yang berbelit-belit, ketidakpercayaan bahwa pihak berwajib dapat menindaklanjuti/ menangani kasus dan keengganan melaporkan kasus-kasus kejahatan tertentu seperti kekerasan dalam rumah tangga, pemerkosaan dan kasus asusila lainnya karena dianggap sebagai aib dalam masyarakat. Oleh karenanya data yang dilaporkan pihak kepolisian menjadi lebih rendah dari yang seharusnya.

Konsekuensi dari kejadian *underreporting* adalah bahwa nilai rata-rata dari kejadian yang diamati lebih kecil dibandingkan dengan nilai rata-rata sebenarnya (Neubauer, 2011). Artinya *underreported count* tidak bisa diabaikan begitu saja dalam sebuah permodelan. Oleh karena itu, diperlukan sebuah pengembangan model yang dapat digunakan untuk dapat menaksir nilai dari rata-rata kejadian yang sebenarnya.

### **Model Poisson Untuk *Underreported Count***

Model untuk *underreported count* pertama kali dikenalkan dalam Moran’s

(1952) “*A Characterization Property of the Poisson Distribution*”. Moran (1952) (dalam Papadatos, 2005) menyatakan bahwa  $N_1$  dan  $N_2$  adalah variabel acak independen *no-generate* yang bernilai non-negatif dan jika distribusi bersyarat dari  $N_1$  dimana  $(N_1 + N_2 = n)$  adalah binomial dengan parameter  $n \in N = \{0,1, \dots\}$  dan peluang sukses  $p \in [0,1]$  untuk semua  $n \in N, P[N_1 + N_2 = n] > 0$  maka distribusi  $N_1, N_2$  dan  $(N_1 + N_2 = n)$  adalah poisson. Dimana untuk  $i \in N, P[N_1 = i] > 0$  dan  $P[N_2 = i] > 0$ .

Bentuk lain dari Moran’s characterization (1952) diperkenalkan oleh Rao & Rubin (1964) dalam makalah “*On a Characterization of the Poisson Distribution*”, dimana kebebasan  $N_1$  dan  $N_2$  disebut sebagai *Rao-Rubin condition* yang didefinisikan sebagai:

$$P[N_2 = n_2 | N_1 = 0] = P[N_2 = n_2], n_2 \in N \quad (2.1)$$

Rao & Rubin (1964) (dalam Papadatos, 2005) menyatakan bahwa jika distribusi dari  $N = N_1 + N_2$  tidak terkonsentrasi pada 0 dan jika untuk semua  $n \in N$  dengan  $P[N = n] > 0$ , maka:

$$P[N_1 = n_1 | N = n] = \binom{n}{n_1} p^{n_1} (1 - p)^{n - n_1}, n_1 = 0, 1, 2, \dots \quad (2.2)$$

untuk  $p \in [0,1]$  maka *Rao-Rubin conditional* pada persamaan (2.1) memberikan pengertian bahwa  $N_1$  dan  $N_2$  berdistribusi poisson dengan parameter  $\lambda p$  dan  $\lambda(1 - p)$  dengan  $\lambda > 0$ .

Dengan kata lain, *Rao-Rubin condition* pada persamaan (2.1) ekuivalen dengan kebebasan  $N_1$  dan  $N_2$  di bawah *binomial damage model* (Papadaos, 2005). Berkaitan dengan *underreported counts*, *Rao-Rubin*

condition menunjukkan  $N_1$  sebagai observasi yang dilaporkan (*undamaged*) sedangkan  $N_2 = N - N_1$  sebagai observasi yang tidak dilaporkan (*damage*) sehingga  $N_1$  dan  $N_2$  adalah bagian dari variable acak diskrit  $N$ . Dengan demikian, jelas bahwa model *underreported count* merupakan model poisson dengan parameter  $\lambda p$  ( $\lambda > 0$ ).

### Model Regresi Poisson Untuk *Underreported Count*

Seperti Moran (1952) dan Rao & Rubin (1964) (dalam Papadatos, 2005), model regresi poisson untuk *underreported counts* yang dikembangkan oleh Winkelmann (1996) juga merupakan model hasil penggabungan antara distribusi poisson dan distribusi binomial. Anggap bahwa  $Y_i^*$  adalah banyaknya kejadian yang diamati pada waktu tertentu untuk unit pengamatan ke- $i$ .  $Y_i^*$  diasumsikan bergantung pada variabel independen  $x$  dan berdistribusi poisson dengan rata-rata:

$$E[Y_i^* | \mathbf{x}_i] = \lambda_i = \exp[\mathbf{x}_i' \boldsymbol{\beta}] \quad (2.3)$$

dengan:  $i = 1, 2, \dots, n$

$$\mathbf{x} = \begin{bmatrix} 1 & 1 & \dots & 1 \\ X_{11} & X_{12} & \dots & X_{1n} \\ X_{21} & X_{22} & \dots & X_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ X_{k1} & X_{k2} & \dots & X_{kn} \end{bmatrix} = \begin{bmatrix} \mathbf{1} \\ \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_k \end{bmatrix} \text{ dan } \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix}$$

Fungsi distribusi peluang  $Y_i^*$  dapat didefinisikan sebagai:

$$P[Y_i^* | \lambda_i] = \frac{[\exp(\mathbf{x}_i' \boldsymbol{\beta})]^{y_i^*} e^{-\exp(\mathbf{x}_i' \boldsymbol{\beta})}}{y_i^*!} \quad (2.4)$$

untuk  $Y_i^* = 0, 1, 2, \dots; i = 1, 2, \dots, n$

Kemudian asumsikan bahwa banyaknya kejadian yang diamati pada waktu tertentu

untuk unit pengamatan ke- $i$  yang dilaporkan adalah  $Y_i$  dan berdistribusi binomial dengan fungsi distribusi peluang sebagai berikut:

$$P(Y_i | Y_i^*, p) = \binom{y_i^*}{y_i} p^{y_i} (1-p)^{y_i^* - y_i} \quad (2.5)$$

untuk  $Y_i^* = 0, 1, 2, \dots; Y_i = 1, 2, \dots; i = 1, 2, \dots, n; 0 \leq p \leq 1$

$p$  = peluang sebuah kejadian dilaporkan

*Underreported count* terjadi jika  $Y_i^* > Y_i$  sehingga fungsi distribusi peluang marginal  $Y_i$  dapat didefinisikan sebagai:

$$\begin{aligned} P(Y_i = y_i) &= \sum_{y_i^* \geq y_i} P(Y_i^* = y_i^* | \lambda) P(Y_i = y_i | y_i^*, p) \\ &= \sum_{y_i^* \geq y_i} \frac{\lambda_i^{y_i^*} e^{-\lambda_i}}{y_i^*!} \binom{y_i^*}{y_i} p^{y_i} (1-p)^{y_i^* - y_i} \\ &= \frac{(\lambda_i p)^{y_i} e^{-\lambda_i p}}{y_i!} \end{aligned} \quad (2.6)$$

dengan  $Y_i = 1, 2, \dots; \lambda_i \geq 0; i = 1, 2, \dots, n; 0 \leq p \leq 1$

$p$  = peluang sebuah kejadian dilaporkan

Berdasarkan persamaan (2.6) dapat dilihat bahwa model regresi untuk *underreported count* merupakan model poisson dengan rata-rata:

$$\lambda_i p = \exp(\mathbf{x}_i' \boldsymbol{\beta}) \cdot p \quad (2.7)$$

dengan  $i = 1, 2, \dots, n; 0 \leq p \leq 1$

Winkelmann, menggunakan model pada persamaan (2.6) untuk mengevaluasi hubungan antara banyaknya ketidakhadiran pekerja yang *underreported* dengan upah, status kesehatan dan jenis pekerjaan. Hasil penelitiannya menunjukkan bahwa tidak terdapat bukti tentang hubungan negatif antara upah dengan ketidakhadiran

pekerja. Hasil ini sesuai dengan kesimpulan yang diberikan oleh Johansson & Palme, namun bertentangan dengan hasil penelitian yang dilakukan oleh Delgado & Kniesner yang menemukan bahwa hubungan antara upah dan ketidakhadiran adalah signifikan.

### **Penaksiran Parameter Model *Underreported Count***

Terdapat dua pendekatan dalam metode penaksiran, yaitu pendekatan statistika klasik dan pendekatan Bayesian. Melalui pendekatan statistika klasik, penaksiran parameter sepenuhnya berdasarkan data sampel yang diambil dari populasi. Sedangkan melalui pendekatan Bayesian, disamping memanfaatkan data sampel juga memperhitungkan suatu distribusi awal yang disebut distribusi *prior*. Pendekatan Bayesian memandang parameter dari sebuah model sebagai variable acak yang memiliki distribusi yang disebut dengan distribusi *prior*.

Informasi dari sampel penelitian sebagai *likelihood* digunakan sebagai bahan untuk memperbaharui informasi prior menjadi sebuah *posterior* yang akan digunakan untuk bahan inferensi. Dengan kata lain metode Bayes merupakan sebuah metode yang optimal tentang bagaimana memperbaharui kepercayaan seseorang tentang parameter populasi setelah diberikannya informasi baru (Hoff, 2009). Distribusi *posterior* diperlukan untuk menentukan estimasi parameter model hasil taksiran Bayesian yang merupakan rata-rata dari distribusi *posterior*. Secara proposional Teorema Bayes (Bolstad, 2007) didefinisikan sebagai:

$$\textit{posterior} \propto \textit{likelihood} \times \textit{prior} \quad (2.8)$$

Dalam penaksiran parameter model *underreported counts*, beberapa peneliti

menggunakan pendekatan statistik klasik dan beberapa lainnya menggunakan pendekatan Bayesian. Olkin et al. (1981) (dalam Moreno, 1998) menggunakan metode maksimum *likelihood* dan metode momen dalam menaksir parameter model *underreported counts*. Penaksir yang dihasilkan sangat tidak stabil, sebuah perubahan kecil pada data dapat menyebabkan perubahan besar pada taksiran parameter model (Moreno, 1998). Sedangkan pendekatan Bayes untuk menaksir parameter model *underreported count* telah diterapkan oleh Draper & Guttman (1971) dan Rafteri (dalam Moreno, 1998; Winkelmann (1996).

Mengenai keuntungan penggunaan pendekatan bayes dalam analisis statistik, Hoff (2009) menyatakan bahwa pendekatan Bayes mampu menyediakan inferensia dan estimasi yang stabil meski pada situasi dimana ukuran sampel sangat kecil. Bolstad (2007) juga menambahkan bahwa pendekatan Bayes memiliki keuntungan karena mengakomodir penggunaan informasi prior yang tidak diakomodir oleh pendekatan klasik (frekuentis). Informasi prior ini sangat logis penerapannya mengingat dalam dunia sains biasanya terdapat informasi terdahulu (prior) dari sebuah fenomena yang sedang diamati.

Berdasarkan persamaan 2.8 dapat dilihat bahwa distribusi *posterior* akan bergantung pada pemilihan distribusi prior. Dengan demikian permasalahan utama adalah bagaimana memilih distribusi prior untuk parameter yang tidak diketahui tetapi memiliki relevansi dengan permasalahan data penelitian. Menurut Pertiwi (2012), terdapat beberapa macam distribusi prior, yaitu (1). *Conjugate* dan *Non-Conjugate Prior*. *Conjugate prior* merupakan prior yang dikaitkan dengan pola model *likelihood* dari

datanya. (2). *Informative* dan *Non-Informative Prior*. Prior dikatakan *Informative* atau *Non-Informative* dilihat dari sudah diketahui pola atau frekuensi dari data observasi atau belum. (3). *Proper* dan *Improper Prior*. Prior dikatakan *improper* apabila fungsi yang digunakan sebagai “densitas peluang prior” memiliki integral yang infinit (tidak dapat diselesaikan). (4). *Pseudo-prior*. Prior terkait dengan pemberian nilainya yang disetarakan dengan hasil elaborasi frekuentis.

Pada penelitian *underreported count* sebelumnya para peneliti memilih distribusi prior yang berbeda-beda. Draper & Guttman (1971) mengasumsikan  $N$  pada persamaan (2.2) berdistribusi *uniform* dan  $p$  berdistribusi *beta*. Rafteri (dalam Moreno, 1998) mengasumsikan  $N$  berdistribusi *poisson* dengan parameter  $\lambda$  dan  $p$  berdistribusi *uniform*. Sedangkan Winkelmann (1996) mengasumsikan  $N$  berdistribusi *poisson* dengan parameter  $\lambda = \exp[x'\beta]$  dimana  $\beta$  diasumsikan berdistribusi *normal* dengan parameter  $\mu$  dan  $\sigma$ , sedangkan  $p$  berdistribusi *uniform*, sehingga dalam sebuah model regresi untuk *underreported counts*, berdasarkan persamaan (2.8) distribusi posterior yang diperoleh Winkelmann (1996) didefinisikan sebagai:

$$P(y^*, p, \beta | y, x) \propto P(y | y^*, p) \times P(y^* | \beta) \times f(\beta) \times f(p) \quad (2.9)$$

Kesulitan dapat terjadi karena kompleksitas saat penentuan distribusi *posterior* pada persamaan (2.9) sehingga untuk menaksir parameter model digunakan teknik *Markov Chain Monte Carlo* (MCMC).

### Markov Chain Monte Carlo (MCMC)

Metode *Markov Chain Monte Carlo* (MCMC) merupakan metode simulasi untuk

penarikan sampel dari distribusi posterior dan perhitungan besaran yang diamati. Metode ini menarik sampel berturut-turut dari suatu distribusi targetnya. MCMC membangkitkan data sampel parameter  $\theta$  yang memiliki distribusi tertentu melalui sebuah algoritma, dimana nilai setiap langkah bergantung pada satu langkah sebelumnya sehingga membentuk suatu rantai Markov. Salah satu algoritma yang sering digunakan dalam MCMC adalah *gibbs sampling*.

### Algoritma Gibbs Sampling

Gibbs sampling adalah kasus khusus dari algoritma *Metropolis-Hasting* dengan peluang penerimaan  $\alpha$  selalu bernilai satu. Dalam algoritma *Metropolis-Hasting* peluang penerimaan  $\alpha$  (Ntzoufras, 2009) didefinisikan sebagai :

$$\alpha = \min \left[ 1, \frac{f(y|\theta^t)f(\theta^t)q(\theta|\theta^t)}{f(y|\theta)f(\theta)q(\theta^t|\theta)} \right]$$

dengan :  $f(y|\theta^t)$  = fungsi *likelihood* parameter  $\theta$  pada iterasi ke- $t$ ;  $f(y|\theta)$  = fungsi *likelihood* parameter  $\theta$ ;  $f(\theta^t)$  = fungsi distribusi *prior* parameter  $\theta$  pada iterasi ke- $t$ ;  $f(\theta)$  = fungsi distribusi *prior* parameter  $\theta$ ;  $q(\theta|\theta^t)$  = fungsi distribusi proposal parameter  $\theta$  bersyarat  $\theta^t$ ; dan  $q(\theta^t|\theta)$  = fungsi distribusi proposal parameter  $\theta^t$  bersyarat  $\theta$

Gibbs sampling bisa diterapkan apabila distribusi bersyarat dari tiap-tiap parameter diketahui. Asumsikan terdapat  $k$  parameter  $\theta$ , maka algoritma gibbs sampling dapat dituliskan sebagai berikut : (1). Tentukan nilai awal  $\theta^0 = (\theta_1^0, \theta_2^0, \dots, \theta_k^0)$ . (2). Ulangi langkah untuk  $t=1, 2, \dots, T$ . Bangkitkan  $\theta_1^{t+1}$  dari  $f_1(\theta_1 | y, \theta_2^t, \theta_3^t, \dots, \theta_k^t)$ . Bangkitkan  $\theta_2^{t+1}$  dari  $f_2(\theta_2 | y, \theta_1^t, \theta_3^t, \dots, \theta_k^t)$ . Bangkitkan  $\theta_k^{t+1}$  dari  $f_k(\theta_k | y, \theta_1^t, \theta_2^t, \dots)$ . Fungsi peluang  $f_1, f_2, \dots, f_k$  disebut dengan distribusi bersyarat penuh dan distribusi peluang yang digunakan

untuk simulasi. Dalam gibbs sampling tidak ada mekanisme penerimaan dan penolakan, semua sample hasil simulasi diterima.

### Konvergensi Algoritma

Langkah penting dalam simulasi MCMC adalah mengetahui kekonvergenan dari algoritma yang digunakan. Terdapat beberapa cara untuk mengevaluasi apakah rantai Markov telah mencapai kondisi konvergen dalam algoritma MCMC (Anifa, 2012), yaitu

*Pertama, Trace Plot.* Salah satu cara pendugaan *burn-in* periode adalah memeriksa *trace plot* nilai simulasi dari komponen atau beberapa fungsi lainnya dari terhadap jumlah iterasi. *Trace plot* merupakan gambaran sebuah plot dari iterasi versus nilai yang telah dibangkitkan. *Trace plot* terutama sekali penting ketika algoritma MCMC dimulai dengan nilai-nilai parameter yang jauh dari pusat distribusi target. Pada kasus seperti itu, nilai-nilai simulasi dari pada awal iterasi algoritma akan menyimpang dari daerah ruang parameter dimana distribusi target dipusatkan. Sebuah trend naik atau turun pada nilai parameter pada *trace plot* menunjukkan bahwa *burn-in period* belum tercapai. Dengan kata lain, jika semua nilai-nilai berada dalam sebuah daerah tanpa keperiodikan yang kuat cenderung dapat diasumsikan konvergen.

*Kedua, Autokorelasi.* Untuk algoritma MCMC, nilai simulasi pada iterasi ke-(t+1) bergantung pada nilai simulasi pada iterasi ke-t. Jika pada rantai terdapat korelasi yang kuat diantara nilai-nilai yang berurutan, maka kedua nilai berurutan tersebut memberikan informasi hanya secara marginal mengenai distribusi target dan bukan nilai dari sebuah simulasi tunggal. Korelasi yang kuat diantara iterasi yang

berurutan menunjukkan bahwa algoritma masih berada pada daerah tertentu dari ruang parameter dan mungkin membutuhkan waktu yang lama untuk penyampelan dari keseluruhan daerah distribusi. Statistik yang umum digunakan untuk mengukur tingkat ketergantungan diantara pengambilan berurutan pada rantai adalah autokorelasi. Statistik autokorelasi tersebut didefinisikan sebagai :

$$r_{iL} = \frac{T'}{T' - L} = \frac{\sum_{i=1}^{T'-L} (\theta_i - \bar{\theta})(\theta_{iL} - \bar{\theta})}{\sum_{i=1}^{T'-L} (\theta_i - \bar{\theta})^2}$$

Dengan :  $\bar{\theta}$  = rata-rata dari nilai simulasi; T = banyaknya iterasi yang dilakukan; L= banyaknya iterasi yang dihilangkan pada burn-in periode;  $T' = T - L$ ; L= selisih sampling iterasi yang terpisah antara dua kumpulan nilai-nilai.

*Ketiga, Ergodic Mean Plot.* Istilah *ergodic mean* menunjukkan nilai mean sampai *current iteration*. *Ergodic mean plot* adalah plot antara nilai *ergodic mean* terhadap iterasinya. Jika *ergodic mean* stabil setelah beberapa iterasi, maka algoritma dapat dianggap konvergen.

Setelah konvergensi dari algoritma yang digunakan dicapai, taksiran parameter  $\theta$  adalah rata-rata dari nilai-nilai sampel hasil simulasi, yaitu

$$\hat{\theta}_i = \frac{\sum_{t=1}^{T'} \theta_i^{(t)}}{T'} \text{ dengan } i = 1, 2, \dots, k$$

### Monte Carlo (MC) Error

*MC error* merupakan ukuran variabilitas dari setiap penaksiran dalam sebuah (Ntzoufras, 2009). Salah satu metode yang dapat digunakan untuk menghitung *MC error* adalah metode *batch mean*. *MC error* dapat ditaksir sebagai standar deviasi dari *batch mean*. Jika nilai yang dihasilkan terlalu besar maka algoritma MCMC sebaiknya dilakukan

kembali dengan menggunakan iterasi yang lebih besar.

### Penentuan Likelihood, Distribusi Prior dan Distribusi Posterior

Secara rinci ketiga hal tersebut dijelaskan sebagai berikut.

*Pertama*, fungsi likelihood variabel  $Y$  didefinisikan sebagai berikut.

$$P(Y_i|Y_i^*, p) = \prod_{i=1}^n \binom{y_i^*}{y_i} p^{y_i} (1-p)^{y_i^*-y_i} \quad (2.10)$$

Untuk  $Y_i^* = 0, 1, 2, \dots; Y_i = 0, 1, 2, \dots; i = 1, 2, \dots, n; 0 \leq p \leq 1$

*Kedua*, fungsi Distribusi prior parameter  $p$ , dimana  $p$  diasumsikan berdistribusi uniform,  $p \sim U[0,1]$  dengan fungsi densitas sebagai berikut:  $f(p) \propto 1$

*Ketiga*, fungsi distribusi prior parameter  $Y^*$  didefinisikan sebagai berikut:

$$P(Y_i^* = y_i^* | \beta) = \frac{\exp[(y_i^* \cdot x_i' \beta) - \exp[x_i' \beta]]}{y_i^*!} \quad (2.11)$$

dengan:  $\exp[x_i' \beta] \geq 0; Y_i^* = 0, 1, 2, \dots; i = 1, 2, \dots, n$

*Keempat*, fungsi distribusi prior parameter  $\beta$  diasumsikan berdistribusi normal  $\beta \sim N_m(\mu, \Sigma)$  dengan fungsi densitas sebagai berikut:

$$f(\beta) = \frac{1}{(\sqrt{2\pi})^m |\Sigma|^{1/2}} \exp\left[-\frac{1}{2}(\beta - \mu)' \Sigma^{-1}(\beta - \mu)\right] \quad (2.12)$$

dengan :  $-\infty \leq \beta \leq \infty; -\infty \leq \mu \leq \infty; |\Sigma| > 0$

*Kelima*, fungsi Distribusi Posterior berdasarkan persamaan (2.9) adalah sebagai berikut:

$$P(Y_i^*, \beta, p_i | Y_i) \propto P(Y_i | Y_i^*, p_i) \times P(Y_i^* | \beta) \times f(\beta) \times f(p_i) \propto \prod_{i=1}^n \binom{y_i^*}{y_i} p_i^{y_i} (1-p_i)^{y_i^*-y_i} \cdot \frac{\exp[(y_i^* \cdot x_i' \beta) - \exp[x_i' \beta]]}{y_i^*!} \cdot \left\{ \frac{1}{(\sqrt{2\pi})^m |\Sigma|^{1/2}} \exp\left[-\frac{1}{2}(\beta - \mu)' \Sigma^{-1}(\beta - \mu)\right] \right\} \quad (2.13)$$

$$\propto \exp\left[-\frac{1}{2}(\beta - \mu)' \Sigma^{-1}(\beta - \mu)\right] \cdot \prod_{i=1}^n \exp[(y_i^* \cdot x_i' \beta) - \exp[x_i' \beta]] \cdot \frac{p_i^{y_i} (1-p_i)^{y_i^*-y_i}}{(y_i^* - y_i)! y_i^*!}$$

dengan :  $Y_i^* = 0, 1, 2, \dots; Y_i = 0, 1, 2, \dots; 0 \leq p_i \leq 1; -\infty \leq \beta \leq \infty; i = 1, 2, \dots, n$

Jelas sulit untuk menentukan distribusi posterior yang diperoleh pada persamaan di atas sehingga untuk memudahkan penaksiran parameter model dilakukan simulasi MCMC. Untuk itu perlu ditentukan fungsi peluang bersyarat dari masing-masing parameter sebagai berikut :

*Keenam*, fungsi peluang bersyarat parameter  $Y^*$  adalah:

$$(Y_i^* - Y_i) \sim \text{Poisson}((1-p_i) \cdot \exp(x_i' \beta)) \quad \text{untuk } i = 1, 2, \dots, n \quad (2.14)$$

*Ketujuh*, fungsi peluang bersyarat parameter  $p$  adalah :

$$P|Y_i, Y_i^* \sim \text{Beta} \left( (y_i + 1), ((y_i^* - y_i) + 1) \right) \quad (2.15)$$

*Kedelapan*, fungsi peluang bersyarat parameter  $\beta$  adalah :

$$P(\beta|Y_i, Y_i^*) \propto \exp \left[ -\frac{1}{2} (\beta - \mu)' \Sigma^{-1} (\beta - \mu) \right] \cdot \prod_{i=1}^n \exp [ (y_i^* \cdot x_i' \beta) - \exp [ x_i' \beta ] ]$$

dengan :  $y_i^* = 0, 1, 2, \dots; -\infty \leq \beta \leq \infty; i = 1, 2, \dots, n$  (2.16)

Oleh karena fungsi peluang bersyarat dari masing-masing parameter model diketahui maka penaksiran dapat dilakukan dengan menggunakan teknik MCMC melalui algoritma gibbs sampling. Langkah-langkahnya adalah sebagai berikut:

*Pertama*, penentuan nilai awal parameter model: (1). Tentukan nilai  $Y^{*(0)}$  dengan  $Y_i^{*(0)} > Y_i$  untuk  $i = 1, 2, \dots, n$ . (2). Tentukan nilai  $p^{(0)} = 0,5$ . (3). Tentukan nilai  $\beta^{(0)} = (1,1)$ . (4). Tentukan nilai  $\tau = 0,5$ . (5). Tentukan nilai  $u^{(prior)} = (0,0)$ . (6). Tentukan nilai  $\Sigma_{\beta}^{(prior)} = \begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix}$

*Kedua*, tentukan ukuran iterasi  $t=5000$

*Ketiga*, lakukan simulasi untuk  $t=1, 2, \dots, 5000$  melalui beberapa langkah berikut: (1). Hitung nilai  $\lambda_i^{(t)}$  untuk  $i = 1, 2, \dots, n$ . (2). Bangkitkan kandidat baru nilai baru  $(Y_i^* - Y_i)^{(t)} \sim \text{Poisson}(\lambda_i^{(t)})$  untuk  $i=1, 2, \dots, n$ . (3). Hitung nilai  $a = (y_i + 1)$  untuk  $i=1, 2, \dots, n$ . (4). Hitung nilai  $b^{(t)} = (y_i^{*(t)} - y_i) + 1$  untuk  $i=1, 2, \dots, n$ . (5). Bangkitkan kandidat nilai baru  $p_i^{(t)} \sim \text{beta}(a, b^{(t)})$  untuk  $i = 1, 2,$

$\dots, n$ . (6). Bangkitkan kandidat baru parameter  $\beta$ .

*Keempat*, perbaharui nilai-nilai parameter model dari nilai-nilai hasil simulasi: (1). Atur  $\beta^{(t)} = \text{Means}(\beta^{(r)})$ . (2). Atur  $p_i^{(t)} = p_i$  untuk  $i = 1, 2, \dots, n$ . (3). Atur  $(Y_i^* - Y_i)^{(t)} = (Y_i^* - Y_i)$  untuk  $i = 1, 2, \dots, n$

*Kelima*, periksa konvergensi algoritma dengan membuat: (1). *Trace plot*. (2). Plot autokorelasi. (3). *Ergodic mean plot*

*Keenam*, tentukan *burn-in period*

*Ketujuh*, dari nilai-nilai simulasi setelah *burn in* sampai dengan iterasi terakhir hitung: (1). Nilai rata-rata dan standar deviasi dari parameter  $\beta$ . (2). Nilai rata-rata dan standar deviasi dari parameter  $p_i$ . (3). Nilai rata-rata dan standar deviasi dari parameter  $Y_i^*$

## PEMBAHASAN

### Deskripsi Data Penelitian

Letak Kota Ambon berada sebagian besar dalam wilayah Pulau Ambon dan secara geografis terletak pada posisi: 3<sup>o</sup>- 4<sup>o</sup> Lintang Selatan dan 128<sup>o</sup>-129<sup>o</sup> Bujur Timur, dimana secara keseluruhan Kota Ambon berbatasan dengan Kabupaten Maluku Tengah. Sesuai Peraturan Pemerintah Nomor 13 Tahun 1979 luas wilayah Kota Ambon seluruhnya 377 Km<sup>2</sup> dan berdasarkan hasil Survei Tata Guna Tanah tahun 1980 luas daratan Kota Ambon tercatat 359,45 km<sup>2</sup>. Sesuai Perda Kota Ambon Nomor 2 Tahun 2006, Kota Ambon memiliki lima kecamatan dengan luas masing-masing: Kecamatan Nusaniwe 8.834,30 Ha, Kecamatan Sirimau 8.681,32 Ha, Kecamatan Teluk Ambon 9.368,00 Ha, Kecamatan Teluk Ambon Baguala 4.011,00 Ha dan Kecamatan Leitimur Selatan dengan luas 5.050 Ha.

Wilayah Kota Ambon secara administratif terbagi menjadi lima kecamatan dengan 50 desa/kelurahan berklasifikasi

swasembada. Dari jumlah tersebut, sesuai dengan perkembangan Lembaga Ketahanan Masyarakat Desa, seluruhnya telah berkategori maju.

Jumlah penduduk Kota Ambon pada pertengahan tahun 2016 berjumlah 427.934 jiwa. Jika dibandingkan dengan jumlah penduduk pada tahun 2015 meningkat sebesar 3,96 persen. Penduduk masih terkonsentrasi di Kecamatan Sirimau dengan tingkat kepadatan penduduk sebesar 2.057 jiwa per Km<sup>2</sup>. Sementara itu Kecamatan yang paling jarang penduduknya adalah Kecamatan Leitimur Selatan dengan tingkat kepadatan penduduk yang hanya sebesar 235 jiwa per Km<sup>2</sup>.

Kota Ambon merupakan daerah di Maluku dengan jumlah kejahatan paling banyak. Pada tahun 2014, tercatat ada 1003 kasus yang di laporkan ke Polres Ambon, namun mengalami penurunan di tahun 2015 menjadi 919 kasus dan kembali meningkat pada tahun 2016 menjadi 1054 kasus. Dari sisi tingkat resiko terkena kejahatan (*crime rate*), Kota Ambon mencatat tingkat resiko terkena kejahatan tertinggi yakni 223 (setiap 100.000 penduduk diperkirakan sebanyak 223 orang yang berisiko terkena tindak kejahatan), disusul oleh Kabupaten Maluku Tenggara (218) dan Kabupaten Maluku Tenggara Barat (121).

Dari 1054 kasus yang dilaporkan, kejahatan yang paling sering dilaporkan adalah kekerasan bersama dan penganiayaan (47 persen). Sedangkan kejahatan yang paling sedikit dilaporkan adalah kejahatan terhadap anak (13 persen). Hal ini sejalan dengan hasil riset yang dilakukan oleh *Bureau of Justice Statistics, U.S. Department of Justice* (2012). Dalam laporannya "*Victimizations Not Reported to the Police*", jika dilihat berdasarkan umur korban, maka peluang tindak kejahatan tidak dilaporkan ke pihak

berwajib yang paling besar adalah ketika yang menjadi korban kejahatan berusia di bawah 17 tahun. Artinya hanya sedikit kejahatan terhadap anak yang tercatat di pihak berwajib, sehingga data tersebut berada dalam keadaan *underreporting*. Kejahatan terhadap anak yang dilaporkan di Polres Ambon sebanyak 140 kasus, dengan rincian sebagaimana tabel 1.

Tabel 1  
Jumlah Tindak Kejahatan Terhadap Anak Yang Dilaporkan Berdasarkan Jenis Kejahatan, Tahun 2016

Jenis Kejahatan	Jumlah
Membawa lari anak di bawah umur	5
KK bersama terhadap anak	16
KK terhadap anak	36
Percabulan/persetubuhan terhadap anak	83
Total	140

Sumber: Polres Ambon, diolah

Jika dilihat berdasarkan tempat kejadian tindak kejahatan, Kecamatan Sirimau merupakan kecamatan dengan jumlah kejahatan terhadap anak yang terbanyak (60 kasus) sedangkan Kecamatan Leitimur Selatan merupakan kecamatan dengan jumlah tindak kejahatan terhadap anak yang paling sedikit (6 kasus). Oleh karena itu, sebagaimana yang telah dijelaskan di atas bahwa data jumlah tindak kejatan terhadap berada dalam keadaan *underreporting* artinya jumlah tindak kejahatan terhadap anak yang sebenarnya lebih besar dari yang dilaporkan (fenomena gunung es). Hal ini tentunya dapat menyebabkan kekeliruan dalam perencanaan penanganan kasus maupun pencegahan tindak kekerasan terhadap anak. Untuk itu perlu dilakukan penaksiran terhadap banyaknya jumlah tindak kejahatan terhadap anak yang sebenarnya.

Tabel 2  
Variabel Penelitian

Nama kecamatan	Jumlah Tindak Kejahatan Terhadap Anak ( $Y_i$ )	Proporsi Keluarga Miskin ( $X_i$ )
Nusaniwe	37	0,33
Sirimau	60	0,31
Leitimur Selatan	6	0,55
Baguala	21	0,46
Teluk Ambon	16	0,32

Sumber: Polres Ambon, BPS Kota Ambon

Tabel 2 di atas merupakan variabel yang digunakan dalam penelitian ini.

### Analisis Data

#### Pemeriksaan Konvergensi Algoritma

Pemeriksaan ini dapat dijelaskan sebagai berikut.

*Pertama*, konvergensi Simulasi parameter  $\beta$ . Penaksiran parameter  $\beta$  dilakukan berdasarkan persamaan (2.16) dengan algoritma yang telah dijelaskan sebelumnya dengan 5000 iterasi. Pemeriksaan konvergensi algoritma gibbs sampling untuk parameter  $\beta$  dilakukan melalui *trace plot*, *autocorrelation plot* dan *ergodic mean plot* yang dapat dilihat pada gambar 2 dan gambar 3.

*Trace plot* pada gambar 2 menunjukkan bahwa plot antara nilai-nilai sampel  $\beta^{(t)}$  terhadap iterasinya ( $t$ ) masing-masing tidak menunjukkan sebuah pola tertentu. Informasi ini belum cukup untuk menyimpulkan bahwa simulasi telah konvergen sehingga selanjutnya dilakukan pemeriksaan melalui *autocorrelation plot*.

*Autocorrelation plot* pada gambar 2 menunjukkan bahwa plot nilai-nilai sampel  $\beta$  terhadap *lag*-nya semakin menurun mendekati nol. Hal ini menunjukkan korelasi di antara iterasi yang beruntun semakin kecil. Informasi ini juga belum cukup untuk menyimpulkan bahwa simulasi telah konvergen sehingga

selanjutnya dilakukan pemeriksaan melalui *ergodic mean plot* pada gambar 3.

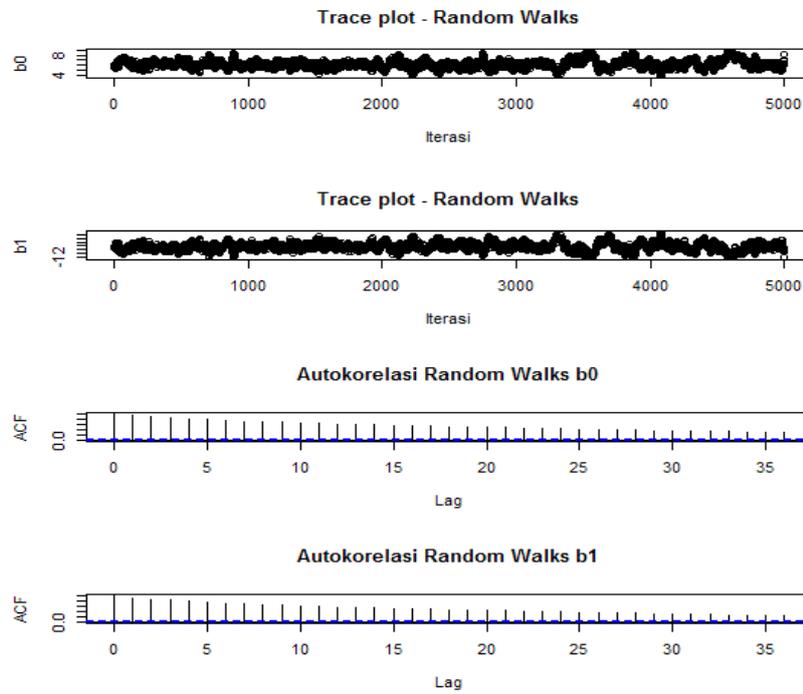
*Ergodic mean plot* pada gambar 3 menunjukkan bahwa plot nilai-nilai sampel  $\beta^{(t)}$  hasil simulasi terhadap iterasinya ( $t$ ) mulai stabil setelah 3000 iterasi pertama. Maka dapat disimpulkan bahwa konvergensi simulasi penaksiran  $\beta$  telah tercapai dengan *burn in period* 3000 iterasi. Dengan demikian taksiran parameter  $\beta$  adalah rata-rata dari nilai-nilai sampel  $\beta^{(t)}$  hasil simulasi yang dihitung dari  $t = 3001, \dots, 5000$ .

*Kedua*, konvergensi Simulasi Parameter  $p$ . Penaksiran peluang satu tindak kejahatan terhadap anak dilaporkan ke pihak kepolisian atau  $p$  dilakukan berdasarkan persamaan (2.15) dengan algoritma gibbs sampling dengan 5000 iterasi. Pemeriksaan konvergensi algoritma gibbs sampling untuk parameter  $\beta$  dilakukan melalui *trace plot*, *autocorrelation plot* dan *ergodic mean plot*.

*Trace plot* pada gambar 4 menunjukkan bahwa plot antara nilai-nilai sampel  $p_i$  untuk setiap kecamatan terhadap iterasinya masing-masing tidak menunjukkan sebuah pola tertentu. Informasi ini belum cukup untuk menyimpulkan bahwa simulasi telah konvergen sehingga selanjutnya dilakukan pemeriksaan melalui *autocorrelation plot*.

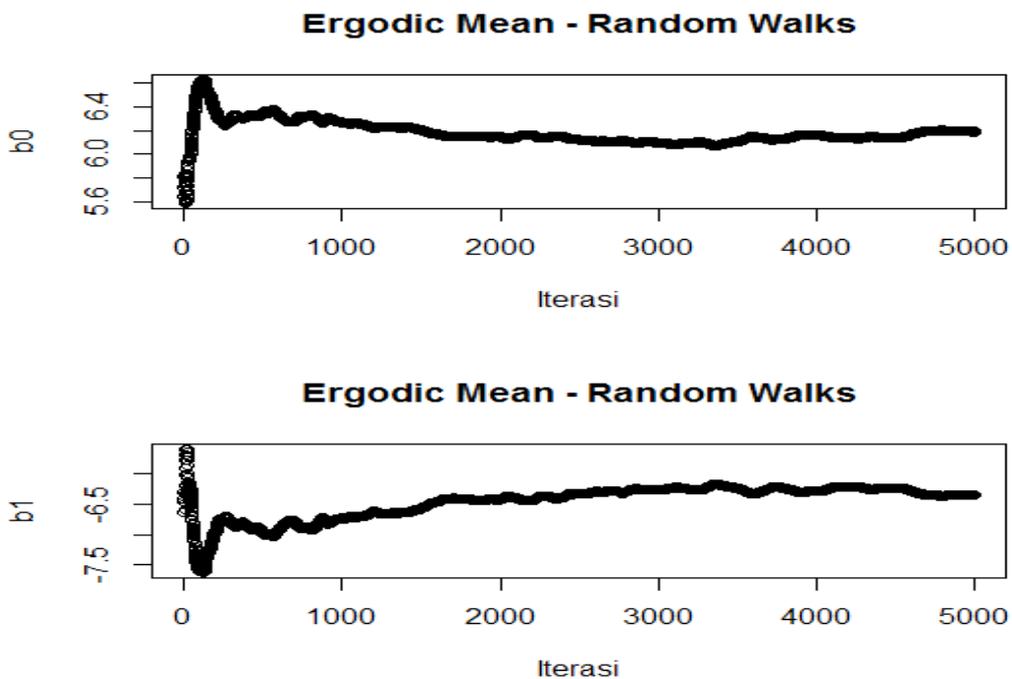
*Autocorrelation plot* pada gambar 4 menunjukkan plot nilai-nilai korelasi di antara

Gambar 2  
Trace Plot dan Autocorrelation Plot Simulasi  $\beta$



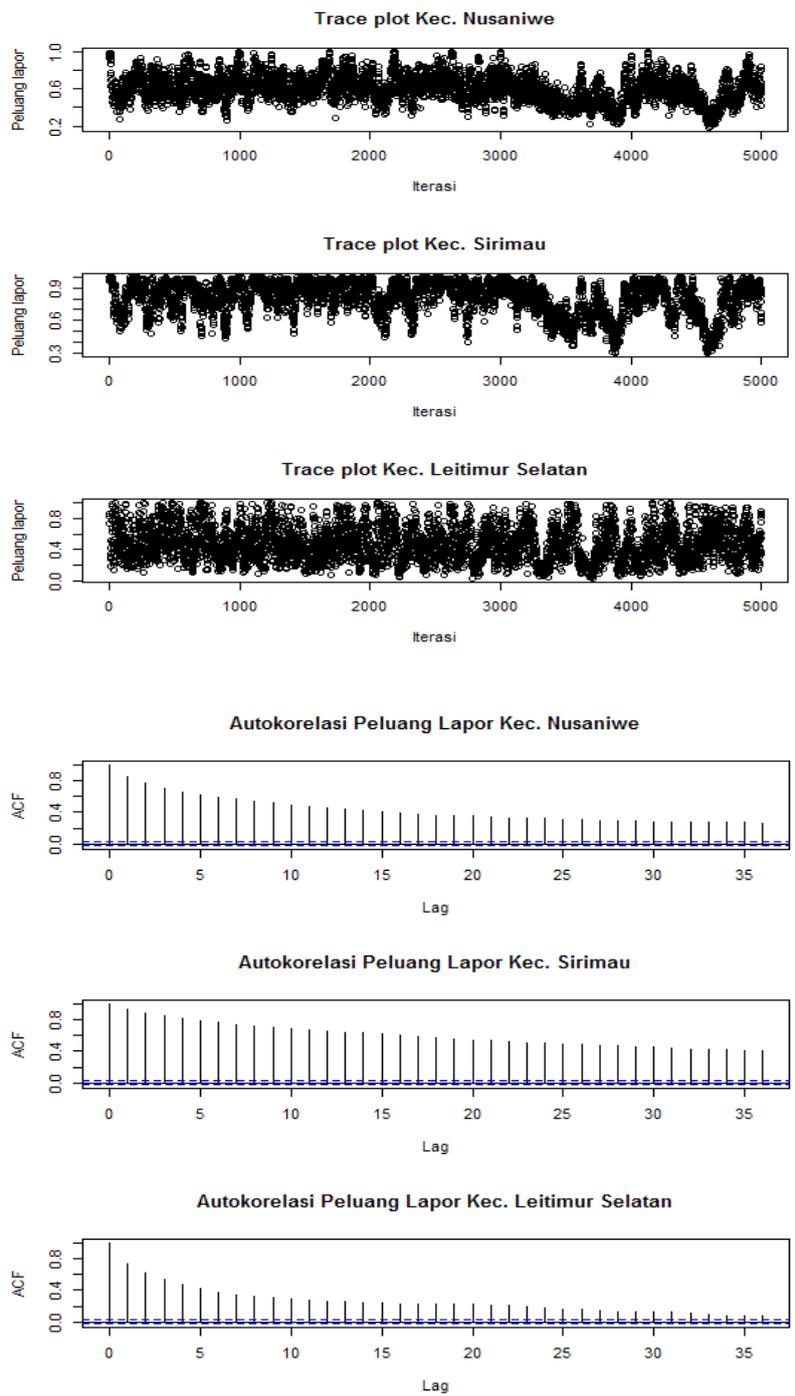
Sumber: Hasil Olah Data

Gambar 3  
Ergodic Mean Plot Simulasi  $\beta$



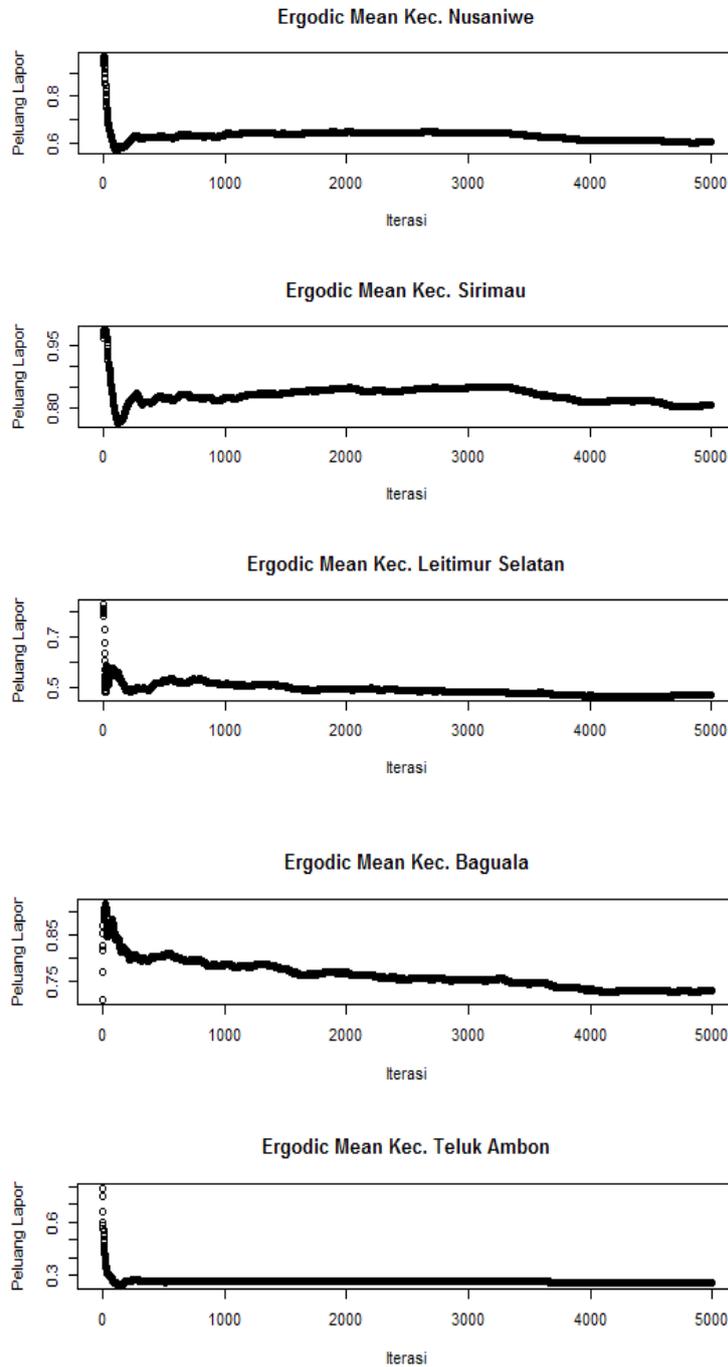
Sumber: Hasil Olah Data

Gambar 4  
*Trace Plot dan Autocorrelation Plot Simulasi  $p_{1t}$ ,  $p_{2t}$ ,  $p_{3t}$ ,  $p_{4t}$  dan  $p_{5t}$*



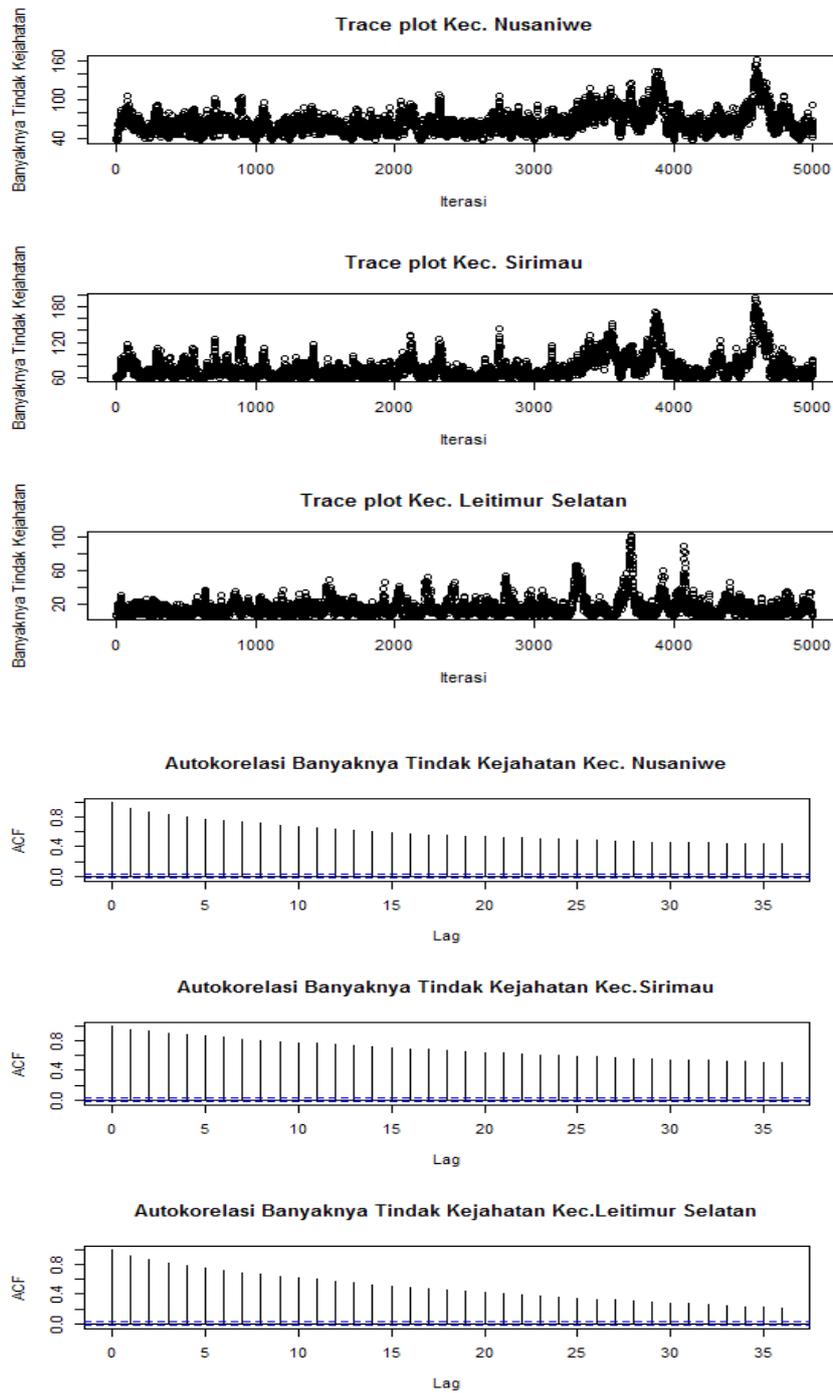
Sumber: Hasil Olah Data

Gambar 5  
Ergodic Mean Plot Simulasi  $p_1, p_2, p_3, p_4$  dan  $p_5$



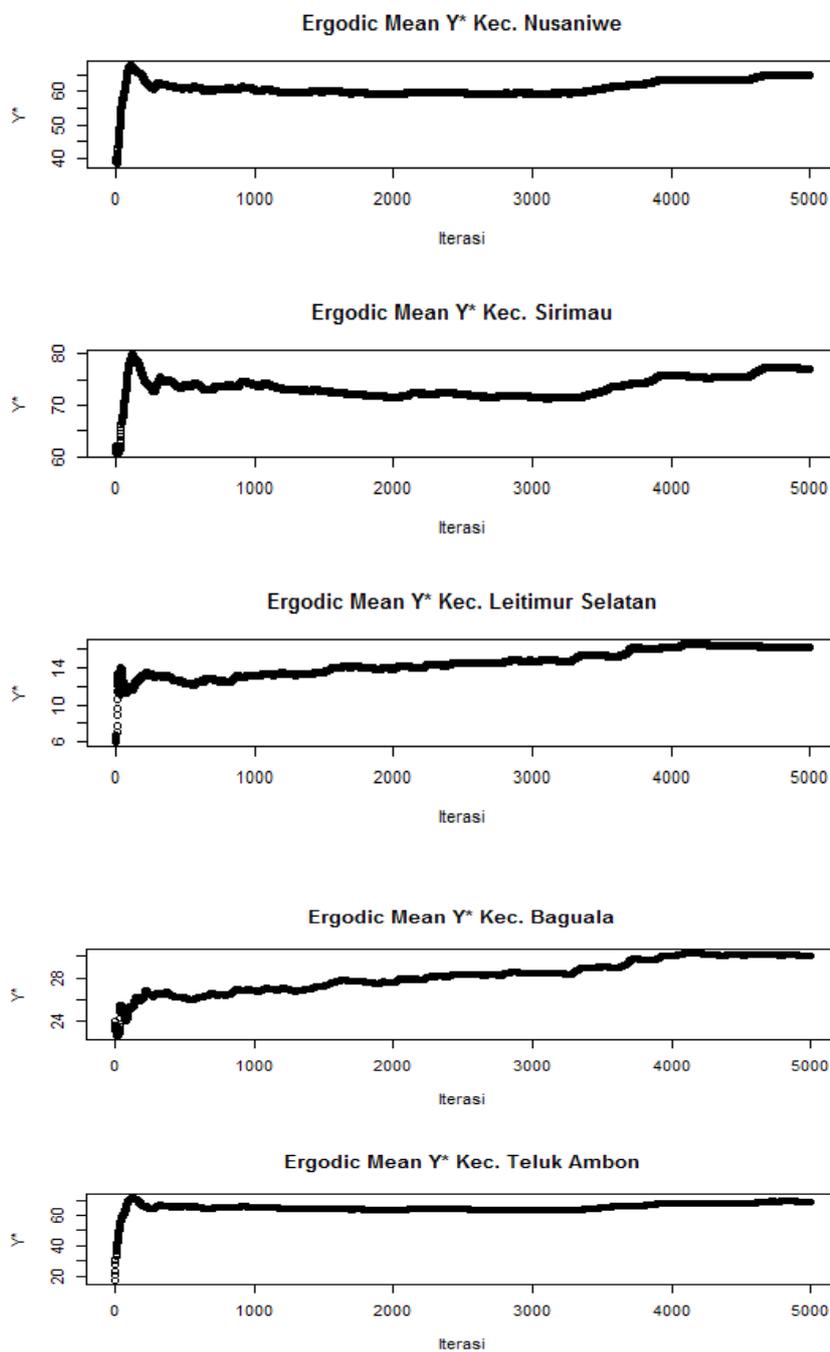
Sumber: Hasil Olah Data

Gambar 6  
Trace Plot dan Autocorrelation Plot Simulasi  $Y^*$



Sumber: Hasil Olah Data

Gambar 7  
Ergodic Mean Plot Simulasi  $Y^*$



Sumber: Hasil Olah Data

nilai sampel  $p_i$  terhadap lag-nya yang semakin kecil. Informasi ini juga belum cukup untuk menyimpulkan bahwa simulasi telah konvergen sehingga selanjutnya dilakukan pemeriksaan melalui *ergodic mean plot* pada gambar 5.

Selanjutnya melalui *Ergodic mean plot* pada gambar 5 ditunjukkan bahwa plot nilai-nilai sampel  $p_i$  untuk setiap kecamatan mulai stabil setelah 3000 iterasi pertama. Maka dapat disimpulkan bahwa konvergensi simulasi penaksiran  $p_i$  untuk  $i = 1, 2, \dots, 5$  telah tercapai dengan *burn in period* 3000 iterasi. Dengan demikian taksiran parameter  $p_i$  adalah rata-rata dari nilai-nilai sampel  $p_i^{(t)}$  hasil simulasi yang dihitung dari  $t = 3001, \dots, 5000$ .

*Ketiga*, konvergensi Simulasi Parameter  $Y^*$ . Penaksiran jumlah tindak kejahatan terhadap anak dilaporkan ke pihak kepolisian yang sebenarnya atau  $Y^*$  dilakukan berdasarkan persamaan (2.14) dengan algoritma gibbs sampling dengan 5000 iterasi.

*Trace plot* pada gambar 6 menunjukkan bahwa plot antara nilai-nilai sampel  $Y^*$  untuk setiap kecamatan terhadap iterasinya masing-masing tidak menunjukkan sebuah pola tertentu. *Autocorrelation plot* pada gambar 6 menunjukkan plot nilai-nilai korelasi di antara nilai sampel jumlah tindak kejahatan terhadap anak terhadap lag-nya yang semakin kecil. Informasi ini belum cukup untuk menyimpulkan bahwa simulasi telah

konvergen sehingga selanjutnya dilakukan pemeriksaan melalui *ergodic mean plot* pada gambar 7.

Selanjutnya melalui *Ergodic mean plot* pada gambar 7 ditunjukkan bahwa plot nilai-nilai sampel jumlah tindak kejahatan terhadap anak yang sebenarnya untuk setiap kecamatan mulai stabil setelah 3000 iterasi pertama. Maka dapat disimpulkan bahwa konvergensi simulasi penaksiran  $p_i$  untuk  $i = 1, 2, \dots, 5$  telah tercapai dengan *burn in period* 3000 iterasi. Dengan demikian taksiran parameter  $Y^*$  adalah rata-rata dari nilai-nilai sampel  $Y_i^{*(t)}$  hasil simulasi yang dihitung dari  $t = 3001, \dots, 5000$ .

### Hasil Penaksiran Parameter

Taksiran parameter regresi antara jumlah tindak kejahatan terhadap anak dengan proporsi rumah tangga miskin di Kota Ambon ditunjukkan dalam tabel 3.

Tabel 3  
Taksiran Parameter Regresi

Parameter	Taksiran
$\beta_0$	6,2188
$\beta_1$	-6,2979

Sumber: Hasil Olah Data

Tabel 3 tersebut menunjukkan bahwa untuk kecamatan ke- $i = 1, 2, \dots, 5$ , rata-rata

Tabel 4  
Taksiran Rata-rata Jumlah Tindak Kejahatan Terhadap Anak Berdasarkan Proporsi Rumah Tangga Miskin di Kota Ambon, Tahun 2016

Nama Kecamatan	Jumlah Tindak Kejahatan Terhadap Anak ( $Y_i$ )	Proporsi Keluarga Miskin ( $X_i$ )	Taksiran Rata-rata ( $\hat{\lambda}_i$ )
Nusaniwe	37	0,33	63
Sirimau	60	0,31	72
Leitimur Selatan	6	0,55	15
Baguala	21	0,46	28
Teluk Ambon	16	0,32	68

Sumber: Hasil Olah Data

Tabel 5  
Hasil Penaksiran Parameter Model

Kecamatan	$y$	$\hat{y}^*$	Std( $\hat{y}^*$ )	$\hat{p}^*$	Std( $\hat{p}^*$ )	$(1 - \hat{p}^*) \times 100\%$	$\hat{y}^* \pm \text{Std}(\hat{y}^*)$
Nusaniwe	37	68	22	0,57	0,21	43	46-90
Sirimau	60	80	20	0,78	0,16	22	60-100
Leitimur Selatan	6	17	17	0,45	0,26	55	0-34
Baguala	21	31	21	0,70	0,25	30	10-52
Teluk Ambon	16	72	20	0,25	0,29	75	52-92
<b>Rata-rata</b>	28	54	30	0,55	0,25	45	24-84

Sumber: Hasil Olah Data

tindak kejahatan terhadap anak tahun 2016 merupakan fungsi dari :

$$\hat{\lambda}_i = E[Y_i^* | X_i] = \exp^{[6,2188 - 6,2979 X_i]}$$

Berdasarkan proporsi rumah tangga miskin, taksiran rata-rata jumlah tindak kejahatan terhadap anak di Kota Ambon pada tahun 2016 dapat dilihat pada tabel 4. Berdasarkan tabel 4 dapat dilihat bahwa *underreported counts* terjadi pada data jumlah tindak kejahatan terhadap anak, dimana taksiran rata-rata ( $\hat{\lambda}_i$ ) memiliki nilai yang lebih besar dari jumlah kasus yang dilaporkan ( $Y_i$ ).

Hasil penaksiran peluang sebuah tindak kejahatan terhadap anak dilaporkan ke pihak kepolisian atau  $p$  dan hasil penaksiran jumlah tindak kejahatan terhadap anak yang sebenarnya atau  $y^*$  dapat dilihat pada tabel 5.

Tabel 5 memberikan pengertian bahwa taksiran jumlah tindak kejahatan terhadap anak yang sebenarnya di Kecamatan Nusaniwe pada tahun 2016 adalah 68 kasus atau berkisar antara 46-90 kasus dengan peluang sebuah kasus dilaporkan sebesar 0,57. Hal ini dapat diinterpretasikan dari 100 tindak kejahatan terhadap anak yang terjadi hanya sekitar 57 kasus yang dilaporkan sedangkan 43 kasus lainnya tidak dilaporkan ke pihak kepolisian.

Hasil ini memperlihatkan bahwa sesungguhnya tindak kejahatan yang terjadi pada anak masih sangat tinggi sehingga membutuhkan kerjasama secara yang komprehensif dari pihak-pihak yang berwajib baik untuk penanganan kasus/penegakan hukum maupun untuk menetapkan strategi yang tepat untuk pencegahan tindak kejahatan terhadap anak, salah satunya dengan menciptakan lingkungan yang ramah dan aman bagi tumbuh kembang anak.

## SIMPULAN

Berdasar penjelasan tersebut di atas dapat ditarik simpulan sebagai berikut.

*Pertama*, penaksiran parameter model untuk *underreported* data kejahatan yang dilakukan melalui pendekatan Bayesian akan bergantung pada penentuan *burn in period* karena dapat mempengaruhi nilai rata-rata *posterior* jika nilai-nilai tersebut berada pada rentang jauh dari daerah probabilitas *posterior*. Penentuan *burn in period* dilakukan sejalan dengan pemeriksaan konvergensi algoritma melalui *trace plot*, *autocorrelation plot* dan *ergodic mean plot*. Setelah *burn in period* ditentukan, maka taksiran parameter model adalah rata-rata nilai sampel iterasi setelah *burn in period* sampai dengan iterasi terakhir.

*Kedua*, tindak kejahatan yang terjadi pada anak sebenarnya masih sangat tinggi jika dibandingkan dengan jumlah kasus yang dilaporkan ke pihak kepolisian sehingga membutuhkan kerjasama secara komprehensif dari pihak-pihak yang berwajib untuk mencari solusi terbaik untuk penanganan masalah ini.

#### DAFTAR PUSTAKA

- Badan Pusat Statistik, 2016, '*Statistik Kriminal 2016*', Jakarta: Badan Pusat Statistik.
- BPS Kota Ambon, 2017, '*Kota Ambon Dalam Angka 2017*', Ambon: Aman Jaya.
- Bolstad, William M., 2007, '*Introduction to Bayesian Statistics*', Second Edition, New Jersey: John Wiley & Son, Inc.
- Bureau of Justice Statistics, U.S. Department of Justice, 2012, '*Victimizations Not Reported to the Police*'.
- Draper, N., and Guttman, L 1971, "Bayesian Estimation of the Binomial Parameter" *Technometrics*, 13, 667-673.
- Hoff, P.D. 2009. "*A First Course in Bayesian Statistical Methods*", New York: Springer.
- Lemhannas, 2000, Ketahanan Nasional, Jakarta: Balai Pustaka.
- Moran, P. A. P., 1952, "A Characterization Property of the Poisson Distribution". *Proc. Camb. Phil. Soc. Math. Phys.* 48, 206-207.
- Moreno, Elias & Javier Giron, 1998, 'Estimating with Incomplete Count Data, A Bayesian Approach'. *Journal of Statistical Planning and Inference*, vol.66, 147-159.
- Morocoima-Black R., Chavarra S. and Lucas C., 2001, '*Selected Intersection Crash Analysis for 1993-1998. Champaign-Urban Urbanized Area Transportation Study (CUUATS)*', Champaign County Regional Planning Commission.
- Mustafa, M., 2016, '*Kajian Sosiologi terhadap Kriminalitas, Perilaku Menyimpang, dan Pelanggaran Hukum*'. Sari Ilmu Pratama.
- Neubauer Gerhard, 2011. "Models for Underreporting: A Bernoulli Sampling Approach for Underreported Counts". *Austrian Journal of Statistics*, Vol. 40, Number 1 & 2, 85-92.
- Ntzoufras, I. 2009. "*Bayesian Modeling Using WinBUGS*". New Jersey: John Wiley & Sons, Inc
- Padmadisastra S., J. Suprijadi, 2014, '*Bayesian Underreported in Disease Mapping*', American Institute of Physics.
- Papadatos, Nickos, 2005, 'Characterizations of Discrete Distributions Using The Rao-Rubin Condition'. *Journal of Statistical Planning and Inference*. Vol. 135, pp. 222-228.
- Pararai, Marvis, 2010, 'Generalized Poisson-Poisson Mixture Model for Misreported Counts with an Application to Smoking Data'. *Journal of Data Science*. Vol.8, pp. 607-617.
- Pertiwi, R. 2012. *Pemodelan Pengeluaran Per Kapita Per Kabupaten/Kota di Kalimantan Barat Menggunakan Metode Hirarki Bayesian*. Tesis, Institut Teknologi Sepuluh Nopember, Surabaya.
- Polres Ambon, 2016, *Laporan Bulanan Tindak Kejahatan*. Ambon: Polres Ambon.
- Rao Radhakrishna C., Herman Rubin, 1964, 'On a Characterization of the Poisson Distribution'. *The Indian Journal of Statistics*, Vol. 26, No. 2/3, pp. 295-298.
- Rian Marlina Reny. 2014. "Pendekatan Bayes Dalam Model Poisson Untuk

*Herlin Venny Johannes, Septiadi Padmadisastra, Bertho Tantular -- Meningkatkan Ketahanan Wilayah Melalui Estimasi Underreported Data Kejahatan Dengan Pendekatan Bayes (Studi Di Kota Ambon, Provinsi Maluku)*

Underreported Counts”. *Tesis*, Universitas Padjajaran, Bandung.

Winkelmann, 1996. “Markov Chain Monte Carlo analysis of Underreported Count

Data With an Application to Worker Absenteeism”. *Empirical Economic* 21: 575-587.