ESTIMASI PARAMETER MODEL *INAR*(1) MENGGUNAKAN METODE BAYES

oleh

NURMALITASARI

M0106054

SKRIPSI

ditulis dan diajukan untuk memenuhi sebagian persyaratan

memperoleh gelar Sarjana Sains Matematika

JURUSAN MATEMATIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS SEBELAS MARET

SURAKARTA

2011

SKRIPSI

ESTIMASI PARAMETER MODEL *INAR*(1) MENGGUNAKAN METODE BAYES

yang disiapkan dan disusun oleh

NURMALITASARI

M0106054

dibimbing oleh

Pembimbing I,	Pembimbing II,
Winita Sulandari, M.Si	Supriyadi Wibowo, M.Si
NIP. 19780814 200501 2 002	NIP. 19681110 199512 1 001

telah dipertahankan di depan Dewan Penguji
pada hari senin, tanggal 10 januari 2011
dan dinyatakan telah memenuhi syarat.

Anggota Tim Penguji	Tanda Tangan
1. Dra. Sri Sulistijowati H, M.Si NIP. 19690116 199402 2 001	1.
2. Dra. Etik Zukhronah, M.Si NIP. 19661213 199203 2 003	2.
3. Sri Kuntari, M.Si NIP. 19730225 199903 2 001	3.

Surakarta, Januari 2011

Disahkan oleh

Fakultas Matematika dan Ilmu Pengetahuan Alam

Dekan	Ketua Jurusan Matematika
Prof. Drs. Sutarno, M.Sc, Ph.D.	Drs. Sutrima, M.Si.
NIP. 19600809 198612 1 001	NIP. 19661007 199302 1 001

MOTO

- *Maka sesungguhnya bersama kesulitan ada kemudahan, sesungguhnya bersama kesulitan ada kemudahan (Al-Insyirah:5-6)*
- *Keikhlasan dan Kesabaran adalah kunci dari setiap permasalahan.*
- *Doa dan Usaha mampu mengubah segalanya.*

PERSEMBAHAN

Karya sederhana ini saya persembahkan kepada:

- *Bapak, Ibu tercinta, sebagai wujud bakti saya atas kasih sayang, cinta, pengorbanan dan doa yang selalu diberikan.*
- *Adik-adikku tercinta, Muhammad & Kyky atas dukungan, semangat dan keceriaannya.*

ABSTRAK

Nurmalitasari, 2011. ESTIMASI PARAMETER MODEL *INAR(1)* MENGGUNAKAN METODE BAYES. Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Sebelas Maret.

Model *Integer-value autoregressive* orde pertama *(INAR*(1)) adalah salah satu model yang digunakan untuk data cacah. Dalam model *INAR*(1) terdapat parameter yang belum diketahui dan perlu diestimasi yaitu probabilitas bertahan dalam suatu proses (α) dan parameter komponen kedatangan (λ). Pada penelitian ini parameter diestimasi menggunakan metode Bayes dengan prior sekawan. Nilai estimasi parameter diperoleh menggunakan metode *Markov Chain Monte Carlo (MCMC)* dengan membangun rantai Markov.

Gibbs sampling merupakan salah satu algoritma dalam *MCMC* yang dapat digunakan untuk estimasi parameter model *INAR*(1). Pada aplikasi *Gibbs sampling* jika distribusi posterior dari masing-masing parameter adalah log-konkav maka estimasi parameter menggunakan algoritma *Adaptive Rejection Sampling (ARS)* dan jika distribusi posterior dari masing-masing parameter adalah log-konveks maka estimasi parameter menggunakan algoritma *Adaptive Rejection Metropolis Sampling (ARMS)*.

Berdasarkan penelitian ini diperoleh hasil estimasi parameter model *INAR*(1) adalah $\alpha \approx -\sum \quad \alpha$ dan $\lambda \approx -\sum \quad \lambda$. Nilai α dan λ merupakan rantai Markov yang dibangkitkan dengan algoritma *ARS* atau *ARMS*, tergantung distribusi posterior dari masing-masing parameter.

Kata kunci: *model INAR(1), metode Bayes, MCMC, Gibbs sampling, ARS, ARMS.*

ABSTRACT

Nurmalitasari, 2011. ESTIMATION OF THE PARAMETERS INAR(1) MODEL USING BAYES METHOD. Mathematics and Natural Sciences Faculty, Sebelas Maret University.

Integer-value autoregressive first orde (INAR(1)) model is one of the model used for count data. In the INAR (1) model there are unknown parameters and these parameters need to be estimated, i.e. the probability of survival of the process (α) and arrival elements parameter (λ). In this research, the parameters will be estimated by Bayes method with conjugate prior. The value of parameters estimation can be obtained using the *Markov Chain Monte Carlo (MCMC)* method by generating Markov chain.

Gibbs sampling is one of MCMC algorithms that can be used to estimate parameters of INAR(1) model. On application of Gibbs sampling, if the posterior distribution of each parameters is log-concave then the parameter estimation use the Adaptive Rejection Sampling (ARS) algorithm and if it's log-convex then the parameter estimation use the Adaptive Rejection Metropolis Sampling (ARMS) algorithm.

Based on this research the parameter estimation of INAR(1) model result are $\alpha \approx -\sum \alpha$ and $\lambda \approx -\sum \lambda$. The value α and λ are Markov chain generated from ARS or ARMS algorithm, depends on the posterior distribution of each parameters.

Key words: *INAR(1) model, Bayes method, MCMC, Gibbs sampling, ARS, ARMS*

KATA PENGANTAR

Segala puji bagi Allah SWT, yang telah memberikan kekuatan dan kemudahan kepada penulis sehingga dapat menyelesaikan skripsi ini dengan baik. Terselesaikannya skripsi ini tidak lepas dari bimbingan dan motivasi dari berbagai pihak. Untuk itu penulis menyampaikan ucapan terima kasih kepada

1. Winita Sulandari, M.Si dan Supriyadi Wibowo, M.Si sebagai Pembimbing I dan Pembimbing II atas kesediaan dan kesabaran yang diberikan dalam membimbing penulis,
2. Dra. Respatiwulan, M.Si atas pengarahan selama proses penulisan,
3. Suli, Budi, Arine, dan Choiril atas kerjasama dan motivasi yang diberikan saat penulis menghadapi kendala dalam penyusunan skripsi ini,
4. semua pihak yang turut membantu kelancaran penulisan skripsi ini.

Semoga karya sederhana ini bermanfaat bagi pembaca.

Surakarta , Januari 2011

Penulis

DAFTAR ISI

DAFTAR NOTASI

α	: probabilitas bertahan dalam suatu proses
λ	: parameter komponen kedatangan
	: variabel random pada periode ke *t*
$\alpha\|x$	: fungsi *likelihood* dari α dengan syarat x diketahui
$p \quad \|$	: fungsi densitas probabilitas model *INAR(1)*, , dengan syarat diketahui
$p\ \alpha\|\lambda,$	: fungsi densitas probabilitas dari α dengan syarat λ dan diketahui.
$h(\alpha\|\lambda,\ )$	: ln $p\ \alpha\|\lambda,$
$\alpha\|\lambda,$	: batas atas dari garis singgung $h\ \alpha\|\lambda,$
$\alpha\|\lambda,$	: batas bawah dari garis singgung $h\ \alpha\|\lambda,$

DAFTAR GAMBAR

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Silva *et al.* (2005) menyebut data cacah merupakan data runtun waktu diskrit. Data cacah adalah data yang dihitung sebagai jumlah kejadian dalam interval waktu atau dalam interval ruang. Misalnya data banyaknya kebakaran yang terjadi di wilayah Surakarta dalam setiap bulan, banyaknya orang yang meninggal akibat penyakit jantung dalam setiap bulan, dan banyaknya orang meninggal akibat kecelakaan dalam setiap bulan.

Data cacah bernilai bulat positif. Distribusi yang digunakan untuk mewakili distribusi data cacah adalah Poisson, binomial, dan negatif binomial (Brannas, 1994). Salah satu model yang digunakan untuk data cacah adalah *Integer-value Autoregressive (INAR).* Menurut Silva (2005) model *INAR* terdiri dari dua komponen yaitu *survivors* dalam proses sebelumnya dan kedatangan. Dalam model *INAR* terdapat parameter yang belum diketahui dan perlu diestimasi yaitu probabilitas bertahan dalam suatu proses (α) dan parameter kedatangan (λ).

Metode estimasi parameter model *INAR* ada dua macam, yaitu metode klasik dan Bayes. Estimasi parameter model *INAR* dengan menggunakan metode klasik telah diteliti oleh Brannas (1994) dan Silva *et al.* (2005). Dalam penelitian tersebut metode klasik yang dibahas adalah *Yule-Walker*, *Conditional Least Squares (CLS)*, *Conditional Maximum Likelihood* dan *Whittle criterion.* Selain metode klasik, pada penelitian Silva *et al.* (2005) juga dibahas estimasi parameter model *INAR* dengan menggunakan metode Bayes.

Pada penelitian ini penulis mengkaji ulang penelitian Silva *et al.* (2005) khususnya estimasi parameter model *INAR*(1) dengan menggunakan metode Bayes. Metode Bayes dipilih sebagai metode untuk estimasi parameter karena metode Bayes memiliki kelebihan dibandingkan metode klasik. Kelebihan tersebut terletak pada penggunaan informasi sampel dan informasi yang tersedia sebelum pengambilan sampel pada metode Bayes, sedangkan pada metode klasik hanya menggunakan informasi sampel.

Dalam estimasi parameter dengan menggunakan metode Bayes terdapat dua komponen yaitu distribusi prior dan distribusi posterior. Distribusi prior digunakan untuk membentuk distribusi posterior. Distribusi posterior diperlukan untuk menentukan nilai estimasi parameter. Menurut Gilks dan Wild (1992), nilai estimasi parameter diperoleh dengan simulasi pengambilan sampel parameter dari distribusi posterior kompleks menggunakan metode *Markov Chain Monte Carlo (MCMC)*. Konsep utama dalam *MCMC* adalah membuat sampel pendekatan dari distribusi posterior parameter, dengan membangkitkan sebuah rantai Markov yang memiliki distribusi limit mendekati distribusi posterior parameter.

Gibbs sampling merupakan algoritma yang terdapat dalam metode *MCMC* yang digunakan untuk pengambilan sampel dari distribusi kompleks berdimensi tinggi. Algoritma *Gibbs sampling* menggunakan sampel sebelumnya untuk membangkitkan nilai sampel berikutnya secara random sehingga didapatkan rantai yang disebut rantai Markov. Algoritma *Gibbs sampling* bias diterapkan apabila distribusi probabilitas bersama dari parameternya tidak diketahui, tetapi distribusi bersyarat dari tiap-tiap variabel diketahui (Walsh, 2004). Dalam aplikasi *Gibbs sampling* pada umumnya distribusi bersyarat dari tiap-tiap variabel mempunyai bentuk *non-familiar* dan mempunyai bentuk aljabar yang rumit. Sehingga dibutuhkan komputasi yang rumit untuk mengevaluasi distribusi bersyarat tersebut. Alternatif dalam aplikasi *Gibbs sampling* tersebut adalah melakukan pengambilan sampel dengan algoritma *Adaptive Rejection Sampling (ARS).* Algoritma *ARS* dapat diterapkan jika fungsi densitas bersyarat penuh dari masing-masing parameter adalah log-konkav, jika fungsi densitas bersyarat penuh dari masing-masing parameter adalah log-konveks maka digunakan algoritma *Adaptive Rejection Metropolis*

Sampling (ARMS) (Gilks dan Wild, 1992). Hasil yang diperoleh pada penelitian ini diperjelas dengan contoh kasus menggunakan data jumlah orang meninggal akibat kecelakaan bunuh diri di wilayah Surakarta dari Januari 2002-Desember 2006.

1.2 Perumusan Masalah

Perumusan masalah dalam penelitian ini adalah bagaimana menentukan estimasi parameter model *INAR*(1) menggunakan metode Bayes.

1.3 Batasan Masalah

Pada penelitian ini, batasan masalah yang digunakan adalah sebagai berikut.

1. Estimasi parameter model *INAR*(1) untuk data yang berdistribusi Poisson.
2. Parameter model *INAR*(1) diestimasi dengan menggunakan distribusi prior sekawan.
3. Nilai awal dalam algoritma *Gibbs sampling* ditentukan menggunakan metode *Conditional Least Squares (CLS)* hasil penelitian Brannas (1994).

1.4 Tujuan

Tujuan dari penelitian ini adalah menentukan estimasi parameter model *INAR*(1) menggunakan metode Bayes.

1.5 Manfaat

Manfaat dari penelitian ini adalah menambah wawasan mengenai model runtun waktu diskrit serta menambah pemahaman tentang penerapan metode Bayes dalam mengestimasi parameter model *INAR*(1).

BAB II

LANDASAN TEORI

2.1 Tinjauan Pustaka

Pada bagian ini diberikan beberapa teori yang mendukung dalam mencapai tujuan penulisan. Teori-teori yang diberikan meliputi, gambaran singkat mengenai konsep dasar statistik, distribusi-distribusi, estimator Bayes, integrasi Monte Carlo, *Markov chain*, *Markov Chain Monte Carlo*, algoritma *Gibbs sampling,* fungsi log-konkav, algoritma *ARS* dan *ARMS*.

2.1.1 <u>Konsep Dasar Statistik</u>

Berikut adalah beberapa definisi konsep dasar statistik yang digunakan dalam estimasi parameter yang diambil dari buku Bain dan Engelhardt (1995).

Definisi 2.1. *Ruang sampel adalah himpunan dari semua hasil observasi yang mungkin dari suatu percobaan dan dinotasikan S.*

Definisi 2.2. *Variabel random X adalah fungsi yang memetakan setiap hasil yang mungkin e pada ruang sampel S ke bilangan real x, sedemikian sehingga X(e)=x.*

Definisi 2.3. *Variabel random X disebut variabel random diskrit jika himpunan semua nilai yang mungkin dari variabel tersebut adalah himpunan yang terhitung yaitu $x\ , x\ , \dots , x$. Fungsi $f\ x$ x dengan x $x\ , x\ , \dots , x$ disebut fungsi densitas probabilitas.*

Definisi 2.4. *Fungsi densitas probabilitas bersama dari variabel random diskrit berdimensi n, yaitu , , … , didefinisikan sebagai*

$$f\ x\ , x\ , \dots , x \qquad x\ , \qquad x\ , \dots , \qquad x$$

untuk seluruh kemungkinan nilai x $x\ , x\ , \dots , x$ dari X.

Definisi 2.5. *Jika dan merupakan variabel random diskrit mempunyai fungsi densitas probabilitas bersama $f\ x\ , x$, maka fungsi densitas probabilitas bersyarat dari dengan diberikan x adalah*

$$f\ x\ |x \qquad \frac{f\ x\ ,x}{f\ x}.$$

Sedangkan fungsi densitas probabilitas bersyarat dari dengan diberikan x *adalah*

$$f\ x\ |x \qquad \frac{f\ x\ ,x}{f\ x}.$$

Definisi 2.6. *Diasumsikan X variabel random diskrit dengan fungsi densitas probabilitas* $f\ x$ *. Harga harapan dari X didefinisikan sebagai*

$$xf\ x\ .$$

2.1.2 Distribusi yang Digunakan dalam Estimasi Parameter

Pada bagian ini diberikan beberapa definisi distribusi yang digunakan dalam penentuan distribusi prior parameter dan fungsi densitas model *INAR*(1) yang diambil dari Bain dan Engelhardt (1995) dan Larson (1982).

Definisi 2.7. *Percobaan Bernoulli merupakan suatu percobaan yang memiliki dua jenis hasil yaitu sukses atau gagal. Jika X adalah variabel random diskrit dengan fungsi densitas*

$$f\ x \qquad p \qquad ,$$

dengan $x \quad 0, 1$, *maka X disebut sebagai variabel random berdistribusi Bernoulli.*

Definisi 2.8. *Distribusi binomial merupakan barisan dari n percobaan Bernoulli yang saling independen, sehingga banyaknya seluruh kemungkinan x kali sukses dari n percobaan adalah* $_{x}$ *. Jika X adalah variabel random diskrit dengan fungsi densitas*

$$f\ x \qquad _{x}\ p \qquad ,$$

dengan $x \quad 0, 1, \ldots, \quad ,$ *maka X disebut sebagai variabel random berdistribusi binomial dan dinotasikan* $\sim BIN \quad , p$.

Definisi 2.9. *Distribusi gamma merupakan distribusi dengan variabel random kontinu. Variabel random kontinu X dikatakan memiliki distribusi gamma dengan parameter* 0 *dan* θ 0 *jika fungsi densitasnya*

$$f\ \ x;\theta, \qquad \frac{1}{\qquad\theta\qquad} x \qquad exp\ -\theta x$$

dengan x 0. *Variabel random yang berdistribusi gamma dinotasikan* $\sim\ A\ \ \theta,\ \ .$

Definisi 2.10. *Jika X adalah variabel random kontinu dengan fungsi densitas*

$$f\ \ x \qquad \frac{\alpha}{\qquad\alpha\qquad} x \qquad 1 -\ x$$

dengan $x, \alpha,$ 0 *dan* ∞ *e* *untuk* $\alpha,$ *, maka X disebut sebagai variabel random berdistribusi beta.*

Definisi 2.11. *Variabel random diskrit X dikatakan mempunyai distribusi Poisson dengan parameter* λ 0 *jika mempunyai fungsi densitas probabilitas*

$$f\ \ x \qquad \frac{exp\ -\lambda\ \lambda}{x!}$$

untuk x $0, 1, 2, \ldots$ *dan memenuhi ketentuan*

$$a \qquad\qquad\qquad \lambda\,.$$

2.1.3 Estimator Bayes

Metode Bayes merupakan metode estimasi dan inferensi dalam statistika yang berbasis pada aturan Bayes yang menggabungkan informasi dari data observasi baru dan informasi yang telah diperoleh sebelumnya. Pada estimasi parameter dengan menggunakan metode Bayes terdapat dua komponen yaitu distribusi prior dan distribusi posterior. Menurut Berger (1980) distribusi prior adalah distribusi awal sebelum melakukan analisis data dengan parameter θ yang merupakan fungsi densitas probabilitas untuk menggambarkan tingkat keyakinan nilai θ, dinotasikan dengan $f\ \ \theta$. Jika pada sebuah observasi diketahui fungsi probabilitas prior dan fungsi *likelihood* dari data sampel ($f\ \ x\ ,x\ ,\ldots,x\ \ |\theta$) maka distribusi posterior ekuivalen dengan perkalian fungsi *likelihood* dan fungsi priornya. Distribusi posterior adalah

$$f\ \theta\ |x\ ,x\ ,\dots,x \qquad \frac{f\ \theta\ \ f\ x\ ,x\ ,\dots,x\ |\theta}{f\ \theta\ \ f\ x\ ,x\ ,\dots,x\ |\theta\quad\theta}$$

dengan $f\ \theta\ \ f\ x\ ,x\ ,\dots,x\ |\theta\quad\theta$ adalah konstan, maka

$$f\ \theta\ |x\ ,x\ ,\dots,x\quad \propto f\ \theta\ \ f\ x\ ,x\ ,\dots,x\ |\theta\ .$$

Menurut Bain dan Engelhardt (1995), estimator Bayes merupakan estimator yang meminimalkan harga harapan dan fungsi resiko. Berikut ini adalah definisi dan teorema yang berkaitan dengan estimator Bayes yang diambil dari Bain dan Engelhardt (1995).

Definisi 2.12. Jika θ adalah estimator dari I θ , maka fungsi kerugian, $\theta|\theta\quad 0$, untuk setiap θ dan $\theta|\theta\quad 0$, untuk $\theta\quad$ I θ .

Definisi 2.13. Fungsi resiko didefinisikan sebagai harga harapan dari fungsi kerugian yaitu

$$\theta \qquad \theta|\theta\ .$$

Teorema 2.1. Jika $x\ ,x\ ,\dots,x$ merupakan sampel random dari fungsi $f\ x|\theta$, maka estimator Bayes adalah estimator yang meminimalkan harga harapan fungsi kerugian dengan memperhatikan distribusi posterior $\theta|x$

$$_{|}\quad \theta\ \theta\ .$$

Teorema 2.2. Estimator Bayes θ dari I θ dengan menggunakan fungsi kerugian eror kuadrat

$$\theta|\theta \qquad \theta - \mathrm{I}\ \theta$$

adalah harga harapan dari I θ berdasarkan distribusi posterior $\theta|x$

$$\theta \qquad _{|}\ \mathrm{I}\ \theta$$

$$\mathrm{I}\ \theta\ \ f\ \theta|x\quad\theta.$$

2.1.4 <u>Integrasi Monte Carlo</u>

Pada bagian ini dijelaskan mengenai integrasi Monte Carlo yang digunakan dalam penentuan estimator Bayes. Berdasarkan Walsh (2004) integrasi Monte Carlo digunakan untuk melakukan pendekatan dalam penghitungan integral. Jika diberikan suatu bentuk integrasi

$$\square\ x \quad x$$

dengan $\square\ x$ merupakan hasil perkalian dari $f\ x$ dan x, dengan $f\ x$ adalah suatu fungsi distribusi probabilitas tertentu, x adalah fungsi densitas tertentu, maka integrasi $\square\ x \quad x$dapat ditentukan dengan mencari nilai rata-rata *n* sampel yang dibangkitkan dari distribusi $f\ x$, yaitu

$$\square\ x \quad x \qquad f\ x \quad x \quad x \qquad\qquad x \ \approx \frac{1}{} \qquad x \ .$$

2.1.5 <u>*Markov Chain*</u>

Menurut Walsh (2004), jika ditentukan yang merupakan variabel random pada waktu *t* dan $, \quad , \ldots ,$ adalah ruang *state* atau nilai yang mungkin dari *X*. Variabel random disebut proses Markov jika probabilitas transisi antara nilai yang berbeda dalam ruang *state* hanya tergantung pada *state* variabel random sekarang yaitu

$$| \qquad , \ldots , \qquad\qquad\qquad | \qquad .$$

Jadi kemungkinan kejadian pada *t+1*, hanya dipengaruhi oleh kejadian pada waktu sebelumnya atau pada waktu *t*. Data observasi yang diperoleh dari proses Markov dikatakan sebagai rantai Markov.

2.1.6 <u>*Markov Chain Monte Carlo*</u>

Metode *MCMC* merupakan metode pendekatan untuk inferensi Bayesian. Menurut Walsh (2004), *MCMC* digunakan untuk mendapatkan nilai estimasi parameter dengan mensimulasikan pengambilan sampel secara langsung dari distribusi posterior yang kompleks. Konsep utama dalam *MCMC* adalah membuat sampel pendekatan dari distribusi posterior parameter, dengan membangkitkan sebuah rantai Markov yang memiliki distribusi limit mendekati distribusi posterior parameter. Distribusi posterior parameter didapatkan dengan menentukan distribusi prior terlebih dahulu.

Pada *MCMC* terdapat dua macam algoritma, yaitu *Metropolis-Hasting* dan *Gibbs sampling.* Algoritma *Metropolis-Hasting* merupakan algoritma untuk membangkitkan barisan sampel menggunakan mekanisme penerimaan dan penolakan. Algoritma *Metropolis-Hasting* digunakan bila terdapat satu parameter yang tidak diketahui. Algoritma *Gibbs sampling* merupakan kasus khusus dari algoritma *Metropolis-Hasting* yang memerlukan semua distribusi bersyarat dari parameter yang dicari. Algoritma *Gibbs sampling* digunakan bila terdapat lebih dari satu parameter yang tidak diketahui. Dalam penelitian ini digunakan algoritma *Gibbs sampling* karena terdapat dua parameter yang tidak diketahui.

2.1.7 Algoritma *Gibbs sampling*

Gilks dan Wild, (1995) menjelaskan *Gibbs sampling* adalah algoritma *MCMC* yang digunakan untuk pengambilan sampel dari distribusi kompleks berdimensi tinggi. Konsep utama dalam *Gibbs sampling* adalah bagaimana menemukan bentuk distribusi bersyarat univariat, dimana dalam distribusi tersebut memuat semua variabel-variabel random dengan satu variabel saja yang akan ditentukan nilainya. Berikut diberikan algoritma *Gibbs sampling*.

1. Menentukan nilai awal parameter $\theta \quad \theta, \theta, \dots, \theta$
2. Mengulangi langkah untuk $0, 1, 2, \dots, N$ dengan N adalah batas akhir iterasi yang ditetapkan,

   membangkitkan θ dari $f \; \theta \,|\theta, \theta, \dots, \theta$ distribusi bersyarat pertama membangkitkan θ dari $f \quad \theta \; |\theta \quad , \theta \quad , \ldots , \theta$ distribusi bersyarat kedua

…

…

…

membangkitkan θ dari $f \quad \theta \; |\theta \quad , \theta \quad , \ldots , \theta$ distribusi bersyarat ke-*k*

3. Mendapatkan hasil nilai-nilai $\theta , \theta , \ldots , \theta$ yang merupakan rantai Markov. Fungsi $f , f , \ldots , f$ adalah distribusi bersyarat yang digunakan untuk simulasi.

Pada aplikasi Gibbs *sampling* jika fungsi densitas bersyarat dari distribusi posteriornya adalah *log-concav* maka menggunakan algoritma *ARS*, dan jika fungsi densitas bersyarat dari distribusi posteriornya adalah *log-convex* maka menggunakan algoritma *ARMS*.

2.1.8 Fungsi Log-konkav

Pada bagian ini diberikan beberapa definisi yang digunakan dalam penentuan algoritma aplikasi *Gibbs sampling* yang diambil dari Bagnoli dan Bergstrom (1989).

Definisi 2.14. Suatu fungsi $f \; \theta$ adalah konkav pada suatu interval (*a*, *b*), jika untuk setiap dua titik θ dan θ di dalam interval (*a*, *b*) dan untuk sembarang $0 \leq \lambda \leq 1$ maka berlaku

$$f \; \lambda\theta \quad 1 - \lambda \; \theta \quad \geq \lambda f \; \theta \quad 1 - \lambda \; f \; \theta \quad .$$

Definisi 2.15. Suatu fungsi $f \; \theta$ adalah log-konkav pada suatu interval (*a*, *b*) jika fungsi $\ln f \; \theta$ adalah fungsi konkav pada (*a*, *b*).

Definisi 2.16. Suatu fungsi $f \; \theta$ adalah log-konkav pada suatu interval (*a*, *b*) jika memenuhi dua kondisi berikut.

1. $f' \; \theta \; / f \; \theta$ adalah monoton turun pada (*a*, *b*).
2. $\ln f \; \theta \quad '' \quad 0$.

Jika tidak memenuhi Definisi 2.14 dan Definisi 2.16 maka disebut fungsi log-konveks pada interval (a, b).

2.1.9 Algoritma *ARS* dan *ARMS*

Menurut Gilks dan Wild (1992), algoritma *ARS* dapat diterapkan jika fungsi densitas bersyarat penuh dari masing-masing parameter adalah log-konkav, jika fungsi densitas bersyarat penuh dari masing-masing parameter adalah log-konveks maka menggunakan algoritma *Adaptive Rejection Metropolis Sampling (ARMS).* Berikut penjelasan dari masing-masing algoritma *ARS* dan *ARMS.*

1. Algoritma *ARS*
    a. Mendefinisikan $h(\theta)=\ln f\ \theta$ dan $h(\theta)$ konkav di setiap D dan mengevaluasi $h(\theta)$ dan $h'(\theta)$ pada $\theta_1 \leq \theta_2 \leq \ldots \leq \theta_k \in D$.
    b. Menginisialisasikan absis dalam T_k , dengan $T_k = \{\theta_1, \theta_2, \ldots, \theta_k\}$, kemudian mendefinisikan fungsi *envelope*, $u_k(\theta)$, yang merupakan batas atas dari garis singgung $h(\theta)$ dan mendefinisikan fungsi *squeezing*, $l_k(\theta)$, yang merupakan batas bawah dari garis singgung $h(\theta)$.
    c. Mengambil sampel θ^* dari $s_k(\theta)$, dengan

$$\theta \qquad \frac{\exp \quad \theta}{\exp \quad \theta \quad \theta},$$

    dan mengambil sampel w dari distribusi uniform (0,1). Jika $\leq$ exp $\theta^* - \theta^*$, maka θ^* diterima, jika tidak maka mengevaluasi □ θ^* dan □′ θ^* . Jika $\leq$ exp □ $\theta^* - \theta^*$, maka θ^* diterima, jika tidak maka θ^* ditolak.
    d. Langkah-langkah tersebut diulang sampai n iterasi hingga diperoleh rata-rata θ^* yang konvergen.
2. Algoritma *ARMS*

    Menurut Gilks *et al.* (1995) algoritma *ARMS* hanya bisa digunakan jika fungsi densitas bersyarat dari distribusi posteriornya mengacu pada distribusi kontinu. Berikut algoritma *ARMS* dalam *Gibbs sampling.*

a. Menginisialisasikan T_k independen terhadap θ , dengan θ merupakan hasil *Gibbs sampling.*

b. Mengambil sampel θ^* dari θ dan membangkitkan sampel dari $\sim$ 0,1 .

c. Jika $p\ \theta^*\ /\exp\ \theta^*$ maka melabeli $\cup\ \theta^*$ dan kembali ke langkah b. Jika tidak maka melabeli θ^* θ

d. Membangkitkan sampel dari $\sim$ 0,1 .

e. Jika $\min\ 1, \frac{\min\ ,}{\min\ ,}$ maka melabeli θ θ .

Jika tidak maka melabeli θ θ .

f. Langkah-langkah tersebut diulang sampai *n* iterasi hingga diperoleh rata-rata θ yang konvergen.

2.2 Kerangka Pemikiran

Metode inferensi Bayes memerlukan distribusi prior untuk menentukan distribusi posterior. Distribusi posterior digunakan untuk menentukan nilai estimasi parameter model *INAR*(1). Nilai estimasi parameter merupakan harga harapan dari distribusi posterior. Penentuan harga harapan dari distribusi posterior digunakan metode *MCMC.* Karena terdapat dua nilai parameter yang tidak diketahui maka digunakan algoritma *Gibbs sampling*. Dalam *Gibbs sampling* jika distribusi bersyarat penuh adalah log-konveks maka digunakan algoritma *ARMS*, tetapi jika distribusi bersyarat penuh dari masing-masing parameter adalah log-konkav maka menggunakan algoritma *ARS.*

BAB III

METODE PENELITIAN

Metode yang digunakan dalam penelitian ini adalah studi literatur, yaitu dengan mempelajari berbagai referensi dari buku dan jurnal-jurnal yang bersesuaian dengan tujuan penelitian.

Adapun langkah-langkah yang ditempuh dalam kaji ulang penelitian Silva *et al.* (2005) khususnya estimasi parameter model *INAR*(1) dengan menggunakan metode Bayes ini sebagai berikut.

1. Menentukan distribusi prior parameter model *INAR*(1).
2. Membentuk fungsi *likelihood* untuk estimasi parameter
3. Membentuk distribusi posterior parameter dengan mengalikan hasil 1 dan 2.
4. Membentuk distribusi posterior parameter α dari hasil langkah 3.
5. Membentuk distribusi posterior parameter λ dari hasil langkah 3.
6. Membentuk algoritma *Gibbs sampling* dengan menggunakan hasil 4 dan 5.
7. Membentuk algoritma *ARS* dan *ARMS* dari hasil 6.
8. Menentukan nilai estimasi parameter dari hasil 7.
9. Menerapkan pada data jumlah orang meninggal akibat kecelakaan bunuh diri di wilayah Surakarta dari Januari 2002-Desember 2006 dengan bantuan *software R 2.11.1*.

BAB IV

PEMBAHASAN

4.1 Model *INAR* (1)

Silva (2005) menjelaskan jika diberikan X adalah variabel random bilangan bulat positif, α adalah probabilitas bertahan dalam suatu proses, dengan $0 \quad \alpha \quad 1$, dan $\;,\;,\;,\ldots,\;$, dengan $\;1, 2, 3, \ldots,\;$, adalah variabel random berdistribusi Bernoulli, $\quad 1 \quad 1 - \quad 0 \quad \alpha,$ *binomial thinning operation* didefinisikan sebagai jumlahan dari variabel random Bernoulli,

$$\alpha \circ \qquad .$$

Variabel random diskrit yang bernilai bilangan bulat positif, dan berdistribusi Poisson, $\;$, dikatakan model *INAR*(1) jika memenuhi persamaan

$$\alpha \circ$$

dengan $\alpha \circ$ merupakan *binomial thinning operation* dan adalah barisan variabel independen yang berdistribusi Poisson dengan parameter λ.

Fungsi densitas dari distribusi bersyarat yang diberikan oleh $\;$, dinotasikan $p \quad |$, adalah hasil konvolusi dari distribusi binomial hasil *binomial thinning operation* dan distribusi Poisson yang merupakan distribusi dari

$$f \qquad \alpha \quad 1 - \alpha$$

$$f \qquad \frac{\lambda \ \exp \ -\lambda}{!}$$

$$p \quad | \qquad f \otimes f$$

$$f \quad f \quad -$$

$$\alpha \quad 1-\alpha \qquad \frac{\lambda \qquad \exp\ -\lambda}{-\quad !}$$

$$\exp\ -\lambda \qquad \frac{\lambda}{-\quad !} \qquad \alpha \quad 1-\alpha \qquad\qquad 4.1$$

4.2 Estimasi Parameter

Pada model *INAR*(1) terdapat parameter yang belum diketahui nilainya dan perlu diestimasi yaitu probabilitas bertahan dalam suatu proses (α) dan parameter laju kedatangan (λ). Terdapat dua metode untuk melakukan estimasi kedua parameter tersebut, yaitu metode klasik dan metode Bayes. Pada penelitian ini membahas tentang estimasi parameter dengan menggunakan metode Bayes. Metode Bayes dipilih sebagai metode untuk estimasi parameter karena dalam metode Bayes dianggap mempunyai informasi yang lengkap yaitu informasi sampel dan informasi yang tersedia sebelum pengambilan sampel. Dalam metode Bayes informasi yang tersedia sebelum pengambilan sampel disebut distribusi prior. Fungsi *likelihood* yang merupakakan informasi sampel dan distribusi prior digunakan untuk menghitung distribusi posterior model *INAR*(1). Berikut adalah penjelasan dari masing-masing distribusi prior dan posterior model *INAR*(1).

4.1.1 Distribusi Prior Parameter Model *INAR*(1)

Distribusi prior merupakan distribusi awal suatu variabel random sebelum dilakukan pengambilan sampel. Pada estimasi parameter dengan menggunakan metode Bayes pemilihan distribusi prior sangat penting dalam menentukan distribusi posterior, karena dengan pemilihan distribusi prior yang tepat akan memudahkan dalam melakukan estimasi. Distribusi posterior lebih mudah diprediksi dengan distribusi prior sekawan, yaitu himpunan distribusi yang tiap anggotanya dapat dikombinasikan dengan fungsi *likelihood* yang dipunyai tanpa menimbulkan kesulitan dalam perhitungan (Berger, 1980).

Pada model *INAR*(1) terdapat dua komponen yaitu *survivors* dalam proses sebelumnya dan kedatangan. Pada komponen *survivors* dalam proses sebelumnya diperoleh data yang dianggap proses Bernoulli. Proses ini diulang sebanyak *n* data observasi sehingga diperoleh fungsi *likelihood* berdistribusi binomial.

$$\alpha|x \qquad \alpha \quad 1-\alpha \qquad \alpha^{\Sigma} \qquad 1-\alpha \quad {}^{\Sigma}$$

Menurut Fink (1997), prior sekawan untuk fungsi *likelihood* yang berupa distribusi binomial adalah prior dengan distribusi beta. Sehingga dari keterangan tersebut dapat ditentukan asumsi probabilitas bertahan dalam suatu proses (α) berdistribusi beta dengan parameter *a* dan *b*, dan dinotasikan $\alpha \sim Beta\ \ a,$.

$$p\ \ \alpha \qquad \frac{a}{a} \alpha \qquad 1-\alpha$$

Pada komponen kedatangan diperoleh data yang dianggap proses Poisson. Data yang dianggap proses Poisson mempunyai fungsi *likelihood* yang sepola dengan distribusi gamma.

$$\lambda|x \quad \propto \lambda^{\Sigma} \quad \exp - \quad \lambda$$

Sehingga dapat ditentukan asumsi parameter laju kedatangan (λ) berdistribusi gamma dengan parameter *c* dan *d*, dan dinotasikan $\lambda \sim \quad A \quad ,$

$$p\ \ \lambda \qquad \frac{1}{\Gamma} \lambda \quad \exp - \ \lambda\ .$$

Distribusi prior parameter model *INAR*(1) dapat diuraikan sebagai berikut.

$$p\ \ \alpha,\lambda \qquad \frac{a}{a} \alpha \qquad 1-\alpha \qquad \frac{1}{\Gamma} \lambda \quad \exp - \ \lambda$$

$$\frac{a}{a} \frac{1}{\Gamma} \alpha \qquad 1-\alpha \qquad \lambda \quad \exp - \ \lambda$$

$$\propto \alpha \qquad 1-\alpha \qquad \lambda \quad \exp - \ \lambda \qquad\qquad 4.2$$

dengan *a*, *b*, *c*, dan *d* adalah parameter yang tidak diketahui.

4.1.2 <u>Distribusi Posterior Parameter Model *INAR*(1)</u>

Distribusi posterior parameter model dapat ditentukan dengan mengalikan fungsi *likelihood* dengan distribusi prior. Fungsi *likelihood* dari model *INAR*(1) dapat ditentukan dari persamaan (4.1)

$$, \alpha, \lambda | \qquad p \quad |$$

$$\exp -\lambda \qquad \frac{\lambda}{- \quad !} \qquad \alpha \quad 1 - \alpha$$

$$\exp - \quad - 1 \ \lambda \qquad \frac{\lambda}{- \quad !} \qquad \alpha \quad 1 - \alpha \qquad 4.3$$

dengan min , .

Distribusi posterior parameter model INAR dapat ditentukan dari persamaan (4.2) dan (4.3) dengan perhitungannya sebagai berikut.

$$p \ \alpha, \lambda | \quad \propto \quad , \alpha, \lambda | \quad p \ \alpha, \lambda$$

$$\exp - \quad - 1 \ \lambda \qquad \frac{\lambda}{- \quad !} \qquad \alpha \quad 1 - \alpha$$

$$\alpha \quad 1 - \alpha \quad \lambda \quad \exp - \ \lambda$$

$$= \exp - \quad - 1 \ \lambda \ \alpha \quad 1 - \alpha \quad \lambda$$

$$\frac{\lambda}{- \quad !} \qquad \alpha \quad 1 - \alpha \qquad 4.4$$

Distribusi posterior pada persamaan (4.4) digunakan untuk menentukan nilai estimasi parameter. Menurut Gilks dan Wild (1992), nilai estimasi parameter diperoleh dengan simulasi pengambilan sampel parameter dari distribusi posterior kompleks menggunakan metode *MCMC*, khususnya dengan algoritma *Gibbs sampling*.

4.1.3 Algoritma *Gibbs sampling*

Algoritma *Gibbs sampling* bisa diterapkan jika distribusi probabilitas bersama dari parameternya tidak diketahui, tetapi distribusi bersyarat dari tiap-tiap parameternya dapat diketahui. Distribusi posterior parameter model *INAR*(1) tidak dapat diketahui, maka akan ditentukan distribusi posterior bersyarat penuh dari masing-masing parameter. Jika diketahui

$$p\ \alpha| \quad \propto \alpha \quad 1-\alpha \quad ,$$

maka distribusi posterior bersyarat penuh dari λ dapat dihitung dari persamaan (4.4) sebagai berikut.

$$p\ \lambda|\alpha, \quad \frac{p\ \alpha,\lambda|}{p\ \alpha|}$$

$$\propto \frac{\exp - \quad - 1\ \lambda\ \alpha \quad 1-\alpha \quad \lambda \quad \prod \quad \sum \frac{\lambda}{-\ !} \quad \alpha\ \ 1-\alpha}{\alpha \quad 1-\alpha}$$

$$\exp - \quad - 1\ \lambda\ \lambda \quad \frac{\lambda}{-\ !} \quad \alpha\ \ 1-\alpha \qquad 4.5$$

Distribusi bersyarat penuh dari λ tersebut merupakan kombinasi linier dari fungsi densitas gamma.

Jika diketahui

$$p\ \lambda| \quad \propto \exp - \quad - 1\ \lambda\ \lambda$$

distribusi posterior bersyarat penuh dari α dapat dihitung dari persamaan (4.4) sebagai berikut.

$$p\ \alpha|\lambda, \quad \frac{p\ \alpha,\lambda|}{p\ \lambda|}$$

$$\propto \frac{\exp - \quad - 1\ \lambda\ \alpha \quad 1-\alpha \quad \lambda \quad \prod \quad \sum \frac{\lambda}{-\ !} \quad \alpha\ \ 1-\alpha}{\exp - \quad - 1\ \lambda\ \lambda}$$

$$\alpha \quad 1-\alpha \quad \prod \quad \sum \frac{}{!} \quad \alpha\ \ 1-\alpha \qquad (4.6)$$

Distribusi bersyarat penuh dari α tersebut merupakan kombinasi linier dari fungsi densitas beta.

Distribusi posterior bersyarat penuh dari masing-masing parameter model *INAR*(1) dapat diketahui. Oleh karena itu algoritma *Gibbs sampling* dapat diterapkan. Nilai awal yang digunakan dalam *Gibbs sampling* adalah hasil estimasi parameter metode *CLS.* Menurut Brannas (1994), jika adalah model *INAR*(1) dengan parameter $\theta \quad \alpha, \lambda$, maka rata-rata bersyarat dari X adalah

$$| \qquad \alpha \qquad \lambda.$$

Estimator CLS dari θ diperoleh dengan meminimalkan Q θ terhadap θ, dengan

$$\theta \qquad - \alpha \qquad - \lambda \; .$$

Estimator CLS dari parameter θ dapat diperoleh sebagai berikut

$$\alpha \quad \frac{\sum \qquad - \sum \quad \sum \qquad /}{\sum \qquad - \sum \qquad /}$$

$$\lambda \quad \frac{\sum \qquad - \alpha \sum}{\qquad}.$$

Berikut adalah algoritma *Gibbs sampling* dari model *INAR*(1).

1. Mengambil nilai $\alpha \; , \lambda$
2. Mencari nilai α dari $p \;\; \alpha | \lambda \;\; ,$
3. Menggunakan nilai α untuk mencari λ dari $p \;\; \lambda | \alpha \;\; ,$
4. Mencari nilai α dari $p \;\; \alpha | \lambda \;\; ,$
5. Menggunakan nilai α untuk mencari λ dari $p \;\; \lambda | \alpha \;\; ,$

   . . .

   . . .

   . . .

6. Mencari nilai α dari $p \;\; \alpha | \lambda \;\; ,$
7. Menggunakan nilai α untuk mencari λ dari $p \;\; \lambda | \alpha \; ,$

4.1.4 Algoritma *ARS*

Distribusi posterior bersyarat penuh untuk α dan λ adalah *log-concav* jika memenuhi:

a. $p'\ \alpha|\lambda, \quad /p\ \alpha|\lambda,$ adalah monoton turun pada (0, 1).

b. $p'\ \lambda|\alpha, \quad /p\ \lambda|\alpha,$ adalah monoton turun pada (0, ∞).

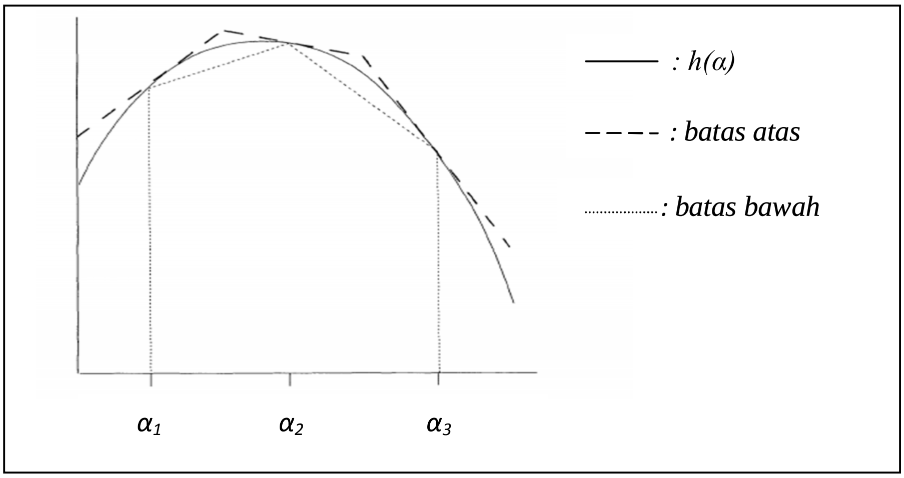
c. $\ln p\ \lambda|\alpha, \quad ''\quad 0$ dan $\ln p\ \alpha|\lambda, \quad ''\quad 0.$

Algoritma *ARS* untuk parameter α adalah sebagai berikut.

1. Langkah Inisialisasi

   Diasumsikan $p\ \alpha|\lambda,$ kontinu dan terdiferensiasi dalam domain *D* yaitu (0, 1). Didefinisikan *h*($\alpha|\lambda,$)=*ln* $p\ \alpha|\lambda,$ dan *h*($\alpha|\lambda,$) konkav disetiap *D.* Sebelum menerapkan algoritma *ARS*, pertama yang harus dilakukan adalah mengevaluasi *h*($\alpha|\lambda,$) dan *h'*($\alpha|\lambda,$) pada $\alpha_1 \le \alpha_2 \le \ldots \le \alpha_k \in D$.

   Menginisialisasikan absis dalam T_k, dengan $T_k = \{\alpha_1, \alpha_2, \ldots, \alpha_k\}$, kemudian mendefinisikan fungsi *envelope*, $\alpha|\lambda,$, yang merupakan batas atas dari garis singgung $h\ \alpha|\lambda,$ dan mendefinisikan fungsi *squeezing*, $\alpha|\lambda,$, yang merupakan batas bawah dari garis singgung *h*($\alpha|\lambda,$), dan dapat digambarkan dalam Gambar 4.1.



Gambar 4.1. Fungsi densitas log-konkav, dengan batas atas dan batas bawah yang didasarkan pada tiga titik absis ($\alpha\ ,\alpha\ ,\alpha\ $)

Garis singgung pada α_j dan α_{j+1} berpotongan di titik

$$\frac{\square\ \alpha \quad |\lambda, \quad - \square\ \alpha\ |\lambda, \quad - \alpha \quad {}_1\square'\ \alpha \quad |\lambda, \qquad \alpha\ \square'\ \alpha\ |\lambda,}{\square'\ \alpha\ |\lambda, \quad - \square'\ \alpha \quad |\lambda,}$$

dengan j = 1, ..., k-1 dan adalah batas bawah dari D, adalah batas atas dari D. Fungsi batas atas dapat didefinisikan sebagai berikut.

$$\alpha|\lambda, \qquad \square\ \alpha\ |\lambda, \qquad \alpha - \alpha \quad \square'\ \alpha\ |\lambda,$$

dengan $x \in$, . Fungsi batas bawah dapat didefinisikan sebagai berikut.

$$\alpha|\lambda, \qquad \frac{\alpha \quad - \alpha\ \square\ \alpha\ |\lambda, \qquad \alpha - \alpha\ \ \square\ \alpha \quad |\lambda,}{\alpha \quad - \alpha}$$

dengan $\alpha \in\ \ \alpha\ , \alpha \quad {}_1$.

2. Langkah Penyampelan

Mengambil sampel α^* dari $s_k(\alpha|\lambda, \ )$, dengan

$$\alpha|\lambda, \qquad \frac{\exp \quad \alpha|\lambda,}{\exp \quad \alpha|\lambda, \quad \alpha},$$

dan mengambil sampel u dari distribusi uniform (0,1).

a) Uji *squeezing*

Jika $\ \leq \exp \quad \alpha^*|\lambda, \quad - \quad \alpha^*|\lambda,$, maka α^* diterima, jika tidak maka mengevaluasi $\square\ \alpha^*|\lambda,$ dan $\square'\ \alpha^*|\lambda,$ kemudian dilakukan uji *rejection.*

b) Uji *rejection*

Jika $\ \leq \exp\ \square\ \alpha^*|\lambda, \quad - \quad \alpha^*|\lambda,$, maka α^* diterima, jika tidak maka α^* ditolak.

3. Langkah Pembaruan

Jika $\square\ \alpha^*|\lambda,$ dan $\square'\ \alpha^*|\lambda,$ dievaluasi dalam setiap langkah penyampelan, maka memasukkan nilai α^* ke dalam T_k untuk membentuk T_{k+1}. Merekonstruksikan fungsi $\alpha^*|\lambda,$, $\alpha^*|\lambda,$, dan $\alpha^*|\lambda,$ dari T_{k+1} yang digunakan untuk iterasi selanjutnya.

Langkah-langkah dalam algoritma *ARS* tersebut diulang sampai n iterasi hingga diperoleh rata-rata α^* yang konvergen. Algoritma yang sama digunakan dalam penentuan λ dengan mengganti unsur α dengan λ.

4.1.5 Algoritma *ARMS* dalam *Gibbs sampling*.

Algoritma *ARMS* hanya bisa digunakan jika fungsi densitas bersyarat dari distribusi posteriornya mengacu pada distribusi kontinu. Berikut algoritma *ARMS* dalam *Gibbs sampling.*

Algoritma *ARMS* model *INAR*(1) untuk parameter α dapat ditulis sebagai berikut.

1. Menginisialisasikan T_k independen terhadap α , dengan α merupakan hasil *Gibbs sampling.*
2. Mengambil sampel α^* dari α dan membangkitkan sampel dari $\sim$ 0,1 .
3. Jika $p\ \alpha^*|\lambda, x\ / exp\ \ \alpha^*|\lambda,$ maka melabeli $\cup\ \alpha^*$ dan kembali kelangkah 2. Jika tidak maka melabeli α^* α
4. Membangkitkan sampel dari w$\sim$ 0,1 .
5. Jika $m\ \ 1, \frac{|\ ,\quad |\ ,\ ,\quad |\ ,}{|\ ,\quad |\ ,\ ,\quad |\ ,}$ maka melabeli α α . Jika tidak maka melabeli α α .

Langkah-langkah dalam algoritma *ARMS* tersebut di n iterasi hingga diperoleh rata-rata α^* yang konvergen. Algoritma yang sama digunakan dalam penentuan λ dengan mengganti unsur α dengan λ.

4.1.6 Estimator Bayes untuk Parameter Model *INAR*(1)

Informasi pada distribusi posterior bersyarat penuh dari masing-masing parameter dapat digunakan untuk menentukan estimator untuk parameter. Semua estimator untuk parameter model merupakan fungsi dari data hasil observasi. Jika θ adalah parameter model, maka θ adalah merupakan estimator untuk parameter θ. Pada model *INAR(1)* terdapat dua nilai yang belum diketahui yaitu α dan λ.

Ditentukan θ α, λ yang merupakan himpunan nilai yang belum diketahui, dengan demikian θ α, λ adalah estimator untuk θ α, λ . Misalkan I θ merupakan fungsi dari parameter θ α, λ , jika diambil θ I θ maka θ merupakan estimator dari I θ . Estimator Bayes merupakan estimator yang meminimalkan fungsi resiko θ , dengan θ merupakan harga harapan dari fungsi kerugian, $\theta|\theta$. Estimasi Bayes dapat ditentukan dengan menggunakan Teorema 2.2 dan diperoleh

$$\alpha \quad |, \ I\ \alpha \quad \mathrm{I}\ \alpha\ f\ \alpha|\lambda, \quad \alpha \quad \alpha\ f\ \alpha|\lambda, \quad \alpha \qquad 4.7$$

$$\infty \qquad \infty$$

$$\lambda \quad |, \ \mathrm{I}\ \lambda \quad \mathrm{I}\ \lambda\ f\ \lambda|\alpha, \quad \lambda \quad \lambda\ f\ \lambda|\alpha, \quad \lambda \qquad 4.8$$

Perhitungan integrasi pada persamaan (4.7) dan (4.8) sangat sulit dilakukan, sehingga digunakan konsep integrasi Monte Carlo. Konsep integrasi Monte Carlo adalah dengan membangkitkan sampel random dari distribusi $f\ \alpha|\lambda,$ dan $f\ \lambda|\alpha,$ kemudian menghitung rata-rata dari sampel yang telah dibangkitkan dari masing-masing fungsi tersebut. Perhitungan untuk harga harapan estimator parameter model dapat dituliskan sebagai berikut.

$$\alpha \quad \alpha\ f\ \alpha|\lambda, \quad \alpha \approx \frac{1}{} \quad \alpha$$

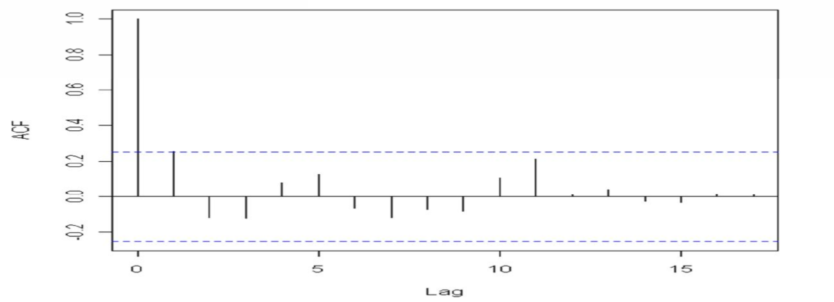
$$\lambda \quad \lambda\ f\ \lambda|\alpha, \quad \lambda \approx \frac{1}{} \quad \alpha$$

Penyampelan dari distribusi probabilitas $f\ \alpha|\lambda,$ dan $f\ \lambda|\alpha,$ dapat dilakukan dengan proses Markov. Proses Markov dilakukan dengan membuat rantai Markov dengan distribusi stasionernya mendekati distribusi probabilitas $f\ \alpha|\lambda,$ dan $f\ \lambda|\alpha,$. Pembuatan rantai Markov tersebut dapat dilakukan dengan menggunakan algoritma *ARS* jika fungsi densitas bersyarat penuh dari masing-masing parameter adalah log-konkav dan menggunakan algoritma *ARMS* jika fungsi densitas bersyarat penuh dari masing-masing parameter adalah log-konveks. Estimasi Bayes untuk parameter α dapat diperoleh dengan menghitung rata-rata barisan α. Estimasi Bayes untuk parameter λ dapat diperoleh dengan menghitung rata-rata barisan λ.

4.3 Contoh Kasus

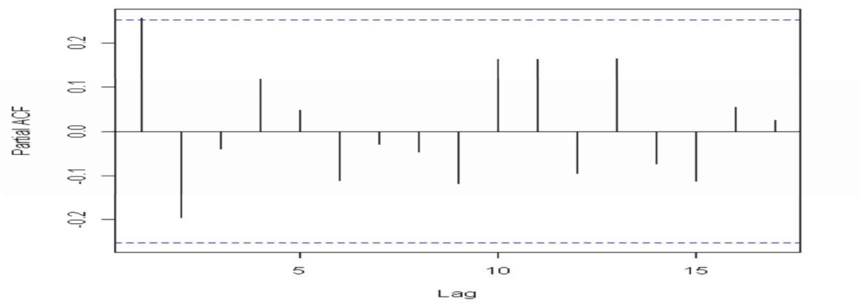
Lampiran 1 adalah data jumlah orang meninggal akibat kecelakaan bunuh diri di wilayah Surakarta dari Januari 2002-Desember 2006 dari BPS. Data berdistribusi Poisson karena memiliki mean dan variansi yang hampir sama yaitu sebesar 0.416667 dan 0.687853. Identifikasi model awal data adalah model *INAR*(1) dengan plot ACF dapat digambarkan pada Gambar 4.2. dan PACF dapat digambarkan pada Gambar 4.3.

Data Kecelakaan



Gambar 4.2. Plot ACF data jumlah orang meninggal akibat kecelakaan bunuh diri di wilayah Surakarta

Data Kecelakaan



Gambar 4.3.Plot PACF data jumlah orang meninggal akibat kecelakaan bunuh diri di wilayah Surakarta

Distribusi prior parameter model adalah sekawan maka dapat ditentukan dengan persamaan (1) yaitu

$$\alpha \qquad 1-\alpha \qquad \lambda \qquad \exp - \quad \lambda$$

dengan a=b=c=d=10 . Distribusi posterior bersyarat penuh untuk parameter α dapat ditentukan dengan persamaan (4.5) yaitu

$$p \;\; \alpha|\lambda, \qquad Beta \;\; a, \quad \lambda \qquad 1-\alpha \qquad \frac{\lambda}{6} \quad \frac{\lambda}{8}$$

$$\frac{\lambda}{8} \;\; 1-\alpha \qquad \frac{\lambda}{2}\alpha \;\; 1-\alpha \qquad \frac{3}{2}\lambda \;\; \alpha \quad 1-\alpha \qquad \lambda\alpha$$

$$\frac{\lambda}{2} \;\; 1-\alpha \qquad 2\lambda\alpha \;\; 1-\alpha \qquad \alpha \qquad \lambda \;\; 1-\alpha \qquad \alpha$$

$$\frac{\lambda}{2} \;\; 1-\alpha \;\; \lambda\alpha \quad \lambda \;\; 1-\alpha \qquad 2\alpha \;\; 1-\alpha \qquad\qquad 4.9$$

Sedangkan distribusi posterior untuk parameter λ, dapat ditentukan dengan menggunakan persamaan (4.6) yaitu

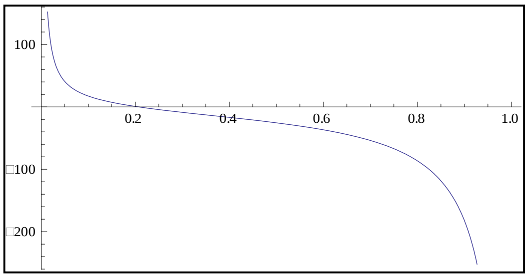
$$p \;\; \lambda|\alpha, \qquad amma \;\; , \qquad 119 \qquad 1-\alpha \qquad \frac{\lambda}{6} \quad \frac{\lambda}{8}$$

$$\frac{\lambda}{8} \;\; 1-\alpha \qquad \frac{\lambda}{2}\alpha \;\; 1-\alpha \qquad \frac{3}{2}\lambda \;\; \alpha \quad 1-\alpha \qquad \lambda\alpha$$

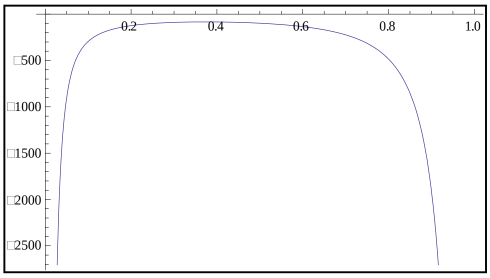
$$\frac{\lambda}{2} \;\; 1-\alpha \qquad 2\lambda\alpha \;\; 1-\alpha \qquad \alpha \qquad \lambda \;\; 1-\alpha \qquad \alpha$$

$$\frac{\lambda}{2} \;\; 1-\alpha \;\; \lambda\alpha \quad \lambda \;\; 1-\alpha \qquad 2\alpha \;\; 1-\alpha \qquad\qquad 4.10$$

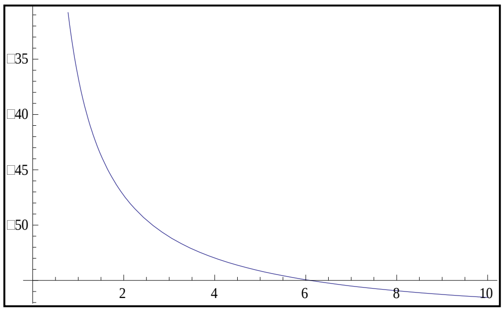
Estimator parameter model *INAR*(1) dapat ditentukan dengan menggunakan pendekatan *MCMC*. Distribusi posterior dari masing-masing parameter ditunjukkan terlebih dahulu log-konkav atau log-konveks dengan melihat Gambar 4.4- 4.7.



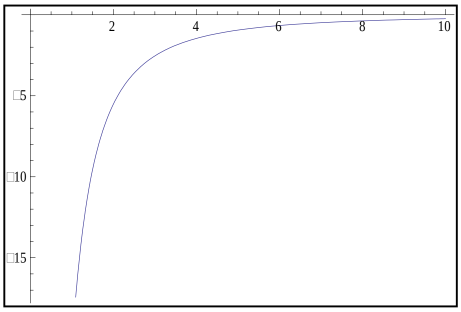
Gambar 4.4. Plot fungsi $p' \;\; \alpha|\lambda, \quad /p \;\; \alpha|\lambda,$



Gambar 4.5. Plot turunan kedua dari fungsi log $p \;\; \alpha|\lambda,$



Gambar 4.6. Plot fungsi $p'\ \lambda|\alpha,\quad /p\ \lambda|\alpha,$



Gambar 4.7. Plot turunan kedua dari fungsi log $p\ \lambda|\alpha,$

Dari Gambar 4.4- 4.7 fungsi $\ln p\ \lambda|\alpha,\quad ''\quad 0$ dan $\ln p\ \alpha|\lambda,\quad ''\quad 0$, selain itu $p'\ \alpha|\lambda,\quad /p\ \alpha|\lambda,$ dan $p'\ \lambda|\alpha,\quad /p\ \lambda|\alpha,$ adalah monoton turun, sehingga dapat disimpulkan fungsi densitas bersyarat penuh dari masing-masing parameter adalah log-konkav.

Persamaan (4.9) dan (4.10) digunakan untuk membangkitkan rantai Markov dengan algoritma *ARS*. Karena distribusi bersyarat penuh dari masing-masing parameter berdimensi tinggi, maka dalam pembangkitan rantai Markov menggunakan bantuan *software R 2.11.1*. Tabel 4.1 adalah nilai-nilai estimasi parameter α dan λ.

Tabel 4.1. Nilai-nilai estimasi parameter model *INAR(1)* untuk data jumlah orang meninggal akibat kecelakaan bunuh diri di wilayah Surakarta

n	100	250	500	750	1000
α	0.253174	0.2374823	0.2551190	0.2453593	0.2457684
Var(α	0.009237554	0.007994338	0.00791093	0.007418821	0.007775721
λ	0.3478246	0.3636025	0.3665941	0.3626544	0.3626545
Var(λ	0.007144554	0.00652917	0.006624199	0.006087363	0.005894286

Dari Tabel 4.1 dapat diperoleh nilai α=0.2474 dan λ=0.30607, yang artinya bahwa setiap individu pada bulan sebelumnya memiliki probabilitas bertahan hidup sampai bulan berikutnya sebesar 0.2474, dan rata-rata banyaknya orang yang meninggal dalam waktu tiga bulan adalah satu orang.

BAB V

PENUTUP

5.1 Kesimpulan

Berdasar hasil pembahasan diperoleh kesimpulan bahwa estimasi Bayes untuk parameter model *INAR*(1) adalah

$$\alpha \quad {}_{|\,,} \quad \mathrm{I}\ \alpha \quad \mathrm{I}\ \alpha \quad f \ \alpha|\lambda, \quad \alpha \quad \alpha \ f \ \alpha|\lambda, \quad \alpha$$

$$\lambda \quad {}_{|\,,} \quad \mathrm{I}\ \lambda \quad {}^{\infty}\mathrm{I}\ \lambda \quad f \ \lambda|\alpha, \quad \lambda \quad {}^{\infty}\lambda \ f \ \lambda|\alpha, \quad \lambda \ .$$

Integrasi tersebut diselesaikan mengggunakan metode *MCMC* dan diperoleh $\alpha \approx -\sum \quad \alpha$ dan $\lambda \approx -\sum \quad \lambda$. Nilai α dan λ merupakan rantai Markov yang dibangkitkan dari algoritma *ARS* atau *ARMS.*

5.2 Saran

Penulis mengestimasi parameter model *INAR*(1) menggunakan metode Bayes untuk data yang berdistribusi Poisson. Bagi pembaca yang tertarik dengan topik ini, dapat mengembangkan lebih lanjut untuk mengestimasi parameter model *INAR*(1) untuk data berdistribusi negatif binomial atau binomial .