

USULAN TUGAS AKHIR

1. IDENTITAS PENGUSUL

NAMA : Nuriana Rahma Setyani
NRP : 5110100069
DOSEN WALI : Dr.Ir. Siti Rochimah, MT.
DOSEN PEMBIMBING : 1. Ahmad Saikhu, S.SI., MT.
2. Rully Soelaiman, S.Kom, M.Kom

2. JUDUL TUGAS AKHIR

“Implementasi Model Statistik *Multivariate Zero-Inflated* yang Melibatkan Prediktor Laten”

3. LATAR BELAKANG

Dalam menganalisis hubungan antara beberapa variabel, bentuk dari variabel respon biasanya berbentuk biner dan diskrit. Namun sering dihadapi pula peristiwa dimana variabel responnya banyak mengandung nilai nol sehingga analisis dengan pendekatan distribusi diskrit atau kontinu tidak lagi membuahkan hasil yang tepat. Model *Multivariate Zero-Inflated* dengan efek *random* disajikan untuk membahas fitur khusus dari data multivariat yang banyak mengandung nilai nol. Model multivariat ini terdiri dari dua komponen yaitu komponen yang membahas *excess zero* (data mengandung banyak nilai nol) dan komponen yang membahas pemodelan multivariat yang menggambarkan parameter yang dibutuhkan. Namun, pemodelan *Multivariate Zero-Inflated* ini masih sulit digunakan untuk menganalisis data dan merepresentasikan hasil karena persentase nilai nol lebih tinggi dibandingkan dengan apa yang diharapkan. Oleh sebab itu dibutuhkan pendekatan algoritma *Markov Chain Monte Carlo* (MCMC) untuk mengestimasi parameter yang dibutuhkan[1].

Pemodelan multivariat ini menggunakan Distribusi Poisson dan Distribusi Gamma. Distribusi Poisson digunakan untuk menghitung parameter yang bersifat diskrit sedangkan Distribusi Gamma digunakan untuk menghitung parameter yang bersifat kontinu. Prediktor laten digunakan untuk mengembangkan pemodelan

Multivariate Zero-Inflated. Algoritma MCMC digunakan untuk mengestimasi parameter dari model *Multivariate Zero-Inflated* dengan membangun rantai Markov pada Distribusi Poisson dan Gamma.

Hasil yang diharapkan dari tugas akhir ini adalah algoritma MCMC dapat mengestimasi parameter pemodelan multivariat yang didukung prediktor laten sebagai pengembangan pemodelan *Multivariate Zero-Inflated*.

4. RUMUSAN MASALAH

Permasalahan yang dibahas dalam Tugas Akhir ini meliputi hal-hal sebagai berikut.

- a. Memahami konsep metode *Multivariate Zero-Inflated* yang melibatkan prediktor laten.
- b. Mengimplementasi algoritma *Markov Chain Monte Carlo* (MCMC) pada metode *Multivariate Zero-Inflated*.
- c. Menguji dengan beberapa parameter berdasarkan Distribusi Poisson dan Distribusi Gamma untuk mengukur kinerja dari algoritma tersebut.

5. BATASAN MASALAH

Batasan masalah yang ada pada Tugas Akhir ini antara lain

- a. Implementasi tugas akhir ini menggunakan aplikasi Matlab.
- b. Menggunakan data sintetik yang berasal dari efek *random* berdasarkan Distribusi Poisson dan Gamma.
- c. Menggunakan metode MCMC untuk mengembalikan apakah parameter pada pemodelan *Multivariate Zero-Inflated* sesuai atau tidak.

6. TUJUAN PEMBUATAN TUGAS AKHIR

Tujuan dari Tugas Akhir ini antara lain

- a. Mempelajari karakteristik pemodelan *Multivariate Zero-Inflated* yang melibatkan prediktor laten.
- b. Mengimplementasikan algoritma MCMC untuk mengestimasi parameter pemodelan *Multivariate Zero-Inflated*.

7. MANFAAT TUGAS AKHIR

Manfaat yang diharapkan dari Tugas Akhir ini yaitu implementasi metode ini dapat diaplikasikan pada dunia bisnis untuk mengetahui optimasi dari *feedback behaviour* perusahaan.

8. TINJAUAN PUSTAKA

Model *Multivariate Zero-Inflated*

Model *Multivariate Zero-Inflated* dengan efek *random* disajikan untuk membahas fitur khusus dari data multivariat yang banyak mengandung nilai nol.

Model multivariat ini terdiri dari dua komponen yaitu komponen yang membahas data yang mengandung banyak nilai nol dan komponen yang membahas pemodelan multivariat yang menggunakan Distribusi Poisson dan Distribusi Gamma. Distribusi Gamma digunakan untuk mengukur parameter yang bersifat diskrit sedangkan Distribusi Gamma digunakan untuk mengukur parameter yang bersifat kontinu[2].

Exploratory Factor Analysis

Exploratory Factor Analysis (EFA) merupakan analisis faktor yang memiliki prosedur, yaitu memeriksa data dan memberikan penelitian dengan informasi kira-kira berapa banyak faktor yang diperlukan untuk menggambarkan data[4]. Pada studi kasus ini Variabel laten penjelas digunakan untuk memperjelas parameter yang dibutuhkan[2].

Algoritma Markov Chain Monte Carlo

Algoritma *Markov Chain Monte Carlo* (MCMC) merupakan algoritma yang digunakan untuk melakukan *sampling* dari distribusi probabilitas dengan membangun rantai Markov pada suatu distribusi tertentu yang sulit melakukan simulasi secara langsung[3]. Dalam Tugas Akhir ini algoritma MCMC yang digunakan adalah algoritma Metropolis-Hastings. Algoritma Metropolis-Hastings merupakan algoritma untuk membangkitkan barisan sampel menggunakan mekanisme penerimaan dan penolakan (*accept-reject*) dari suatu distribusi probabilitas yang sulit untuk dilakukan penarikan sampel. Langkah-langkah algoritma Metropolis-Hastings antara lain[5]:

- a Dimulai dengan sebarang $x^{(t)}$ pada setiap iterasi $t = 1, 2, \dots T$.
- b Bangkitkan sampel $Y_t \sim g(y)$ dengan $g(y) = q(y|x)$
- c Lakukan tes ratio

$$TR = \frac{f(Y_t)g(x^{(t)})}{f(x^{(t)})g(Y_t)} \quad (1)$$

- d Hitung probabilitas *accept-reject*

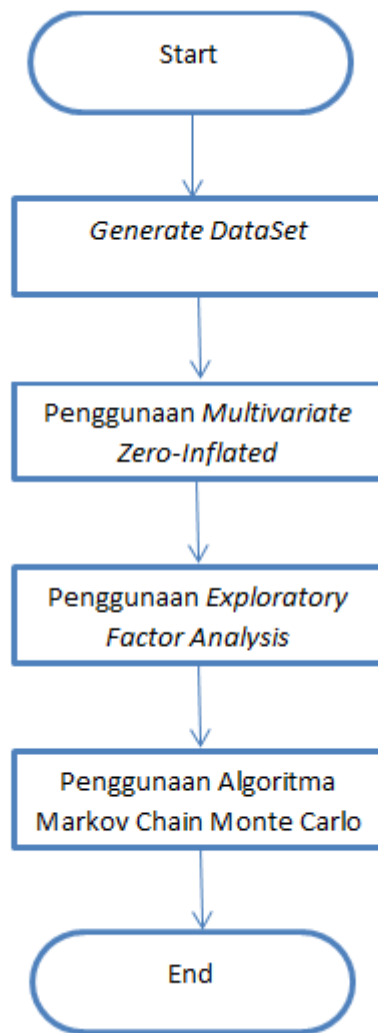
$$a_{ij} = \min\{TR, 1\} \quad (2)$$

$$X^{(t+1)} = \begin{cases} Y_t & \text{dengan probabilitas } a_{ij} \\ x^{(t)} & \text{selainnya} \end{cases} \quad (3)$$

Pada Tugas Akhir ini algoritma MCMC digunakan untuk mengestimasi parameter pemodelan multivariat yang didukung prediktor laten sebagai penjelas parameter[2].

9. RINGKASAN ISI TUGAS AKHIR

Tugas akhir ini akan melakukan eksplorasi terhadap pemodelan *Multivariate Zero-Inflated* beserta algoritma *Markov Chain Monte Carlo*. Proses eksplorasi ini melalui tahapan-tahapan yang sebagaimana dijelaskan dalam diagram alur pada Gambar 1.



Gambar 1 Diagram Alur untuk Pemodelan Multivariate Zero-Inflated yang Melibatkan Prediktor Laten Beserta Algoritma MCMC

Berikut penjelasan beberapa tahapan Pemodelan Multivariate Zero-Inflated yang Melibatkan Prediktor Laten Beserta Algoritma MCMC yang dilalui yaitu:

1. *Generate DataSet*
Pada tahapan ini dilakukan *generate* data dengan efek *random* berdasarkan Distribusi Poisson dan Gamma. Distribusi Poisson digunakan untuk merepresentasikan parameter yang bersifat diskrit. Sedangkan Distribusi Gamma digunakan untuk menggambarkan parameter yang bersifat kontinu.
2. *Penggunaan Multivariate Zero-Inflated*
Pada tahapan ini dilakukan penggabungan antara Distribusi Poisson dan Distribusi Gamma.
3. *Penggunaan Exploratory Factor Analysis*

Exploratory Factor Analysis (EFA) merupakan analisis faktor yang berguna untuk memberikan informasi penjabar kira-kira berapa banyak faktor yang diperlukan untuk menjelaskan data.

4. Penggunaan Algoritma *Markov Chain Monte Carlo*

Algoritma *Markov Chain Monte Carlo* (MCMC) digunakan untuk mengestimasi parameter yang dibutuhkan oleh pemodelan *Multivariate Zero-Inflated* yang melibatkan prediktor laten.

10. METODOLOGI

a. **Penyusunan proposal tugas akhir**

Proposal Tugas Akhir ini ditulis untuk mengajukan ide mengenai pemodelan *Multivariate Zero-Inflated* beserta prediktor laten. Untuk mengestimasi parameter-parameter yang dibutuhkan diperlukan pendekatan algoritma *Markov Chain Monte Carlo* (MCMC).

b. **Studi literatur**

Pada proses ini dilakukan studi lebih lanjut terhadap konsep-konsep yang terdapat pada jurnal, buku, artikel, dan literatur lain yang menunjang. Studi dilakukan untuk mendalami konsep algoritma baik *Multivariate Zero-Inflated*, Algoritma MCMC maupun EFA dan konsep lain yang berguna untuk menyelesaikan permasalahan yang muncul pada proses pengerjaan Tugas Akhir ini.

c. **Implementasi algoritma**

Implementasi merupakan tahap untuk membangun sistem tersebut. Terdapat tiga macam metode serta algoritma yang akan diimplementasikan yaitu metode *Multivariate Zero-Inflated*, algoritma MCMC dan algoritma EFA. Implementasi diproses menggunakan Matlab.

d. **Pengujian dan evaluasi**

Pengujian yang akan dilakukan dalam Tugas Akhir ini yaitu dilakukan uji parameter menggunakan algoritma MCMC sehingga nantinya dapat diketahui parameter yang dibutuhkan oleh pemodelan *Multivariate Zero-Inflated*.

e. **Penyusunan Buku Tugas Akhir**

Pada tahap ini dilakukan penyusunan laporan yang menjelaskan dasar teori dan metode yang digunakan dalam Tugas Akhir ini serta hasil dari implementasi aplikasi perangkat lunak yang telah dibuat. Sistematika penulisan buku tugas akhir secara garis besar antara lain:

1. Pendahuluan

- a. Latar Belakang
- b. Rumusan Masalah
- c. Batasan Tugas Akhir
- d. Tujuan

- e. Metodologi
- f. Sistematika Penulisan
2. Tinjauan Pustaka
3. Desain dan Implementasi
4. Pengujian dan Evaluasi
5. Kesimpulan dan Saran
6. Daftar Pustaka

11. JADWAL KEGIATAN

Berikut dijelaskan pada tabel 1 mengenai jadwal kegiatan yang diajukan dalam pengerjaan Tugas Akhir.

Tabel 1. Jadwal Kegiatan

Tahapan	2014																							
	Januari				Februari				Maret				April				Mei				Juni			
Penyusunan Proposal																								
Studi literatur																								
Implementasi																								
Pengujian dan evaluasi																								
Penyusunan buku																								

12.DAFTAR PUSTAKA

- [1] Klein Entink, R.H., Fox, J.P., van der Linden, W.J., A Multivariate Multilevel Approach to The Modeling of Accuracy and Speed of Test Takers. *Psychometrika*, vol. 74, pp. 21–48, 2009.
- [2] Fox, Jean-Paul, Multivariate Zero-Inflated Modeling with Latent Predictors: Modeling *feedback* behavior, *Computational Statistics and Data Analysis*, vol. 68, pp. 361-374, 2013.
- [3] Tunaru, Radu, Hierarchical Bayesian Models for Multiple Count Data, *Austrian Journal of Statistics*, Vol. 31, pp. 221-229, 2002.
- [4] Williams, Brett, Exploratory Factor Analysis: A five-step guide for novices, *Australasian Journal of Emergency Primary Health Care (JEPHC)*, Vol.8, pp. 1-13, 2010.
- [5] Chib, Siddhartha, Greenberg, Edward, Understanding the Metropolis-Hastings Algorithm, *American Statistical Association*, Vol. 49, pp. 327-335, 2001.