

Penerapan Distribusi *Random-Clumped Binomial* Untuk Mengatasi Masalah Overdispersi Dalam Data Biner

Implementation Of Random-Clumped Binomial Distribution For Overdispersion Problems In Binary Data

¹ Verdina Ambarwati, ² Nusar Hajarisman

^{1,2}Prodi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Islam Bandung, Jl. Tamansari No.1 Bandung 40116

Email ¹ verdinambarwati4@gmail.com, ²nusarhajarisman@yahoo.com

Abstract. Binary data is the generated data from the numerical mapping of an event that has only two possible events. Where the event is considered a "success" is symbolized by the number 1 whereas the event "fail" is symbolized by the number 0. The number of "success" events is followed by the binomial distribution if it meets the assumption that response variables are independent and have equal "success" chances. In practice, there are often violations of assumptions. This is an indication of an overdispersion problem that is the condition in which the variance of the T response (the number of successful events) exceeds the variance of the Binomial. Data can be expressed as overdispersion when the ratio between devides or chi squares of Pearson to its degree of freedom is more than 1. The Random-Clumped Binomial distribution discussed in this thesis is one method that can be used to overcome the overdispersion problem in the data following the binomial distribution. Random-Clumped Binomial Distribution is then applied to find out whether the distribution can overcome the problem of overdispersion in the data of Student Pass Status Year 2016/2017 in Statistics Majority, Science and Mathematics Faculty, Bandung Islamic Univeristy at Courses in Semester 1 and Semester 2. Results obtained that the distribution of Random-Clumped Binomial can overcome the problem of overdispersion that occurs in the data, and obtained the opportunity that the possibilty of Student Pass Status Year 2016/2017 in Statistics Majority, Science and Mathematics Faculty, Bandung Islamic Univeristy get grades not graduated in the semester 1 and 2 courses is 0.1016 with the correlation between the value of one course with the value of the eye other lectures in semesters 1 and 2 of Student Pass Status Year 2016/2017 in Statistics Majority, Science and Mathematics Faculty, Bandung Islamic Univeristy amounted 0.5661.

Keywords: Binary Data, Binomial Distribution, Random-Clumped Binomial Distribution, Maximum Likelihood Method, Devians and Chi Squares Pearson.

Abstrak. Data biner merupakan data yang dibangkitkan dari pemetaan numerik dari suatu kejadian yang hanya memiliki dua kemungkinan kejadian. Dimana kejadian yang dianggap sebagai kejadian "sukses" disimbolkan dengan angka 1 sedangkan kejadian "gagal" disimbolkan dengan angka 0. Banyaknya peristiwa "sukses" dinyatakan mengikuti distribusi binomial apabila memenuhi asumsi bahwa variabel responnya saling bebas dan mempunyai peluang "sukses" yang sama. Pada praktiknya sering terjadi pelanggaran asumsi. Hal tersebut merupakan salah satu indikasi adanya masalah overdispersi yaitu kondisi dimana varians dari respon T (banyaknya kejadian sukses) melebihi varians dari Binomial. Data dapat dinyatakan mengalami masalah overdispersi ketika rasio antara nilai devians atau chi kuadrat *Pearson* terhadap derajat bebasnya lebih dari 1. Distribusi *Random-Clumped Binomial* yang dibahas dalam skripsi ini merupakan salah satu metode yang dapat digunakan untuk mengatasi masalah overdispersi pada data yang mengikuti distribusi binomial. Distribusi *Random-Clumped Binomial* kemudian diaplikasikan untuk mengetahui apakah distribusi tersebut dapat mengatasi masalah overdispersi pada data Status Kelulusan Mahasiswa Tahun 2016/2017 Program Studi Statistika FMIPA Universitas Islam Bandung pada Mata Kuliah di Semester 1 dan Semester 2. Hasil yang diperoleh bahwa distribusi *Random-Clumped Binomial* dapat mengatasi masalah overdispersi yang terjadi pada data, dan diperoleh peluang bahwa Mahasiswa Tahun 2016/2017 Program Studi Statistika FMIPA Universitas Islam Bandung memperoleh nilai tidak lulus pada mata kuliah semester 1 dan 2 adalah 0.1016 dengan korelasi antara nilai mata kuliah yang satu dengan nilai mata kuliah yang lainnya pada semester 1 dan 2 dari Mahasiswa Tahun 2016/2017 Program Studi Statistika FMIPA Universitas Islam Bandung sebesar 0.5661.

Kata kunci: Data biner, Distribusi Binomial, Distribusi Random-Clumped Binomial, Metode Kemungkinan Maksimum, Devians dan Chi Kuadrat Pearson.

A. Pendahuluan

Penelitian pada dasarnya adalah cara ilmiah untuk mendapatkan data dengan tujuan dan kegunaan tertentu. Pada kenyataannya sering ditemukan data hasil penelitian yang berdistribusi diskrit. Salah satunya adalah data hasil penelitian yang mempunyai struktur biner. Dimana data biner merupakan data yang dibangkitkan dari pemetaan numerik dari satu tindakan atau percobaan yang menghasilkan hanya dua kemungkinan kejadian. Dalam data biner kejadian yang dianggap sebagai kejadian “sukses” disimbolkan dengan angka 1 dan untuk kejadian “gagal” diberi angka 0.

Banyaknya peristiwa “sukses” dapat dinyatakan mengikuti distribusi binomial apabila memenuhi asumsi bahwa variabel acaknya saling bebas dan mempunyai peluang “sukses” yang sama (konstan). Namun pada praktiknya tak jarang terjadi pelanggaran asumsi. Pelanggaran asumsi tersebut merupakan suatu indikasi adanya masalah overdispersi pada pengamatan data biner. Sebagaimana yang diungkapkan oleh Morel dan Nagaraj (2011) bahwa, jika diasumsikan satuan unsur saling bebas dengan peluang sukses yang sama, model binomial sesuai untuk menganalisis data. Pengklasteran menyebabkan adanya korelasi antar satuan unsur dan peluang “sukses” dalam kelompok menjadi tidak sama (bervariasi). Akibatnya, data biasanya menunjukkan varians yang lebih besar daripada varians yang diizinkan oleh model binomial. Fenomena ini dikenal sebagai overdispersi atau *extra variation*.

Masalah overdispersi yang terjadi pada pengamatan data diskrit ini dapat dijelaskan oleh dua hal, yaitu adanya korelasi antar variabel acak dan adanya keragaman dari peluang “sukses” pada variabel acak tersebut. Keduanya merupakan kejadian yang saling berhubungan, sehingga ketika terdapat korelasi antar variabel acak, maka akan terdapat juga keragaman dari peluang “sukses” dan begitupun sebaliknya. McCullagh dan Nelder (1989) menyebutkan bahwa kedua kejadian tersebut dapat terjadi karena adanya pengklasteran di dalam populasi. Sedangkan menurut Collet (1991) kejadian-kejadian tersebut muncul karena sejumlah unit percobaan yang diamati beberapa kali pada kondisi yang sama, sehingga akan diperoleh suatu peluang pada variabel acak yang berbeda dari suatu percobaan ke percobaan lainnya.

Munculnya masalah overdispersi dalam data diskrit juga akan berdampak pada nilai taksiran untuk variansnya. Di mana, ketika nilai penaksir varians ini digunakan dalam menghitung selang kepercayaan (*confident interval*), maka akan menghasilkan rata-rata yang kecil sehingga menyebabkan selang kepercayaan yang terlalu pendek. Selain itu juga, apabila nilai penaksir parameter ini digunakan untuk pengujian hipotesis maka hasilnya akan cenderung terima hipotesis nol.

Berdasarkan uraian di atas, maka perlu dicari suatu metode statistika yang tepat untuk dapat digunakan dalam menentukan taksiran peluang bagi parameter pada data yang mengikuti distribusi binomial, dimana ada masalah overdispersi di dalamnya.

Ada beberapa macam model distribusi yang telah diusulkan oleh para penulis untuk mengatasi masalah overdispersi pada data yang mengikuti distribusi binomial, seperti distribusi *Zero-Inflated Binomial* yang biasanya digunakan ketika ada kelebihan angka nol pada data yang akan diamati, *A Generalization of the Binomial Distribution* yang digunakan ketika hasil bernoulli terdistribusi secara identik namun tidak saling bebas diusulkan oleh Altham (1978) dan yang lebih sering digunakan yaitu distribusi beta-binomial yang merupakan gabungan dari distribusi beta dan distribusi binomial dimana distribusi beta binomial adalah distribusi binomial dengan peluang sukses pada setiap percobaan tidak tetap tetapi acak dan mengikuti distribusi beta yang pertama kali digunakan oleh Skellam (1948).

Dari sekian banyak model yang diusulkan, yang akan dibahas dalam skripsi ini adalah penyelesaian masalah overdispersi pada data Status Kelulusan Mahasiswa Tahun 2016/2017 Program Studi Statistika FMIPA Universitas Islam Bandung pada Mata Kuliah Semester 1 dan 2 dengan menggunakan model distribusi *Random-Clumped* binomial yang merupakan campuran dari dua buah distribusi binomial yang diboboti.

Berdasarkan latar belakang yang telah diuraikan, maka perumusan masalah dalam penelitian ini sebagai berikut: “Apakah terdapat masalah overdispersi pada data status kelulusan Mahasiswa Tahun 2016/2017 Program Studi Statistika FMIPA Universitas Islam Bandung?”. Selanjutnya, tujuan dalam penelitian ini diuraikan dalam pokok-pokok sbb.

1. Untuk mengetahui apakah terdapat masalah overdispersi pada data status kelulusan Mahasiswa Tahun 2016/2017 Program Studi Statistika FMIPA Universitas Islam Bandung.
2. Untuk mengetahui bahwa distribusi *Random-Clumped* binomial dapat mengatasi masalah overdispersi pada data status kelulusan Mahasiswa Tahun 2016/2017 Program Studi Statistika FMIPA Universitas Islam Bandung.

B. Landasar Teori

Distribusi binomial adalah distribusi peluang diskret dengan jumlah keberhasilan dalam m buah percobaan ya/tidak (berhasil/gagal) yang variabel acaknya (Y) saling bebas dan mempunyai peluang “sukses” sebesar π . Metode kemungkinan maksimum adalah salah satu metode yang dapat digunakan untuk menaksir parameter yang mempunyai peluang terbesar dari data pengamatan, akan lebih mudah memaksimumkan fungsi ln-kemungkinan karena fungsi ln merupakan fungsi dalam bentuk penjumlahan (bukan perkalian). Adapun fungsi ln-kemungkinan maksimum dari distribusi binomial adalah $l(\pi) = \sum Y_i \ln(\pi) + (m^2 - \sum Y_i) \ln(1 - \pi)$ dengan turunan pertama terhadap parameter π adalah

$$\frac{\partial l(\pi)}{\partial(\pi)} = \frac{\sum Y_i}{\pi} - \frac{(m^2 - \sum Y_i)}{1 - \pi} = \frac{\sum Y_i - m^2 \pi}{\pi(1 - \pi)}$$

Menurut Morel dan Nagaraj (2011) pengklasteran menyebabkan adanya korelasi pada variabel acak dan peluang “sukses” dalam klaster menjadi bervariasi. Akibatnya, data biasanya menunjukkan varians yang lebih besar daripada varians yang diizinkan oleh distribusi binomial. Dimana, fenomena ini dikenal sebagai overdispersi atau extra variation. Rasio antara nilai chi-kuadrat Pearson terhadap derajat bebas yang bernilai lebih dari 1 menunjukkan adanya masalah overdispersi (McCullagh dan Nelder, 1989).

Rumus dari nilai chi-kuadrat Pearson adalah $\chi^2 = \sum_{i=0}^n \frac{(O_i - E_i)^2}{E_i}$, dimana O_i adalah nilai observasi ke i dan E_i adalah frekuensi harapan dari observasi ke i . Adapun derajat bebas untuk statistik chi-kuadrat *Pearson* adalah $(n - 1 - p)$ dimana p adalah banyaknya parameter.

Distribusi *Random-Clumped* binomial merupakan suatu distribusi yang dihasilkan dalam rangka memodelkan keragaman ekstra (overdispersi) dari suatu model binomial biasa. Distribusi *Random-clumped* binomial yang diusulkan oleh Morel dan Nagaraj (1993), adalah campuran dari dua buah distribusi binomial dan karenanya sesuai dengan pemilihan distribusi diskrit seperti sebelumnya. Misalkan T merupakan variabel acak yang mengikuti distribusi *Random-Clumped* binomial, maka fungsi

peluangnya adalah $Pr(T = t) = \pi Pr(X_1 = t) + (1 - \pi) Pr(X_2 = t), t = 0, 1, \dots, m$. Dimana $X_1 \sim \text{binomial}\{m; (1 - \rho)\pi + \rho\}$ dan $X_2 \sim \text{binomial}\{m; (1 - \rho)\pi\}$. Dan fungsi ln-kemungkinan maksimumnya adalah $(\pi, \rho) = \ln \left\{ \prod_{i=1}^m [(\pi(\pi - \rho\pi + \rho))^t (1 - \pi + \rho\pi - \rho)^{m-t}] + ((1 - \pi)(\pi - \rho\pi)^t (1 - \pi + \rho\pi)^{m-t}) \right\}$.

C. Metode Penelitian

Data yang digunakan untuk mengaplikasikan metode yang akan dibahas dalam skripsi ini adalah data sekunder yang diperoleh dari pihak Administrasi Pengelolaan Data Universitas Islam Bandung. Dimana data yang digunakan merupakan data nilai mata kuliah semester satu dan semester dua dari mahasiswa tahun 2016/2017 Program Studi Statistika FMIPA Universitas Islam Bandung.

Seperti yang diketahui, bahwa nilai akhir dari suatu mata kuliah yang diperoleh oleh mahasiswa dinyatakan dengan nilai mutu yaitu: *A, B, C, D, E, T* dan *K*. Nilai mutu sebagai variabel *Y* di kelompokkan menjadi dua yaitu 0 untuk nilai mutu *A, B* dan *C* dengan keterangan “LULUS” dan 1 untuk nilai mutu *D, E, T* (komponen penilaian tidak lengkap) dan *K* (komponen penilaian tidak ada/kosong) dengan keterangan “TIDAK LULUS”.

Tabel. Data Status Kelulusan Mahasiswa Tahun 2016/2017 Program Studi Statistika FMIPA Universitas Islam Bandung pada Mata Kuliah Semester 1 dan 2

Mahasiswa	<i>Y</i> ₁	<i>Y</i> ₂	<i>Y</i> ₃	<i>Y</i> ₄	<i>Y</i> ₅	<i>Y</i> ₆	<i>Y</i> ₇	<i>Y</i> ₈	<i>Y</i> ₉	<i>Y</i> ₁₀	<i>Y</i> ₁₁	<i>Y</i> ₁₂	<i>Y</i> ₁₃	<i>Y</i> ₁₄	<i>Y</i> ₁₅	<i>Y</i> ₁₆	<i>T</i>	
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
98	0	1	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	3
99	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Dimana *Y*₁= status kelulusan untuk mata kuliah PKN, *Y*₂ = Kalkulus 1, *Y*₃ = Bahasa Indonesia, *Y*₄ = Bahasa Inggris, *Y*₅=Matematika Dasar, *Y*₆ = Aljabar Linier Elementer, *Y*₇=Vektor dan Matriks, *Y*₈ = Metode Statistika 1, *Y*₉ = Pengantar Peluang, *Y*₁₀= Statistika Kependudukan, *Y*₁₁= Metode Statistika 2, *Y*₁₂=Manajemen Data, *Y*₁₃= Analisis Input Output, *Y*₁₄= Praktikum Managemen Data, *Y*₁₅= Praktikum Metode Statistika, *Y*₁₆= PAI Aqidah dan *T*= Jumlah Mata Kuliah yang Tidak Lulus.

Adapun tahapan pelaksanaan dari penelitian ini adalah sebagai berikut:Pencarian data(1), Menentukan penaksir peluang dari distribusi Binomial melalui metode penaksir kemungkinan maksimum (2), Memeriksa ada tidaknya masalah overdispersi dalam data melalui ukuran rasio antara statistik chi-kuadrat *Pearson* berdasarkan derajat bebasnya (3), Apabila hasil tidak menunjukkan adanya masalah overdispersi, maka nilai penaksir peluang pada distribusi Binomial cocok untuk data tersebut (stop), tetapi apabila hasil

menunjukkan adanya masalah overdispersi maka lakukan langkah-langkah selanjutnya (4), Melakukan penanganan overdispersi dengan menggunakan distribusi *Random-Clumped* Binomial (5), Berhenti (6).

D. Hasil Penelitian dan Pembahasan

Data status kelulusan Mahasiswa Tahun 2016/2017 Program Studi Statistika FMIPA Universitas Islam Bandung pada mata kuliah di semester 1 dan semester 2 pada Tabel 1 kemudian disajikan kedalam bentuk tabel frekuensi. Dan berikut ini adalah Tabel Frekuensi dari data status kelulusan Mahasiswa Tahun 2016/2017 Program Studi Statistika FMIPA Universitas Islam Bandung pada mata kuliah di semester 1 dan semester 2 yang berjumlah 99 Mahasiswa:

Tabel 2. Data Jumlah Mata Kuliah dari Mahasiswa Tahun 2016/2017 Program Studi Statistika FMIPA Universitas Islam Bandung yang “TIDAK LULUS” pada Semester 1 dan Semester 2

T	0	1	2	3	4	Total
Frekuensi	75	13	6	4	1	99

Tabel 2 menunjukkan bahwa dari 99 Mahasiswa Tahun 2016/2017 Program Studi Statistika FMIPA Universitas Islam Bandung diantaranya ada sebanyak 75 Mahasiswa yang memperoleh nilai lulus pada semua mata kuliah di semester 1 dan semester 2, 13 Mahasiswa yang memperoleh 1 nilai tidak lulus pada mata kuliah di semester 1 dan semester 2, 6 Mahasiswa yang memperoleh 2 nilai tidak lulus pada mata kuliah di semester 1 dan semester 2, 4 Mahasiswa yang memperoleh 3 nilai tidak lulus pada mata kuliah di semester 1 dan semester 2, 1 Mahasiswa yang memperoleh 4 nilai tidak lulus pada mata kuliah di semester 1 dan semester 2.

Nilai taksiran parameter pada distribusi binomial dan *Random-Clumped* Binomial diperoleh melalui program *nlmixed* pada perangkat lunak SAS 9.4. Adapun nilai taksiran parameter pada distribusi binomial akan disajikan kedalam tabel 1.

Tabel 3. Nilai Taksiran Parameter dari Distribusi Binomial

Parameter	
	π_{Bin}
Penaksir	0.1035

Dari hasil pada Tabel 3 di atas dapat dilihat bahwa nilai taksiran dari parameter π_{Bin} sebesar 0.1035 yang berarti bahwa peluang Mahasiswa Tahun 2016/2017 Program Studi Statistika FMIPA Universitas Islam Bandung memperoleh nilai tidak lulus pada mata kuliah di semester 1 dan 2 adalah sebesar 0.1035.

Berikut ini adalah nilai chi-kuadrat Pearson dari distribusi binomial.

Tabel 4. Hasil Chi- Kuadrat Pearson dari Distribusi Binomial

t	Frekuensi		Binomial	
	Observasi (O_i)		Probability Value	Frekuensi Harapan
0	75		0.646	63.954
1	13		0.298	29.502
2	6		0.052	5.148
3	4		0.004	0.396
4	1		0.0001	0.0099
χ^2			143.099	
db			3	
χ^2/db			47.6996667	

Rasio antara nilai chi-kuadrat *Pearson* terhadap derajat bebas pada distribusi binomial yang bernilai 47.6996667 (lebih besar dari 1) menunjukkan adanya masalah overdispersi pada data status kelulusan Mahasiswa Tahun 2016/2017 Program Studi Statistika FMIPA Universitas Islam Bandung pada mata kuliah di semester 1 dan semester 2, sehingga distribusi *Random-Clumped* binomial dapat digunakan untuk mengatasi masalah overdispersi. Adapun nilai taksiran parameter dan nilai chi kuadrat *Pearson* dari distribusi *Random-Clumped* binomial adalah sbb:

Tabel 5. Nilai Taksiran Parameter dari Distribusi *Random-Clumped* Binomial

Penaksir	Parameter	
	π_{RCB}	ρ_{RCB}
	0.1016	0.5661

Dari hasil di atas dapat dilihat bahwa nilai taksiran dari parameter π_{RCB} sebesar 0.1016 yang berarti bahwa peluang Mahasiswa Tahun 2016/2017 Program Studi Statistika FMIPA Universitas Islam Bandung memperoleh nilai tidak lulus pada mata kuliah di semester 1 dan semester 2 adalah sebesar 0.1016, nilai taksiran dari parameter ρ_{RCB} sebesar 0.5661 yang berarti bahwa terdapat korelasi sebesar 0.5661 diantara mata kuliah yang satu dengan mata kuliah yang lainnya pada semester 1 dan semester 2 dari Mahasiswa Tahun 2016/2017 Program Studi Statistika FMIPA Universitas Islam Bandung.

Tabel 6. Hasil Chi-Kuadrat Pearson dari Distribusi *Random-Clumped* Binomial

T	Frekuensi Observasi (O_i)	Random-Clumped Binomial	
		Probability Value	Frekuensi Harapan
0	75	0.752	74.448
1	13	0.153	15.147
2	6	0.044	4.356
3	4	0.036	3.564
4	1	0.014	1.386
χ^2		1.615	
db		2	
χ^2/db		0.8075	

Nilai chi-kuadrat *Pearson* terhadap derajat bebas pada distribusi *Random-Clumped* binomial yang bernilai 0.8075 (kurang dari 1) menunjukkan bahwa distribusi *Random-Clumped* binomial telah berhasil mengatasi masalah overdispersi yang ada dalam data status kelulusan Mahasiswa Tahun 2016/2017 Program Studi Statistika FMIPA Universitas Islam Bandung pada mata kuliah di semester 1 dan semester 2.

E. Kesimpulan dan Saran

Kesimpulan

Berdasarkan pembahasan dalam penelitian ini, peneliti menyimpulkan beberapa hasil penelitian sebagai berikut:

1. Diperoleh rasio antara nilai chi kuadrat *Pearson* terhadap derajat bebasnya pada distribusi binomial yang lebih besar dari 1 yaitu 47.6996667 menunjukkan bahwa terdapat masalah overdispersi pada data status kelulusan Mahasiswa Tahun 2016/2017 Program Studi Statistika FMIPA Universitas Islam Bandung pada mata kuliah di semester 1 dan semester 2, sehingga distribusi binomial tidak sesuai untuk data ini.
2. Distribusi *Random-Clumped* binomial dapat mengatasi masalah overdispersi pada data status kelulusan Mahasiswa Tahun 2016/2017 Program Studi Statistika FMIPA Universitas Islam Bandung pada mata kuliah di semester 1 dan semester 2, karena rasio antara nilai chi kuadrat *Pearson* terhadap derajat bebasnya yang kurang dari 1 yaitu sebesar 0.8075. Dimana pada distribusi *Random-Clumped* binomial diperoleh juga nilai-nilai sebagai berikut:
 - Nilai taksiran dari π_{RCB} sebesar 0.1016 dan π_{Bin} sebesar 0.1035 yang artinya peluang bahwa Mahasiswa tahun 2016/2017 Program Studi Statistika FMIPA Universitas Islam Bandung memperoleh nilai tidak lulus pada mata kuliah di semester 1 dan semester 2 adalah sebesar 0.1016.
 - Nilai taksiran dari ρ_{RCB} sebesar 0.5661 yang artinya bahwa terdapat korelasi antara nilai mata kuliah yang satu dengan nilai mata kuliah yang lainnya pada semester 1 dan 2 dari Mahasiswa tahun 2016/2017 Program Studi Statistika FMIPA Universitas Islam Bandung sebesar 0.5661.

- Nilai galat baku taksiran, yang menjelaskan besarnya kesalahan, untuk parameter π_{RCB} sebesar 0.02104 dan ρ_{RCB} sebesar 0.07799.
- Nilai t sebesar 4.83 dengan p-value sebesar 0.0048 untuk parameter π_{RCB} dan nilai t sebesar 7.26 dengan p-value sebesar 0.0008 untuk parameter ρ_{RCB} .
- Selang kepercayaan bagi parameter dengan tingkat kepercayaan sebesar 95% diperoleh nilai sebagai berikut: nilai π_{RCB} berada diantara 0.04754 sampai 0.1557 dan nilai ρ_{RCB} berada diantara 0.3656 sampai 0.7666.

Saran

1. Selain distribusi *Random-Clumped* Binomial masih banyak distribusi-distribusi lain yang dapat digunakan dalam mengatasi masalah overdispersi dalam data biner. Untuk itu para pembaca dapat menggunakan distribusi-distribusi tersebut dalam mengatasi masalah overdispersi atau ekstra variasi khusus untuk kasus penelitian data biner agar bisa menjadi perbandingan.
2. Dalam skripsi ini penulis tidak membahas model, sehingga dapat dikembangkan dan diimplementasikan kedalam model-model untuk overdispersi agar dapat mengetahui pula faktor-faktor apa saja yang berpengaruh terhadap kasus yang sedang diteliti.

F. Daftar Pustaka

- Agresti, A. (1990). *Categorical Data Analysis*. New York: John Wiley and Sons.
- Agresti, A. (2007). *An Introduction to Categorical Data Analysis (Second Edition)*. New York: John Wiley and Sons.
- Altham, Patricia M.E. (1978). "Two Generalizations of the Binomial Distribution". *Applied Statistics*, 27, 162-167.
- Collet, D. (1991). *Modeling Binary Data*. London: Chapman and Hall.
- Hajarisman, N. (1998). *Kajian Perbandingan Model Regresi Beta Binomial dengan Model Regresi Logistik dan Penerapannya untuk Menduga Pola Kelulusan Mahasiswa TPB-IPB*. Bogor: Institut Pertanian Bogor
- Hajarisman, N. (2009). *Analisis Data Kategorik*. Bandung: Pusat Penerbitan Unisba.
- McCullagh, P., and J.A. Nelder. (1989). *Generalized Linear Models. (Second Edition)*. New York: Chapman and Hall. Wilson
- Morel, Jorge G., and Neerchal, Nagaraj K. (1993). A Finite Mixture Distribution for Modelling Multinomial Extra Variation. *Biometrika*, 80, 363-371.
- Morel, Jorge G., and Neerchal, Nagaraj K. (2011). *Overdispersion Models in SAS*. Cary, NC, USA: SAS Institute Inc.
- Sunendiari, S. (2012). *Statistika Matematika II*. Bandung: Pusat Penerbitan Unisba.