UNIVERSITAS INDONESIA

PENAKSIRAN PARAMETER DISTRIBUSI BINOMIAL NEGATIF PADA KASUS OVERDISPERSI

SKRIPSI

SHAFIRA
0706261915

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
PROGRAM STUDI MATEMATIKA
DEPOK
JULI 2011

UNIVERSITAS INDONESIA

PENAKSIRAN PARAMETER DISTRIBUSI BINOMIAL NEGATIF PADA KASUS OVERDISPERSI

SKRIPSI

Skripsi ini diajukan sebagai syarat untuk memperoleh gelar Sarjana Sains

SHAFIRA
0706261915

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
PROGRAM STUDIMATEMATIKA
DEPOK
JULI 2011

HALAMAN PERNYATAAN ORISINALITAS

Skripsi ini adalah hasil karya saya sendiri,

dan semua sumber baik yang dikutip maupun dirujuk

telah saya nyatakan dengan benar.

Nama	: Shafira
NPM	: 0706261915
Tanda Tangan	:
Tanggal	: 13 Juli 2011

HALAMAN PENGESAHAN

Skripsi ini diajukan oleh

Nama	: Shafira
NPM	: 0706261915
Program Studi	: Matematika
Judul Skripsi	: Penaksiran Parameter Distribusi Binomial Negatif pada Kasus Overdispersi

Telah berhasil dipertahankan di hadapan Dewan Penguji dan diterima sebagai persyaratan yang diperlukan untuk memperoleh gelar Sarjana Sains pada Program Studi Matematika, Fakultas Ilmu Pengetahuan Alam Universitas Indonesia

DEWAN PENGUJI

Pembimbing	: 1. Dra. Siti Nurrohmah, M. Si	( )
	2. Dra. Saskya Mary Soemartojo, M. Si	( )
Penguji	: 1. Dra. Netty Sunandi, M.Si	(Netty S)
	2. Dra. Rianti Setiadi, M.Si	( )
	3. Mila Novita, S.Si, M.Si	( )

Ditetapkan di	: Depok
Tanggal	: 14 Juni 2011

KATA PENGANTAR

Alhamdulillah, puji dan syukur penulis panjatkan kepada Allah SWT yang telah memberikan rahmat, taufik dan hidayah-Nya, sehingga penulisan skripsi dengan judul "Penaksiran Parameter Distribusi Binomial Negatif pada Kasus Overdispersi" ini dapat terselesaikan dengan baik. Penulisan skripsi ini dilakukan dalam rangka memenuhi salah satu syarat untuk mencapai gelar Sarjana Sains Jurusan Matematika di Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Indonesia.

Penulis menyadari bahwa penyelesaian skripsi ini tidak terlepas dari bantuan serta dukungan dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis ingin mengucapkan terima kasih kepada :

1. Orang tua penulis yang telah mengasuh dan mendidik penulis sampai saat ini dan selalu memberikan dukungan baik moril maupun materiil, serta tak pernah berhenti mendoakan penulis.
2. Ibu Dra. Siti Nurrohmah, M.Si dan ibu Dra. Saskya Mary, M.Si selaku dosen pembimbing. Terima kasih sebesar-besarnya untuk semua bantuan, saran, kritik dan dorongan semangat serta motivasi yang luar biasa yang diberikan kepada penulis dalam menyelesaikan skripsi ini.
3. Ibu Dra. Ida Fithriani, M.Si selaku pembimbing akademis. Terima kasih untuk bantuan, dorongan, saran, bimbingan, dan perhatian untuk penulis selama menjalani masa kuliah.
4. Bapak Dr. Yudi Satria, M.T. selaku ketua departemen, Ibu Rahmi Rusin, S.Si., M.Sc.Tech selaku sekretaris departemen, dan Ibu Dr. Dian Lestari selaku koordinator pendidikan yang telah membantu penulis selama masa perkuliahan dan proses penyelesaian skripsi ini.
5. Ibu Rianti, Ibu Netty, Ibu Sarini, Ibu Mila, Ibu Fevi, Ibu Sri, dan seluruh staf pengajar departemen Matematika UI yang tanpa mengurangi rasa hormat tidak dapat disebutkan namanya satu per satu. Terima kasih atas segala masukan, saran, kritikan serta semangat yang diberikan kepada penulis selama proses

penyelesaian skripsi ini dan terimakasih atas ilmu yang telah diberikan. Semoga penulis dapat menggunakan ilmu tersebut dengan sebaik-baiknya.

6. Mas Salman, Mba Rusmi, Mba Santi, Mas Ansori, Mba Fia dan seluruh karyawan departemen Matematika UI, terima kasih atas segala bantuan dan kemudahan yang telah diberikan untuk penulis.
7. Kak Isat, Fina, Fairuz, terima kasih atas doa dan dukungannya.
8. Widya, Nedi, Hikmah, Manda, Andi, Danar, Adit (Aca), dan Hanif, terima kasih sahabat untuk semua saat-saat indah yang kita lalui bersama.
9. Ashari, terima kasih atas bantuannya dalam penyelesaian program matlab pada tugas akhir ini.
10. Teman-teman seperjuangan di Statistik, Gamgam, Petos, Riyantoto, Winwin, Shafa, Sica, Anggun, Anjar, Adi, Ciput, Para, Teteh, Wiwi , serta mahasiswa Matematika UI angkatan 2007 lainnya, Afni, Arip, Ashar, Bow, Citi, Dita, Fauzan, Ferdi, Lois, Nora, Putuwira, Farah, Iki, Daw, Siska, Tatep, Widi, Anis, Yos, dan Zul. Terima kasih atas semuanya, masa kuliah ini tak akan bisa dilupakan.
11. Angkatan 2006, khususnya Kak Teguh, Kak Dian dan Kak Tino, terima kasih untuk semua nasihat, saran, dukungan semangat, motivasi, dan do'a yang diberikan kepada penulis.
12. Angkatan 2009, khususnya Handa, Cepi, Tika, Faisal, Agung, Soleman, terima kasih untuk dukungan semangat, motivasi dan doa yang kalian berikan.
13. Semua mahasiswa Matematika UI angkatan 2005 dan 2008.
14. Semua pihak yang telah membantu pengerjaan skripsi ini, yang namanya tidak bisa disebutkan satu per satu, penulis mengucapkan terima kasih.

Penulis menyadari bahwa masih terdapat banyak kekurangan pada skripsi ini, penulis mohon maaf. Semoga skripsi ini dapat memberikan manfaat bagi yang membacanya.

Penulis

2011

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik Universitas Indonesia, saya yang bertanda tangan di bawah ini :

Nama	: Shafira
NPM	: 0706261915
Program Studi	: S1
Departemen	: Matematika
Fakultas	: MIPA
Jenis karya	: Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Indonesia Hak Bebas Royalti Nonekslusif (*Non-exclusive Royalty Free Right*) atas karya ilmiah saya yang berjudul :

Penaksiran Parameter Distribusi Binomial Negatif pada Kasus Overdispersi

beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalty Nonekslusif ini Universitas Indonesia berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data(*database*), dan mempublikasikan skripsi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Depok

Pada tanggal : 13 Juli 2011

Yang menyatakan

(Shafira)

ABSTRAK

Nama	: Shafira
Program Studi	: Matematika
Judul	: Penaksiran Parameter Distribusi Binomial Negatif pada Kasus Overdispersi

Data *count* sering diasumsikan berdistribusi Poisson yang hanya memiliki satu parameter dengan mean dan variansi sama. Namun pada kenyataannya, sering ditemukan data *count* dengan variansi yang lebih besar dari mean, keadaan ini dikenal dengan overdispersi. Saat terjadi overdispersi, maka data *count* tidak berdistribusi Poisson, sehingga perlu dicari distribusi lain yang dapat digunakan untuk menganalisis data *count.* Salah satu distribusi yang dapat digunakan adalah distribusi Binomial Negatif yang merupakan distribusi Campuran Poisson-Gamma (*Mixture Poisson-Gamma Distribution*). Pada tugas akhir ini akan dijelaskan mengenai perumusan distribusi Binomial Negatif sebagai distribusi Campuran Poisson-Gamma (*Mixture Poisson-Gamma Distribution*), penaksiran parameter pada distribusi Binomial Negatif yang merupakan distribusi campuran Poisson-Gamma, serta mempelajari sifat-sifat taksirannya.

Kata Kunci	: data *count*, distribusi Poisson, overdispersi, distribusi binomial negatif, distribusi Campuran Poisson-Gamma.
xi + 79 halaman	; 4 tabel , 2 gambar
Bibliografi	: 18 (1962 – 2011)

ABSTRACT

Name : Shafira
Program Study : Mathematics
Tittle : Estimating Parameter of Negative Binomial Distribution on Overdispersion Case.

Counts data, often assumed follows a Poisson distribution has same mean and variance value. But in fact, count data often has variance value greater than mean value, this condition is called by overdispersion. When overdispersion occured, data count doesn’t have a Poisson distribution, so need to find another distribution which can be applied for data count analyzing. One of distribution which often be applied is Negative Binomial distribution which is form as a mixture Poisson-Gamma distribution. In this minithesis, estimators of Negative Binomial distribution (Mixture Poisson-Gamma distribution) and their characteristics will be explained.

Keyword : *count data*, Poisson distribution, overdispersion, Negative Binomial distribution, Mixture Poisson-Gamma distribution.
xi + 79 pages ; 4 tables, 2 figure
Bibliography : 18 (1962 – 2011)

DAFTAR ISI

DAFTAR TABEL

DAFTAR GAMBAR

DAFTAR LAMPIRAN

BAB I
PENDAHULUAN

Bab ini membahas mengenai latar belakang masalah, perumusan masalah, tujuan penulisan, pembatasan masalah dan sistematika penulisan dari tugas akhir ini.

1.1 Latar Belakang

Data *count* adalah data hasil percobaan acak yang nilai-nilainya berupa bilangan bulat non negatif {0, 1, 2, …}. Misalkan Y adalah suatu peubah acak yang nilainya berupa data *count*. Distribusi yang biasa digunakan untuk memodelkan Y adalah distribusi Poisson. Distribusi Poisson memiliki suatu karakteristik khusus, yaitu memiliki nilai mean dan variansi yang sama, atau dapat dinyatakan sebagai berikut :

$$E(Y) = \mathrm{var}(Y).$$

Pada kenyataanya, sering ditemukan data *count* dengan variansi yang lebih besar dari mean, keadaan ini dikenal dengan overdispersi. Overdispersi dapat disebabkan oleh adanya korelasi positif antar observasi atau individu yang diamati atau terdapat nilai variansi yang besar pada data *count* (Joseph M Hilbe, 2011).

Terjadinya overdispersi pada data *count* mengindikasikan bahwa nilai variansi lebih besar dari mean , hal ini dapat menimbulkan beberapa masalah, antara lain menyebabkan taksiran parameter yang didapat menjadi tidak efisien walaupun tetap konsisten, sehingga akan memberikan informasi yang tidak sesuai (Joseph M Hilbe, 2011).

Saat terjadi overdispersi, asumsi kesamaan mean dan variansi pada distribusi Poisson dilangar, sehingga perlu dicari distribusi lain yang dapat digunakan untuk menganalisis data *count* saat terjadi overdispersi. Salah satu distribusi yang dapat digunakan adalah distribusi Binomial Negatif yang merupakan distribusi Campuran Poisson-Gamma (*Mixture Poisson-Gamma distribution*).

Pada tugas akhir ini akan dijelaskan mengenai perumusan distribusi Binomial Negatif sebagai distribusi Campuran Poisson-Gamma (*Mixture Poisson-Gamma distribution*), metode penaksiran parameter pada distribusi Binomial Negatif yang merupakan distribusi Campuran Poisson-Gamma, serta mempelajari sifat-sifat taksirannya.

1.2 Perumusan Masalah

1. Bagaimana cara merumuskan distribusi Binomial Negatif sebagai distribusi Campuran Poisson-Gamma (*Mixture Poisson-Gamma distribution*) pada kasus overdispersi pada data *count* ?
2. Bagaimana mengestimasi parameter-parameter pada distribusi Binomial Negatif saat terjadi overdispersi ?
3. Bagaimana sifat-sifat taksiran parameter tersebut?

1.3 Tujuan Penulisan

Tujuan dari tugas akhir ini adalah :

1. Membahas dan mempelajari distribusi Binomial Negatif sebagai distribusi Campuran Poisson-Gamma (*Mixture Poisson-Gamma distribution*) saat terjadi overdispersi.
2. Membahas tentang estimasi parameter pada distribusi Binomial Negatif saat terjadi overdispersi dengan menggunakan metode maksimum *likelihood*.
3. Mempelajari sifat-sifat taksiran parameter dari distribusi Binomial Negatif.

1.4 Pembatasan Masalah

Pembatasan masalah pada tugas akhir ini yaitu :

1. Distribusi Binomial Negatif yang dibahas pada tugas akhir ini adalah distribusi Binomial Negatif sebagai distribusi Campuran Poisson-Gamma (*Mixture Poisson-Gamma distribution*) pada kasus overdispersi.
2. Penaksiran parameter dari distribusi Binomial Negatif dilakukan dengan menggunakan metode maksimum *likelihood*.
3. Penentuan solusi dari fungsi *likelihood* untuk menentukan taksiran parameter dispersi dari distribusi Binomial Negatif diselesaikan dengan pendekatan numerik dan metode numerik yang akan digunakan adalah metode *Newton-Rhapson*.

1.5 Sistematika Penulisan

Penulisan tugas akhir ini dibagi menjadi lima bab, yaitu :

Bab I : Pendahuluan

Bab ini berisi latar belakang masalah, perumusan masalah, tujuan penulisan, pembatasan masalah dan sistematika penulisan.

Bab II : Landasan Teori

Bab ini membahas tentang teori-teori dasar yang melandasi penulisan tugas akhir ini yaitu, peubah acak dan jenisnya yaitu peubak acak diskrit dan kontinu, transformasi peubah acak diskrit, distribusi Binomial, distribusi Poisson, distribusi Gamma, distribusi campuran (*Mixture distribution*), distribusi Binomial Negatif, metode Bayes, *prior natural conjugate*, metode taksiran titik salah satunya denga metode maksimum *likelihood*, sifat-sifat taksiran titik antara lain tak bias, konsisten dan efisien, metode *Newton-Rhapson*, uji rasio *likelihood*.

Bab III : Penaksiran Parameter pada Distribusi Binomial Negatif untuk Mengatasi Overdispesi

Bab ini membahas tentang terjadinya overdispersi pada distribusi Poisson, kemudian distribusi Binomial Negatif pada kasus overdispersi untuk data *count*, distribusi Gamma sebagai *prior natural conjugate* dari distribusi Poisson, mean dan variansi dari distribusi Binomial Negatif, taksiran parameter dari distribusi Binomial Negatif , metode maksimum *likelihood*, serta pengujian parameter dispersi.

Bab IV : Aplikasi Data

Bab ini membahas tentang aplikasi data yang meliputi penjelasan mengenai sumber data, analisis data. Analisis data yang dilakukan yaitu uji *Kolmogorov*-Smirnov, penaksiran parameter, dan uji rasio *likelihood*.

Bab V : Penutup

Bab ini berisi kesimpulan dan saran dari tugas akhir ini.

BAB II
LANDASAN TEORI

Bab ini membahas tentang teori-teori dasar yang melandasi penulisan tugas akhir ini yaitu, peubah acak dan jenisnya yaitu peubak acak diskrit dan kontinu, transformasi peubah acak diskrit, distribusi Binomial, distribusi Poisson, distribusi Gamma, distribusi campuran (*Mixture distribution*), distribusi Binomial Negatif, metode Bayes, *prior natural conjugate*, metode taksiran titik salah satunya dengan metode maksimum *likelihood*, sifat-sifat taksiran titik antara lain tak bias, konsisten dan efisien, metode *Newton-Rhapson,* uji rasio *likelihood.*

2.1 Peubah Acak dan Jenisnya

Ruang sampel adalah himpunan semua kemungkinan hasil dari suatu percobaan acak. Anggota sampel adalah setiap kemungkinan hasil dalam suatu ruang sampel. Kejadian merupakan himpunan bagian (*subset*) dari ruang sampel. Ruang sampel tidak selalu beranggotakan bilangan numerik, sehingga perlu didefinisikan suatu fungsi yang memetakan anggota-anggota ruang sampel ke bilangan real. Fungsi yang memetakan anggota-anggota ruang sampel ke bilangan real disebut peubah acak.

Misalkan terdapat suatu percobaan acak memiliki ruang sampel $\mathcal{C}$. Suatu fungsi X yang memetakan setiap elemen $c \in \mathcal{C}$ ke tepat satu bilangan riil x, yaitu $X(c) = x$ disebut peubah acak. Ruang nilai dari X adalah himpunan bilangan $\mathcal{A} = \{x : x = X(c), c \in \mathcal{C}\}$.

Misalkan C adalah subset dari $\mathcal{C}$($C \subset \mathcal{C}$) berhubungan dengan A adalah subset dari $\mathcal{A}$ ($A \subset \mathcal{A}$), yaitu :

$$C = \{c : c \in \mathcal{C} \text{ dan } X(c) \in A\}$$

maka Pr ($X \in A$) = $P(C)$, dimana Pr ($X \in A$) menyatakan probabilitas dari kejadian A.

2.1.1 Peubah Acak Diskrit

Misalkan X adalah peubah acak dengan ruang sampel berdimensi satu $\mathscr{A}$, dimana $\mathscr{A}$ memuat bilangan yang berhingga, atau $\mathscr{A}$ memuat bilangan-bilangan tak berhingga yang dapat berkorespondensi satu-satu dengan bilangan bulat positif. Maka ruang sampel $\mathscr{A}$ disebut himpunan bilangan diskrit.

Misalkan suatu fungsi $f(x)$, dimana $f(x) > 0$ untuk $x \in \mathscr{A}$ dan memenuhi:

1. $\sum_{\mathscr{A}} f(x) = 1$
2. $P(A) = \Pr(X \in A) = \sum_{A} f(x)$

dengan $P(A)$ adalah fungsi himpunan probabilitas, dan $A \subset \mathscr{A}$. Maka X disebut peubah acak bertipe diskrit dan $f(x)$ disebut fungsi probabilitas dari X.

2.1.2 Peubah Acak Kontinu

Misalkan X adalah peubah acak dengan ruang nilai $\mathscr{A}$ yang merupakan ruang berdimensi satu , yang memuat suatu interval atau gabungan dari beberapa interval. Misalkan $f(x)$ adalah suatu fungsi non negatif, sehingga

1. $\int_{\mathscr{A}} f(x)dx = 1$
2. $P(A) = \Pr(X \in A) = \int_{A} f(x)dx$

dimana $P(A)$ adalah fungsi himpunan probabilitas, dengan $A \subset \mathscr{A}$. Maka X disebut sebagai peubah acak bertipe kontinu dengan $f(x)$ disebut sebagai fungsi probabilitas dari X.

2.2 Transformasi Peubah Acak Diskrit

Misalkan $X_1, X_2, \ldots, X_n$ adalah peubah acak diskrit yang memiliki fungsi probabilitas bersama yaitu $f(x_1, x_2, \ldots, x_n)$ dan $\mathscr{A}$ adalah himpunan diskrit yang merupakan ruang nilai dari $X_1, X_2, \ldots, X_n$ sedemikian sehingga $f(x_1, x_2, \ldots, x_n) > 0$ untuk setiap $(x_1, x_2, \ldots, x_n) \in \mathscr{A}$. Misalkan $y_1 = u_1(x_1, x_2, \ldots, x_n)$, $y_2 = u_2(x_1, x_2, \ldots, x_n)$, …, $y_n = u_n(x_1, x_2, \ldots, x_n)$ mendefinisikan suatu transformasi satu-satu yang memetakan $\mathscr{A}$ ke $\mathscr{B}$, dengan $\mathscr{B} = \{(y_1, y_2, \ldots, y_n) : y_1 = u_1(x_1, x_2, \ldots, x_n), y_2 = u_2(x_1, x_2, \ldots, x_n), \ldots, y_n = u_n(x_1, x_2, \ldots, x_n), (x_1, x_2, \ldots, x_n) \in \mathscr{A}\}$. Sehingga diperoleh transformasi invers $x_1 = w_1(y_1, y_2, \ldots, y_n)$, $x_2 = w_2(y_1, y_2, \ldots, y_n)$, …, $x_n = w_n(y_1, y_2, \ldots, y_n)$ yang memetakan $\mathscr{B}$ ke $\mathscr{A}$. Berdasarkan fungsi probabilitas bersama dari $X_1, X_2, \ldots, X_n$, didapat fungsi probabilitas bersama dari $Y_1, Y_2, \ldots, Y_n$, yaitu :

$$\begin{aligned} g(y_1, y_2, \ldots, y_n) &= f[w_1(y_1, y_2, \ldots, y_n), w_2(y_1, y_2, \ldots, y_n), \ldots w_n(y_1, y_2, \ldots, y_n)] \quad , y \in \mathscr{B} \\ &= 0 \quad , \text{lainnya} \end{aligned} \tag{2.1}$$

2.3 Distribusi Binomial

Suatu percobaan acak yang memiliki dua kemungkinan hasil disebut percobaan Bernoulli. Jika X adalah peubah acak yang berkaitan dengan percobaan Bernoulli, maka X dikatakan berdistribusi Bernoulli jika memiliki fungsi probabilitas sebagai berikut:

$$\begin{aligned} f(x) &= \theta^x (1-\theta)^{1-x} \quad , x = 0,1 \\ &= 0 \quad , x \text{ lain} \end{aligned} \tag{2.2}$$

dimana θ merupakan probabilitas terjadinya sukses pada percobaan Bernoulli.

Jika dilakukan n percobaan Bernoulli dibawah kondisi yang sama dengan setiap pengulangan yang saling bebas dan probabilitas sukses dari setiap percobaan konstan yaitu θ, dan X adalah peubah acak dimana X didefinisikan sebagai banyaknya sukses. Maka X dikatakan berdistribusi Binomial dengan fungsi probabilitas sebagai berikut:

$$f(x) = \binom{n}{x}\theta^x(1-\theta)^{n-x}, \; x = 0,1,2,...$$
$$= 0 \qquad\qquad , x \text{ lain} \tag{2.3}$$

Distribusi Binomial dengan parameter n dan p, dimana n dan p konstan biasa dinotasikan dengan b(n,p). *Moment Generating Function* (MGF) dari distribusi Binomial adalah

$$M(t) = \left[(1-p) + pe^t\right]^n \text{,untuk seluruh bilangan riil t.} \tag{2.4}$$

Dengan persamaan MGF tersebut, dapat ditentukan mean dan variansi dari distribusi Binomial.

μ = M'(0) = np

σ^2 = M''(0) – μ^2

= np(1 – p)

2.4 Distribusi Poisson

Peubah acak Y dikatakan berdistribusi Poisson jika memiliki fungsi probabilitas sebagai berikut :

$$f(y;\lambda) = \Pr(Y = y;\lambda)$$
$$= \frac{e^{-\lambda}\lambda^y}{y!} \quad , y = \; 0, 1, 2, ... \text{ dan } \lambda > 0 \tag{2.5}$$
$$= 0 \qquad , y \text{ lain}$$

dimana λ adalah suatu parameter yang menyatakan rata – rata banyaknya kejadian dalam selang atau interval tertentu. Distribusi Poisson dengan parameter λ biasa dinotasikan dengan P(λ).

Distribusi Poisson memiliki suatu karakteristik khusus yaitu nilai mean dan variansi yang sama. Nilai mean dan variansi dari peubah acak Y yang distribusi Poisson dengan parameter λ adalah

$$E(Y)=\sum_{y=0}^{\infty} y \cdot f(y,\lambda)$$
$$=\sum_{y=0}^{\infty} y \cdot \frac{\lambda^{y} e^{-\lambda}}{y!}$$
$$=\sum_{y=1}^{\infty} y \cdot \frac{\lambda^{y} e^{-\lambda}}{y!}$$
$$=\sum_{y=1}^{\infty} y \cdot \frac{\lambda^{y-1} \lambda e^{-\lambda}}{y.(y-1)!}$$
$$=\lambda \sum_{y=1}^{\infty} \frac{\lambda^{y-1} e^{-\lambda}}{(y-1)!}$$

Misalkan $x = y - 1$, maka akan didapat

$$\mathrm{E}(Y)=\lambda \sum_{x=o}^{\infty} \frac{\lambda^{x} e^{-\lambda}}{(x)!}$$
$$=\lambda \cdot \sum_{x=o}^{\infty} f(x,\lambda)$$
$$=\lambda \cdot 1=\lambda$$

Sehingga mean dari distribusi Poisson adalah parameternya yaitu λ. Selanjutnya akan dihitung nilai variansi dari distribusi Poisson, yaitu :

$$E(Y^{2})=\sum_{y=0}^{\infty} y^{2} \cdot f(y,\lambda)$$
$$=\sum_{y=0}^{\infty} y^{2} \cdot \frac{\lambda^{y} e^{-\lambda}}{y!}$$
$$=\sum_{y=1}^{\infty} y^{2} \cdot \frac{\lambda^{y} e^{-\lambda}}{y!}$$
$$=\sum_{y=1}^{\infty} y^{2} \cdot \frac{\lambda^{y-1} \lambda e^{-\lambda}}{y.(y-1)!}$$
$$=\lambda \sum_{y=1}^{\infty} \left((y-1)+1\right). \frac{\lambda^{y-1} e^{-\lambda}}{(y-1)!}$$
$$=\lambda \sum_{y=1}^{\infty} \left\{\left[(y-1). \frac{\lambda^{y-1} e^{-\lambda}}{(y-1)!}\right]+\left[\frac{\lambda^{y-1} e^{-\lambda}}{(y-1)!}\right]\right\}$$
$$=\lambda\left[\sum_{y=1}^{\infty}(y-1). \frac{\lambda^{y-2} \lambda e^{-\lambda}}{(y-1)(y-2)!}+\sum_{y=1}^{\infty} \frac{\lambda^{y-1} e^{-\lambda}}{(y-1)!}\right]$$

$$= \lambda \left[\lambda \sum_{y=1}^{\infty} \frac{\lambda^{y-2} e^{-\lambda}}{(y-2)!} + \sum_{y=1}^{\infty} \frac{\lambda^{y-1} e^{-\lambda}}{(y-1)!} \right]$$

Misalkan $x = y - 1$, dan $z = y - 2$, maka akan didapat

$$\mathrm{E}\left(Y^2\right) = \lambda \left[\lambda \sum_{y=1}^{\infty} \frac{\lambda^{y-2} e^{-\lambda}}{(y-2)!} + \sum_{y=1}^{\infty} \frac{\lambda^{y-1} e^{-\lambda}}{(y-1)!} \right]$$

$$= \lambda \left[\lambda \sum_{y=1}^{\infty} f(z,\lambda) + \sum_{y=1}^{\infty} f(x,\lambda) \right]$$

$$= \lambda [\lambda.1 + 1] = \lambda^2 + \lambda$$

$$\mathrm{Var}(Y) = E\left(Y^2\right) - \left[E(Y)\right]^2$$

$$= \lambda^2 + \lambda - \lambda^2$$

$$= \lambda$$

Sehingga variansi dari distribusi Poisson juga merupakan parameternya dan bernilai sama dengan mean yaitu λ.

Dalam penerapannya distribusi Poisson sering digunakan sebagai distribusi untuk menghitung jumlah kesalahan ketik per halaman dari suatu laporan, jumlah kecelakaan di suatu jalan tertentu, banyaknya *claim* pada perusahaan asuransi, dan jumlah mahkluk hidup di suatu wilayah tertentu.

2.5 Distribusi Gamma

Suatu peubah acak X dikatakan berdistribusi Gamma dengan dua parameter yaitu α dan β, jika memiliki bentuk fungsi distribusi probabilitas (pdf) sebagai berikut :

$$f(x) = \begin{cases} \dfrac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{-x/\beta} & , 0 < x < \infty \\ = 0 & , x \text{ lain} \end{cases} \tag{2.6}$$

dimana definisi dari fungsi Gamma dari α atau $\Gamma(\alpha)$ pada persamaan (2.6) yaitu

$$\Gamma(\alpha) = \int_0^{\infty} y^{\alpha-1} e^{-y} dy \tag{2.7}$$

untuk $\alpha > 0$ dan nilai dari integral tersebut adalah bilangan positif.

Untuk $\alpha = 1$, maka persamaan $\Gamma(\alpha)$ menjadi $\Gamma(1) = \int_0^\infty y^{1-1}e^{-t}dt = 1$. Sementara

untuk $\alpha > 1$, persamaan $\Gamma(\alpha)$ menjadi $\Gamma(\alpha) = \int_0^\infty y^{\alpha-1}e^{-t}dt = (\alpha-1)\Gamma(\alpha-1)$.

Untuk α bernilai integer, akan didapat persamaan $\Gamma(\alpha) = (\alpha-1)!$. Selain itu fungsi Gamma memiliki sifat lain, yaitu

$$\frac{\Gamma(\alpha+c)}{\Gamma(c)} = c(c+1)...(c+\alpha-1) \tag{2.8}$$

Distribusi Gamma dengan parameter α dan β biasa dinotasikan dengan $\Gamma(\alpha,\beta)$, dimana α merupakan *scale* parameter dan β merupakan *shape* parameter.

Moment Generating Function (MGF) dari distribusi Gamma adalah

$$M(t) = \frac{1}{(1-\beta t)^\alpha}, t < \frac{1}{\beta}, \text{untuk } t \in \mathbb{R}. \tag{2.9}$$

dengan persamaan MGF tersebut, dapat ditentukan mean dan variansi dari distribusi Gamma yaitu

μ = M'(0) = $\alpha\beta$

σ^2 = M''(0) – μ^2

= $\alpha\beta^2$

Pada aplikasinya, distribusi Gamma sering digunakan sebagai distribusi untuk memodelkan waktu tunggu (*waiting time*), pendapatan per tahun, dan besarnya *claim* pada perusahaan asuransi.

2.6 Distribusi Campuran (*Mixture Distribution*)

Misalkan X adalah peubah acak yang bergantung pada paremeter λ yang merupakan nilai dari suatu peubah acak Λ dengan fungsi probabilitas $u_\Lambda(\lambda)$. Fungsi probabilitas bersyarat dari X diberikan $\Lambda = \lambda$ adalah $f_{X|\Lambda}(x|\lambda)$. Distribusi Campuran (*Mixture Distribution*) didefinisikan sebagai berikut :

$$f_X(x) = \int_{\lambda \in \Lambda} f_{X|\Lambda}(x|\lambda) \cdot u_\Lambda(\lambda)\, d\lambda \tag{2.10}$$

dengan distribusi dari Λ disebut sebagai distribusi pencampur atau *Mixing Distribution* (Norman L. Johnson, Samuel Kotz, and Adrianne W. Kemp, 1992).

Pada distribusi Campuran (*Mixture Distribution*), nilai mean dan variansi dari peubah acak *X* dapat dinyatakan dengan

$$E(X) = E\left(E(X|\Lambda)\right) \tag{2.11}$$

$$\begin{aligned}
\mathrm{Var}(X) &= E\left(X^2\right) - \left[E(X)\right]^2 \\
&= E\left[E\left(X^2|\Lambda\right)\right] - \left\{E\left[E\left(X|\Lambda\right)\right]\right\}^2 \\
&= E\left\{\mathrm{Var}\left(X|\Lambda\right) + \left[E\left(X|\Lambda\right)\right]^2\right\} - \left\{E\left[E\left(X|\Lambda\right)\right]\right\}^2 \\
&= E\left[\mathrm{Var}\left(X|\Lambda\right)\right] + \mathrm{Var}\left[E\left(X|\Lambda\right)\right]
\end{aligned} \tag{2.12}$$

2.7 Distribusi Binomial Negatif

Distribusi Binomial Negatif merupakan distribusi yang memiliki banyak cara dalam penurunannya. Boswell dan Patil (1970) menunjukkan bahwa terdapat dua belas cara untuk mendapatkan distribusi Binomial Negatif, diantaranya dapat diturunkan sebagai distribusi Campuran Poisson-Gamma, sebagai distribusi *Compund* Poisson, sebagai barisan percobaan Bernoulli, sebagai model *Polya–Eggenberger urn*, atau sebagai invers dari distribusi Binomial. Penurunan klasik dari distribusi Binomial Negatif yang sering digunakan adalah distribusi Binomial Negatif sebagai barisan percobaan Bernoulli, yaitu jumlah percobaan Bernoulli yang dibutuhkan sampai terjadi *k* buah sukses, dengan setiap pengulangan yang saling bebas, dimana probabilitas sukses dari setiap percobaan konstan yaitu θ, dan probabilitas gagal yaitu $1 - \theta$. Jika peubah acak *X* menyatakan jumlah percobaan yang dibutuhkan sampai terjadi *k* buah sukses, maka distribusi probabilitas dari peubah acak *X*, disebut berdistribusi Binomial Negatif dengan fungsi probabilitas sebagai berikut :

$$\Pr\left(X=x\right)=\binom{x-1}{k-1}\theta^k(1-\theta)^{x-k} \quad , x \ = \ k,\ k+1,\ k+2,\ \ldots \quad \text{dan } 0\le\theta\le 1$$
$$=0 \qquad\qquad\qquad , x \text{ lain} \tag{2.13}$$

dimana Pr($X = x$) adalah probabilitas terjadi sukses ke – k pada percobaan ke x.

Distribusi probabilitas dari peubah acak X, dapat dinotasikan menjadi bentuk lain, dengan menggunakan transformasi $Y = X – k$, dimana Y menyatakan jumlah kegagalan sebelum terjadi k buah sukses. Distribusi probabilitas dari peubah acak Y dapat dinyatakan sebagai berikut :

$$\Pr\left(Y=y\right)=\binom{y+k-1}{k-1}\theta^k(1-\theta)^{y} \quad , y \ = \ 0,\ 1,\ 2,\ \ldots \quad \text{dan } 0\le\theta\le 1$$
$$=0 \qquad\qquad\qquad , y \text{ lain} \tag{2.14}$$

dimana Pr ($Y = y$) adalah probabilitas terjadi sukses ke – k setelah y kegagalan.

Pada persamaan (2.13) dan (2.14), k merupakan suatu bilangan bulat positif. Saat k merupakan suatu bilangan bulat positif, maka distribusi Binomial Negatif biasa disebut sebagai Distribusi Pascal (Norman L. Johnson, Samuel Kotz, and Adrianne W. Kemp, 1992).

Pada tahun 2008, Piet De Jong dan Gillian Z. H menjelaskan bahwa distribusi Binomial Negatif dapat didefinisikan untuk setiap nilai positif dari k dengan menggunakan fungsi Gamma sebagai pengganti dari faktorial atau kombinasi, yaitu :

$$\Pr\left(Y=y\right)=\frac{\Gamma\left(y+k\right)}{\Gamma\left(k\right)y!}\theta^k(1-\theta)^{y} \quad , y \ = \ 0,\ 1,\ 2,\ \ldots \quad \text{dan } 0\le\theta\le 1$$
$$=0 \qquad\qquad\qquad , y \text{ lain} \tag{2.15}$$

2.8 Kaidah Bayes

Metode Bayesian merupakan metode yang menggabungkan pengetahuan terdahulu (informasi *prior*) dari parameter yang akan ditaksir dengan informasi yang didapat dari sampel. Penggabungan dari kedua informasi tersebut akan menghasilkan informasi *posterior* yang selanjutnya akan digunakan dalam penaksiran parameter.

Misalkan ruang sampel $\Im$ dipartisi menjadi K kejadian yang *mutually exclusive* dan *exhaustive* $A_1, A_2, \ldots, A_K$, dimana $P(A_i) \neq 0$ untuk $i = 1, 2, \ldots, K$. Misalkan terdapat kejadian B didalam ruang sampel $\Im$ sehingga P(B) > 0 dan

$$B = (B \cap A_1) \cup (B \cap A_2) \cup \ldots \cup (B \cap A_K) \tag{2.16}$$

karena $(B \cap A_1)$, $(B \cap A_2)$, …, $(B \cap A_K)$ saling lepas, maka

$$P(B) = P(B \cap A_1) + P(B \cap A_2) + \ldots + P(B \cap A_K) \tag{2.17}$$

Kaidah Bayes :

Jika $A_1, A_2, \ldots, A_K$ merupakan kejadian *mutually exclusive* dan *exhaustive* dari ruang sampel $\Im$, dengan $P(A_i) \neq 0$ untuk $i = 1, 2, \ldots, K$ maka untuk sembarang kejadian B, dimana $P(B) \neq 0$,

$$P(A_i | B) = \frac{P(B | A_i) P(A_i)}{P(B | A_1) P(A_1) + P(B | A_2) P(A_2) + \ldots + P(B | A_K) P(A_K)} \tag{2.18}$$

Bukti :

$$P(B \cap A_i) = P(B | A_i) P(A_i) \qquad , i = 1, 2, \ldots, K$$
$$P(A_i \cap B) = P(A_i | B) P(B) \qquad , i = 1, 2, \ldots, K$$

karena $P(B \cap A_i) = P(A_i \cap B)$, maka

$$P(B | A_i) P(A_i) = P(A_i | B) P(B)$$

bagi kedua ruas dengan $P(B)$, didapat

$$P(A_i | B) = \frac{P(B | A_i) P(A_i)}{P(B)}$$

dimana

$$\begin{aligned} P(B) &= P(B \cap A_1) + P(B \cap A_2) + \ldots + P(B \cap A_K) \\ &= P(B | A_1) P(A_1) + P(B | A_2) P(A_2) + \ldots + P(B | A_K) P(A_K) \end{aligned}$$

sehingga persamaannya menjadi

$$P(A_i|B)=\frac{P(B|A_i)P(A_i)}{P(B|A_1)P(A_1)+P(B|A_2)P(A_2)+...+P(B|A_K)P(A_K)}$$ (terbukti)

Misalkan X adalah peubah acak yang memiliki distribusi probabilitas yang bergantung kepada $\boldsymbol{\theta}$, dimana $\boldsymbol{\theta}$ merupakan suatu peubah acak yang memiliki distribusi probabilitas tertentu. Jika $f(x|\theta)$ merupakan pdf bersyarat dari peubah acak X diberikan nilai $\boldsymbol{\theta} = \theta$ dan $h(\theta)$ merupakan pdf dari peubah acak $\boldsymbol{\theta}$, maka distribusi *posterior* dari $\boldsymbol{\theta}|x$ dapat ditentukan dengan menggunakan metode Bayes.

Misalkan $X_1, X_2, ..., X_n$ adalah sampel acak dari peubah acak X, maka pdf bersama dari $X_1, X_2, ..., X_n$ diberikan $\boldsymbol{\theta}$, yaitu :

$$f\left(x_1,x_2,...,x_n|\theta\right)=f(x_1|\theta)f(x_2|\theta)...f(x_n|\theta) \tag{2.19}$$

Pdf bersama $X_1, X_2, ..., X_n$, dan $\boldsymbol{\theta}$ adalah

$$f\left(x_1,x_2,...,x_n,\theta\right)=f\left(x_1,x_2,...,x_n|\theta\right)h(\theta) \tag{2.20}$$

Jika $\boldsymbol{\theta}$ merupakan peubah acak kontinu, maka pdf marginal bersama dari $X_1, X_2, ..., X_n$ yaitu:

$$\begin{aligned} f(x_1,x_2,...,x_n) &= \int_{-\infty}^{\infty} f(x_1,x_2,...,x_n,\theta)\, d\theta \\ &= \int_{-\infty}^{\infty} f\left(x_1,x_2,...,x_n|\theta\right)h(\theta)\, d\theta \end{aligned} \tag{2.21}$$

Jika $\boldsymbol{\theta}$ merupakan peubah acak diskrit, maka pdf marginal bersama dari $X_1, X_2, ..., X_n$ yaitu:

$$\begin{aligned} f(x_1,x_2,...,x_n) &= \sum_{\theta} f(x_1,x_2,...,x_n,\theta) \\ &= \sum_{\theta} f(x_1,x_2,...,x_n|\theta)\, h(\theta) \end{aligned} \tag{2.22}$$

Pdf bersyarat dari $\boldsymbol{\theta}$ diberikan $X_1 = x_1, X_2 = x_2, ..., X_n = x_n$ yaitu :

$$f(\theta | x_1, x_2, ..., x_n) = \frac{f(x_1, x_2, ..., x_n, \theta)}{f(x_1, x_2, ..., x_n)}$$
$$= \frac{f(x_1, x_2, ..., x_n | \theta)\; h(\theta)}{f(x_1, x_2, ..., x_n)} \qquad (2.23)$$

Persamaan (2.23) merupakan bentuk lain dari aturan Bayes. Dimana fungsi $f(\theta | x_1, x_2, ..., x_n)$ disebut pdf *posterior* dari $\boldsymbol{\theta}$, sementara fungsi $f(x_1, x_2, ..., x_n | \theta)$ disebut *likelihood* dari $X_1, X_2, ..., X_n$, dan fungsi $h(\theta)$ disebut sebagai pdf *prior* dari $\boldsymbol{\theta}$. Aturan Bayes sering ditulis sebagai berikut :

$$f(\theta | x_1, x_2, ..., x_n) \propto f(x_1, x_2, ..., x_n | \theta)\; h(\theta) \text{ atau} \qquad (2.24)$$

Posterior ∝ *likelihood* * *prior*

(Andrew G., John B.C, Hal S.S & Donald B.R, 2000).

2.9 *Prior Natural Conjugate*

Misalkan F adalah kelas dari distribusi sampling $f(y|\theta)$ dan P adalah kelas dari distribusi *prior* θ, maka kelas P disebut *conjugate* untuk kelas F jika fungsi probabilitas *posterior* $h(\theta|y)$ memiliki distribusi yang sama dengan fungsi probabilitas *prior* $h(\theta)$ untuk seluruh $f(y|\theta) \in F$.

Pemilihan *conjugate* yang lebih spesifik dapat dilakukan dengan memilih kelas distribusi *prior* P yang merupakan himpunan dari seluruh fungsi kepadatan yang memiliki bentuk fungsional yang sama dengan fungsi *likelihood* dari $f(y|\theta)$. Pada kasus ini P disebut *natural conjugate* dari F (Andrew G., John B.C, Hal S.S & Donald B.R, 2000).

Suatu kelas distribusi probabilitas *prior* $h(\theta)$ dikatakan *prior natural conjugate* dari kelas fungsi *likelihood* $f(y|\theta)$ jika kelas distribusi probabilitas *prior* $h(\theta)$ merupakan *conjugate* dari kelas fungsi *likelihood* dan fungsi probabilitas *prior* $h(\theta)$ memiliki bentuk fungsional yang sama dengan fungsi *likelihood* $f(y|\theta)$.

Berikut adalah tabel beberapa *prior natural conjugate* dari beberapa fungsi *likelihood*.

Tabel 2.1. ***Prior Natural Conjugate*** **dari Beberapa Fungsi** ***Likelihood***

Likelihood	***Prior Natural Conjugate***
Binomial	Beta
Multinomial	Dirichlet
Poisson	Gamma
Normal	
μ tidak diketahui, σ^2 diketahui	Normal
μ diketahui, σ^2 tidak diketahui	Inverse Chi-Square
Multivariat Normal	
$\boldsymbol{\mu}$ tidak diketahui, $\boldsymbol{V}$ diketahui	Multivariat Normal
$\boldsymbol{\mu}$ diketahui, $\boldsymbol{V}$ tidak diketahui	Inverse Wishart

[Sumber : Introduction to Bayesian Analysis, B. Walsh, 2002]

2.10 Metode Taksiran Titik

Misalkan X adalah peubah acak yang memiliki fungsi probabilitas (pdf) yang bergantung pada suatu parameter θ yang tidak diketahui, dimana θ merupakan anggota dari himpunan Ω, yang merupakan ruang parameter. Hal ini dapat dinotasikan dengan bentuk $f(x;\theta)$, $\theta \in \Omega$. Keluarga dari fungsi probabilitas dapat dinotasikan dengan symbol $\{f(x;\theta) : \theta \in \Omega\}$. Untuk setiap nilai θ, dimana $\theta \in \Omega$, berhubungan dengan satu anggota dari keluarga fungsi probabilitas.

Misalkan salah satu anggota dari keluarga distribusi probabilitas akan dipilih sebagai fungsi probabilitas dari peubah acak X, sehingga dibutuhkan taksiran titik dari θ.

Selanjutnya akan dipilih sampel acak $X_1, X_2, \ldots, X_n$ dari distribusi yang memiliki fungsi probabilitas yang merupakan salah satu dari anggota keluarga $\{f(x\,;\theta) : \theta \in \Omega\}$. Sehingga sampel berasal dari distribusi yang memiliki fungsi probabilitas $f(x\,;\theta) : \theta \in \Omega$. Permasalahannya adalah bagaimana mendefinisikan statistik $Y_1 = u_1(X_1, X_2, \ldots, X_n)$ sehingga jika $x_1, x_2, \ldots, x_n$ adalah nilai dari $X_1, X_2, \ldots, X_n$, maka nilai $y_1 = u_1(x_1, x_2, \ldots, x_n)$ akan menjadi taksiran titik yang baik untuk θ.

Terdapat beberapa metode yang dapat digunakan untuk mencari taksiran titik dari suatu parameter yang tidak diketahui pada fungsi probabilitas dari peubah acak X, diantaranya adalah metode *moment* dan metode maksimum *likelihood*.

2.10.1 Metode Maksimum *Likelihood*

Misalkan $X_1, X_2, \ldots, X_n$ adalah sampel acak dari distribusi yang memiliki distribusi probabilitas $f(x\,;\theta) : \theta \in \Omega$, dimana θ adalah parameter yang tidak diketahui dan Ω adalah ruang parameter. Fungsi *likelihood* didefinisikan sebagai pdf bersama dari $X_1, X_2, \ldots, X_n$ yang merupakan fungsi dari θ, dan biasa dinotasikan dengan $L(\theta)$.

$$\begin{aligned} L(\theta; x_1, x_2, \ldots, x_n) &= L(\theta) \\ &= f(x_1, x_2, \ldots, x_n; \theta) \\ &= f(x_1;\theta) f(x_2;\theta) \ldots f(x_n;\theta) \\ &= \prod_{i=1}^{n} f(x_i;\theta) \end{aligned}$$

Misalkan dapat ditemukan suatu fungsi nontrivial dari $x_1, x_2, \ldots, x_n$, misalkan disebut $u(x_1, x_2, \ldots, x_n)$ sedemikian sehingga ketika θ diganti dengan $u(x_1, x_2, \ldots, x_n)$, maka fungsi *likelihood* $L(\theta)$ akan bernilai maksimum. Maka statistik $u(X_1, X_2, \ldots, X_n)$ disebut penaksir maksimum *likelihood* dan dinotasikan dengan $\hat{\theta} = u(X_1, X_2, \ldots, X_n)$. Mencari nilai θ yang memaksimumkan fungsi *likelihood* ($L(\theta)$) akan memberikan nilai yang sama dengan mencari nilai θ yang memaksimumkan fungsi ln *likelihood* (ln $L(\theta) = l(\theta)$). Sehingga dalam mencari taksiran titik dengan metode maksimum *likelihood*, dapat diperoleh dengan menyelesaikan persamaan berikut :

$$\frac{\partial L(\theta)}{\partial \theta} = 0 \text{ atau } \frac{\partial l(\theta)}{\partial \theta} = 0$$

2.11 Sifat – Sifat Taksiran Titik

Setelah mendapatkan suatu taksiran titik dari parameter θ yang merupakan suatu statistik $u(X_1, X_2, \ldots, X_n)$, perlu diketahui apakah taksiran titik tersebut merupakan suatu taksiran yang cukup baik untuk menaksir parameter θ. Untuk mengetahui hal tersebut, perlu dilihat sifat-sifat dari taksiran titik dari parameter θ. Beberapa sifat dari taksiran titik yang biasa digunakan sebagai indikator apakah taksiran titik yang didapat sudah cukup baik antara lain sifat tak bias, konsisten, dan efisien.

2.11.1 Taksiran Tak Bias

Suatu penaksir dari θ, yaitu $\hat{\theta}$ dikatakan sebagai penaksir yang tak bias untuk θ, jika ekspektasi atau nilai harapan dari statistik $\hat{\theta}$ sama dengan parameter θ. Atau dapat dinotasikan sebagai berikut :

$$E(\hat{\theta}) = \theta \text{ , } \theta \in \Omega$$

2.11.2 Taksiran Konsisten

Suatu penaksir dari parameter θ, yaitu $\hat{\theta}$ dikatakan sebagai penaksir yang konsisten untuk θ jika $\hat{\theta}$ konvergen dalam probabilitas ke parameter θ.

Suatu barisan peubah acak $X_1, X_2, \ldots, X_n$ dikatakan konvergen dalam probabilitas ke peubah acak X, jika untuk setiap $\varepsilon > 0$, berlaku

$$\lim_{n\to\infty} \Pr\{|X_n - X| < \varepsilon\} = 1$$

atau ekuivalen dengan

$$\lim_{n\to\infty} \Pr\{|X_n - X| \geq \varepsilon\} = 0$$

Sehingga $\hat{\theta}$ dikatakan konvergen dalam probabilitas ke θ, jika untuk setiap $\varepsilon > 0$, berlaku

$$\lim_{n\to\infty} \Pr\left(\left|\hat{\theta} - \theta\right| < \varepsilon\right) = 1 \tag{2.25}$$

atau ekuivalen dengan

$$\lim_{n\to\infty} \Pr\left(\left|\hat{\theta} - \theta\right| \geq \varepsilon\right) = 0 \tag{2.26}$$

2.11.3 Taksiran Efisien

Misalkan X adalah peubah acak dengan fungsi probabilitas $f(x;\theta)$, $\theta \in \Omega$, dimana ruang parameter Ω merupakan suatu interval. Suatu penaksir tak bias dari parameter θ, yaitu $\hat{\theta}$ dikatakan sebagai penaksir yang efisien untuk θ, jika dan hanya jika variansi dari $\hat{\theta}$ mencapai batas bawah dari *Rao-Cramér*, atau dapat dinyatakan sebagai berikut :

$$\sigma_{\hat{\theta}}^2 \geq \frac{1}{nI(\theta)} \tag{2.27}$$

dimana $I(\theta)$ adalah *Fisher information*, yang dinyatakan dengan

$$I(\theta) = \int_{-\infty}^{\infty} \left[\frac{\partial \ln f(x;\theta)}{\partial \theta} \right]^2 f(x;\theta)\, dx \tag{2.28}$$

atau

$$I(\theta) = -\int_{-\infty}^{\infty} \frac{\partial^2 \ln f(x;\theta)}{\partial \theta^2} f(x;\theta)\, dx \tag{2.29}$$

2.12 Metode *Newton-Rhapson*

Metode Newton-Rhapson adalah metode yang digunakan untuk mencari akar-akar persamaan dari suatu fungsi *non-linear* ($f(x) = 0$), dengan menggunakan satu titik awal dan kemudian mendekatinya dengan memperhatikan *slope* atau kemiringan dari fngsi di titik tersebut. Selain untuk mencari akar-akar persamaan dari suatu fungsi, metode ini juga berguna untuk mencari nilai maksimum atau minimum dari suatu fungsi, dimana suatu fungsi $f(x)$ akan maksimum atau minimum jika $f\,'(x) = 0$. Titik pendekatan ke-n dari metode *Newton-Rhapson* ditulis dengan :

$$p_n = p_{n-1} - \frac{f(p_{n-1})}{f'(p_{n-1})}, n \geq 1 \tag{2.30}$$

Metode ini merupakan salah satu metode yang sering digunakan dan cukup baik karena dengan metode ini kekonvergenan lebih cepat tercapai.

Teorema Newton-Raphson

Asumsikan $f \in C^2[a,b]$, dimana $C^2[a,b]$ merupakan himpunan dari setiap fungsi yang memiliki turunan kedua yang kontinu pada interval tutup $[a, b]$. Misalkan terdapat suatu bilangan $p \in [a,b]$, dimana $f(p) = 0$. Jika $f\,'(p) \neq 0$, maka terdapat $\delta > 0$ sedemikian sehingga barisan $\{p_k\}_{k=0}^{\infty}$ didefinisikan dengan iterasi

$$p_{k+1} = g(p_k) = p_k - \frac{f(p_k)}{f'(p_k)}, \; k = 0,1,...$$

akan konvergen ke p untuk setiap aproksimasi awal $p_0 \in [p - \delta, p + \delta]$.

2.13 Uji Rasio *Likelihood*

Penggunaan besaran dari rasio dua pdf sebagai dasar dari pengujian terbaik atau UMPT (*Uniformly Most Powerfull* Test) dapat dimodifikasi dan memberikan metode untuk membentuk suatu pengujian dari suatu hipotesis majemuk terhadap hipotesis alternatif majemuk atau membentuk suatu pengujian dari suatu hipotesis sederhana terhadap hipotesis alternatif majemuk ketika UMPT tidak ada. Metode ini membawa kepada suatu pengujian yang disebut uji rasio *likelihood.* Uji rasio *likelihood* tidak perlu merupakan suatu UMPT, namun pengujian ini seringkali mempunyai sifat-sifat yang diinginkan.

Misalkan $X_1, X_2, ..., X_n$ menyatakan n peubah acak yang masing-masing mempunyai pdf $f_i(x_i; \theta_1, \theta_2, ..., \theta_m)$, dengan i = 1, 2, …, n. Himpunan yang terdiri dari semua titik parameter $(\theta_1, \theta_2, ..., \theta_m)$ dinotasikan dengan Ω yang biasa disebut ruang parameter. Misalkan ω adalah subset dari ruang parameter Ω. Misalkan ingin melakukan pengujian hipotesis (sederhana atau majemuk) dengan $H_0 : (\theta_1, \theta_2, ..., \theta_m) \in \omega$ terhadap semua hipotesis alternatif. Definisikan fungsi *likelihood* sebagai berikut :

$$L(\omega) = \prod_{i=1}^{n} f_i(x_i; \theta_1, \theta_2, ..., \theta_m) \quad , \quad (\theta_1, \theta_2, ..., \theta_m) \in \omega \tag{2.31}$$

dan

$$L(\Omega) = \prod_{i=1}^{n} f_i(x_i; \theta_1, \theta_2, ..., \theta_m) \quad , \quad (\theta_1, \theta_2, ..., \theta_m) \in \Omega \tag{2.32}$$

Dimana $L(\omega)$ merupakan fungsi *likelihood* dibawah asumsi H_0 benar dan $L(\Omega)$ merupakan fungsi *likelihood* dibawah asumsi H_1 benar. Misalkan $L(\widehat{\omega})$ dan $L(\widehat{\Omega})$ adalah fungsi *likelihood* maksimum yang diasumsikan ada untuk kedua fungsi *likelihood* tersebut. Rasio dari $L(\hat{\omega})$ terhadap $L(\hat{\Omega})$ disebut sebagai rasio *likelihood* dan dinotasikan dengan

$$\delta(x_1, x_2, ..., x_n) = \delta = \frac{L(\hat{\omega})}{L(\hat{\Omega})} \tag{2.33}$$

Misalkan δ_0 adalah suatu bilangan pecahan positif. Prinsip uji rasio *likelihood* menyatakan bahwa hipotesis $H_0 : (\theta_1, \theta_2, ..., \theta_m) \in \omega$ ditolak jika dan hanya jika $\delta(x_1, x_2, ..., x_n) = \delta = \dfrac{L(\hat{\omega})}{L(\hat{\Omega})} \leq \delta_0$.

Fungsi δ mendefinisikan suatu peubah acak $\delta(X_1, X_2, ..., X_n)$ dan tingkat signifikansi dari pengujian diberikan oleh

$$\alpha = \Pr\left[\delta(X_1, X_2, ... X_n) \leq \delta_0 ; H_0\right] \tag{2.34}$$

Teorema

Misalkan $X_1, X_2, ..., X_n$ adalah peubah acak yang memiliki distribusi identik dan saling bebas dengan pdf $f(x_i ; \boldsymbol{\theta})$, $\boldsymbol{\theta} \in \boldsymbol{\Omega}$, dimana $\boldsymbol{\Omega}$ adalah subset dari $\mathbb{R}^r$ dengan dimensi r, dan misalkan $\boldsymbol{\omega}$ adalah subset dari $\boldsymbol{\Omega}$ yang berdimensi m. Misalkan himpunan dimana pdf bernilai positif, tidak bergantung pada $\boldsymbol{\theta}$**.** Maka dibawah beberapa kondisi tambahan yang regular, distribusi *asymptotic* dari $A = -2 \log \delta$ adalah χ^2_{r-m}, untuk $\boldsymbol{\theta} \in \boldsymbol{\Omega}$, dan saat $n \rightarrow \infty$ (George G Roussas, 1997).

Bukti dari teorema diatas terdapat pada lampiran1.

BAB III
PENAKSIRAN PARAMETER PADA DISTRIBUSI BINOMIAL NEGATIF UNTUK MENGATASI OVERDISPERSI

Bab ini membahas tentang terjadinya overdispersi pada distribusi Poisson, kemudian distribusi Binomial Negatif pada kasus overdispersi untuk data *count*, distribusi Gamma sebagai *prior natural conjugate* dari distribusi Poisson, mean dan variansi dari distribusi Binomial Negatif, taksiran parameter dari distribusi Binomial Negatif , metode maksimum *likelihood*, serta pengujian parameter dispersi.

3.1 Overdispersi pada Distribusi Poisson

Misalkan *Y* merupakan suatu peubah acak yang berdistribusi Poisson dengan parameter λ. Pada distribusi Poisson terdapat beberapa asumsi yang harus dipenuhi, salah satunya yaitu asumsi kesamaan mean dan variansi dari peubah acak *Y* atau dinyatakan dengan,

$$E(Y) = Var(Y) = \lambda$$

keadaan ini disebut keadaan ekuidispersi.

Namun pada kenyataannya, dalam analisis statistik, sering dijumpai data *count* dengan variansi yang lebih besar dari mean, atau dinyatakan dengan,

$$Var(Y) > E(Y)$$

keadaan ini disebut keadaan overdispersi.

Pada distribusi Poisson, mean (λ) diasumsikan konstan atau dinotasikan dengan $E(Y)=\lambda$ untuk setiap nilai dari *Y*. Namun dalam kenyataannya suatu kejadian sering dipengaruhi oleh faktor lain yang tidak teramati yang menyebabkan mean dalam suatu populasi berbeda-beda antar observasi. Hal ini menunjukkan bahwa populasi tersebut heterogen, dan bergantung pada parameter

λ. Pelanggaran terhadap asumsi tersebut akan mengakibatkan taksiran parameter yang diperoleh dari distribusi Poisson menjadi tidak efisien walaupun tetap konsisten, dan akan memberikan taksiran yang tidak sesuai (Joseph M Hilbe, 2011). Sehingga diperlukan distribusi lain yang lebih cocok untuk menganalisis data tersebut.

3.2 Distribusi Binomial Negatif pada Kasus Overdispersi untuk Data *Count*

Misalkan *Y* adalah peubah acak yang bergantung pada paremeter λ yang merupakan nilai dari suatu peubah acak Λ dengan fungsi probabilitas $h(\lambda)$. $Y|\lambda$ merupakan peubah acak yang berdistribusi Poisson dengan fungsi probabilitas $f(y|\lambda)$. Misalkan $h(\lambda)$ merupakan fungsi probabilitas dari Λ yang berdistribusi Gamma dengan parameter α dan β, sehingga fungsi probabilitas dari Λ adalah

$$\begin{aligned} h(\lambda) &= \frac{1}{\Gamma(\alpha)(\beta)^{\alpha}} \lambda^{\alpha-1} \exp\left(\frac{-\lambda}{\beta}\right) \\ &= \frac{1}{\Gamma(\alpha)(\beta)^{\alpha}} \exp\left[(\alpha-1)\ln(\lambda)\right] \exp\left(\frac{-\lambda}{\beta}\right) \end{aligned} \tag{3.1}$$

Jika Λ berdistribusi Gamma dengan parameter α dan β, maka akan didapat nilai harapan serta variansi dari Λ adalah

E (Λ) = $\alpha\beta$

Var (Λ) = $\alpha\,\beta^2$

Pada umumnya, nilai mean dikenal dengan notasi μ, atau dapat ditulis sebagai

E (Λ) = $\alpha\beta = \mu$

sehingga didapat $\beta = \dfrac{\mu}{\alpha}$. Misalkan $\dfrac{1}{\alpha}$ dinotasikan dengan *a* atau $\dfrac{1}{\alpha} = a$, maka $\beta = \mu\, a$, dan $\alpha = \dfrac{1}{a}$. Dengan kata lain dapat dikatakan bahwa, Λ berdistribusi Gamma dengan parameter 1/*a* dan $\mu\, a$, dengan fungsi probabilitas sebagai berikut

$$h(\lambda)=\frac{1}{\Gamma\left(\frac{1}{a}\right)(\mu a)^{\frac{1}{a}}}\lambda^{\frac{1}{a}-1}\exp\left(\frac{-\lambda}{a\mu}\right) \tag{3.2}$$

Distribusi Gamma dipilih karena distribusi Gamma merupakan *prior natural conjugate* dari distribusi Poisson, hal ini akan ditunjukkan di subbab berikutnya. Fungsi probabilitas bersama antara $Y|\lambda$ dan Λ adalah

$$\begin{aligned} f(y|\lambda)h(\lambda) &= \frac{e^{-\lambda}\lambda^{y}}{y!}\cdot\frac{1}{\Gamma\left(\frac{1}{a}\right)(\mu a)^{\frac{1}{a}}}\lambda^{\frac{1}{a}-1}\exp\left(\frac{-\lambda}{a\mu}\right) \\ &= \frac{\lambda^{y+\frac{1}{a}-1}\exp\left(-\lambda+\frac{-\lambda}{a\mu}\right)}{\Gamma\left(\frac{1}{a}\right)(\mu a)^{\frac{1}{a}}\,y!} \end{aligned} \tag{3.3}$$

Maka menurut persamaa (2.10) fungsi probabilitas marginal dari Y adalah

$$f(y)=\int_{\lambda} f(y|\lambda)h(\lambda)d\lambda$$

$$\begin{aligned} \Pr(Y=y) &= \int_{0}^{\infty} f(y|\lambda)h(\lambda)d\lambda \\ &= \int_{0}^{\infty}\Pr(Y=y|\lambda)h(\lambda)d\lambda \\ &= \int_{0}^{\infty}\frac{\lambda^{y+\frac{1}{a}-1}\exp\left(-\lambda+\frac{-\lambda}{a\mu}\right)}{\Gamma\left(\frac{1}{a}\right)(\mu a)^{\frac{1}{a}}\,y!}d\lambda \\ &= \frac{1}{\Gamma\left(\frac{1}{a}\right)(\mu a)^{\frac{1}{a}}\,y!}\int_{0}^{\infty}\lambda^{y+\frac{1}{a}-1}\exp\left(-\lambda+\frac{-\lambda}{a\mu}\right)d\lambda \\ &= \frac{1}{y!\,\Gamma\left(\frac{1}{a}\right)(\mu a)^{\frac{1}{a}}}\int_{0}^{\infty}\exp\left(\frac{-\lambda(a\mu+1)}{a\mu}\right)\lambda^{y+\frac{1}{a}-1}d\lambda \end{aligned}$$

$$= \frac{1}{y!\Gamma\left(\frac{1}{a}\right)(\mu a)^{\frac{1}{a}}} \int_0^{\infty} \exp\left(\frac{-\lambda(a\mu+1)}{a\mu}\right) \lambda^{y+\frac{1}{a}-1} \left(\frac{a\mu+1}{a\mu+1}\right)^{y+\frac{1}{a}-1} \left(\frac{a\mu}{a\mu}\right)^{y+\frac{1}{a}-1} d\lambda$$

$$= \frac{1}{y!\Gamma\left(\frac{1}{a}\right)(\mu a)^{\frac{1}{a}}} \int_0^{\infty} \exp\left(\frac{-\lambda(a\mu+1)}{a\mu}\right) \left(\frac{\lambda(a\mu+1)}{a\mu}\right)^{y+\frac{1}{a}-1} \left(\frac{a\mu}{a\mu+1}\right)^{y+\frac{1}{a}-1} d\lambda$$

$$\Pr(Y=y) = \frac{1}{y!\Gamma\left(\frac{1}{a}\right)(\mu a)^{\frac{1}{a}}} \left(\frac{a\mu}{a\mu+1}\right)^{y+\frac{1}{a}-1} \int_{-\infty}^{\infty} \exp\left(\frac{-\lambda(a\mu+1)}{a\mu}\right) \left(\frac{\lambda(a\mu+1)}{a\mu}\right)^{y+\frac{1}{a}-1} d\lambda$$

misalkan $t = \frac{\lambda(a\mu+1)}{a\mu}$, maka $\frac{dt}{d\lambda} = \frac{a\mu+1}{a\mu}$ sehingga $d\lambda = \frac{a\mu}{a\mu+1} dt$.

Sehingga persamaan Pr(*Y* = *y*) menjadi

$$\Pr(Y=y) = \frac{1}{y!\Gamma\left(\frac{1}{a}\right)(\mu a)^{\frac{1}{a}}} \left(\frac{a\mu}{a\mu+1}\right)^{y+\frac{1}{a}-1} \int_0^{\infty} \exp(-t) t^{y+\frac{1}{a}-1} \left(\frac{a\mu}{a\mu+1}\right) dt$$

$$= \frac{1}{y!\Gamma\left(\frac{1}{a}\right)(\mu a)^{\frac{1}{a}}} \left(\frac{a\mu}{a\mu+1}\right)^{y+\frac{1}{a}-1} \left(\frac{a\mu}{a\mu+1}\right) \int_0^{\infty} \exp(-t) t^{y+\frac{1}{a}-1} dt$$

$$= \frac{1}{y!\Gamma\left(\frac{1}{a}\right)(\mu a)^{\frac{1}{a}}} \left(\frac{a\mu}{a\mu+1}\right)^{y+\frac{1}{a}} \int_0^{\infty} \exp(-t) t^{y+\frac{1}{a}-1} dt$$

$$= \frac{1}{y!\Gamma\left(\frac{1}{a}\right)(\mu a)^{\frac{1}{a}}} \left(\frac{a\mu}{a\mu+1}\right)^{y+\frac{1}{a}} \Gamma\left(y+\frac{1}{a}\right)$$

$$= \frac{1}{y!\Gamma\left(\frac{1}{a}\right)(\mu a)^{\frac{1}{a}}} \left(\frac{a\mu}{a\mu+1}\right)^{y} \left(\frac{a\mu}{a\mu+1}\right)^{\frac{1}{a}} \Gamma\left(y+\frac{1}{a}\right)$$

$$= \frac{\Gamma\left(y+\frac{1}{a}\right)}{\Gamma\left(\frac{1}{a}\right)y!}\left(\frac{a\mu}{a\mu+1}\right)^{y}\left(\frac{1}{a\mu+1}\right)^{\frac{1}{a}}$$

$$= \frac{\Gamma\left(y+\frac{1}{a}\right)}{\Gamma\left(\frac{1}{a}\right)y!}\left(\frac{\mu}{\mu+\frac{1}{a}}\right)^{y}\left(\frac{\frac{1}{a}}{\mu+\frac{1}{a}}\right)^{\frac{1}{a}} \tag{3.4}$$

Persamaan (3.4) diatas memiliki bentuk yang similar dengan bentuk fungsi probabilitas dari distribusi Binomial Negatif pada persamaan (2.14) dengan parameter k = 1/a dan $\theta = \frac{1}{a\mu+1}$, dengan nilai k = 1/a pada persamaan diatas merupakan suatu bilangan positif. Penurunan distribusi Binomial Negatif diatas tidak berhubungan dengan penurunan klasik dari distribusi Binomial Negatif sebagai barisan dari percobaan Bernoulli yang dijelaskan pada subbab 2.7 (Piet De Jong & Gillian Z Heller, 2008)). Bukti bahwa persamaan (3.4) merupakan suatu fungsi probabilitas dapat dilihat pada lampiran 4.

Distribusi Binomial Negatif dengan fungsi probabilitas seperti persamaan (2.14) mengasumsikan bahwa k adalah suatu bilangan bulat positif, dimana definisi peubah acak dari distribusi Binomial Negatif adalah jumlah kegagalan yang terjadi sebelum k buah sukses. Distribusi Binomial Negatif dengan fungsi probabilitas pada persamaan (3.4) disebut sebagai distribusi Campuran Poisson-Gamma (*Mixture* Poisson-Gamma *Distribution*). Pada distribusi Binomial Negatif sebagai distribusi Campuran Poisson-Gamma, nilai dari parameter k = 1/a bukan merupakan suatu bilangan bulat positif, melainkan suatu bilangan real positif. Sehingga definisi peubah acak dari distribusi Binomial Negatif sebelumnya tidak dapat digunakan (Joseph M Hilbe, 2011). Definisi peubah acak yang digunakan pada distribusi Binomial Negatif sebagai distribusi Campuran Poisson-Gamma adalah jumlah kejadian dengan dispersi pada data sebesar 1/a (Norman L. Johnson, Samuel Kotz, and Adrianne W. Kemp, 1992). Definisi tersebut memiliki kesamaan dengan definisi peubah acak pada distribusi Poisson, namun pada distribusi Binomial Negatif terdapat parameter tambahan yang menjelaskan mengenai dispersi dari data. Distribusi Binomial Negatif dengan fungsi probabilitas pada persamaan (3.4) memiliki dua parameter yaitu μ dan *a*, dimana μ merupakan parameter yang menjelaskan tentang mean dan *a* merupakan parameter dispersi.

Berdasarkan uraian diatas dapat disimpulkan bahwa distribusi Binomial Negatif merupakan salah satu distribusi alternatif yang dapat digunakan untuk menganalisis data *count* saat terjadi overdispersi, karena pada distribusi ini terdapat parameter yang digunakan untuk mengamati dispersi yang terjadi pada data.

3.3 Distribusi Gamma sebagai *Prior* Natural Conjugate dari Distribusi Poisson

Metode Bayesian merupakan metode yang digunakan untuk menggabungkan pengetahuan terdahulu (*prior*) dari parameter yang akan ditaksir dengan informasi yang didapat dari data sampel. Pada kondisi overdispersi, paremeter λ dari distribusi Poisson tidak lagi konstan atau bervariasi antar observasi pada populasi dan merupakan nilai dari suatu peubah acak Λ yang memiliki suatu distribusi tertentu. Distribusi dari Λ tersebut disebut distribusi *prior*. Sedangkan distribusi dari $\Lambda|Y$ setelah diketahui informasi dari *prior* dan dilakukan metode Bayes disebut distribusi *posterior*.

$Y|\lambda$ adalah suatu peubah acak yang bergantung pada parameter λ dari suatu populasi yang berdistribusi Poisson dan dimana *Y* didefinisikan sebagai jumlah kejadian pada suatu populasi dalam suatu interval waktu atau area tertentu. Telah dijelaskan sebelumnya bahwa saat terjadi overdispersi pada distribusi Poisson, λ tidak lagi konstan atau bervariasi antar observasi. Hal itu menunjukkan bahwa λ merupakan nilai dari suatu peubah acak Λ yang memiliki suatu distribusi tertentu. Misalkan dipilih distribusi Gamma sebagai distribusi dari Λ, dengan parameter 1/*a* dan $\mu\, a$ atau dapat ditulis $\Lambda \sim \Gamma(1/a, \mu a)$. Alasan pemilihan tersebut adalah karena distribusi Gamma merupakan *prior natural conjugate* dari distribusi Poisson dan terdapat kesamaan bentuk fungsional antara pdf dari distribusi Poisson, yaitu

$$f(y) = \frac{e^{-\lambda}\lambda^{y}}{y!}, \ \lambda > 0, \ y = 0,1,...$$
$$= 0 \qquad , \text{lainnya}$$

dan pdf dari distribusi Gamma pada persamaan (3.2), yaitu

$$h(\lambda) = \frac{1}{\Gamma\left(\frac{1}{a}\right)(\mu a)^{\frac{1}{a}}} \lambda^{\frac{1}{a}-1} \exp\left(\frac{-\lambda}{a\mu}\right), \ 0 < \lambda < \infty$$
$$= 0 \qquad , \text{lainnya}$$

Sebelumnya pada subbab 2.9 telah dijelaskan bahwa suatu kelas distribusi probabilitas *prior* $h(\theta)$ dikatakan *prior natural conjugate* dengan kelas fungsi *likelihood* $f(y|\theta)$ jika kelas distribusi probabilitas *prior* $h(\theta)$ merupakan *conjugate* dari kelas fungsi *likelihood* , dan fungsi probabilitas *prior* $h(\theta)$ memiliki bentuk fungsional yang sama dengan fungsi *likelihood* $f(y|\theta)$. Selanjutnya akan dibuktikan bahwa distribusi Gamma merupakan *prior natural conjugate* dari distribusi Poisson.

Misalkan $Y|\lambda$ adalah peubah acak yang bergantung pada parameter λ yang berdistribusi Poisson, dimana distribusi Poisson yang merupakan anggota dari keluarga eksponensial (bukti terdapat pada lampiran 3), sehingga fungsi probabilitasnya adalah

$$f(y|\lambda) = \frac{e^{-\lambda}\lambda^{y}}{y!}, \ \lambda > 0, \ y = 0,1,...$$

Selanjutnya, misalkan distribusi *prior*nya adalah distribusi Gamma dengan parameter $1/a$ dan μ a, sehingga fungsi probabilitas dari Λ adalah persamaan (3.2), yaitu

$$h(\lambda) = \frac{1}{\Gamma\left(\frac{1}{a}\right)(\mu a)^{\frac{1}{a}}} \lambda^{\frac{1}{a}-1} \exp\left(\frac{-\lambda}{a\mu}\right)$$

Pada subbab sebelumnya telah diketahui bahwa persamaan (3.4) merupakan fungsi probabilitas marginal dari peubah acak Y, yaitu

$$f(y)=\frac{\Gamma\left(y+\frac{1}{a}\right)}{\Gamma\left(\frac{1}{a}\right)y!}\left(\frac{\mu}{\mu+\frac{1}{a}}\right)^{y}\left(\frac{\frac{1}{a}}{\mu+\frac{1}{a}}\right)^{\frac{1}{a}}$$

Selanjutnya dengan menggunakan persamaan (2.23) akan didapat fungsi probabilitas dari distribusi *posterior*nya, yaitu :

$$h(\lambda|y)=\frac{p(y|\lambda)\,p(\lambda)}{p(y)}$$

$$=\frac{\frac{e^{-\lambda}\lambda^{y}}{y!}\cdot\frac{1}{\Gamma\left(\frac{1}{a}\right)(\mu a)^{\frac{1}{a}}}\lambda^{\frac{1}{a}-1}\exp\left(\frac{-\lambda}{a\mu}\right)}{\frac{\Gamma\left(y+\frac{1}{a}\right)}{\Gamma\left(\frac{1}{a}\right)y!}\left(\frac{\mu}{\mu+\frac{1}{a}}\right)^{y}\left(\frac{\frac{1}{a}}{\mu+\frac{1}{a}}\right)^{\frac{1}{a}}}$$

$$=\frac{\exp\left(-\lambda\left(1+\frac{1}{a\mu}\right)\right)\lambda^{y+\frac{1}{a}-1}}{\Gamma\left(y+\frac{1}{a}\right)\left(\frac{\mu}{\mu+\frac{1}{a}}\right)^{y}(\mu a)^{\frac{1}{a}}\left(\frac{\frac{1}{a}}{\mu+\frac{1}{a}}\right)^{\frac{1}{a}}}$$

$$=\frac{\exp\left(-\lambda\left(\frac{a\mu+1}{a\mu}\right)\right)\lambda^{y+\frac{1}{a}-1}}{\Gamma\left(y+\frac{1}{a}\right)\left(\frac{a\mu}{a\mu+1}\right)^{y+\frac{1}{a}}} \qquad (3.5)$$

Fungsi probabilitas posterior yang didapat pada persamaan (3.5) diatas merupakan fungsi probabilitas dari distribusi Gamma dengan parameter yaitu $y+\frac{1}{a}$ dan $\frac{a\mu}{a\mu+1}$. Distribusi *posterior* yang dihasilkan ternyata juga berdistribusi Gamma seperti distribusi *priornya*, sehingga terbukti bahwa distribusi Gamma merupakan *prior natural conjugate* dari distribusi Poisson.

3.4 Mean dan Variansi dari Distribusi Binomial Negatif

Distribusi Binomial Negatif dengan dua parameter, yaitu μ dan a, memilik fungsi probabilitas sebagai berikut:

$$\Pr(Y=y)=\frac{\Gamma\left(y+a^{-1}\right)}{\Gamma\left(a^{-1}\right)y!}\left(\frac{\mu}{\mu+a^{-1}}\right)^{y}\left(\frac{a^{-1}}{\mu+a^{-1}}\right)^{a^{-1}}\ ,\ \mu>0,\ a^{-1}>0\ ,\ y=0,1,2,...$$

$$=\binom{y+a^{-1}-1}{a^{-1}-1}\left(\frac{\mu}{\mu+a^{-1}}\right)^{y}\left(\frac{a^{-1}}{\mu+a^{-1}}\right)^{a^{-1}}\ ,\ \mu>0\ ,\ a^{-1}>0\ ,\ y=0,1,2,...$$

kombinasi $\binom{y+\frac{1}{a}-1}{\frac{1}{a}-1}$ pada persamaan diatas dapat dijabarkan menjadi

$$\binom{y+\frac{1}{a}-1}{\frac{1}{a}-1}=\frac{\left(y+\frac{1}{a}-1\right)!}{\left(\frac{1}{a}-1\right)!\left(y+\frac{1}{a}-1-\frac{1}{a}+1\right)!}$$

$$=\frac{\left(y+\frac{1}{a}-1\right)!}{\left(\frac{1}{a}-1\right)!(y)!}$$

$$=\frac{\left(y+\frac{1}{a}-1\right)\left(y+\frac{1}{a}-2\right)...\left(\frac{1}{a}+2\right)\left(\frac{1}{a}+1\right)\left(\frac{1}{a}\right)}{(y)!}$$

Kemudian perhatikan kombinasi

$$\frac{\left(y+\frac{1}{a}-1\right)...\left(\frac{1}{a}+1\right)\left(\frac{1}{a}\right)}{(y)!}=\frac{(-1)^{n}\left(-\frac{1}{a}\right)\left(-\frac{1}{a}-1\right)...\left(-\frac{1}{a}-y+1\right)}{(y)!}$$

$$=(-1)^{n}\binom{-\frac{1}{a}}{y}$$

Sehingga didapat persamaan

$$\binom{y_i + \frac{1}{a} - 1}{\frac{1}{a} - 1} = (-1)^{y_i} \binom{-\frac{1}{a}}{y_i} \tag{3.6}$$

dimana jika persamaan (3.6) di substitusikan ke persamaan (3.4), akan didapat fungsi probabilitas sebagai berikut :

$$\begin{aligned}
\Pr(Y = y) &= \binom{y + a^{-1} - 1}{a^{-1} - 1} \left(\frac{\mu}{\mu + a^{-1}}\right)^{y} \left(\frac{a^{-1}}{\mu + a^{-1}}\right)^{a^{-1}} \ , \ \mu > 0 \ , \ a^{-1} > 0 \ , \ y = 0,1,2,\ldots \\
&= (-1)^{y} \binom{-a^{-1}}{y} \left(\frac{\mu}{\mu + a^{-1}}\right)^{y} \left(\frac{a^{-1}}{\mu + a^{-1}}\right)^{a^{-1}} \ , \ \mu > 0 \ , \ a^{-1} > 0 \ , \ y = 0,1,2,\ldots \\
&= \binom{-a^{-1}}{y} \left(\frac{-\mu}{\mu + a^{-1}}\right)^{y} \left(\frac{a^{-1}}{\mu + a^{-1}}\right)^{a^{-1}} \ , \ \mu > 0 \ , \ a^{-1} > 0 \ , \ y = 0,1,2,\ldots
\end{aligned} \tag{3.7}$$

Selanjutnya akan dilihat bagaimana sifat-sifat dari distribusi Binomial Negatif antara lain bagaimana bentuk *moment generating function* (MGF), mean serta variansinya. Pertama, akan dilihat bentuk dari *moment generating function* (MGF) dari distribusi Binomial Negatif. Karena *Y* berdistribusi Binomial Negatif adalah peubah acak diskrit, maka *moment generating function* (MGF) dapat dinyatakan dengan

$$\begin{aligned}
M(t) &= E\left(e^{tY}\right) \\
&= \sum_{y} e^{ty} f(y)
\end{aligned}$$

Sehingga *moment generating function* (MGF) dari distribusi Binomial Negatif adalah

$$\begin{aligned}
M(t) &= \sum_{y} e^{ty} f(y) \\
&= \sum_{y=0}^{\infty} e^{ty} \binom{-a^{-1}}{y} \left(\frac{-\mu}{\mu + a^{-1}}\right)^{y} \left(\frac{a^{-1}}{\mu + a^{-1}}\right)^{a^{-1}} \\
&= \sum_{y=0}^{\infty} \binom{-a^{-1}}{y} \left(\frac{-\mu \cdot e^{t}}{\mu + a^{-1}}\right)^{y} \left(\frac{a^{-1}}{\mu + a^{-1}}\right)^{a^{-1}} \\
&= \left(\frac{a^{-1}}{\mu + a^{-1}}\right)^{a^{-1}} \sum_{y=0}^{\infty} \binom{-a^{-1}}{y} \left(\frac{-\mu \cdot e^{t}}{\mu + a^{-1}}\right)^{y}
\end{aligned}$$

$$\mathrm{M}(t)=\left(\frac{a^{-1}}{\mu+a^{-1}}\right)^{a^{-1}}\left(1-e^{t}\left(\frac{\mu}{\mu+a^{-1}}\right)\right)^{-a^{-1}} \tag{3.8}$$

dimana $t<-\ln\left(\frac{\mu}{\mu+a^{-1}}\right),\ \mu>0,\ a^{-1}>0$.

Setelah mengetahui bentuk dari *moment generating function* (MGF), akan dicari mean serta variansi dari distribusi Binomial Negatif, dimana

$$\text{Mean}=E(Y)=M'(0)$$

$$\begin{aligned}\text{Variansi}&=E(Y^2)-\left[E(Y)\right]^2\\&=M''(0)-\left[E(Y)\right]^2\end{aligned}$$

Selanjutnya akan dicari turunan pertama dan turunan kedua dari *moment generating function* (MGF), yang akan digunakan dalam menentukan mean serta variansi dari distribusi Binomial Negatif.

$$\begin{aligned}M'(t)&=\frac{dM(t)}{dt}\\&=\left(\frac{a^{-1}}{\mu+a^{-1}}\right)^{a^{-1}}\left(1-\left(\frac{\mu}{\mu+a^{-1}}\right)e^{t}\right)^{-a^{-1}-1}\left(\left(\frac{\mu a^{-1}}{\mu+a^{-1}}\right)e^{t}\right)\end{aligned}$$

(3.9)

$$\begin{aligned}M''(t)&=\frac{d^2M(t)}{dt^2}\\&=\left[\left(\frac{a^{-1}}{\mu+a^{-1}}\right)^{a^{-1}}\left(1-\left(\frac{\mu}{\mu+a^{-1}}\right)e^{t}\right)^{-a^{-1}-1}\left(\left(\frac{\mu a^{-1}}{\mu+a^{-1}}\right)e^{t}\right)\right]+\\&\left[\left(\frac{a^{-1}}{\mu+a^{-1}}\right)^{a^{-1}}\left(\left(\frac{\mu a^{-1}}{\mu+a^{-1}}\right)e^{t}\right)\left(-a^{-1}-1\right)\left(1-\left(\frac{\mu}{\mu+a^{-1}}\right)e^{t}\right)^{-a^{-1}-2}\left(\left(\frac{-\mu}{\mu+a^{-1}}\right)e^{t}\right)\right]\end{aligned} \tag{3.10}$$

Sehingga didapat mean dan variansinya adalah

$$\text{Mean} = M'(0)$$

$$= \left(\frac{a^{-1}}{\mu+a^{-1}}\right)^{a^{-1}} \left(1-\left(\frac{\mu}{\mu+a^{-1}}\right)\right)^{-a^{-1}-1} \left(\frac{\mu a^{-1}}{\mu+a^{-1}}\right)$$

$$= \left(\frac{a^{-1}}{\mu+a^{-1}}\right)^{a^{-1}} \left(\frac{\mu+a^{-1}}{a^{-1}}\right)^{a^{-1}} \left(\frac{\mu+a^{-1}}{a^{-1}}\right)\left(\frac{\mu a^{-1}}{\mu+a^{-1}}\right)$$

$$= \mu \tag{3.11}$$

$$\text{Variansi} = M''(0) - \mu^2$$

$$= \left[\left(\frac{a^{-1}}{\mu+a^{-1}}\right)^{a^{-1}} \left(1-\left(\frac{\mu}{\mu+a^{-1}}\right)\right)^{-a^{-1}-1} \left(\frac{\mu a^{-1}}{\mu+a^{-1}}\right)\right] +$$

$$\left[\left(\frac{a^{-1}}{\mu+a^{-1}}\right)^{a^{-1}} \left(\frac{\mu a^{-1}}{\mu+a^{-1}}\right)\left(-a^{-1}-1\right)\left(1-\left(\frac{\mu}{\mu+a^{-1}}\right)\right)^{-a^{-1}-2} \left(\frac{-\mu}{\mu+a^{-1}}\right)\right] - \mu^2$$

$$= \left(\frac{a^{-1}}{\mu+a^{-1}}\right)^{a^{-1}} \left(1-\left(\frac{\mu}{\mu+a^{-1}}\right)\right)^{-a^{-1}-1} \left(\frac{\mu a^{-1}}{\mu+a^{-1}}\right)$$

$$\left[1+\left(-a^{-1}-1\right)\left(1-\left(\frac{\mu}{\mu+a^{-1}}\right)\right)^{-1} \left(\frac{-\mu}{\mu+a^{-1}}\right)\right] - \mu^2$$

$$= \left(\frac{a^{-1}}{\mu+a^{-1}}\right)^{a^{-1}} \left(\frac{\mu+a^{-1}}{a^{-1}}\right)^{a^{-1}} \left(\frac{\mu+a^{-1}}{a^{-1}}\right)\left(\frac{\mu a^{-1}}{\mu+a^{-1}}\right)$$

$$\left[1+\left(-a^{-1}-1\right)\left(\frac{\mu+a^{-1}}{a^{-1}}\right)\left(\frac{-\mu}{\mu+a^{-1}}\right)\right] - \mu^2$$

$$= \mu\left[1+\left(-a^{-1}-1\right)\left(\frac{-\mu}{a^{-1}}\right)\right] - \mu^2$$

$$= \mu\left[1+\mu+\frac{\mu}{a^{-1}}\right] - \mu^2$$

$$= \mu + \frac{\mu^2}{a^{-1}}$$

$$\text{Variansi} = \mu + a\mu^2 \tag{3.12}$$

Dari persamaan (3.11) dan (3.12) didapat bahwa mean dari distribusi Binomial Negatif adalah μ, dan variansinya adalah $\mu + a\mu^2$. Dengan nilai a > 0, maka dapat disimpulkan bahwa variansi dari distribusi Binomial Negatif lebih besar dari nilai meannya. Nilai mean dan variansi dari distribusi Binomial Negatif dapat dinyatakan sebagai

$$E(Y) = \mu$$
$$= \frac{1}{a} \cdot a\mu$$
$$= E(\lambda) \tag{3.13}$$

$$\text{var}(Y) = \mu + a\mu^2$$
$$= \frac{1}{a} \cdot a\mu + \frac{1}{a} \cdot (a\mu)^2$$
$$= E(\lambda) + \text{var}(\lambda) \tag{3.14}$$

Persamaan (3.13) dan (3.14) diatas sesuai dengan mean dan variansi dari distribusi campuran (*Mixture Distribution*) seperti pada persamaan (2.11) dan (2.12).

3.5 Taksiran Parameter dari distribusi Binomial Negatif

Pada distribusi Binomial Negatif terdapat dua parameter yang nilainya tidak diketahui yaitu μ dan *a*, sehingga kedua parameter ini akan ditaksir. Terdapat beberapa metode penaksiran yang dapat digunakan, namun yang akan dibahas adalah metode maksimum *likelihood*.

3.5.1 Metode Maksimum *Likelihood*

Misalkan $Y_1, Y_2, \ldots, Y_n$ adalah sampel acak dari distribusi Binomial Negatif yang memiliki fungsi probabilitas, $f\,(y_i\,;\ \mu\,,\,a)$ dimana,

$$f(y_i;\mu,a) = \frac{\Gamma\left(y_i + a^{-1}\right)}{\Gamma\left(a^{-1}\right) y_i\,!}\left(\frac{\mu}{\mu + a^{-1}}\right)^{y_i}\left(\frac{a^{-1}}{\mu + a^{-1}}\right)^{a^{-1}}, \quad \mu > 0,\ a^{-1} > 0\ ,\ y_i = 0,1,2,\ldots$$

dengan μ dan *a* merupakan paremeter yang nilainya tidak diketahui. Sehingga nilai kedua parameter tersebut dapat ditaksir. Pada metode maksimum *likelihood* akan digunakan persamaan *likelihood* dalam mencari taksiran dari parameter μ dan a. Persamaan *likelihood* didefinisikan sebagai pdf bersama dari $Y_1, Y_2, \ldots, Y_n$ dan biasa dinotasikan dengan $L(a, \mu)$. Persamaan *likelihood* dari distribusi Binomial Negatif yaitu :

$$L(a,\mu)=\prod_{i=1}^{n}\frac{\Gamma\left(y_i+a^{-1}\right)}{y_i!\,\Gamma\left(a^{-1}\right)}\left(\frac{1}{1+a\mu}\right)^{a^{-1}}\left(\frac{a\mu}{1+a\mu}\right)^{y_i}$$

$$=\prod_{i=1}^{n}\frac{\Gamma\left(y_i+\left(\frac{1}{a}\right)\right)}{y_i!\,\Gamma\left(\left(\frac{1}{a}\right)\right)}\left(\frac{1}{1+a\mu}\right)^{\left(\frac{1}{a}\right)}\left(\frac{a\mu}{1+a\mu}\right)^{y_i}$$

$$=\prod_{i=1}^{n}\frac{\frac{1}{a}\left(\frac{1+a}{a}\right)\ldots\left(\frac{1+ay_i-a}{a}\right)}{y_i!}\left(\frac{1}{1+a\mu}\right)^{\left(\frac{1}{a}\right)}\left(\frac{a\mu}{1+a\mu}\right)^{y_i}$$

$$L(a,\mu)=\prod_{i=1}^{n}\frac{\frac{1}{a^{y_i}}(1+a)\ldots(1+ay_i-a)}{y_i!}\left(\frac{1}{1+a\mu}\right)^{\left(\frac{1}{a}\right)}\left(\frac{a\mu}{1+a\mu}\right)^{y_i}$$

$$=\prod_{i=1}^{n}\frac{(1+a)\ldots(1+ay_i-a)}{a^{y_i}.y_i!}\left(\frac{1}{1+a\mu}\right)^{\left(\frac{1}{a}\right)}\left(\frac{a\mu}{1+a\mu}\right)^{y_i}$$

$$=\prod_{i=1}^{n}\frac{\left\{\prod_{r=1}^{y_i-1}(1+ar)\right\}}{a^{y_i}.y_i!}\left(\frac{1}{1+a\mu}\right)^{\left(\frac{1}{a}\right)}\left(\frac{a\mu}{1+a\mu}\right)^{y_i} \tag{3.15}$$

Untuk dapat memperoleh taksiran parameter a dan μ dengan metode maksimum *likelihood*, perlu dicari nilai a dan μ yang akan memaksimumkan fungsi *likelihood*, hal ini sama dengan jika mencari nilai a dan μ yang akan memaksimumkan fungsi log *likelihood*. Karena lebih mudah mencari nilai a dan μ yang akan memaksimumkan fungsi log *likelihood*, sehingga perlu dicari persamaan log *likelihood*, yaitu :

$$L(a,\mu) \quad = \prod_{i=1}^{n} \frac{\left\{\prod_{r=1}^{y_i-1}(1+ar)\right\}}{a^{y_i}.y_i!}\left(\frac{1}{1+a\mu}\right)^{\left(\frac{1}{a}\right)}\left(\frac{a\mu}{1+a\mu}\right)^{y_i}$$

$$\log L(a,\mu) = \log\left[\prod_{i=1}^{n} \frac{\left\{\prod_{r=1}^{y_i-1}(1+ar)\right\}}{a^{y_i}.y_i!}\left(\frac{1}{1+a\mu}\right)^{\left(\frac{1}{a}\right)}\left(\frac{a\mu}{1+a\mu}\right)^{y_i}\right]$$

$$\begin{aligned}
l(a,\mu) \quad &= \sum_{i=1}^{n}\left\{\sum_{r=1}^{y_i-1}\log(1+ar)\right\} - y_i\log a - \log y_i! + \frac{1}{a}\log 1 \\
&\quad -\frac{1}{a}\log(1+a\mu) + y_i\log a\mu - y_i\log(1+a\mu) \\
&= \sum_{i=1}^{n}\left\{\sum_{r=1}^{y_i-1}\log(1+ar)\right\} - y_i\log a - \log y_i! + y_i\log a\mu \\
&\quad -\left(y_i+\frac{1}{a}\right)\log(1+a\mu) \\
&= \sum_{i=1}^{n}\left\{\sum_{r=1}^{y_i-1}\log(1+ar)\right\} - y_i\log a - \log y_i! + y_i\log a + y_i\log\mu \\
&\quad -\left(y_i+\frac{1}{a}\right)\log(1+a\mu) \\
&= \sum_{i=1}^{n}\left(\sum_{r=1}^{y_i-1}\log(1+ar)\right) - \log(y_i!) + y_i\log(\mu) - \left(y_i+\frac{1}{a}\right)\log(1+a\mu) \\
&= \sum_{i=1}^{n}\left\{\left(\sum_{r=1}^{y_i-1}\log(1+ar)\right) - \log(y_i!) + y_i\log(\mu) - \left(y_i+a^{-1}\right)\log(1+a\mu)\right\}
\end{aligned} \tag{3.16}$$

Setelah mendapatkan persamaan log *likelihood*, selanjutnya dapat dicari nilai a dan μ yang memaksimumkan fungsi tersebut. Tahap awal yang dilakukan adalah dengan menghitung turunan parsial pertama terhadap parameter a dan μ, yang selanjutnya akan disamakan dengan nol.

$$\frac{\partial l(a,\mu)}{\partial \mu} = \sum_{i=1}^{n}\frac{y_i}{\mu} - \frac{\left(y_i+a^{-1}\right)a}{(1+a\mu)} = 0 \tag{3.17}$$

$$\frac{\partial l(a,\mu)}{\partial a}=\sum_{i=1}^{n}\left[\left(\sum_{r=1}^{y_i-1}\frac{r}{\left(1+ar\right)}\right)+\frac{1}{a^2}\log(1+a\mu)-\frac{\left(y_i+\frac{1}{a}\right)\mu}{(1+a\mu)}\right]=0 \qquad (3.18)$$

Nilai *a* dan μ yang diperoleh dari hasil penyelesaian persamaan (3.17) dan (3.18) akan memaksimumkan $l(a,\mu)$ jika

1. $$\Delta=\left\{\left(\frac{\partial l^2(a,\mu)}{\partial\mu^2}\right)\left(\frac{\partial l^2(a,\mu)}{\partial a^2}\right)-\left(\frac{\partial l^2(a,\mu)}{\partial a\partial\mu}\right)^2\right\}>0 \qquad (3.19)$$

2. $$\frac{\partial l^2(a,\mu)}{\partial\mu^2}<0 \text{ dan } \frac{\partial l^2(a,\mu)}{\partial a^2}<0 \qquad (3.20)$$

Dari persamaan (3.16), turunan parsial kedua terhadap parameter *a* dan μ adalah

$$\frac{\partial l^2(a,\mu)}{\partial\mu^2}=\sum_{i=1}^{n}-\frac{y_i}{\mu^2}+\frac{\left(y_ia+1\right)a}{(1+a\mu)^2} \qquad (3.21)$$

$$\frac{\partial l^2(a,\mu)}{\partial a^2}=\sum_{i=1}^{n}\left[\begin{array}{l}\left(\sum_{r=1}^{y_i-1}-\left(\frac{r}{\left(1+ar\right)}\right)^2\right)-\frac{2}{a^3}\log(1+a\mu)+\frac{1}{a^2}\frac{\mu}{1+a\mu}+\\ \frac{y_i\mu^2}{(1+a\mu)^2}-\frac{-\left(\frac{\mu}{a^2}(1+a\mu)+\frac{\mu^2}{a}\right)}{(1+a\mu)^2}\end{array}\right] \qquad (3.22)$$

$$\frac{\partial l(a,\mu)}{\partial a\partial\mu}=\sum_{i=1}^{n}\frac{1}{a^2}\frac{a}{(1+a\mu)}-\left[\frac{y_i(1+a\mu)-y_ia\mu}{(1+a\mu)^2}\right]-\left[\frac{\frac{1}{a}(1+a\mu)-\frac{1}{a}a\mu}{(1+a\mu)^2}\right]$$

$$=\sum_{i=1}^{n}\frac{1}{a(1+a\mu)}-\left(\frac{y_i+\frac{1}{a}}{(1+a\mu)^2}\right) \qquad (3.23)$$

Telah ditunjukkan pada lampiran 5 bahwa $\Delta>0$, dan $\frac{\partial l^2(a,\mu)}{\partial\mu^2}<0$ dan $\frac{\partial l^2(a,\mu)}{\partial a^2}<0$, sehingga nilai *a* dan μ yang diperoleh sebagai hasil dari persamaan (3.17) dan (3.18) akan memaksimumkan $l(a, \mu)$, dan disebut sebagai taksiran *likelihood*.

Penaksiran Parameter μ

Untuk mencari nilai μ yang memaksimumkan fungsi log *likelihood* dapat dilakukan dengan menurunkan persamaan log *likelihood* terhadap μ, dengan parameter a dianggap konstan atau memiliki nilai tertentu. Kemudian menyamakan persamaan turunan log *likelihood* terhadap μ tersebut dengan 0. Sehingga didapat taksiran untuk μ yaitu :

$$l(a,\mu)=\sum_{i=1}^{n}\left(\sum_{r=1}^{y_i-1}\log\left(1+ar\right)\right)-\log\left(y_i!\right)+y_i\log(\mu)-\left(y_i+a^{-1}\right)\log(1+a\mu)$$

$$\frac{\partial l(a,\mu)}{\partial\mu}=\sum_{i=1}^{n}\frac{y_i}{\mu}-\frac{\left(y_i+a^{-1}\right)a}{(1+a\mu)}$$

$$=\sum_{i=1}^{n}\frac{y_i}{\mu}-\frac{y_ia+1}{1+a\mu}=0$$

$$\sum_{i=1}^{n}\left(\frac{y_i}{\mu}-\frac{y_ia+1}{1+a\mu}\right)=0$$

$$\frac{1}{\mu}\sum_{i=1}^{n}y_i-\frac{a}{1+a\mu}\sum_{i=1}^{n}y_i-\frac{n}{1+a\mu}=0$$

$$\sum_{i=1}^{n}y_i\left(\frac{1}{\mu}-\frac{a}{1+a\mu}\right)-\frac{n}{1+a\mu}=0$$

$$\sum_{i=1}^{n}y_i\left(\frac{1}{\mu}-\frac{a}{1+a\mu}\right)-\frac{n}{1+a\mu}=0$$

$$\sum_{i=1}^{n}y_i\left(\frac{1+a\mu-a\mu}{\mu\left(1+a\mu\right)}\right)-\frac{n}{1+a\mu}=0$$

$$\sum_{i=1}^{n}y_i\left(\frac{1}{\mu\left(1+a\mu\right)}\right)-\frac{n\mu}{\mu\left(1+a\mu\right)}=0$$

$$\frac{\sum_{i=1}^{n} y_i - n\mu}{\mu(1+a\mu)} = 0$$

$$\sum_{i=1}^{n} y_i - n\mu = 0$$

$$\mu = \frac{\sum_{i=1}^{n} y_i}{n}$$

$$\mu = \bar{y}$$

$$\hat{\mu} = \bar{Y} \tag{3.24}$$

Maka dengan menggunakan metode maksimum *likelihood* didapat taksiran dari μ yaitu $\hat{\mu} = \bar{Y}$. Taksiran ini merupakan penaksir yang tak bias, konsisten dan efisien, pembuktiannya terdapat pada lampiran 2.

Penaksiran Parameter *a*

Untuk mencari nilai *a* yang memaksimumkan fungsi log *likelihood* dapat dilakukan dengan menurunkan persamaan log *likelihood* terhadap *a*, dengan parameter μ dianggap konstan atau memiliki nilai tertentu. Kemudian menyamakan persamaan turunan log *likelihood* terhadap *a* tersebut dengan 0. Sehingga didapat taksiran untuk *a* yaitu:

$$\begin{aligned} l(a,\mu) &= \sum_{i=1}^{n}\left(\sum_{r=1}^{y_i-1}\log\left(1+ar\right)\right) - \log\left(y_i!\right) + y_i\log(\mu) - \left(y_i + a^{-1}\right)\log(1+a\mu) \\ &= \sum_{i=1}^{n}\left(\sum_{r=1}^{y_i-1}\log\left(1+ar\right)\right) - \log\left(y_i!\right) + y_i\log(\mu) - \left(y_i + \frac{1}{a}\right)\log(1+a\mu) \end{aligned}$$

$$\begin{aligned} \frac{\partial l(a,\mu)}{\partial a} &= \sum_{i=1}^{n}\left(\sum_{r=1}^{y_i-1}\frac{r}{\left(1+ar\right)}\right) + \frac{1}{a^2}\log(1+a\mu) - \frac{\left(y_i + \frac{1}{a}\right)\mu}{(1+a\mu)} \\ &= \sum_{i=1}^{n}\left(\sum_{r=1}^{y_i-1}\frac{r}{\left(1+ar\right)}\right) + a^{-2}\log(1+a\mu) - \frac{\left(y_i + a^{-1}\right)\mu}{(1+a\mu)} \\ &= 0 \end{aligned}$$

$$\sum_{i=1}^{n}\left\{\left(\sum_{r=1}^{y_i-1}\frac{r}{(1+ar)}\right)+a^{-2}\log(1+a\mu)-\frac{\left(y_i+a^{-1}\right)\mu}{(1+a\mu)}\right\}=0$$

$$\left(\sum_{i=1}^{n}\sum_{r=1}^{y_i-1}\frac{r}{(1+ar)}\right)+\sum_{i=1}^{n}a^{-2}\log(1+a\mu)-\sum_{i=1}^{n}\frac{\left(y_i+a^{-1}\right)\mu}{(1+a\mu)}=0$$

$$\left(\sum_{i=1}^{n}\sum_{r=1}^{y_i-1}\frac{r}{(1+ar)}\right)+\frac{n}{a^2}\log(1+a\mu)-\frac{\mu}{(1+a\mu)}\sum_{i=1}^{n}\left(y_i+a^{-1}\right)=0$$

$$\left(\sum_{i=1}^{n}\sum_{r=1}^{y_i-1}\frac{r}{(1+ar)}\right)+\frac{n}{a^2}\log(1+a\mu)-\frac{\mu}{(1+a\mu)}\left(\frac{n}{a}+\sum_{i=1}^{n}y_i\right)=0 \qquad (3.25)$$

Solusi eksak dari persamaan diatas tidak dapat diperoleh, sehingga akan dilakukan pendekatan numerik dengan menggunakan metode *Newton-Rhapson* untuk mencari solusi dari persamaan tersebut.

3.7 Pengujian Parameter Overdispersi

Pada pembahasan sebelumnya telah dijelaskan bahwa distribusi Binomial Negatif dapat digunakan untuk mengatasi kasus overdispersi pada data *count*. Sehingga dibutuhkan pengujian untuk memeriksa ada atau tidaknya overdispersi pada data *count*. Pengujian yang dilakukan dengan menggunakan uji rasio *likelihood* dengan hipotesis yaitu:

H_0 : $a = 0$

H_1 : $a > 0$

Misalkan ω adalah ruang parameter di bawah asumsi H_0 benar, dan $L = L(\hat{\omega})$ adalah nilai maksimum dari fungsi *likelihood* sampel yang dibatasi oleh ω, dimana

$$L(a,\mu) \quad =\prod_{i=1}^{n}\frac{\left\{\prod_{r=1}^{y_i-1}(1+ar)\right\}}{a^{y_i}.y_i!}\left(\frac{1}{1+a\mu}\right)^{\left(\frac{1}{a}\right)}\left(\frac{a\mu}{1+a\mu}\right)^{y_i}$$

$$= \prod_{i=1}^{n} \frac{\left\{ \prod_{r=1}^{y_i - 1} (1 + ar) \right\}}{a^{y_i}} \frac{1}{y_i!} \left(\frac{1}{1 + a\mu} \right)^{\left(\frac{1}{a}\right)} \left(\frac{a}{1 + a\mu} \right)^{y_i} \mu^{y_i}$$

$$= \prod_{i=1}^{n} \left[\left\{ \prod_{r=1}^{y_i - 1} \frac{1 + ar}{1 + a\mu} \right\} \left(\frac{1}{1 + a\mu} \right)^{\left(\frac{1}{a}\right)} \mu^{y_i} \frac{1}{y_i!} \right]$$

$$= \prod_{i=1}^{n} \left[\left\{ \prod_{r=1}^{y_i - 1} \frac{\frac{1}{a} + r}{\frac{1}{a} + \mu} \right\} \left(\frac{\frac{1}{a}}{\frac{1}{a} + \mu} \right)^{\left(\frac{1}{a}\right)} \mu^{y_i} \frac{1}{y_i!} \right]$$

Misalkan $b = 1/a$, sehingga persamaannya menjadi

$$L(a, \mu) = \prod_{i=1}^{n} \left[\left\{ \prod_{r=1}^{y_i - 1} \frac{b + r}{b + \mu} \right\} \left(\frac{b}{b + \mu} \right)^{(b)} \mu^{y_i} \frac{1}{y_i!} \right]$$

$$= \prod_{i=1}^{n} \left[\left\{ \prod_{r=1}^{y_i - 1} \frac{1 + \frac{r}{b}}{1 + \frac{\mu}{b}} \right\} \left(\frac{1}{1 + \frac{\mu}{b}} \right)^{(b)} \mu^{y_i} \frac{1}{y_i!} \right]$$

Saat hipotesis H_0 benar, maka saat nilai $a \approx 0$, sehingga nilai $b \approx \infty$

$$\lim_{b \to \infty} L(a, \mu) = \lim_{b \to \infty} \left\{ \prod_{i=1}^{n} \left[\left\{ \prod_{r=1}^{y_i - 1} \frac{1 + \frac{r}{b}}{1 + \frac{\mu}{b}} \right\} \left(\frac{1}{1 + \frac{\mu}{b}} \right)^{(b)} \mu^{y_i} \frac{1}{y_i!} \right] \right\}$$

$$= \prod_{i=1}^{n} \lim_{b \to \infty} \left[\left\{ \prod_{r=1}^{y_i - 1} \frac{1 + \frac{r}{b}}{1 + \frac{\mu}{b}} \right\} \left(\frac{1}{1 + \frac{\mu}{b}} \right)^{(b)} \mu^{y_i} \frac{1}{y_i!} \right]$$

$$= \prod_{i=1}^{n} \left[\prod_{i=1}^{n} \lim_{b \to \infty} \left(\frac{1 + \frac{r}{b}}{1 + \frac{\mu}{b}} \right) \cdot \lim_{b \to \infty} \left(\frac{1}{1 + \frac{\mu}{b}} \right)^{(b)} \cdot \lim_{b \to \infty} \left(\mu^{y_i} \frac{1}{y_i!} \right) \right]$$

$$= \prod_{i=1}^{n} \left(\mu^{y_i} \frac{1}{y_i!} \right) \left[\prod_{i=1}^{n} 1 \cdot \lim_{b \to \infty} \left(\frac{1}{1 + \frac{\mu}{b}} \right)^{(b)} \right]$$

$$= \prod_{i=1}^{n} \left(\mu^{y_i} \frac{1}{y_i!} \right) \left(e^{-\mu} \right)$$

Sehingga persamaan dari $L(\hat{\omega})$ dapat ditulis sebagai berikut :

$$L(\hat{\omega}) = \prod_{i=1}^{n} \left(\hat{\mu}^{y_i} \frac{1}{y_i!} \right) \left(e^{-\hat{\mu}} \right)$$

$$= \frac{\hat{\mu}^{\sum_{i=1}^{n} y_i} \cdot e^{-n\hat{\mu}}}{\prod_{i=1}^{n} y_i!} \qquad (3.26)$$

Definisikan $L = L(\hat{\Omega})$ adalah nilai maksimum dari fungsi *likelihood* sampel, dimana parameter dapat diambil dari sembarang nilai spesifik dalam gabungan H_0 dan H_1, sehingga didapat

$$L(\hat{\Omega}) = \prod_{i=1}^{n} \frac{\left\{ \prod_{r=1}^{y_i - 1} (1 + \hat{a} r) \right\}}{\hat{a}^{y_i} . y_i!} \left(\frac{1}{1 + \hat{a}\hat{\mu}} \right)^{\left(\frac{1}{\hat{a}}\right)} \left(\frac{\hat{a}\hat{\mu}}{1 + \hat{a}\hat{\mu}} \right)^{y_i}$$

$$= \prod_{i=1}^{n} \frac{\left\{ \prod_{r=1}^{y_i - 1} (1 + \hat{a} r) \right\}}{y_i!} \left(\frac{1}{1 + \hat{a}\hat{\mu}} \right)^{\left(\frac{1}{\hat{a}}\right)} \left(\frac{\hat{\mu}}{1 + \hat{a}\hat{\mu}} \right)^{y_i}$$

$$= \left(\frac{\hat{\mu}}{1 + \hat{a}\hat{\mu}} \right)^{\sum_{i=1}^{n} y_i} \left(\frac{1}{1 + \hat{a}\hat{\mu}} \right)^{\left(\frac{n}{\hat{a}}\right)} \prod_{i=1}^{n} \frac{\left\{ \prod_{r=1}^{y_i - 1} (1 + \hat{a} r) \right\}}{y_i!} \quad , \quad \hat{a} > 0 \qquad (3.27)$$

dimana $\hat{a}$ dan $\hat{\mu}$ adalah taksiran maksimum *likelihood* dari distribusi Binomial Negatif. Rasio *likelihood* dinotasikan dengan

$$\delta = \frac{L(\hat{\omega})}{L(\hat{\Omega})}$$

$$\delta = \frac{\dfrac{\hat{\mu}^{\sum_{i=1}^{n} y_i} \cdot e^{-n\hat{\mu}}}{\prod_{i=1}^{n} y_i!}}{\left(\dfrac{\hat{\mu}}{1 + \hat{a}\hat{\mu}} \right)^{\sum_{i=1}^{n} y_i} \left(\dfrac{1}{1 + \hat{a}\hat{\mu}} \right)^{\left(\frac{n}{\hat{a}}\right)} \prod_{i=1}^{n} \dfrac{\left\{ \prod_{r=1}^{y_i - 1} (1 + \hat{a} r) \right\}}{y_i!}} \quad , \hat{a} > 0$$

$$= \frac{e^{-n\hat{\mu}}}{\left(\dfrac{1}{1 + \hat{a}\hat{\mu}} \right)^{\sum_{i=1}^{n} y_i} \left(\dfrac{1}{1 + \hat{a}\hat{\mu}} \right)^{\left(\frac{n}{\hat{a}}\right)} \prod_{i=1}^{n} \left\{ \prod_{r=1}^{y_i - 1} (1 + \hat{a} r) \right\}} \quad , \hat{a} > 0$$

$$= \frac{e^{-n\hat{\mu}}}{\left(\frac{1}{1+\hat{a}\hat{\mu}}\right)^{\left(\frac{n}{\hat{a}}\right)+\sum_{i=1}^{n} y_i} \prod_{i=1}^{n}\left\{\prod_{r=1}^{y_i-1}(1+\hat{a}r)\right\}} \quad , \hat{a}>0$$

$$= \frac{e^{-n\hat{\mu}} \cdot (1+\hat{a}\hat{\mu})^{\left(\frac{n}{\hat{a}}\right)+\sum_{i=1}^{n} y_i}}{\prod_{i=1}^{n}\left\{\prod_{r=1}^{y_i-1}(1+\hat{a}r)\right\}} \quad , \hat{a}>0 \tag{3.28}$$

Rasio *likelihood* yang dinotasikan dengan δ tersebut tidak memiliki suatu distribusi yang spesifik, namun -2 kali dari logaritma natural dari rasio *likelihood* (δ) di bawah asumsi H_0 benar akan mengaproksimasi distribusi *chi-square* dengan derajat bebas satu atau $\chi^2_{(1)}$, sehingga statistik uji yang akan digunakan dalam pengujian rasio *likelihood* yaitu:

$$A = -2\log\left(\frac{L(\hat{\omega})}{L(\hat{\Omega})}\right)$$

$$= -2\log(\delta)$$

$$= -2\log\left(\frac{e^{-n\hat{\mu}} \cdot (1+\hat{a}\hat{\mu})^{\left(\frac{n}{\hat{a}}\right)+\sum_{i=1}^{n} y_i}}{\prod_{i=1}^{n}\left\{\prod_{r=1}^{y_i-1}(1+\hat{a}r)\right\}}\right)$$

$$= -2\left[\log\left(e^{-n\hat{\mu}} \cdot (1+\hat{a}\hat{\mu})^{\left(\frac{n}{\hat{a}}\right)+\sum_{i=1}^{n} y_i}\right) - \log\left(\prod_{i=1}^{n}\left\{\prod_{r=1}^{y_i-1}(1+\hat{a}r)\right\}\right)\right]$$

$$= -2\left[\left(-n\hat{\mu} + \left\{\left(\frac{n}{\hat{a}}\right) + \sum_{i=1}^{n} y_i\right\}\log(1+\hat{a}\hat{\mu})\right) - \left(\sum_{i=1}^{n}\sum_{r=1}^{y_i-1}\log(1+\hat{a}r)\right)\right] \quad , \hat{a}>0 \tag{3.29}$$

Dengan menggunakan informasi dimana saat H_0 benar, maka berdasarkan teorema pada subbab 2.13, statistik uji A akan berdistribusi *chi-square* dengan derajat bebas satu atau $\chi^2_{(1)}$, dapat ditentukan aturan keputusan yang akan digunakan dalam pengujian ini. Aturan keputusan : pada tingkat signifikansi α, H_0 akan ditolak jika nilai statistik uji $A > \chi^2_{\alpha,1}$. Penolakan H_0 pada tingkat singnifikansi α, dapat disimpulkan bahwa terjadi overdispersi pada data *count* pada tingkat signifikansi α.

BAB IV
CONTOH APLIKASI

Tingginya tingkat kecelakaan yang terjadi di berbagai ruas jalan mulai meresahkan masyarakat. Hal ini memaksa pihak yang berwenang, dalam kasus ini adalah kepolisian atau dinas lalu lintas untuk melakukan analisis mengenai hal tersebut. Salah satu analisis yang dapat diterapkan adalah analisis statistik untuk mengetahui penyebaran serta rata-rata dari jumlah kecelakaan tersebut di ruas-ruas jalan tertentu.

Bab ini membahas tentang contoh kasus overdispersi pada data count beserta penaksiran dan pengujian parameternya.

4.1 Sumber Data

Pada contoh aplikasi ini, diberikan 50 data yang merupakan data jumlah kecelakaan yang terjadi di suatu jalan tertentu, dengan pengemudi yang dipilih secara acak. Misalkan Y adalah suatu peubah acak yang menjelaskan jumlah kecelakaan yang terjadi pada suatu jalan tertentu. Untuk melakukan analisis pada data kecelakaan dalam analisis keselamatan lalu lintas, dibutuhkan suatu distribusi statistik yang cocok untuk data tersebut. Jumlah kecelakaan di suatu jalan tertentu diasumsikan berdistribusi Poisson, yang memiliki satu parameter, yang menyatakan mean dan variansi yang sama. Hal ini mengindikasikan bahwa rata-rata tingkat kecelakaan (λ) dalam populasi konstan.

Namun pada kenyataannya, setiap individu memiliki tingkat kecendrungan kecelakaan yang berbeda. Tingkat kecenderungan yang berbeda antar individu mengindikasikan bahwa nilai λ bervariasi antar individu, sehingga apabila dilakukan analisis statistika dan penaksiran parameter dengan menggunakan distribusi Poisson pada data saat terjadi overdispersi, akan didapat taksiran parameter yang tidak efisien dan akan memberikan informasi yang tidak sesuai.

Karena alasan tersebut, distribusi Poisson tidak cocok digunakan pada data saat terjadi overdispersi. Distribusi Binomial Negatif merupakan salah satu alternatif yang dapat digunakan dalam melakukan analisis statistik pada data jumlah kecelakaan saat terjadi overdispersi.

Tabel 4.1 Data Jumlah Kecelakaan

Data Jumlah Kecelakaan di Suatu Ruas Jalan									
8	6	8	6	3	0	0	4	3	2
4	4	9	6	2	0	5	0	9	0
0	0	0	5	5	8	4	4	3	12
0	3	2	5	0	5	0	4	6	6
1	0	3	0	3	2	0	0	3	3
3	0	0	3	6	5	1	2	0	2
0	1	4	10	7	0	1	0	1	4
0	1	0	7	6	0	4	0	0	5
8	1	0	0	3	6	1	0	0	7
6	4	0	0	4	5	6	11	3	0
5	2	0	0	4	4	0	1	0	10
4	0	14	0	0	5	6	4	0	0
3	2	0	5	0	15	2	3	0	0
4	0	1	0	8	0	4	0	0	4
3	6	3	1	0	5	0	0	0	7
5	10	4	1	0	0	0	0	1	3
8	5	5	1	9	1	6	0	1	0
3	3	0							

4.2 Analisis Data

Dari data diatas selanjutnya akan dilakukan analisis terhadap data tersebut. Tujuan dari analisis data ini adalah untuk mengetahui apakah terjadi overdispersi pada data tersebut dan mencari distribusi statistik yang dapat mewakili data jumlah kecelakaan serta mengetahui penyebaran dan rata-rata dari jumlah kecelakaan. Terdapat beberapa tahap analisis yang akan dilakukan pada data tersebut antara lain uji *Kolmogorov-Smirnov*, penaksiran parameter, dan uji rasio *likelihood*.

4.2.1 Uji Kolmogorov-Smirnov

Hal pertama yang akan dilakukan dalam analisis adalah memeriksa apakah data tersebut berdistribusi Poisson atau tidak, pengujian yang akan dilakukan menggunakan uji *Kolmogorov Smirnov*.

- Hipotesis

  Untuk melakukan pengujian ini digunakan hipotesis sebagai berikut :

  H_0 : Data berdistribusi Poisson

  H_1 : Data tidak berdistribusi Poisson

- Tingkat Singnifikansi

  Pada pengujian ini akan digunakan tiingkat signifikansi $\alpha = 0.05$

- Aturan Keputusan

  H_0 ditolak jika $\hat{\alpha} < \alpha = 0.05$

- Keputusan

  Dengan menggunakan program SPSS, akan didapat output sebagai berikut :

Tabel 4.2 Uji Satu Sampel Kolmogorov-Smirnov

		y
N		173
Poisson Parameter[a]	Mean	2.92
Most Extreme Differences	Absolute	.304
	Positive	.304
	Negative	-.115
Kolmogorov-Smirnov Z		4.004
Asymp. Sig. (2-tailed)		.000

a. Test distribution is Poisson.

Dari table 4.2 diatas, didapat bahwa nilai $\hat{\alpha} = 0.000 < \alpha = 0.05$, sehingga H_0 ditolak.

- Kesimpulan

  H_0 ditolak, hal ini dapat disimpulkan bahwa dengan tingkat kepercayaan 95%, kita percaya data tersebut tidak berdistribusi Poisson.

Dari pengujian didapat bahwa ternyata data tersebut tidak berdistribusi Poisson. Hal ini menunjukkan bahwa keadaan ekuidispersi atau keadaan dimana mean dan variansi bernilai sama tidak terpenuhi pada data tersebut. Terdapat dua kemungkinan dari data tersebut yaitu data mengalami underdispersi atau overdispersi. Jika dilihat nilai rasio antara variansi dan mean yang bernilai lebih dari 1, maka diduga terjadi overdispersi pada data tersebut. Distribusi Binomial Negatif merupakan salah satu alternatif distribusi yang dapat digunakan saat terjadi overdispersi pada data.

4.2.2 Penaksiran Parameter

Pada distribusi Binomial Negatif, terdapat dua parameter yang nilainya tidak diketahui, yaitu parameter μ dan a. Selanjutnya akan digunakan distribusi Binomial Negatif dalam melakukan analisis pada data tersebut, dan akan dilakukan penaksiran dari kedua parameter tersebut dengan menggunakan metode maksimum *likelihood*.

4.2.2.1 Penaksiran Parameter μ

Pada bab sebelumnya telah dijelaskan bagaimana mencari taksiran dari parameter μ dengan menggunakan metode maksimum *likelihood*, dan didapat bahwa taksiran dari parameter μ, yaitu:

$$\hat{\mu} = \bar{Y}$$

Hasil output deskriptif statistik dengan mengunakan SPSS, sebagai berikut :

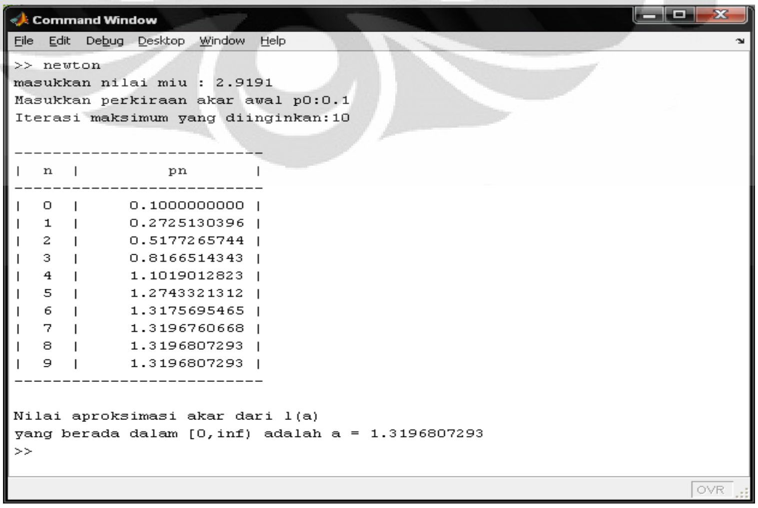
Tabel 4.3 Statistik Deskriptif

	N	Mean	Std. Deviation	Minimum	Maximum
y	173	2.92	3.148	0	15

Dari tabel 4.3 diatas didapat bahwa taksiran parameter μ adalah $\hat{\mu} = 2.92$

4.2.2.2 Penaksiran Parameter *a*

Pada bab sebelumnya telah dibahas bahwa penaksiran parameter a dengan menggunakan metode maksimum *likelihood* tidak dapat diselesaikan dengan cara yang biasa seperti pada penaksiran parameter μ karena persamaan yang didapat bukan merupakan suatu fungsi linear, sehingga akan digunakan metode *Newton-Rhapson* pada Matlab dalam penyelesaiannya.Hasil output yang didapat untuk mencari taksiran parameter a dengan menggunakan matlab yaitu:




```
>> newton
masukkan nilai miu : 2.9191
Masukkan perkiraan akar awal p0:0.1
Iterasi maksimum yang diinginkan:10

--------------------------
|  n  |        pn        |
--------------------------
|  0  |     0.1000000000 |
|  1  |     0.2725130396 |
|  2  |     0.5177265744 |
|  3  |     0.8166514343 |
|  4  |     1.1019012823 |
|  5  |     1.2743321312 |
|  6  |     1.3175695465 |
|  7  |     1.3196760668 |
|  8  |     1.3196807293 |
|  9  |     1.3196807293 |
--------------------------

Nilai aproksimasi akar dari l(a)
yang berada dalam [0,inf) adalah a = 1.3196807293
>>
```

Gambar 4.1 Nilai Taksiran Parameter *a*

Dari hasil output diatas didapat bahwa $\hat{a} = 1.3196807293$

4.2.3 Uji Rasio *Likelihood*

Setelah mengetahui nilai taksiran dari kedua parameter tersebut, akan dilakukan pengujian terhadap parameter dispersi pada distribusi Binomial Negatif, yaitu parameter *a*. Pegujian dilakukan untuk memeriksa asumsi terjadinya overdispersi pada data.

- Hipotesis

  Untuk melakukan pengujian ini digunakan hipotesis sebagai berikut :

  H_0 : $a = 0$

  H_1 : $a > 0$

- Tingkat Singnifikansi

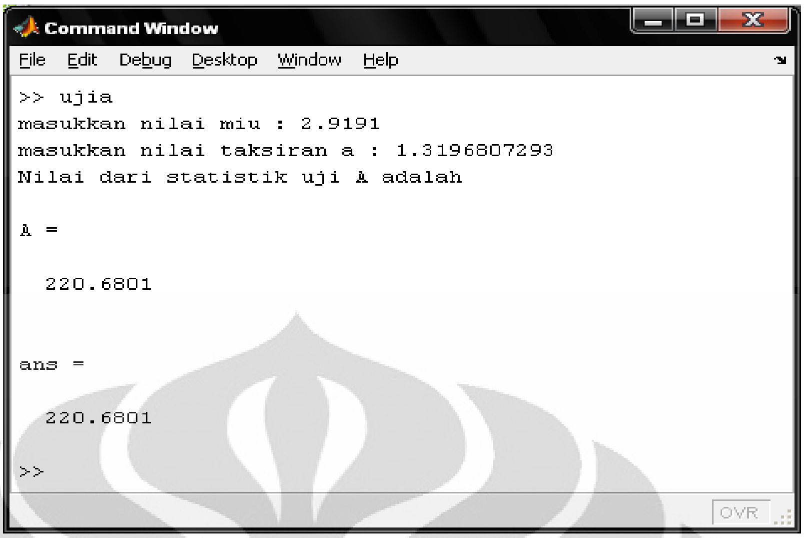
  Pada pengujian ini akan digunakan tingkat signifikansi α = 0.05

- Aturan Keputusan

  H_0 akan ditolak jika nilai statistik uji $A > \chi^2_{0.05,1}$

- Keputusan

  Selanjutnya untuk mengetahui keputusan apakah H_0 ditolak atau tidak akan dihitung nilai dari statistik uji *A* dengan menggunakan bantuan matlab. Hasil yang diperoleh adalah sebagai berikut :



Gambar 4.2 Nilai Statistik Uji *A*

Dari hasil perhitungan diatas didapat bahwa jika nilai statistik uji A = 220.6801 > $\chi^2_{0.05,1} = 3.8415$, sehingga H_0 ditolak.

- Kesimpulan

  Dengan tingkat signifikansi 95%, kita percaya bahwa terjadi overdispersi pada data yang kita gunakan sehingga distribusi Binomial Negatif lebih cocok digunakan untuk melakukan analisis pada data tersebut dibandingkan dengan distribusi Poisson.

BAB V
PENUTUP

Pada bab ini, akan diberikan kesimpulan serta saran dari penulisan tugas akhir ini.

5.1 Kesimpulan

Data *count* adalah data hasil percobaan acak yang nilai-nilainya berupa bilangan bulat non negatif. Distribusi yang biasa digunakan untuk memodelkan data *count* adalah distribusi Poisson yang memiliki mean dan variansi yang sama. Saat terjadi overdispersi, yaitu keadaan dimana variansi dari peubah acak bernilai lebih besar dari nilai mean dari peubah acak tersebut, asumsi mean dan variansi sama pada distribusi Poisson tidak lagi terpenuhi. Oleh karena itu diperlukan distribusi lain untuk menganalisis data *count* tersebut.

Pada kondisi overdispersi, nilai mean bervariasi antar individu sehingga mean merupakan suatu peubah acak. Pemilihan distribusi Gamma sebagai distribusi dari mean akan menghasilkan distribusi Binomial Negatif yang merupakan distribusi Campuran Poisson-Gamma (*Mixture Poisson-Gamma distribution*). Distribusi Binomial Negatif ini merupakan salah satu distribusi alternatif yang dapat digunakan saat terjadi overdispersi, karena pada distribusi ini terdapat tambahan parameter yang menjelaskan tentang besar dispersi dari data.

Metode maksimum *likelihood* dan metode *Newton-Rhapson* digunakan untuk melakukan penaksiran terhadap parameter mean dan parameter dispersi. Taksiran dari parameter mean merupakan penaksir yang tak bias, konsisten, dan efisien.

5.2 Saran

Selain metode maksimum *likelihood*, penaksiran parameter dari distribusi Binomial Negatif dapat dilakukan dengan metode lain diantaranya adalah metode Quasi *Likelihood* dan metode *Bootstrapped* Maksimum *Likelihood.*

DAFTAR PUSTAKA

Cameron, A. C., dan Trivedi, P. K. (1998). *Regression Analysis of Count Data*. Cambridge: Cambridge University Press.

Gelman, A., et al. (2000). *Bayesian Data Analysis*. New York: Chapman & Hall/CRC.

Goovaerts M.J., et al. (2002). *Modern Actuarial Risk Theory*. Netherlands: Kluwer Academic Publishers.

Hilbe, Joseph.M. (2011). *Negative Binomial Regression*. Cambridge: Cambridge University Press.

Hogg, R.V., dan Craig, A.T. (1995). *Introduction to Mathematical Statistics* (5^{th} ed.). New Jersey: Prentice – Hall Internatioal.

Jong, Piet De., dan Heller, Gillian. Z. (2008). *Generalized Linear Models for Insurance Data*. Cambridge: Cambridge University Press.

Ismail, N., dan Jemain, A. A. (2007). *Handling Overdispersion with Negative Binomial and Generalzed Poisson Regression Models*. Virginia: Casualty Actuarial Society Forum.

Johnson, Norman.L., Kotz, Samuel., dan Kemp, Adrienne.W. (1992). *Univariate Discrete Distribution*. New York: The Willey Interscience Publication.

Panjer, Harry.H., and Willmot, Gordo.E. (1992). *Insurance Risk Models*. USA : Society of Actuaries.

Piegorsch, Walter.W. (1990). *Maximum Likelihood Estimation for the Negative Binomial Dispersion Parameter*. U.S.A: International Biometric Society.

Richard, D. (2008). *Bayesian Analysis*. Paper of lecturer Notes.

Ross, G.J.S. and Preece, D.A. (1985). *The Negative Binomial Distribution*. Royal Statistical Society, U.K, 323-336.

Roussas, George.G. (1997). *A Course in Mathematical Statistics* (2^{nd} Edition). USA: Academic Press.

Schmock, Uwe. (2008). *Modelling Dependent Credit Risk with Extensions of Credit Risk and It's Application to Operational Risk*. Vienna University of Technology. Vienna, Australia.

Setiawani, Susi. (2006). Analisis Distribusi Binomial Negatif. Saintifika Vol 7. Jember, Indonesia.

Simon, Leroy, J. (1962). An Intoduction to Negative Binomial Distribution and It's Aplication. Vol XLIX Part 1.

Walsh, B. (2002). *Introduction to Bayesian Analysis*. Paper of Lecturer Notes for EEB 596z.

Willsan, Linda.J. (1981). *Estimation and Testing Procedures for the Parameters of Negative Binomial Distribution*. Thesis for Doktor of Philosophy, Faculty of Graduate Collage. Oklahoma: Oklahoma State University.

Lampiran 1. Bukti Teorema Uji Rasio *Likelihood*

Akan dicari distribusi dari statistik uji $A = -2\left[l(\hat{\omega}) - l(\hat{\Omega}) \right]$, dimana $l(\hat{\omega})$ adalah nilai maksimum dari fungsi log *likelihood* sampel di bawah asumsi H_0 benar yang dibatasi oleh ω , dan $l(\hat{\Omega})$ adalah nilai maksimum dari fungsi log *likelihood* sampel, dimana parameter dapat diambil dari sembarang nilai spesifik dalam gabungan H_0 dan H_1 benar. Misalkan diambil nilai taksiran parameter $\hat{\boldsymbol{\theta}}$ dari ruang taksiran parameter $\hat{\omega}$ dan taksiran parameter $\hat{\boldsymbol{\theta}}_{maks}$ dari ruang taksiran parameter $\hat{\Omega}$, maka statistik uji A dapat ditulis sebagai berikut :

$$A = -2\log \frac{L(\hat{\boldsymbol{\theta}})}{L(\hat{\boldsymbol{\theta}}_{maks})} = -2\left[l(\hat{\boldsymbol{\theta}}) - l(\hat{\boldsymbol{\theta}}_{maks}) \right] \qquad (1)$$

dimana $l(\hat{\boldsymbol{\theta}}_{maks})$ adalah nilai maksimum fungsi log *likelihood* dalam gabungan H_0 dan H_1 benar dengan jumlah parameter r yang dihitung saat $\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}}_{maks}$ dan $l(\hat{\theta})$ adalah nilai maksimum fungsi log *likelihood* di bawah asumsi H_0 benar dengan jumlah parameter m yang dihitung saat $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$.

Pendekatan deret Taylor orde 2 untuk mencari nilai $l(\hat{\boldsymbol{\theta}})$, saat nilai $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$ adalah :

$$l(\boldsymbol{\theta}) \approx l(\hat{\boldsymbol{\theta}}) + (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}) \left. \frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} + \frac{1}{2}(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})^T \left. \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2} \right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}) \qquad (2)$$

dimana $\left. \frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} = 0$, karena $\hat{\boldsymbol{\theta}}$ merupakan taksiran maksimum *likelihood* , dan $l(\hat{\boldsymbol{\theta}})$ merupakan suatu kostanta, sehingga persamaannya menjadi

$$l(\boldsymbol{\theta}) \approx K + \frac{1}{2}(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})^T \left. \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2} \right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}) \qquad (3)$$

(lanjutan)

jika persamaan diatas dinyatakan dalam bentuk eksponensial akan menjadi

$$p(\boldsymbol{\theta}|x) \approx C \exp\left[\frac{1}{2}(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}})^T \left.\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2}\right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} (\boldsymbol{\theta}-\hat{\boldsymbol{\theta}})\right] \tag{4}$$

dengan *C* adalah suatu konstanta.

Perhatikan peubah acak X berdistribusi Multivariate Normal ($\boldsymbol{\mu}$, $\boldsymbol{A}^{-1}$), pdf nya adalah :

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{n/2}\sqrt{|\mathbf{A}^{-1}|}} \exp\left(-\frac{(\mathbf{x}-\boldsymbol{\mu})^T \mathbf{A}(\mathbf{x}-\boldsymbol{\mu})}{2}\right) \tag{5}$$

Jika kedua persamaan tersebut dibandingkan, maka dapat disimpulkan bahwa persamaan $p(\theta|x)$ memiliki bentuk fungsional yang sama dengan distribusi Multivariate Normal, dimana :

$$\boldsymbol{\mu} = \hat{\boldsymbol{\theta}}$$

$$\mathbf{A}^{-1} = \left(\left[-\left.\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2}\right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}\right]\right)^{-1}$$

Sehingga distribusi Multivariate Normal sering digunakan untuk mengaproksimasi distribusi $p(\boldsymbol{\theta}|x)$,

$$p(\boldsymbol{\theta}|x) \approx N\left(\hat{\boldsymbol{\theta}}, [I(\boldsymbol{\theta})]^{-1}\right)$$

Dimana $I(\boldsymbol{\theta})$ adalah informasi terobservasi dengan $I(\boldsymbol{\theta}) = -\left.\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2}\right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}$, karena $\boldsymbol{\theta}$ adalah vektor parameter, maka $I(\boldsymbol{\theta})$ adalah matriks.

Dengan menggunakan informasi tersebut, maka dengan menggunakan dalil limit pusat, akan didapat

$$\frac{\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}}{[I(\boldsymbol{\theta})]^{-1/2}} \approx N(\mathbf{0}, \mathbf{I})$$

(lanjutan)

sehingga

$$\left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}\right)^T I(\boldsymbol{\theta})\left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}\right) \sim \chi_m^2$$

Persamaan (2) dapat ditulis sebagai berikut :

$$l(\boldsymbol{\theta})-l\left(\hat{\boldsymbol{\theta}}\right) \approx -\frac{1}{2}\left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}\right)^T I(\boldsymbol{\theta})\left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}\right)$$

$$-2\left[l(\boldsymbol{\theta})-l\left(\hat{\boldsymbol{\theta}}\right)\right] \approx \left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}\right)^T I(\boldsymbol{\theta})\left(\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}\right)$$

sehingga akan didapat $-2\left[l(\boldsymbol{\theta})-l\left(\hat{\boldsymbol{\theta}}\right)\right] \sim \chi_m^2$. (6)

Dari persamaan (1), didapat :

$$\begin{aligned} A &= -2\left[l\left(\hat{\boldsymbol{\theta}}\right)-l\left(\hat{\boldsymbol{\theta}}_{maks}\right)\right] \\ &= 2\left[l\left(\hat{\boldsymbol{\theta}}_{maks}\right)-l\left(\hat{\boldsymbol{\theta}}\right)\right] \\ &= 2\left[\left(l\left(\hat{\boldsymbol{\theta}}_{maks}\right)-l(\boldsymbol{\theta}_{maks})\right)-\left(l\left(\hat{\boldsymbol{\theta}}\right)-l(\boldsymbol{\theta})\right)+\left(l(\boldsymbol{\theta}_{maks})-l(\boldsymbol{\theta})\right)\right] \\ &= 2\left(l\left(\hat{\boldsymbol{\theta}}_{maks}\right)-l(\boldsymbol{\theta}_{maks})\right)-2\left(l\left(\hat{\boldsymbol{\theta}}\right)-l(\boldsymbol{\theta})\right)+2\left(l(\boldsymbol{\theta}_{maks})-l(\boldsymbol{\theta})\right) \qquad (7) \end{aligned}$$

Berdasarkan persamaan (6), kita mengetahui bahwa suku pertama pada persamaan (7) diatas berdistribusi *chi-square* dengan derajat bebas *r*, sementara suku kedua berdistribusi *chi-square* dengan derajat bebas *m*, dan suku ketiga merupakan suatu konstanta. Konstanta tersebut akan mendekati nol jika model dengan *m* parameter (*fitted model*) dapat menggambarkan data seperti pada *r* parameter (*saturated model*). Sehingga diperoleh

$$A = -2\left[l\left(\hat{\boldsymbol{\theta}}\right)-l\left(\hat{\boldsymbol{\theta}}_{maks}\right)\right] \sim \chi_{r-m}^2 \qquad (8)$$

Lampiran 2. Sifat – Sifat Taksiran Parameter μ

1. Penaksir Tak bias

Untuk melihat apakah taksiran parameter dari metode maksimum *likelihood* merupakan taksiran yang tak bias, perlu dilihat apakah nilai ekspektasi dari taksiran tersebut sama dengan nilai dari parameter yang ditaksir.

$$
\begin{aligned}
E(\hat{\mu}) = E(\bar{Y}) &= E\left(\sum_{i=1}^{n} \frac{Y_i}{n}\right) \\
&= \frac{1}{n} E\left(Y_1 + Y_2 + \ldots + Y_n\right) \\
&= \frac{1}{n}\left[E(Y_1) + E(Y_2) + \ldots + E(Y_n)\right] \\
&= \frac{1}{n}\left(\mu + \mu + \ldots + \mu\right) = \frac{1}{n} n\mu \\
E(\hat{\mu}) &= \mu
\end{aligned}
$$

sehingga $\hat{\mu} = \bar{Y}$ adalah penaksir yang tak bias untuk μ.

2. Penaksir Konsisten

Untuk melihat apakah taksiran parameter a dan μ dari metode maksimum *likelihood* merupakan taksiran yang konsisten perlu dilihat apakah taksiran tersebut konvergen dalam probabilitas ke parameter yang ditaksir.

Suatu barisan peubah acak $X_1, X_2, \ldots, X_n$ dikatakan konvergen dalam probabilitas ke peubah acak X, jika untuk setiap $\varepsilon > 0$, sedemikian sehingga

$$\lim_{n\to\infty} \Pr\{|X_n - X| < \varepsilon\} = 1$$

atau ekuivalen dengan

$$\lim_{n\to\infty} \Pr\{|X_n - X| \geq \varepsilon\} = 0$$

Taksiran parameter dari μ adalah $\hat{\mu} = \bar{Y}$, dimana $\bar{Y}$ merupakan mean dari sampel acak berukuran n dari suatu distribusi Binomial Negatif yang memiliki

(lanjutan)

mean μ dan variansi μ+a μ^2. Sehinggan mean dan variansi dari $\bar{Y}$ yaitu μ dan (μ+a μ^2)/n. untuk setiap ε > 0,

$$\Pr\left(\left|\bar{Y}-\mu\right|\geq\varepsilon\right)=\Pr\left(\left|\bar{Y}-\mu\right|\geq\frac{k\left(\mu+a\mu^2\right)}{\sqrt{n}}\right)$$

dimana $k=\dfrac{\varepsilon\sqrt{n}}{\left(\mu+a\mu^2\right)}$. Dengan menggunakan pertidaksamaan *Chebyshev*,

didapat :

$$\Pr\left(\left|\bar{Y}-\mu\right|\geq k\frac{\sqrt{\mu+a\mu^2}}{\sqrt{n}}\right)\leq\frac{1}{k^2}$$

$$\Pr\left(\left|\bar{Y}-\mu\right|\geq\frac{\varepsilon\sqrt{n}}{\sqrt{\mu+a\mu^2}}\frac{\sqrt{\mu+a\mu^2}}{\sqrt{n}}\right)\leq\frac{1}{\left(\dfrac{\varepsilon\sqrt{n}}{\sqrt{\mu+a\mu^2}}\right)^2}$$

$$\Pr\left(\left|\bar{Y}-\mu\right|\geq\varepsilon\right)\leq\frac{\mu+a\mu^2}{\varepsilon^2 n}$$

$$\lim_{n\to\infty}\Pr\left(\left|\bar{Y}-\mu\right|\geq\varepsilon\right)\leq\lim_{n\to\infty}\frac{\mu+a\mu^2}{\varepsilon^2 n}$$

$$\lim_{n\to\infty}\Pr\left(\left|\bar{Y}-\mu\right|\geq\varepsilon\right)\leq 0$$

Karena nilai dari suatu probabilitas tidak mungkin negatif, maka didapat

$$\lim_{n\to\infty}\Pr\left(\left|\bar{Y}-\mu\right|\geq\varepsilon\right)=0$$

Sehingga diperoleh $\hat{\mu}=\bar{Y}$ konvergen dalam probabilitas ke parameter μ dan taksiran $\hat{\mu}=\bar{Y}$ dikatakan penaksir yang konsisten untuk μ.

3. Penaksir Efisien

Untuk melihat apakah taksiran unbiased $\hat{\mu}=\bar{Y}$ merupakan suatu penaksir yang efisien untuk μ, perlu dilihat apakah variansi dari statistik $\hat{\mu}=\bar{Y}$ mencapai batas bawah dari *Rao-Cramér*. Atau dapat dinyatakan sebagai berikut :

(lanjutan)

$$\sigma^2_{\hat{\mu}=\bar{Y}} \geq \frac{1}{nI\left(\mu\right)}$$

Sebelum memeriksa apakah $\hat{\mu} = \bar{Y}$ merupakan penaksir yang efisien, akan dicari persamaan *Fisher Information* atau I(μ) terlebih dahulu.

$$I(\mu) = \int_{-\infty}^{\infty} \left[\frac{\partial \ln f\left(y_i;\mu,a\right)}{\partial \mu} \right]^2 f\left(y_i;\mu,a\right) dy$$
$$= E\left\{ \left[\left(\frac{\partial \ln f\left(y_i;\mu,a\right)}{\partial \mu} \right)^2 \right] \right\}$$

atau

$$I\left(\mu\right) = -\int_{-\infty}^{\infty} \frac{\partial^2 \ln f\left(y_i;\mu,a\right)}{\partial \mu^2} f\left(y_i;\mu,a\right) dx$$
$$= -E\left\{ \left[\left(\frac{\partial^2 \ln f\left(y_i;\mu,a\right)}{\partial \mu^2} \right) \right] \right\}$$

Sehingga dengan

$$\Pr(Y_i = y_i) = f(y_i)$$
$$= \frac{\Gamma\left(y_i + a^{-1}\right)}{\Gamma\left(a^{-1}\right) y_i!} \left(\frac{\mu}{\mu + a^{-1}} \right)^{y_i} \left(\frac{a^{-1}}{\mu + a^{-1}} \right)^{a^{-1}} , \quad \mu > 0,\, a^{-1} > 0 \ ,\ y_i = 0,1,2,...$$

dimana

$$\ln f(y_i) = \ln\left(\Gamma\left(y_i + a^{-1}\right)\right) - \ln\left(\Gamma\left(a^{-1}\right)\right) - \ln\left(y_i!\right) + y_i \ln\left(\mu\right) - y_i \ln\left(\mu + a^{-1}\right)$$
$$a^{-1} \ln\left(a^{-1}\right) - a^{-1} \ln\left(\mu + a^{-1}\right)$$

akan didapat persamaan *Fisher Information* dari μ adalah

$$\frac{\partial \ln f(y_i)}{\partial \mu} = \frac{y_i}{\mu} - \left(\frac{y_i + \frac{1}{a}}{\mu + \frac{1}{a}} \right)$$
$$= \frac{\frac{1}{a}\left(y_i - \mu\right)}{\mu\left(\frac{1 + a\mu}{a}\right)}$$

(lanjutan)

$$\left(\frac{\partial \ln f(y_i)}{\partial \mu}\right)^2 = \frac{\frac{1}{a^2}(y_i-\mu)^2}{\mu^2\left(\frac{1+a\mu}{a}\right)^2}$$

$$= \frac{(y_i-\mu)^2}{\mu^2(1+a\mu)^2}$$

$$E\left[\left(\frac{\partial \ln f(y_i)}{\partial \mu}\right)^2\right] = E\left[\frac{(y_i-\mu)^2}{\mu^2(1+a\mu)^2}\right]$$

$$I(\mu) = \frac{1}{\mu^2(1+a\mu)^2} E\left[(y_i-\mu)^2\right]$$

$$I(\mu) = \frac{1}{\mu^2(1+a\mu)^2}\left(\mu + a\mu^2\right)$$

$$I(\mu) = \frac{1}{\mu(1+a\mu)}$$

Sebelumnya telah diketahui bahwa variansi dari $\hat{\mu} = \bar{Y}$ adalah $(\mu + a\mu^2)/n$, maka untuk memeriksa apakah $\hat{\mu} = \bar{Y}$ merupakan penaksir yang efisien, akan dilihat apakah variansi dari $\hat{\mu} = \bar{Y}$ akan mencapai batas bawah dari persamaan *Rao-Cramér*.

$$\frac{1}{nI(\mu)} = \frac{1}{n \cdot \frac{1}{\mu(1+a\mu)}}$$

$$= \frac{\mu(1+a\mu)}{n}$$

$$\sigma^2_{\hat{\mu}=\bar{Y}} \geq \frac{1}{nI(\mu)}$$

$$\frac{\mu(1+a\mu)}{n} \geq \frac{\mu(1+a\mu)}{n}$$

Sehingga diperoleh variansi dari $\hat{\mu} = \bar{Y}$ menyentuh batas bawah persamaan *Rao-Cramér*, dan taksiran $\hat{\mu} = \bar{Y}$ dikatakan penaksir yang efisien untuk μ.

Lampiran 3. Distribusi Poisson sebagai Anggota Keluarga Eksponensial

Suatu distribusi dikatakan sebagai anggota dari keluarga eksponensial jika memiliki bentuk fungsi probabilitas seperti berikut :

$$\begin{aligned} f(x;\theta) &= \exp\left[p(\theta)K(x)+S(x)+q(\theta)\right], \quad x = a_1, a_2, a_3, ... \\ &= 0 \qquad\qquad\qquad\qquad\qquad\qquad , \text{lainnya} \end{aligned} \tag{1}$$

Selanjutnya akan dibuktikan bahwa distribusi Poisson memiliki bentuk seperti persamaan (1) diatas. Misalkan Y adalah peubah acak yang berdistribusi Poisson dan bergantung pada parameter λ, maka fungsi probabilitasnya adalah

$$\begin{aligned} f(y) &= \frac{\lambda^y e^{-\lambda}}{y!} \quad , \lambda > 0 \;, y = 0,1,2,... \\ &= 0 \qquad\quad , \text{lainnya} \end{aligned} \tag{2}$$

Fungsi probabilitas pada persamaan (2) diatas dapat dinyatakan dalam bentuk eksponensial, sehingga persamaannya menjadi

$$\begin{aligned} f(y,\lambda) &= \exp\left[y \ln\lambda + \ln\left(\frac{1}{y!}\right) - \lambda\right] \;, \lambda > 0 \;, y = 0,1,2,... \\ &= 0 \qquad\quad , \text{lainnya} \end{aligned} \tag{3}$$

Persamaan pada persamaan (3) diatas memiliki bentuk yang similar dengan persamaan (1), dimana

$$S(y) = \ln\left(\frac{1}{y!}\right)$$

$$q(\lambda) = (-\lambda)$$

$$p(\lambda) = \ln(\lambda)$$

$$K(y) = y$$

Sehingga distribusi Poisson merupakan anggota dari keluarga eksponensial.

Lampiran 4. Bukti Persamaan (3.4) sebagai Fungsi Probabilitas

Suatu fungsi *f*(*x*) dikatakan sebagai fungsi probabilitas jika memenuhi :

- $f(x) \geq 0, \quad \forall x$
- $\sum_x f(x) = 1$

Akan dibuktikan bahwa persamaan (3.4) memenuhi kedua syarat diatas, sehingga dapat dikatakan sebagai fungsi probabilitas. Yang pertama akan dibuktikan bahwa persamaan (3.4) bernilai non-negatif untuk semua *y*.

$$f(y) = \frac{\Gamma\left(y+\frac{1}{a}\right)}{\Gamma\left(\frac{1}{a}\right)y!}\left(\frac{\mu}{\mu+\frac{1}{a}}\right)^{y}\left(\frac{\frac{1}{a}}{\mu+\frac{1}{a}}\right)^{\frac{1}{a}}, \quad \mu > 0, \ \frac{1}{a} > 0 \ , \ y = 0,1,2,... \qquad (3.4)$$

Dengan $\left(\frac{\mu}{\mu+\frac{1}{a}}\right)+\left(\frac{\frac{1}{a}}{\mu+\frac{1}{a}}\right)=1$. Dimana $\Gamma(\alpha) = \int_0^\infty e^{-t}t^{\alpha-1}dt$ akan bernilai positif

$\{\Gamma(\alpha)\} > 0$ untuk nilai α > 0. Karena nilai $\mu, y \geq 0$, $\frac{1}{a} > 0$ serta $\Gamma\left(y+\frac{1}{a}\right) > 0$

dan $\Gamma\left(\frac{1}{a}\right) > 0$, maka dapat

disimpulkan bahwa $f(y) \geq 0, \ \forall y$.

Selanjutnya akan akan dibuktikan bahwa jumlah dari persamaan (3.4) untuk seluruh nilai *y* bernilai satu.

$$\sum_{y=0}^{\infty} f(y) = \sum_{y=0}^{\infty}\frac{\Gamma\left(y+\frac{1}{a}\right)}{\Gamma\left(\frac{1}{a}\right)y!}\left(\frac{\mu}{\mu+\frac{1}{a}}\right)^{y}\left(\frac{\frac{1}{a}}{\mu+\frac{1}{a}}\right)^{\frac{1}{a}} \qquad (1)$$

(lanjutan)

Dalam fungsi Gamma, dapat dibuktikan bahwa :

$$\Gamma(\alpha)=\int_0^\infty t^{\alpha-1}e^{-t}dt$$

$$=\lim_{M\to\infty}\int_0^M t^{\alpha-1}e^{-t}dt \qquad (2)$$

Misalkan $u=t^{\alpha-1}$ dan $dv=e^{-t}dt$, sehingga akan didapat $v=-e^{-t}$ dan $du=(\alpha-1)t^{\alpha-2}$. Persamaan (2) akan menjadi

$$\Gamma(\alpha)=\lim_{M\to\infty}\left\{\left(t^{\alpha-1}\right)\left(-e^{-t}\right)\Big]_0^M-\int_0^M -e^{-t}(\alpha-1)t^{\alpha-2}dt\right\}$$

$$=\lim_{M\to\infty}\left\{\left(-M^{\alpha-1}e^{-M}\right)+(\alpha-1)\int_0^M e^{-t}t^{\alpha-2}dt\right\}$$

$$=\lim_{M\to\infty}\left(-M^{\alpha-1}e^{-M}\right)+\lim_{M\to\infty}(\alpha-1)\int_0^M e^{-t}t^{\alpha-2}dt$$

$$=0+(\alpha-1)\int_0^\infty t^{\alpha-2}e^{-t}dt$$

$$=(\alpha-1)\Gamma(\alpha-1)$$

Sehingga bagian $\dfrac{\Gamma\left(y+\frac{1}{a}\right)}{\Gamma\left(\frac{1}{a}\right)y!}$ pada persamaan (3.4) dapat dinyatakan dalam bentuk

$$\frac{\Gamma\left(y+\frac{1}{a}\right)}{\Gamma\left(\frac{1}{a}\right)y!}=\frac{\left(y+\frac{1}{a}-1\right)\left(y+\frac{1}{a}-2\right)...\left(\frac{1}{a}\right)\left(\frac{1}{a}-1\right)\left(\frac{1}{a}-2\right)...}{y!\left(\frac{1}{a}-1\right)\left(\frac{1}{a}-2\right)...}$$

$$=\frac{\left(y+\frac{1}{a}-1\right)\left(y+\frac{1}{a}-2\right)...\left(\frac{1}{a}\right)}{y!} \qquad (3)$$

Suatu kombinasi dapat didefinsikan untuk setiap bilangan riil a ,

$$\binom{a}{y}=\begin{cases}\dfrac{(a)(a-1)...(a-y+1)}{y!} & , \ y=1,2,...\\ 1 & , \ y=0\end{cases}$$

(lanjutan)

Sehingga persamaan (3) dapat ditulis menjadi,

$$\frac{\Gamma\left(y+\frac{1}{a}\right)}{\Gamma\left(\frac{1}{a}\right)y!}=\frac{\left(y+\frac{1}{a}-1\right)\left(y+\frac{1}{a}-2\right)...\left(\frac{1}{a}\right)}{y!}$$

$$=\begin{pmatrix} y+\frac{1}{a}-1 \\ y \end{pmatrix}$$

Maka persamaan (1) menjadi sebagai berikut :

$$\sum_{y=0}^{\infty} f(y)=\sum_{y=0}^{\infty}\begin{pmatrix} y+\frac{1}{a}-1 \\ y \end{pmatrix}\left(\frac{\mu}{\mu+\frac{1}{a}}\right)^{y}\left(\frac{\frac{1}{a}}{\mu+\frac{1}{a}}\right)^{\frac{1}{a}} \tag{4}$$

Kemudian perhatikan persamaan berikut

$$\frac{\left(y+\frac{1}{a}-1\right)...\left(\frac{1}{a}+1\right)\left(\frac{1}{a}\right)}{(y)!}=\frac{(-1)^{y}\left(-\frac{1}{a}\right)\left(-\frac{1}{a}-1\right)...\left(-\frac{1}{a}-y+1\right)}{(y)!}$$

$$=(-1)^{y}\begin{pmatrix} -\frac{1}{a} \\ y \end{pmatrix}$$

Sehingga didapat persamaan

$$\begin{pmatrix} y+\frac{1}{a}-1 \\ y \end{pmatrix}=(-1)^{y}\begin{pmatrix} -\frac{1}{a} \\ y \end{pmatrix}$$

maka persamaan (4) dapat ditulis menjadi:

$$\sum_{y=0}^{\infty} f(y)=\sum_{y=0}^{\infty}(-1)^{y}\begin{pmatrix} -\frac{1}{a} \\ y \end{pmatrix}\left(\frac{\mu}{\mu+\frac{1}{a}}\right)^{y}\left(\frac{\frac{1}{a}}{\mu+\frac{1}{a}}\right)^{\frac{1}{a}}$$

$$=\sum_{y=0}^{\infty}\begin{pmatrix} -\frac{1}{a} \\ y \end{pmatrix}\left(\frac{-\mu}{\mu+\frac{1}{a}}\right)^{y}\left(\frac{\frac{1}{a}}{\mu+\frac{1}{a}}\right)^{\frac{1}{a}}$$

(lanjutan)

$$= \left(\frac{\frac{1}{a}}{\mu + \frac{1}{a}} \right)^{\frac{1}{a}} \sum_{y=0}^{\infty} \binom{-\frac{1}{a}}{y} \left(\frac{-\mu}{\mu + \frac{1}{a}} \right)^{y} \tag{5}$$

Persamaan $\sum_{y=0}^{\infty} \binom{-\frac{1}{a}}{y} \left(\frac{-\mu}{\mu + \frac{1}{a}} \right)^{y}$ merupakan suatu deret kuasa.

Selanjutnya akan diperiksa apakah deret tersebut konvergen. Uji rasio mutlak akan digunakan untuk mencari interval konvergensi dari deret tersebut. Suku ke-n dan n+1 dari deret tersebut adalah

$$U_n = \frac{-\frac{1}{a}\left(-\frac{1}{a}-1\right)\ldots\left(-\frac{1}{a}-n+1\right)}{n!} \left(-\frac{\mu}{\mu + \frac{1}{a}} \right)^{n}$$

$$U_{n+1} = \frac{-\frac{1}{a}\left(-\frac{1}{a}-1\right)\ldots\left(-\frac{1}{a}-(n+1)+1\right)}{(n+1)!} \left(-\frac{\mu}{\mu + \frac{1}{a}} \right)^{n+1}$$

sehingga akan didapat

$$\lim_{n\to\infty} \left| \frac{U_{n+1}}{U_n} \right| = \lim_{n\to\infty} \left| \frac{\frac{\left(-\frac{1}{a}\right)\left(-\frac{1}{a}-1\right)\ldots\left(-\frac{1}{a}-n+1\right)\left(-\frac{1}{a}-n\right)}{(n+1)n!} \left(-\frac{\mu}{\mu + \frac{1}{a}} \right)^{n+1}}{\frac{-\frac{1}{a}\left(-\frac{1}{a}-1\right)\ldots\left(-\frac{1}{a}-n+1\right)}{n!} \left(-\frac{\mu}{\mu + \frac{1}{a}} \right)^{n}} \right|$$

$$= \lim_{n\to\infty} \left| \frac{\left(-\frac{1}{a}-n\right)}{(n+1)} \left(-\frac{\mu}{\mu + \frac{1}{a}} \right) \right|$$

$$= \lim_{n\to\infty} \left| \frac{\left(-\frac{1}{a}-n\right)}{(n+1)} \right| \left| -\frac{\mu}{\mu+\frac{1}{a}} \right|$$

$$= \lim_{n\to\infty} \frac{\left(\frac{1}{a}+n\right)}{(n+1)} \left| -\frac{\mu}{\mu+\frac{1}{a}} \right|$$

$$= 1. \left| -\frac{\mu}{\mu+\frac{1}{a}} \right| < 1$$

Maka deret tersebut akan konvergen pada $\left| -\frac{\mu}{\mu+\frac{1}{a}} \right| < 1$. Karena $\mu \geq 0$ dan $a > 0$, maka didapat $-1 < -\frac{\mu}{\mu+\frac{1}{a}} < 0$, sehingga deret tersebut konvergen.

Setelah mengetahui deret tersebut konvergen, maka dapat dihitung jumlah tak hingga dari deret tersebut, dimana deret kuasa tersebut dapat didefinisikan sebagai berikut :

$$(1+x)^p = 1 + px + \frac{p(p-1)}{2!}x^2 + \ldots + \frac{p(p-1)(p-2)\ldots(p-n+1)}{n!}x^n + \ldots$$
$$= \sum_{n=0}^{\infty} \binom{p}{n} x^n \ , \ |x| < 1 \tag{6}$$

dimana p dan x merupakan bilangan riil serta n adalah bilangan bulat.

Untuk menghitung jumlah tak hingga dari deret $\sum_{y=0}^{\infty} \binom{-\frac{1}{a}}{y} \left(\frac{-\mu}{\mu+\frac{1}{a}} \right)^y$, akan digunakan persamaan deret kuasa diatas dengan $x = -\frac{\mu}{\mu+\frac{1}{a}}$, $p = -1/a$, dan $n = y$ pada persamaan (6) , sehingga persamaan (6) dapat ditulis menjadi

(lanjutan)

$$\left(1-\frac{\mu}{\mu+\frac{1}{a}}\right)^{-\frac{1}{a}}=1+-\frac{1}{a}\left(-\frac{\mu}{\mu+\frac{1}{a}}\right)+\frac{-\frac{1}{a}\left(-\frac{1}{a}-1\right)}{2!}\left(-\frac{\mu}{\mu+\frac{1}{a}}\right)^{2}+\ldots+$$

$$\frac{-\frac{1}{a}\left(-\frac{1}{a}-1\right)\ldots\left(-\frac{1}{a}-y+1\right)}{y!}\left(-\frac{\mu}{\mu+\frac{1}{a}}\right)^{y}+\ldots$$

$$=\sum_{y=0}^{\infty}\binom{-\frac{1}{a}}{y}\left(-\frac{\mu}{\mu+\frac{1}{a}}\right)^{y} \tag{7}$$

Pada deret di atas, karena nilai $a > 0$, sehingga didapat $-\frac{1}{a} < 0$. Deret tersebut akan konvergen jika $-1 < -\frac{\mu}{\mu+\frac{1}{a}} < 1$, dengan $\mu \geq 0$ dan $a > 0$.

Sehingga deret kuasa pada persamaan (7) akan konvergen ke suatu nilai yaitu :

$$\sum_{y=0}^{\infty}\binom{-\frac{1}{a}}{y}\left(-\frac{\mu}{\mu+\frac{1}{a}}\right)^{y}=\left(1-\frac{\mu}{\mu+\frac{1}{a}}\right)^{-\frac{1}{a}}$$

$$=\left(\frac{\frac{1}{a}}{\mu+\frac{1}{a}}\right)^{-\frac{1}{a}} \tag{8}$$

Sehingga dengan mensubtitusikan persamaan (8) ke dalam persamaan (5) akan didapat

$$\sum_{y=0}^{\infty}f(y)=\left(\frac{\frac{1}{a}}{\mu+\frac{1}{a}}\right)^{\frac{1}{a}}\sum_{y=0}^{\infty}\binom{-\frac{1}{a}}{y}\left(\frac{-\mu}{\mu+\frac{1}{a}}\right)^{y}$$

(lanjutan)

$$= \left(\frac{\frac{1}{a}}{\mu + \frac{1}{a}} \right)^{\frac{1}{a}} \left(\frac{\frac{1}{a}}{\mu + \frac{1}{a}} \right)^{-\frac{1}{a}}$$

$$= 1$$

Karena telah dibuktikan bahwa $f(y) \geq 0, \;\; \forall y$ dan $\sum_{y=0}^{\infty} f(y) = 1$, maka terbukti bahwa persamaan (3.4) merupakan suatu fungsi probabilitas.

Lampiran 5. Menunjukkan $\Delta = \left\{ \left(\frac{\partial l^2(a,\mu)}{\partial \mu^2} \right) \left(\frac{\partial l^2(a,\mu)}{\partial a^2} \right) - \left(\frac{\partial l^2(a,\mu)}{\partial a \partial \mu} \right)^2 \right\} > 0$

dan $\frac{\partial l^2(a,\mu)}{\partial \mu^2} < 0$ dan $\frac{\partial l^2(a,\mu)}{\partial a^2} < 0$

Sebelumnya telah diketahui bahwa turunan parsial kedua terhadap a dan μ adalah :

$$\frac{\partial l^2(a,\mu)}{\partial \mu^2} = \sum_{i=1}^{n} -\frac{y_i}{\mu^2} + \frac{(y_i a + 1)\, a}{(1+a\mu)^2}$$

$$\frac{\partial l^2(a,\mu)}{\partial a^2} = \sum_{i=1}^{n} \left[\begin{array}{l} \left(\sum_{r=1}^{y_i - 1} - \left(\frac{r}{(1+ar)} \right)^2 \right) - \frac{2}{a^3} \log(1+a\mu) + \frac{1}{a^2} \frac{\mu}{1+a\mu} + \\ \frac{y_i \mu^2}{(1+a\mu)^2} - \frac{-\left(\frac{\mu}{a^2}(1+a\mu) + \frac{\mu^2}{a} \right)}{(1+a\mu)^2} \end{array} \right]$$

$$\frac{\partial l(a,\mu)}{\partial a \partial \mu} = \sum_{i=1}^{n} \frac{1}{a(1+a\mu)} - \left(\frac{y_i + \frac{1}{a}}{(1+a\mu)^2} \right)$$

Nilai μ yang diperoleh sebagai hasil dari persamaan (3.20) adalah $\bar{y}$ sementara nilai a sebagai solusi eksak hasil persamaan (3.21) tidak dapat diperoleh.

Selanjutnya akan dipeiksa apakah $\frac{\partial l^2(a,\mu)}{\partial \mu^2} < 0$ dan $\frac{\partial l^2(a,\mu)}{\partial a^2} < 0$ serta

$$\Delta = \left\{ \left(\frac{\partial l^2(a,\mu)}{\partial \mu^2} \right) \left(\frac{\partial l^2(a,\mu)}{\partial a^2} \right) - \left(\frac{\partial l^2(a,\mu)}{\partial a \partial \mu} \right)^2 \right\} > 0 \text{ saat } \mu = \bar{y}.$$

(lanjutan)

- Memeriksa $\frac{\partial l^2(a,\mu)}{\partial \mu^2} < 0$

$$
\begin{aligned}
\frac{\partial l^2(a,\mu)}{\partial \mu^2} &= \sum_{i=1}^{n} -\frac{y_i}{\mu^2} + \frac{\left(y_i a + 1\right)a}{(1+a\mu)^2} \\
&= \sum_{i=1}^{n} -\frac{y_i}{\bar{y}^2} + \frac{\left(y_i a + 1\right)a}{(1+a\bar{y})^2} \\
&= -\frac{1}{\bar{y}^2}\sum_{i=1}^{n} y_i + \frac{a}{(1+a\bar{y})^2}\left[\sum_{i=1}^{n} 1 + \sum_{i=1}^{n} y_i a\right] \\
&= -\frac{n\bar{y}}{\bar{y}^2} + \frac{a}{(1+a\bar{y})^2}\left[n + an\bar{y}\right] \\
&= -\frac{n}{\bar{y}} + \frac{na}{(1+a\bar{y})^2}\left(1+a\bar{y}\right) \\
&= -\frac{n}{\bar{y}} + \frac{na}{(1+a\bar{y})} \\
&= n\left(-\frac{1}{\bar{y}} + \frac{a}{(1+a\bar{y})}\right) \\
&= n\left(-\frac{a\bar{y} - 1 - a\bar{y}}{\bar{y}(1+a\bar{y})}\right) \\
&= n\left(-\frac{1}{\bar{y}(1+a\bar{y})}\right) < 0
\end{aligned}
$$

- Memeriksa $\frac{\partial l^2(a,\mu)}{\partial a^2} < 0$

$$
\frac{\partial l^2(a,\mu)}{\partial a^2} = \sum_{i=1}^{n} \left[\begin{array}{l} \left(\sum_{r=1}^{y_i-1} -\left(\frac{r}{(1+ar)}\right)^2\right) - \frac{2}{a^3}\log(1+a\mu) + \frac{1}{a^2}\frac{\mu}{1+a\mu} + \\ \frac{y_i\mu^2}{(1+a\mu)^2} - \frac{-\left(\frac{\mu}{a^2}(1+a\mu) + \frac{\mu^2}{a}\right)}{(1+a\mu)^2} \end{array} \right]
$$

(lanjutan)

$$= \sum_{i=1}^{n} \left[\begin{array}{l} \left(\sum_{r=1}^{y_i-1} - \left(\dfrac{r}{(1+ar)} \right)^2 \right) - \dfrac{2}{a^3} \log(1+a\bar{y}) + \dfrac{1}{a^2} \dfrac{\bar{y}}{1+a\bar{y}} + \\ \dfrac{y_i \bar{y}^2}{(1+a\bar{y})^2} - \dfrac{-\left(\dfrac{\bar{y}}{a^2}(1+a\bar{y}) + \dfrac{\bar{y}^2}{a} \right)}{(1+a\bar{y})^2} \end{array} \right]$$

$$= \sum_{i=1}^{n} \left[\begin{array}{l} \left(\sum_{r=1}^{y_i-1} - \left(\dfrac{r}{(1+ar)} \right)^2 \right) - \dfrac{2}{a^3} \log(1+a\bar{y}) + \dfrac{1}{a^2} \dfrac{\bar{y}}{1+a\bar{y}} \\ - \dfrac{-\left(\dfrac{\bar{y}}{a^2}(1+a\bar{y}) + \bar{y}^2 \left(y_i + \dfrac{1}{a} \right) \right)}{(1+a\bar{y})^2} \end{array} \right]$$

$$= \sum_{i=1}^{n} \left[\begin{array}{l} \left(\sum_{r=1}^{y_i-1} - \left(\dfrac{r}{(1+ar)} \right)^2 \right) - \dfrac{2}{a^3} \log(1+a\bar{y}) + \dfrac{1}{a^2} \dfrac{\bar{y}}{1+a\bar{y}} \\ + \dfrac{1}{a^2} \dfrac{\bar{y}}{1+a\bar{y}} + \dfrac{\bar{y}^2 \left(y_i + \dfrac{1}{a} \right)}{(1+a\bar{y})^2} \end{array} \right]$$

$$= \sum_{i=1}^{n} \left[\begin{array}{l} \left(\sum_{r=1}^{y_i-1} - \left(\dfrac{r}{(1+ar)} \right)^2 \right) - \dfrac{2}{a^3} \log(1+a\bar{y}) + 2 \dfrac{1}{a^2} \dfrac{\bar{y}}{1+a\bar{y}} + \\ \dfrac{\bar{y}^2 \left(y_i + \dfrac{1}{a} \right)}{(1+a\bar{y})^2} \end{array} \right]$$

$$= \left[\begin{array}{l} \sum_{i=1}^{n} \left(\sum_{r=1}^{y_i-1} - \left(\dfrac{r}{(1+ar)} \right)^2 \right) - \sum_{i=1}^{n} \dfrac{2}{a^3} \log(1+a\bar{y}) + \sum_{i=1}^{n} \dfrac{2\bar{y}}{a^2 (1+a\bar{y})} + \\ \sum_{i=1}^{n} \dfrac{\bar{y}^2 \left(y_i + \dfrac{1}{a} \right)}{(1+a\bar{y})^2} \end{array} \right]$$

$$= \left[\sum_{i=1}^{n} \left(\sum_{r=1}^{y_i-1} - \left(\dfrac{r}{(1+ar)} \right)^2 \right) - \dfrac{2n}{a^3} \log(1+a\bar{y}) + \dfrac{2n\bar{y}}{a^2 (1+a\bar{y})} + \dfrac{\bar{y}^2 \left(\sum_{i=1}^{n} y_i + \dfrac{n}{a} \right)}{(1+a\bar{y})^2} \right]$$

(lanjutan)

$$= \left[\sum_{i=1}^{n} \left(\sum_{r=1}^{y_i - 1} -\left(\frac{r}{(1+ar)} \right)^2 \right) - \frac{2n}{a^3} \log(1+a\bar{y}) + \frac{2n\bar{y}}{a^2(1+a\bar{y})} + \frac{\bar{y}^2 \left(n\bar{y} + \dfrac{n}{a} \right)}{(1+a\bar{y})^2} \right]$$

$$= \left[\sum_{i=1}^{n} \left(\sum_{r=1}^{y_i - 1} -\left(\frac{r}{(1+ar)} \right)^2 \right) - \frac{n}{a^2} \left[\frac{2\bar{y} + a\bar{y}^2}{1+a\bar{y}} - \frac{2\ln(1+a\bar{y})}{a} \right] \right]$$

$$= \left[\sum_{i=1}^{n} \left(\sum_{r=1}^{y_i - 1} -\left(\frac{r}{(1+ar)} \right)^2 \right) - \frac{n}{a^2} \left[\frac{2a\bar{y} + a^2\bar{y}^2 - 2(1+a\bar{y})\ln(1+a\bar{y})}{a(1+a\bar{y})} \right] \right]$$

$$= \left[\sum_{i=1}^{n} \left(\sum_{r=1}^{y_i - 1} -\left(\frac{r}{(1+ar)} \right)^2 \right) - \frac{n}{a^3(1+a\bar{y})} \left[a\bar{y} + a\bar{y}(1+a\bar{y}) - 2(1+a\bar{y})\ln(1+a\bar{y}) \right] \right]$$

$$= \left\{ \sum_{i=1}^{n} - \left[\left(\sum_{r=1}^{y_i - 1} \left(\frac{r}{(1+ar)} \right)^2 \right) + \frac{n}{a^3(1+a\bar{y})} \left[a\bar{y} + a\bar{y}(1+a\bar{y}) - 2(1+a\bar{y})\ln(1+a\bar{y}) \right] \right] \right\} < 0$$

- Memeriksa $\Delta = \left\{ \left(\dfrac{\partial l^2(a,\mu)}{\partial \mu^2} \right) \left(\dfrac{\partial l^2(a,\mu)}{\partial a^2} \right) - \left(\dfrac{\partial l^2(a,\mu)}{\partial a \partial \mu} \right)^2 \right\} > 0$

Sebelum mengitung Δ, akan dihitung $\dfrac{\partial l(a,\mu)}{\partial a \partial \mu}$ terlebih dahulu, yaitu :

$$\frac{\partial l(a,\mu)}{\partial a \partial \mu} = \sum_{i=1}^{n} \frac{1}{a(1+a\mu)} - \left(\frac{y_i + \dfrac{1}{a}}{(1+a\mu)^2} \right)$$

$$= \sum_{i=1}^{n} \frac{1}{a(1+a\bar{y})} - \left(\frac{y_i + \dfrac{1}{a}}{(1+a\bar{y})^2} \right)$$

$$= \sum_{i=1}^{n} \frac{1}{a(1+a\bar{y})} - \sum_{i=1}^{n} \left(\frac{y_i + \dfrac{1}{a}}{(1+a\bar{y})^2} \right)$$

(lanjutan)

$$= \frac{n}{a(1+a\bar{y})} - \frac{1}{(1+a\bar{y})^2}\left(\sum_{i=1}^{n} y_i + \sum_{i=1}^{n}\frac{1}{a}\right)$$

$$= \frac{n}{a(1+a\bar{y})} - \frac{\left(n\bar{y} + \frac{n}{a}\right)}{(1+a\bar{y})^2}$$

$$= \frac{n(1+a\bar{y}) - n\left(a\bar{y} + \frac{a}{a}\right)}{a(1+a\bar{y})^2}$$

$$= 0$$

Karena nilai $\frac{\partial l^2(a,\mu)}{\partial \mu^2} < 0$ dan $\frac{\partial l^2(a,\mu)}{\partial a^2} < 0$, maka persamaan $\left(\frac{\partial l^2(a,\mu)}{\partial \mu^2}\right)\left(\frac{\partial l^2(a,\mu)}{\partial a^2}\right) > 0$ atau bernilai positif. Sebelumnya telah diketahui bahwa $\frac{\partial l(a,\mu)}{\partial a \partial \mu} = 0$, sehingga akan didapat bahwa

$$\Delta = \left\{\left(\frac{\partial l^2(a,\mu)}{\partial \mu^2}\right)\left(\frac{\partial l^2(a,\mu)}{\partial a^2}\right) - \left(\frac{\partial l^2(a,\mu)}{\partial a \partial \mu}\right)^2\right\} > 0 .$$

Lampiran 6. Algoritma *Newton-Rhapson*

```
%%% Newton-Rhapson %%%
function newton
miu=input('masukkan nilai miu : ');
F=fungsi(miu);
l=diff(F);
TOL=10^-6;

p0=input('Masukkan perkiraan akar awal p0:');
N=input('Iterasi maksimum yang diinginkan:');
fprintf('\n--------------------------\n');
fprintf('|%3s  | %10s       |\n','n','pn');
fprintf('--------------------------\n');
fprintf('|  0  | %16.10f |\n',p0);

for(i=1:N)
   p=p0-subs(F,p0)/subs(l,p0);
   fprintf('|%3.d  | %16.10f |\n',i,p);
   if(abs(p-p0)<TOL)
      fprintf('--------------------------\n\n');
      fprintf('Nilai aproksimasi akar dari l(a)\n');
      fprintf('yang berada dalam [0,inf) adalah a = %.10f\n',p);
      break;
   end
   p0=p;
end
```

Lampiran 7. Algoritma Dari Fungsi Log *Likelihood*

```
function [F]=fungsi(miu)
load data.mat Y
syms a;
N=length(Y);
m=0;
for i=1:N
   for r=1:(Y(i)-1)
      m=m+r/(1+a*r);
   end
end
F=m+(N/a^2)*log(1+a*miu)-miu/(1+a*miu)*(N/a+sum(Y));
```

Lampiran 8. Algoritma Dari Statistik Uji *A*

```
function [A]=fungsiA(miu)
load datay.mat datay
N=length(datay);
m=0;
miu=input('masukkan nilai miu : ');
a=input('masukkan nilai taksiran a : ');
for i=1:N
   for r=1:(datay(i)-1)
      m=m+log(1+a*r);
   end
end
fprintf('Nilai dari statistik uji A adalah \n');
A=-2*((-N*miu)+((N/a)+sum(datay))*log(1+a*miu)-m)
```