

PENYUAIAN TABURAN BAGI DATA KEMALANGAN
JALAN RAYA DI MALAYSIA

THONG SZE KEE

PERPUSTAKAAN
UNIVERSITI MALAYSIA SABAH

DISERTASI INI DIKEMUKAKAN UNTUK MEMENUHI
SEBAHAGIAN DARIPADA SYARAT MEMPEROLEHI IJAZAH
SARJANA MUDA SAINS DENGAN KEPUJIAN

PROGRAM MATEMATIK DENGAN EKONOMI
SEKOLAH SAINS DAN TEKNOLOGI
UNIVERSITI MALAYSIA SABAH

April 2007

BORANG PENGESAHAN STATUS TESIS@

JUDUL: PENYUAIAN TABURAN BAGI DATA KEMALANGAN JALAN RAYA
DI MALAYSIA

Ijazah: Sarjana Muda Sains dengan Kejuruteraan Matematik Dengan Ekonomi

SESI PENGAJIAN: 2004/2005

Saya THONG SZE KEE

(HURUF BESAR)

mengaku membenarkan tesis (~~LPS~~ Sarjana/Doktor Falsafah)* ini disimpan di Perpustakaan Universiti Malaysia Sabah dengan syarat-syarat kegunaan seperti berikut:

1. Tesis adalah hakmilik Universiti Malaysia Sabah.
2. Perpustakaan Universiti Malaysia Sabah dibenarkan membuat salinan untuk tujuan pengajian sahaja.
3. Perpustakaan dibenarkan membuat salinan tesis ini sebagai bahan pertukaran antara institusi pengajian tinggi.
4. **Sila tandakan (/)

☐

SULIT

(Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia seperti yang termaktub di dalam AKTA RAHSIA RASMI 1972)

☒

TERHAD

(Mengandungi maklumat TERHAD yang telah ditentukan oleh organisasi/badan di mana penyelidikan dijalankan)

☐

TIDAK TERHAD

Disahkan oleh

Thong Sze Kee
(TANDATANGAN PENULIS)

P. Darnesah Gabda
(TANDATANGAN PUSTAKAWAN)

Alamat Tetap: 79A, JLN NEGAT,
83000 BATU PAHAT,
JOHOR.

PN. DARNESAH GABDA
Nama Penyelia

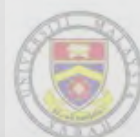
Tarikh: 23.4.2007

Tarikh: _____

CATATAN: * Potong yang tidak berkenaan.

** Jika tesis ini SULIT atau TERHAD, sila lampirkan surat daripada pihak berkuasa/organisasi berkenaan dengan menyatakan sekali sebab dan tempoh tesis ini perlu dikelaskan sebagai SULIT dan TERHAD.

@ Tesis dimaksudkan sebagai tesis bagi Ijazah Doktor Falsafah dan Sarjana secara penyelidikan, atau disertasi bagi pengajian secara kerja kursus dan penyelidikan, atau Laporan Projek Sarjana Muda (LPSM).

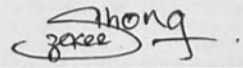


UMS
UNIVERSITI MALAYSIA SABAH

PENGAKUAN

Saya akui karya ini adalah hasil kerja saya sendiri kecuali nukilan dan ringkasan yang setiap satunya telah dijelaskan sumbernya.

23 April 2007



THONG SZE KEE

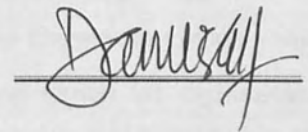
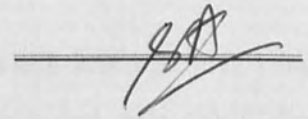
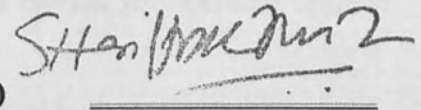
HS2004-1186



UMS
UNIVERSITI MALAYSIA SABAH

PERAKUAN PEMERIKSA**DIPERAKUKAN OLEH**

Tandatangan

1. PENYELIA**(Pn. Darmesah Gabda)****2. PEMERIKSA****(Cik Suriani Hassan)****3. DEKAN****(Supt/Ks Prof Madya Dr. Shariff A.K Omang)**

PENGHARGAAN

Puji dan syukur kepada Tuhan kerana dengan kasihNya telah dapat saya melengkapkan kajian penyelidikan dan penulisan disertasi ini.

Saya ingin mengambil kesempatan ini untuk merakamkan setinggi-tinggi ribuan terima kasih kepada penyelia penulisan disertasi saya, iaitu Puan Darmesah Gabda yang banyak membantu, memberikan panduan dan nasihat sepanjang kajian ini dijalankan. Atas bimbingan beliau, saya dapat menjalankan kajian disertasi dengan lancarnya serta berjaya menyiapkan kajian penyelidikan dan penulisan disertasi saya.

Selain itu, saya ingin mengucapkan terima kasih kepada ahli keluarga saya yang sentiasa banyak memberi dorongan kepada saya serta sokongan dari segi moral dan kewangan. Akhir sekali kepada rakan-rakan saya kerana banyak memberikan bantuan kepada saya apabila saya menghadapi masalah.

ABSTRAK

Tujuan kajian ini dijalankan adalah untuk mencari menyuaikan model bagi data kemalangan jalan raya di Malaysia yang diperolehi dari Polis Di Raja Malaysia untuk tahun 1972 hingga 2005. Dua model yang dipertimbangkan adalah regresi poisson dan regresi binomial negatif. Kaedah kebolehjadian maksimum digunakan untuk menganggar parameter model dan ujian serakan dijalankan untuk menentukan kewujudan serakan berketerlaluan dalam data. Ujian penerimaan model yang melibatkan ujian statistik devians dan khi-kuasa dua Pearson digunakan untuk menilai kebagusan model. Kriteria AIC dan Schwarz merupakan kriteria pemilihan model di mana ia digunakan untuk menentukan model yang paling sesuai untuk sampel. Kewujudan serakan berketerlaluan menunjukkan model regresi binomial negatif ialah model yang paling sesuai untuk data tersebut.



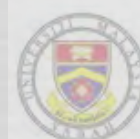
DETERMINE FITTING DISTRIBUTION FOR MALAYSIA'S ROAD ACCIDENT DATA

ABSTRACT

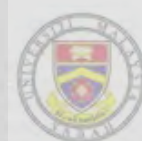
This study is to determine the best fit distribution for Malaysia's road accident data which have been gotten from Royal Malaysian Police for year 1972 to 2005. Two types of models which have been considered were poisson regression and negative binomial regression. Maximum likelihood method was used to estimate the unknown parameters of these models. Dispersion test was used to determine the existence of over-dispersion of the data. The model acceptability test which involved the Deviance and Pearson chi-square statistic test have been used to evaluate the goodness of fit of the models. The AIC and Schwarz criterion are the criteria of model selection which was used to determine the best fit model for the sample. The existence of the over-dispersion indicated that the regression negative binomial model fits the data best.

KANDUNGAN

	Muka Surat
PENGAKUAN	ii
PENGESAHAN	iii
PENGHARGAAN	iv
ABSTRAK	v
ABSTRACT	vi
SENARAI KANDUNGAN	vii
SENARAI JADUAL	ix
SENARAI RAJAH	x
SENARAI SIMBOL	xi
SENARAI SINGKATAN	xii
BAB 1 PENDAHULUAN	
1.1 Pengenalan	1
1.2 Definisi Jalan	3
1.3 Model bagi Data Diskret	5
1.4 Jenis Taburan	5
1.4.1 Taburan Poisson	5
1.4.2 Taburan Binomial Negatif	6
1.5 Kepentingan Kajian	6
1.6 Objektif Kajian	7
1.7 Skop Kajian	7
BAB 2 ULASAN LITERATUR	8
BAB 3 METODOLOGI	
3.1 Pengenalan	16
3.2 Data	16
3.3 Model Linear Teritlak	17
3.3.1 Model Regresi Poisson	19
3.3.1.1 Taburan Poisson	19
3.3.1.2 Model Poisson	21
3.3.1.3 Spesifikasi Model Poisson	21
3.3.2 Model Regresi Binomial Negatif	22



3.3.2.1 Taburan Binomial Negatif	22
3.3.2.2 Spesifikasi Model Binomial Negatif	24
3.4 Penganggar Parameter	24
3.4.1 Kaedah Kebolehjadian Maksimum	25
3.4.1.1 Fungsi Kebolehjadian Poisson	25
3.4.1.2 Fungsi Kebolehjadian Binomial Negatif	26
3.5 Kaedah Newton-Raphson	27
3.6 Ujian Kebagusan Penyuaian	28
3.6.1 Kebagusan Penyuaian Devians	28
3.6.2 Kebagusan Penyuaian Khi-kuasa Dua Pearson	29
3.7 Ujian Serakan Berketerlaluan	29
3.8 Pemilihan Model	31
3.8.1 Kriteria Informasi Akaike	31
3.8.2 Kriteria Schwarz	32
3.9 Bantuan Perisian Komputer	33
BAB 4 HASIL KAJIAN	
4.1 Pengenalan	35
4.2 Model-model Regresi	36
4.3 Penganggar Berparameter	39
4.4 Ujian Kebagusan Penyuaian	41
4.5 Ujian Serakan Berketerlaluan	42
4.6 Ujian Pemilihan Model	44
BAB 5 PERBINCANGAN DAN KEPUTUSAN	
5.1 Perbincangan	46
5.1 Kesimpulan	48
5.2 Cadangan	49
RUJUKAN	50
LAMPIRAN	53



SENARAI JADUAL

No. Jadual	Muka Surat
4.1 Jumlah kemalangan jalan raya yang berlaku di Malaysia pada tahun 1972 hingga 2005	37
4.2 Purata dan sisihan piawai kemalangan jalan raya yang berlaku pada tahun 1972 hingga 2005.	39
4.3 Penganggaran parameter dan nilai $\ln L(\theta)$ yang maksimum bagi taburan poisson dan binomial negatif.	40
4.4 Ujian Kebagusan Penyuaian bagi Model Regresi Poisson	41
4.5 Ujian Kebagusan Penyuaian bagi Model Regresi Binomial Negatif	42
4.6 Statistik kriteria informasi akaike dan Schwarz bagi model regresi poisson dan model regresi binomial negatif.	45



SENARAI RAJAH

No. Rajah		Muka Surat
3.1	Taburan kebarangkalian poisson dengan min 0.5, 1, 2, dan 5.	20
3.2	Taburan kebarangkalian binomial negatif dengan alfa 0.01, 0.5, 1, dan 5	23
4.1	Plot sebaran bagi jumlah kemalangan jalan raya yang berlaku pada tahun 1972 hingga 2005.	38



SENARAI SIMBOL

α	parameter serakan berketerlaluan
$Prob$	kebarangkalian
y_i	kekerapan cerapan
e	eksponen
x_i	regresor
λ_i	min bagi taburan poisson
β	parameter dalam model regresi
∂	pembezaan separa
E	jangkaan bilangan cerapan
var	varians
ε	pembolehubah tak diketahui
Γ	fungsi gamma
$\ln L$	fungsi logaritma kebolehjadian
n	bilangan cerapan
Σ	hasil tambah
$!$	faktorial
∂^2	pembezaan separa kedua
r_d	reja devians
X^2	Khi-kuasa dua
r_p	reja bagi khi-kuasa dua pearson
H_0	hipotesis nol
H_1	hipotesis alternatif
k	parameter serakan berketerlaluan dalam ujian serakan berketerlaluan
H	bilangan parameter
L	fungsi kebolehjadian maksimum
$=$	sama dengan
$\geq$	lebih besar daripada atau sama dengan

SENARAI SINGKATAN

AIC	Kriteria Informasi Akaike
BIC	Kriteria Informasi Bayesian
exp	eksponen
ESS	hasil tambah kuasa dua ralat
log	logaritma
ML	kabolehjadian maksimum
PDRM	Polis Di Raja Malaysia
SC	Schwarz
ZIP	Poisson Inflat sifar



BAB 1

PENDAHULUAN

1.1 PENGENALAN

Kini, pembangunan Malaysia semakin bertambah pesat dan hal ini telah banyak memberi kesan positif dan kesan negatif kepada negara kita. Salah satu kesan negatif adalah peningkatan jumlah kemalangan jalan raya yang berlaku setiap tahun. Peningkatan kemalangan ini berpunca daripada pertambahan bilangan kenderaan di Malaysia. Menurut kepada statistik Polis Di Raja Malaysia (PDRM), bilangan kenderaan yang berdaftar pada tahun 1972 adalah sebanyak 802, 831 buah. Manakala, pada tahun 2005, bilangan kenderaan yang berdaftar ialah 1020108, iaitu pertambahan sebanyak 27.06 peratus. Menurut Zullyadini Abdul Rahman dan Shaharuddib Ahmad (1999), perjalanan menggunakan kenderaan persendirian telah pun menjadi jenis pengangkutan yang digemari oleh pengguna untuk membuat pergerakan di antara bandar-bandar di Malaysia. Di Kuala Lumpur misalnya, perjalanan yang menggunakan pengangkutan awam adalah rendah berbanding dengan menggunakan kenderaan persendirian. Merujuk kepada sampel data kemalangan yang diperolehi, bilangan kemalangan yang berlaku bertambah dari tahun ke tahun.

Sehubungan dengan peningkatan dalam jumlah kenderaan, kadar kematian turut meningkat setiap tahun sekalipun pada kadar yang sederhana. Jika dibandingkan kadar kemalangan maut akibat kemalangan jalan raya di antara Malaysia dengan beberapa negara membangun dan sedang membangun yang lain, didapati Malaysia berada pada kedudukan pertengahan. (Zullyadini Abdul Rahman dan Shaharuddib Ahmad, 1999) Namun begitu, hal ini perlu diambil perhatian kerana kadar kematian per 10 000 kenderaan masih berada pada tahap yang tinggi jika dibandingkan dengan negara-negara membangun. Kadar kematian per 10 000 kenderaan di negara-negara membangun berada pada satu tahap yang konsisten iaitu kurang dari dua kematian per 10 000 kenderaan. (Zullyadini *et al.*, 1999) .

Selain itu, terdapat faktor-faktor lain yang menyebabkan kemalangan jalan raya iaitu memandu di bawah pengaruh alkohol, memecut, mengabaikan penggunaan alat keselamatan seperti penahan kanak-kanak, tali pinggang dan topi keledar, perancangan jalan raya yang lemah, reka cipta kenderaan yang tidak selamat, pelaksanaan langkah-langkah keselamatan jalan raya yang lemah dan perkhidmatan kesihatan kecemasan yang lemah. Kehilangan nyawa akibat kemalangan jalan raya memang mempunyai banyak implikasi terhadap keluarga mangsa secara khususnya dan kepada negara pada keseluruhannya. Implikasi yang paling jelas ialah dari segi kos kehilangan tenaga kerja.

Sebagai langkah menurunkan kadar kemalangan, pemodelan telah digunakan untuk menganalisis faktor-faktor yang mengakibatkan kemalangan jalan raya. Hubungan di antara kemalangan jalan raya dengan faktor-faktornya seperti jenis kemalangan, lokasi dan waktu kejadian kemalangan, keadaan persekitaran semasa

kemalangan, keadaan cuaca semasa, sifat dan keadaan jalan dan lain-lain lagi telah dikaji dengan menggunakan analisis regresi (Oh *et al.*, 2003). Analisis regresi yang digunakan biasanya ialah linear regresi dan linear regresi berganda. Anggapan linear regresi berganda ini adalah tidak realistik kerana analisis yang dijalankan adalah berkaitan dengan keadaan lalu lintas yang mungkin melibatkan kemalangan yang berjumlah sedikit.

Dengan itu, beberapa model telah digunakan untuk pemodelan data kemalangan jalan raya contohnya model regresi poisson, regresi binomial, regresi gamma, regresi multinomial dan lain-lain lagi (Winkelmann, 1997). Untuk membolehkan data kemalangan jalan raya Malaysia dapat dianalisis dengan model yang terbaik, maka satu kertas kerja yang melibatkan perbandingan model telah dijalankan.

1.2 DEFINISI JALAN

Dalam beberapa kes, terdapat kejadian kemalangan yang berlaku bukan di atas jalan sebagaimana yang ditaksir oleh seksyen (2)(a) Akta Pengangkutan Jalan, yang menakrif 'jalan' sebagai:

Mana-mana jalan awam dan mana-mana jalan lain yang awam mempunyai akses dan termasuk jambatan, terowong, hentian sebelah, kemudian feri, jalan bertingkat, bulatan jalan, pulau lalu lintas, pembahagi jalan, segala lorong lalu lintas, lorong laju, lorong perlahan, bahu jalan, garis tengah, jejambat, jalan tembok, jalan tuju, landasan masuk dan keluar, plaza tol, kawasan khidmat dan struktur dan lekapan

lain bagi memberi kesan sepenuh kepada penggunaanya; tetapi tidak termasuk mana-mana jalan persendirian jambatan, terowong atau apa-apa yang berhubung dengan jalan itu yang disenggara dan dijaga oleh orang atau badan persendirian.

Secara umumnya, jalan raya jalan besar yang menghubungkan satu kawasan dengan kawasan yang lain. Biasanya jalan besar ini mempunyai ciri-ciri berikut (Zullyadini *et al.*, 1999):

- (i) Digunakan untuk kenderaan bermotor.
- (ii) Digunakan untuk orang awam
- (iii) Dibiayai oleh badan awam
- (iv) Penggunaannya tertakluk kepada undang-undang pengangkutan

Sistem jalan raya bagi sesebuah negara diklasifikasikan kepada:

- (i) Lebuhraya
- (ii) Jalan raya kebangsaan ataupun laluan persekutuan
- (iii) Jalan raya negeri atau wilayah
- (iv) Jalan perbandaran
- (v) Jalan-jalan lain termasuk jalan kampung dan jalan estet

Di sini harus diingatkan bahawa bukan semua jalan yang dapat dilalui oleh kenderaan bermotor itu jalan raya. Contohnya denai-denai di dalam estet. Di Malaysia jalan raya yang sah haruslah diwartakan oleh pihak berkuasa.

1.3 MODEL BAGI DATA DISKRET

Data kemalangan jalan raya merupakan salah satu data diskret. Data diskret merujuk kepada pembolehubah yang boleh diukur dengan bilangan dan frekuensi data. Setiap cerapan bagi kejadian yang berlaku mempunyai domain yang tetap (fixed domain). Domain tetap ini dirujuk kepada masa (hari, bulan, tahun) atau tempat (unit geografi, individu dan lain-lain). Cerapan data adalah tak negatif, nilai integer, dan mungkin mempunyai sebahagian nilai yang bernombor sifar. Model bagi data ini adalah digunakan untuk membuat penjelasan dan ramalan. Model yang biasa digunakan adalah model regresi poisson dan binomial negatif.

1.4 JENIS TABURAN

Taburan-taburan yang umum digunakan untuk data kemalangan jalan raya adalah taburan poisson dan taburan binomial negatif (Greene, 2003). Di bawah merupakan pengenalan bagi taburan tersebut.

1.4.1 Taburan Poisson

Taburan poisson ialah salah satu taburan kebarangkalian diskret yang sangat berguna di mana pembolehubah rawaknya mewakili bilangan peristiwa yang berlaku secara merdeka pada suatu kadar yang malar. Contohnya: bilangan kemalangan yang berlaku di suatu jalan raya dalam sehari, bilangan panggilan telefon yang diterima di sebuah sekolah dari pukul 8 pagi hingga pukul 8.30 pagi, bilangan kesilapan menaip pada satu muka surat, bilangan tuntutan insurans yang dibuat dalam sebulan, bilangan

bakteria dalam 1ml air tasik dan lain-lain. Tambahan pula, ia memberi penghampiran yang baik kepada taburan binomial apabila p adalah kecil dan n adalah besar (di mana p ialah kebarangkalian 'kejayaan' yang berlaku dan n merupakan nombor cerapan) (Winkelmann, 1997).

1.4.2 Taburan Binomial Negatif

Taburan binomial negatif memberi suatu keadaan iaitu ia menggambarkan satu senario binomial di mana satu urutan percubaan yang tak bersandar telah diperhatikan dan kebarangkalian 'kejayaan' bagi setiap percubaan yang berlaku adalah malar iaitu p . Dengan menggantikan bilangan percubaan di n dan perhatikan bilangan 'kejayaan' yang berlaku serta meneruskan percubaan tak bersandar itu sehingga k 'kejayaan' benar-benar berlaku. Sekarang, pembolehubah rawak ialah bilangan percubaan bagi 'kejayaan', dan perlu memerhatikan sehingga k 'kejayaan' berlaku (Winkelmann, 1997).

1.5 KEPENTINGAN KAJIAN

Disebabkan peningkatan kadar kemalangan jalan raya, kini beberapa kajian perlu dijalankan untuk menganalisis faktor-faktor yang menyebabkan kemalangan jalan raya. Untuk menganalisis faktor-faktor tersebut, pemodelan ramalan telah dijalankan. Untuk membolehkan data kemalangan jalan raya Malaysia dapat dikaji secara efektif dan tepat, perbandingan model ramalan di antara model-model perlu dijalankan. Pemilihan model yang paling suai bagi data kemalangan jalan raya Malaysia dapat membantu jurutera atau pengkaji yang lain menjimatkan masa dan perbelanjaan

dengan menggunakan model yang paling suai semasa menjalankan kajian ke atas kemalangan jalan raya serta dapat mengkaji samada sasaran keselamatan telah dicapai.

1.6 OBJEKTIF KAJIAN

Objektif kajian ini adalah untuk membandingkan penyuaian model regresi poisson dan model regresi binomial negatif untuk data kemalangan jalan raya di Malaysia.

1.7 SKOP KAJIAN

Sampel yang digunakan adalah data kemalangan jalan raya di seluruh Malaysia. Data yang digunakan dalam kajian ini adalah berdasarkan jumlah kemalangan jalan raya yang berlaku selama 34 tahun, iaitu dari tahun 1972 hingga tahun 2005.

BAB 2

ULASAN LITERATUR

2.1 PENGENALAN

Mula-mula, pemodelan linear regresi berganda banyak digunakan dalam analisis kemalangan jalan raya. Berdasarkan andaian model regresi linear berganda, di mana pembolehubah bersandar model adalah selanjut dan tertabur secara normal dengan varians yang malar maka ia mempunyai kekurangan sifat taburan yang diperlukan untuk menjelaskan kerawakan dan diskret bagi kemalangan jalan raya yang berlaku (Miaou, 1994). Dengan itu, model regresi linear tidak boleh digunakan untuk membuat pernyataan kebarangkalian bagi kemalangan jalan raya. Tambahan pula, statistik ujian yang terdapat daripada model tersebut tidak dapat ditentukan benar (Miaou, 1994). Ketidakpuasan bagi sifat model regresi linear terhadap kemalangan jalan raya telah membawa kepada kajian lanjutan terhadap model poisson dan model regresi binomial negatif. (Maycock dan Hall 1984; Joshua dan Garber 1990).

McCullagh dan Nelder (1983) berpendapat penggunaan model poisson dapat dikembangkan dengan menggunakan teknik model linear teritlak (*Generalized linear models*). Walaupun model poisson mempunyai perhitungan yang tepat serta pencapaian pemodelan yang baik, tetapi masih mempunyai kelemahannya iaitu masalah yang berkaitan dengan fenomena “serakan berketerlaluan” telah berlaku di mana andaian bagi struktur ralat poisson tulen (*Pure Poisson Error*) ditunjukkan tidak begitu baik.

Dengan itu, Maycock dan Hall (1984) telah menunjukkan model binomial negatif dapat digunakan sebagai lanjutan bagi poisson tulen. Wedderburn (1974) pula mencadangkan model kuasiPoisson (*quasi-poisson*) untuk membetulkan model poisson tulen bagi masalah serakan berketerlaluan. Menurut Geirt dan Nuyts (2005), model ramalan dapat digunakan ke atas lebuhraya Flemish bagi kemalangan yang melibatkan jenis kemalangan kecederaan dan kematian. Data kemalangan yang terdapat akan diandai tertabur secara poisson atau binomial negatif. Menurut kepada Geirt dan Nuyts (2005), dengan menggunakan taburan binomial negatif, satu parameter serakan berketerlaluan (*over-dispersion*) telah digunakan dimana serakan berketerlaluan merupakan fenomena varians data melebihi min data tersebut. Model regresi binomial negatif sesuai digunakan untuk data serakan berketerlaluan. Jika serakan berketerlaluan, α bersama dengan sifar, model binomial negatif akan diturunkan kepada model poisson. Pekali bagi model binomial negatif dianggar dengan menggunakan kaedah kebolehjadian log maksimum (Oh *et al.*, 2005).

Dalam taburan binomial negatif, parameter serakan berketerlaluan adalah diperlukan. Berdasarkan kepada Famoye *et al.* (2004), terdapat dua cara untuk

mentafsir parameter ini; satu cara adalah penyalingan nilai lain. Untuk menghalang berlakunya kesalahan semasa penjelasan suatu nilai, penggunaan serakan berketerlaluan (*over-dispersion*) parameter α dijalankan di dalam SAS GENMOD (*Statistics Analysis Software Generalised Model*) manual. Mereka menggunakan prosedur GENMOD dalam SAS 8.02 untuk mendapatkan model. Famoye *et al.* (2004) berpendapat jika serakan berketerlaluan, α sama dengan sifar ($\alpha = 0$), model regresi poisson teritlak (*generalization poisson regression model*) akan diturunkan ke model regresi poisson. Jika varians adalah tidak sama dengan min, penganggar model poisson regresi adalah masih konsisten tetapi tidak cekap di mana ia akan menyebabkan ketidaksesuaian bagi penganggaran ralat piawai.

Oh *et al.* (2005) berpendapat bahawa perbandingan model dapat dijalankan di antara model poisson, model binomial negatif dan model gamma bagi data kemalangan yang berlaku di lintasan jalan raya utama dengan landasan keretapi. Untuk membuat pemilihan di antara taburan binomial negatif atau taburan poisson bagi data kemalangan tersebut, Geirt dan Nuyts (2005) telah mengira statistik-t bagi beberapa subset dalam pangkalan data yang terdapat. Oleh kerana model regresi mempunyai masalah heteroskedasiti, pekali model poisson adalah dikira melalui kaedah kebolehjadian maksimum (Geirt dan Nuyts, 2005). Untuk membandingkan model semasa proses pemodelan, Geirt dan Nuyts (2005) mencadangkan penggunaan Kriteria Informasi Akaike (*Akaike Information Criterion*) dan Kriteria Informasi Bayesian (*Bayesian Information Criterion*).

Famoye *et al.* (2004) mencadangkan satu kajian statistik iaitu kebagusan penyuaian (*goodness-of-fit*) untuk membuat pemilihan model di antara model regresi

poisson teritlak dengan model regresi poisson. Penggunaan kebolehjadian log (*log-likelihood*) ke atas model regresi telah dijalankan. Nilai kebolehjadian log bagi model regresi yang bernilai besar adalah lebih baik jika berbanding dengan nilai kebolehjadian log model yang kecil. Selain itu, model binomial negatif dapat digunakan untuk memodelkan serakan berketerlaluan (Maher dan Summergill, 1995). Pada pendapat Maher dan Summergill (1995), penggunaan kuasiPoisson dan model binomial negatif bagi data kemalangan tidak mempunyai perbezaan yang besar dalam pengiraan parameter-paramater. Dengan itu, nilai suaian (*fitted value*) adalah hampir secaman, keputusan ramalan daripada penggunaan kedua-dua model itu juga adalah secaman. Satu plot kebolehjadian log maksimum, $\log L_{\max}(n)$ yang berlawan dengan nilai n telah digunakan untuk memilih model yang lebih sesuai digunakan bagi data kemalangan di antara model kuasipoisson dan binomial negatif. Pemilihan model adalah bergantung kepada sampel saiz data bagi jenis kemalangan yang terlibat.

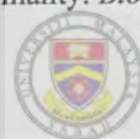
Dengan menggunakan model kuasipoisson atau binomial negatif, pemodelan serakan berketerlaluan akan menyebabkan perbezaan di antara penggunaan data agregat atau tak agregat (*disaggregate*). Selain itu, penggunaan data tak agregat, semua ralat tidak dapat dikatakan tak bersandar kerana kemungkinan struktur ralat mengandungi ralat dalam kumpulan (*within group*) atau di antara kumpulan (*between group*). Maka Maher dan Summergill (1995) berpendapat bahawa penggunaan data agregat adalah lebih baik bagi pemodelan.

Beberapa ujian telah dijalankan untuk membuat pengesahan tentang penyesuaian sesuatu model. Kenormalan reja dapat digunakan dengan ujian Shapiro-Wilk (Shapiro dan Wilk, 1965). Fungsi varians dapat diperiksa dengan ujian grafik.

RUJUKAN

- Amoros, E., Martin, J.L. dan Laumon, B., 2003. Comparison of road crashes incidence and severity between some French counties. *Accident Analysis and prevention* **35**, 537-547.
- Dobson, A.J., 1990. *An introduction to generalized linear models*. Chapman and Hall, London.
- Evans, A.W., 2002. Estimating transport fatality risk from past accident data. *Accident Analysis and prevention* **35**, 459-472.
- Famoye, F., Wulu, J.T. dan Singh, K.P., 2004. On the generalized poisson regression model with an application to accident data. *Journal of Data Science* **2**, 287-295.
- Geirt, F.V. dan Nuyts, E., 2005. *Cross-sectional accident models on Flemish motorways based on infrastructural design*. Provincial College of Limburg, Agoralaan.
- Greene, W.H., 2003. *Econometric Analysis*. Ed.ke-5. Prentice Hall, New Jersey.
- Hamilton, L.C., 1998. *Statistics with STATA 5*. Duxbury Press, London.
- Hemenway, D., 1998. The decomposition method. *Epidemiology* **19**(4), 369-370.
- Joshua, S.C. dan Garber, N.J., 1990. Estimating truck accident rate and involvements using linear and poisson regression models. *Transport Planning and Technology* **15**, 41-58.
- Koop, K., 2003. *Bayesian Econometrics*. Wiley, England.
- Lambert, D., 1992. Zero-inflated poisson regression, with an application to defects in manufacturing. *Technometrics* **34**, 1-14.

- Maher, M.J. dan Summersgill, I., 1996. A comprehensive methodology for the fitting of predictive accident models. *Accident Analysis and prevention* **28**(3), 281-296.
- Maycock, G. dan Hall, R.D., 1984. Accidents at 4-am roundabouts. *Transport and Road Research Laboratory*, report 1120.
- McCullagh, P. dan Nelder, J.A., 1989. *Generalized linear models*. Ed. ke-2. Chapman and Hall, London.
- Mercer, G.W., 1987. Influences on passenger vehicle casualty accident frequency and severity: unemployment, driver gender, driver age, drinking driving and restraint device use. *Accident Analysis and prevention* **19**(3), 231-236.
- Miaou, S.P., 1994. The relationship between truck accident and geometric design of road sections: poisson versus negative binomial regressions. *Accident Analysis and prevention* **26**(4), 471-482.
- Miaou, S.P. dan Lum, H., 1993. Modeling vehicle accidents and highway geometric design relationships. *Accident Analysis and prevention* **25**, 689-709.
- Oh, J., Lyon, C., Washington, S.P., Persaud, B.N. dan Bared, J., 2003. Validation of the FHWA crash models for rural intersections: lessons learned. *Transport Research Record* **1840**, 41-49.
- Oh, J., Washington, S.P. dan Nam, D., 2005. Accident prediction model for railway-highway interfaces. *Accident Analysis and prevention* **38**, 346-356.
- Ramu, R. 2003. *Introductory Econometrics with Applications*. Ed. ke-4. The Dryden Press Harcourt Brace College Publishers, New York.
- Shapiro, S. dan Wilk, M., 1965. An analysis of variance test for normality. *Biometrika* **52**, 591-611.



- Thomas, R.L., 1997. *Modern Econometrics*. Addison-Wesley Longman, New York.
- Wedderburn, R.W.M., 1974. Quasi-likelihood functions, generalized linear models, and the Gauss-Newton method. *Biometrika* **54**, 439-447.
- Winkelmann, R., 1997. *Econometric Analysis of Count Data*. Springer-Verlag Berlin Heidelberg, New York.
- Zullyadini Abdul Rahman dan Shaharuddin Ahmad, 1999. *Analisis Taburan dan punca-punca kemalangan jalan raya di negeri Melaka*. Jabatan Geologi, Universiti Kebangsaan Malaysia.

