

ISBN : 978-979-16353-1-8



PROSIDING SEMINAR NASIONAL MATEMATIKA DAN PENDIDIKAN MATEMATIKA

“Peningkatan Kualitas Penelitian dan Pembelajaran
Matematika untuk Mencapai *World Class University*”

Yogyakarta, 28 November 2008



Penyelenggara :

Jurusan Pendidikan Matematika FMIPA UNY
Kerjasama dengan

Himpunan Matematika Indonesia (Indo-MS)
wilayah Jateng dan DIY

Jurusan Pendidikan Matematika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Negeri Yogyakarta
2008

M - 13.	Pemodelan Persamaan Struktural Dengan Partial Least Square (<i>I Gede Nyoman Mindra Jaya, I Made Sumertajaya</i>)	1 – 118
M - 14.	Penggerombolan Model Parameter Regresi dengan <i>Error-Based Clustering</i> (<i>I Made Sumertajaya, Gusti Adhi Wibawa, I Gede Nyoman Mindra Jaya</i>)	1 – 133
M - 15.	Koreksi Metode Connected Ammi dalam Pendugaan Data Tidak Lengkap (<i>Made Sumertajaya, Ahmad Ansori Matjijik, I Gede Nyoman Mindra Jaya</i>)	1 – 145
M - 16.	Pendekatan Metode Pemulusan Kernel Pada Pendugaan Area Kecil (<i>Small Area Estimation</i>) (<i>Indahwati, Kusman Sadik, Ratih Nurmasari</i>)	1 – 162
M - 17.	Penerapan Metode Pemulusan Kernel Pada Pendugaan Area Kecil (Studi Kasus Pendugaan Pengeluaran Per Kapita Di Kota Bogor Tahun 2005) (<i>Indahwati, Utami Dyah Syafitri, Renita Sukma Mayasari</i>)	1 – 173
M - 18.	Zero Inflated Negative Binomial Models In Small Area Estimation (<i>Irene Muflikh Nadhiroh, Khairil Anwar Notodiputro, Indahwati</i>)	1 – 183
M - 19.	Aplikasi <i>Multidimensional Scaling</i> Untuk Peningkatan Pelayanan Proses Belajar Mengajar (PBM). (<i>Irelandia Ginanjar</i>)	1 – 194
M - 20.	Peranan Formulasi Inversi Pada Fungsi Karakteristik Suatu Variabel Acak (<i>John Maspupu</i>)	1 – 202
M - 21.	Pendugaan Berbasis Model Untuk Kasus Biner Pada <i>Small Area Estimation</i> (<i>Kismiantini</i>)	1 – 209
M - 22.	Pendugaan Komponen Utama Pada Pengaruh Acak Model Linear Campuran Terampat (<i>Mohammad Masjkur</i>)	1 – 216
M - 23.	<i>Distribusi Poisson Tergeneralisasi Tak Terbatas Dan Beberapa Sifat-Sifatnya</i> (Suatu Pengembangan Teori Statistika Matematika) (<i>Mutijah</i>)	1 – 237
M - 24.	Regresi Rasio Prevalensi Dengan Model Log-Binomial: Isu Ketakkonvergenan (<i>Netti Herawati, Alfian Futuhul Hadi, Nusyirwan, Khoirin Nisa</i>)	1 – 249
M - 25.	Pengujian Autokorelasi Terhadap Sisaan Model Spatial Logistik (<i>Utami Dyah Syafitri, Bagus Sartono, Salamattuntzil</i>)	1 – 264
M - 26.	Penerapan Analisis Survival Untuk Menaksir Waktu Bertahan Hidup Bagi Penderita Penyakit Jantung (<i>Yani Hendrajaya, Adi Setiawan dan Hanna A. Parhusip</i>)	1 – 269
M - 27.	Pendekatan Analisis Multilevel Respon Biner dalam Menentukan Faktor-Faktor yang Memengaruhi Imunisasi Lengkap (<i>Bertho Tuntular, I Gede Nyoman Mindra Jaya</i>)	1 – 281

Regresi Rasio Prevalensi dengan Model Log-Binomial: Isu Ketakkonvergenan

Netti Herawati¹⁾
Alfian Futuhul Hadi²⁾
Nusyirwan¹⁾
Khoirin Nisa¹⁾

¹⁾ *Jurusan Matematika FMIPA Universitas Lampung*
Jl. Sumantri Brojonegoro No. 1 Bandar Lampung 35145
netti@unila.ac.id

²⁾ *Jurusan Matematika FMIPA Universitas Jember*
Jl. Kalimantan 37 Jember 68121
Email: afhadi@unej.ac.id

Abstrak

Regresi logistik memainkan peran sentral dalam pengamatan epidemiologi. Regresi logistik memiliki *odds ratio* yang memberikan informasi tentang risiko seseorang didapati sebagai penderita suatu penyakit. Pada kasus tertentu *ratio prevalensi* dapat didekati dari *odds ratio*, dan tidak memberikan perbedaan yang menyesatkan. Namun *odds ratio* tidak selalu diinginkan. Antara lain, karena pada pengamatan *common event* (pengamatan dengan nilai 1 yang banyak), *rasio prevalensi* tidak lagi dapat didekati secara langsung dari *odds ratio*, artinya *odds ratio* tidak lagi mencerminkan informasi yang diinginkan oleh epidemiologi. Untuk itu diperlukan pendekatan lain dalam memperoleh *ratio prevalensi* dari data binomial, antara lain dengan model log-binomial/regresi *relative risk*/*rasio prevalensi*. Regresi log-binomial, dalam terminologi Model Linier Umum adalah model dengan distribusi Y binomial yang menggunakan *link-function* log, sedangkan regresi logistik, menggunakan fungsi link, logit. Pada regresi log-binomial selain masalah ketakkonvergenan, dimungkinkan terjadi pendugaan peluang diluar interval 0-1, karena model ini tidak melakukan restriksi $0 \leq P(Y=1) \leq 1$, seperti pada model logistik. Akibatnya, pada saat memodelkan pengamatan dengan $P(Y=1) \cong 1$ atau *rasio prevalensi* mendekati 1, mungkin terjadi masalah.

Penelitian ini ingin mengatasi masalah ketakkonvergenan pada log-binomial melalui metode copy (Copy Method) melalui suatu studi kasus penderita Demam Berdarah Dengue (DBD) di RSUD Dr. Soebandi Jember.

Tulisan ini membicarakan tentang isu masalah ketakkonvergenan pada log-binomial, dengan alternatif penyelesaian menggunakan metode copy (Copy Method). Studi kasus diberikan sebagai ilustrasi data penderita Demam Berdarah Dengue (DBD) di RSUD Dr. Soebandi Jember.

Kata Kunci: Regresi Rasio Prevalensi, Risiko Relatif, log-binomial, ketakkonvergenan.

PENDAHULUAN

Analisis regresi linier merupakan salah satu metode statistika yang digunakan untuk menyelidiki hubungan antara variabel tak bebas (dependent variable) atau variabel respon dengan variabel bebasnya (independent variable) atau variabel penjelas, selain itu regresi linier baik digunakan untuk data respon kontinu.

Dalam suatu penelitian terkadang variabel respon yang diamati hanya dibagi dalam dua kemungkinan kejadian. Variabel semacam ini lebih dikenal dengan sebutan variabel dikotomi. Variabel dikotomi yaitu variabel indikator yang terdiri atas data biner (0 atau 1). Jika data tersebut membentuk suatu kelompok maka dapat dimodelkan menurut distribusi Binomial.

Sejauh ini model regresi yang paling populer untuk data biner adalah regresi logistik biner. Regresi logistik memainkan peran sentral dalam pengamatan epidemiologi. Hal ini disebabkan oleh dua hal. Pertama, regresi logistik memiliki *odds ratio* yang memberikan informasi tentang risiko seseorang didapati sebagai penderita suatu penyakit. Untuk mendapatkan informasi tentang risiko seringkali faktor penjelas diukur dalam peubah kategorik yang menunjukkan perbedaan tingkat atau level faktor tersebut. Nilai *odds ratio* merupakan perbandingan peluang (risiko) seseorang dengan karakter (faktor penjelas pada level) tertentu didapati sebagai penderita penyakit tertentu terhadap orang lain dengan karakter (faktor penjelas level) yang lain.

Kedua, *odds ratio* yang dimiliki regresi logistik dapat diidentifikasi secara langsung dari suatu studi *case-control* yang sangat lazim dilakukan dalam riset-riset epidemiologi. Rasio prevalensi kemudian dapat diduga melalui *odds ratio*. Pada pengamatan kejadian penyakit-penyakit yang jarang terjadi (*rare event*) nilai *odds ratio* akan mendekati nilai risiko relatif (*relative risk*). Selain itu, regresi logistik dapat pula digunakan untuk memodelkan kejadian penyakit-penyakit yang lazim terjadi (*common event*) dalam penelitian *cross-sectional* atau *longitudinal*. Dalam hal ini risiko relatif pada kasus ini tidak dapat secara langsung dihitung, nilainya pun tidak dapat begitu saja didekati dari *odds ratio*.

Risiko relatif memberikan informasi yang berguna tentang asosiasi atau keterkaitan. Alasan khusus bahwa risiko relatif lebih disukai dalam memberikan ringkasan tentang asosiasi pada data biner diperkirakan adalah kesulitan memberikan penjelasan yang tepat atas interpretasi dari *odds ratio*. Celakanya, meski beberapa peneliti —yang

mengetahui dengan jelas perbedaan antara *odds ratio* dan *relative risk*— akan menemui kesulitan untuk menjelaskan interpretasi yang tepat tentang *odds ratio*, mereka tetap mengkomunikasikannya kepada khalayak umum (Lumley *et al.* 2006)

Model log-binomial telah diusulkan sebagai pendekatan yang bermanfaat untuk menghitung rasio prevalensi. Prevalensi adalah banyaknya subyek yang mengalami kejadian tertentu atau menderita penyakit tertentu pada suatu waktu tertentu. Model log-binomial juga digunakan untuk menganalisa suatu hasil dikotomi. Tidak seperti regresi logistik, angka-angka pada model log binomial dibatasi pada ruang parameter dan MLE mungkin terjadi pada batas ruang parameter. Secara umum, model log-binomial menghasilkan suatu perkiraan dari rasio prevalensi yang takbias dengan selang kepercayaan untuk rasio prevalensi yang mungkin lebih kecil.

Perbedaan antara model logistik dengan model log-binomial menjadi hubungan antara variabel bebas dan dari hasil peluangnya. Pada regresi logistik digunakan fungsi logit sedangkan untuk model log-binomial, fungsi log yang digunakan. Secara umum, model log-binomial menghasilkan suatu perkiraan yang takbias dari rasio prevalensi.

Permasalahan

Penggunaan regresi log-binomial dalam penelitian epidemiologi untuk memperoleh dugaan bukannya tanpa masalah. Secara statistika, sebagaimana Model Linier Umum (*Generalized Linear Models*) pada umumnya, regresi logistik dan log-binomial juga menghadapi masalah ketakkonvergenan karena proses pendugaan parameter yang dilakukan secara iteratif. Ketakkonvergenan dapat pula terjadi akibat defisiensi derajat bebas error, pada regresi yang melibatkan beberapa variabel penjelas berkategori. Bila kekonvergenan tidak tercapai, maka nilai odds ratio maupun risiko relatif tidak dapat diperoleh, karena proses iterasi yang tidak pernah berhenti. Sementara itu, pada model log-binomial pendugaan peluang yang diturunkan dari model (log) mungkin tidak semuanya berada pada interval 0-1.

Demikian penting regresi logistik dan log-binomial untuk menghasilkan odds ratio dan resiko relatif dalam studi epidemiologi. Karena itulah tulisan ini akan membicarakan tentang (i) regresi Risiko Relatif (ratio prevalensi) dengan model log-binomial, (ii) mengatasi kasus ketidakkonvergenan melalui metode copy (*copy methods*) serta (iii) memperoleh nilai risiko relatif (rasio prevalensi) pada studi kasus data penderita Demam Berdarah Dengue (DBD) di RSUD Dr. Soebandi Jember.

MODEL REGRESI LOGISTIK

Logistik merupakan salah satu metode regresi yang digunakan untuk mencari hubungan antara peubah respon kategori dengan satu atau lebih peubah bebas yang kontinu.

Model regresi logistik dapat dinyatakan sebagai :

$$\pi(x_i) = \frac{\exp(\beta_0 + \beta_1 X_{i1} + \dots + \beta_p X_{ip})}{1 + \exp(\beta_0 + \beta_1 X_{i1} + \dots + \beta_p X_{ip})} \quad (2.1)$$

dengan $i = 1, 2, \dots, n$ dan $p =$ banyak peubah bebas.

Bentuk transformasi logitnya adalah

$$\begin{aligned} g(X_i) &= \ln \left(\frac{\pi(X_i)}{1 - \pi(X_i)} \right) \\ &= \ln \left(\frac{\frac{\exp(\beta_0 + \beta_1 + \dots + \beta_p X_{ip})}{1 + \exp(\beta_0 + \dots + \beta_p X_{ip})}}{1 - \frac{\exp(\beta_0 + \beta_1 + \dots + \beta_p X_{ip})}{1 + \exp(\beta_0 + \beta_1 + \dots + \beta_p X_{ip})}} \right) \\ &= \ln \left(\frac{\frac{\exp(\beta_0 + \beta_1 + \dots + \beta_p X_{ip})}{1 + \exp(\beta_0 + \beta_1 + \dots + \beta_p X_{ip})}}{\frac{1}{1 + \exp(\beta_0 + \beta_1 + \dots + \beta_p X_{ip})}} \right) \\ &= \ln \left(\frac{\exp(\beta_0 + \beta_1 + \dots + \beta_p X_{ip})}{1} \right) \\ &= \ln \exp(\beta_0 + \beta_1 + \dots + \beta_p X_{ip}) \\ &= \beta_0 + \beta_1 + \dots + \beta_p X_{ip} \quad (2.2) \end{aligned}$$

dengan $X_{i1}, X_{i2}, \dots, X_{ip} =$ peubah bebas ke 1, 2, ..., p dengan nilai pengamatan ke-i; dan $\beta_0, \beta_1, \dots, \beta_p =$ koefisien regresi untuk setiap peubah bebas.

MODEL LOG BINOMIAL

Model log-binomial telah diusulkan sebagai pendekatan yang bermanfaat untuk menghitung risiko relatif. Model log-binomial digunakan untuk menganalisa suatu hasil dikotomi. Perbedaan antara model logistik dengan model log-binomial menjadi hubungan antara variabel bebas dan dari hasil peluangnya. Pada regresi logistik, fungsi logit digunakan sedangkan untuk model log –binomial, fungsi log yang digunakan. Secara umum, model log-binomial menghasilkan suatu perkiraan yang takbias dari risiko relatif.

Selang kepercayaan untuk risiko relatif yang dihitung mungkin lebih kecil. Model log-binomial tidak konvergen untuk memberikan parameter estimate. Ketidakkonvergenan bisa dilihat dalam hubungannya dengan program software yang mempunyai suatu ukuran ketidakkonvergenan yang tidak cukup. Masalah ini dapat diatasi dengan memberikan tambahan iterasi dalam proses pengepasan model.

Skov et al (1998) merekomendasikan model log-binomial sebagai ganti model logistik untuk studi *cross-sectional*. Misalkan Y (0/1) merupakan tidak/ya tentang peristiwa atau memperlihatkan tanda didalam subjek dengan covarian $X = (x_1, \dots, x_k)$. Kemudian peluang tentang peristiwa dalam subjek adalah

$$p = P(Y = 1 | X) = \exp(\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k) \quad (1.1)$$

Ketika peluang p adalah selalu diantara 0 dan 1, model akan sah hanya jika

$$\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k \leq 0$$

Juga $p = P(Y = 0 | X) = 1 - p = 1 - \exp(\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k)$

Kemudian *likelihood* untuk model log-binomial

$$L(Y | \beta, X) = \prod_{y_i=1} \exp(\beta_0 + \beta_1 x_{1i} + \dots + \beta_k x_{ki}) \times \prod_{y_j=0} (1 - \exp(\beta_0 + \beta_1 x_{1j} + \dots + \beta_k x_{kj})) \quad (1.2)$$

dan fungsi *log-likelihoodnya* adalah

$$\text{Log}(Y | \beta, X) = \sum_{y_i=1} (\beta_0 + \beta_1 x_{1i} + \dots + \beta_k x_{ki}) + \sum_{y_j=0} \log(1 - \exp(\beta_0 + \beta_1 x_{1j} + \dots + \beta_k x_{kj})) \quad (1.3)$$

Untuk masing-masing subjek i , domainnya adalah

$$\Omega_i = \{(\beta_0, \dots, \beta_K) : \beta_0 + \beta_1 x_{1i} + \dots + \beta_K x_{Ki} \leq 0\} \quad (1.4)$$

Oleh karena itu ruang parameter untuk model log-binomial adalah $\Omega = \bigcap_{i=1}^K \Omega_i$

RASIO PREVALENSI

Mempelajari suatu hasil yang umum biasanya lebih tertarik pada penghitungan rasio prevalensi dari pada odds ratio. Odds ratio mungkin lebih disukai karena bisa digunakan untuk menghitung rasio nilai kejadian. Akan tetapi rasio prevalensi biasanya lebih mudah ditafsirkan dibanding odds ratio. Odds ratio bisa digunakan untuk memperkirakan rasio prevalensi, tetapi perkiraan tersebut baik hanya jika prevalensinya rendah. Lee (1994) dan Lee & Chia (1993) merekomendasikan penggunaan model resiko proporsional cox untuk menghitung rasio prevalensi. Skov *et al* (1998) merekomendasikan penggunaan model log-binomial, yang secara langsung memodelkan rasio prevalensi, dan menunjukkan bahwa metode ini lebih baik dibanding metode resiko proporsional cox (Lumley *et al*, 2006).

Model rasio prevalensi berasal dari model Regresi

$$\log P[Y=1|X] = \log \mu = \eta = \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p \quad (1)$$

dimana e^{β_i} , adalah sebuah rasio prevalensi yang membandingkan X_i yang berbeda dari

1. Jika $P[Y=1|X]$ lebih kecil, maka :

$$\log P[Y = 1|X] \approx \log \frac{P[Y = 1|X]}{1 - P[Y = 1|X]} = \log \text{it } P[Y = 1|X],$$

dan jika itu benar untuk semua nilai dari X , maka model rasio prevalensi mendekati model regresi logistik .

$$\log \text{it } P[Y = 1|X] = \log \text{it } \mu = \alpha_0 + \alpha_1 X_1 + \dots + \alpha_p X_p$$

Jika $P[Y=1|X]$ lebih besar daripada 10-15% untuk beberapa nilai pengamatan X (greenland) kemudian α dan β terlihat berbeda dengan $|\alpha|$ dan $|\beta|$.

Seperti Model Regresi Logistik, pada model regresi rasio prevalensi juga merupakan generalized linier model (GLM),(McCullagh & Nelder,1989), dengan hubungan log dan fungsi varians $V(\mu) = \mu(1 - \mu)$. Tidak seperti model regresi logistik, model risiko relatif

memerlukan batasan untuk β yang dipastikan dengan mencoba sisa peluang dalam interval $[0,1]$.

MLE dan semua estimator yang konsisten telah diusulkan, penyelesaian dari persamaan

$$\sum_{i=1}^n \frac{\partial \mu_i}{\partial \beta} w(\mu_i) (Y_i - \mu_i) = \sum_{i=1}^n x_i \mu_i w(\mu_i) (Y_i - \mu_i) = 0$$

dengan pilihan fungsi weight $w(\cdot)$ yang berbeda, sedikitnya dalam situasi dimana solusi untuk persamaannya ada dalam ruang parameter.

Persamaan ini diperkirakan tidak bias untuk beberapa pilihan pada $w(\cdot)$, maka semua estimator akan konsisten., dan akan asimtotik normal sepanjang β adalah dalam bagian ruang parameter yang sedang dipertimbangkan (McCullagh & Nelder, 1989).

Seperti model regresi logistik, model regresi resiko relatif memerlukan batasan-batasan β untuk menjamin bahwa probabilitas yang cocok tetap berada pada interval $[0,1]$ (Lumley *et al*, 2006).

ESTIMASI GLM PADA PROGRAM SAS

Pada SAS estimator log-binomial didapatkan hanya dengan memilih distribusi binomial dan hubungan log. Untuk menghitung *standart error* yang valid untuk fungsi bobot yang lain menambah baris (SAS Institute Inc, 2004) yang memerlukan estimasi yang sesuai dengan data asli dengan satu observasi per individu ini memberikan standar error model *robust* tanpa mengubah estimasi penting resiko relatif. Pada SAS menggunakan proc genmod dengan distribusi log-binomial dan fungsi hubungan log. Tidak seperti model logistik, model log-binomial memberikan batasan-batasan pada ruang parameter, dan penghitungan MLE bisa terjadi pada batas ruang parameter, dalam hal ini prog genmod tidak akan konvergen penghitungan yang tepat. Metode dengan perluasan data menjadi sejumlah copy-copy data asli yang diatur dengan satu copy data asli dengan kasus kontrol yang dibalik untuk mendapatkan penghitungan yang tepat pada standart error rasio prevalensi. Metode SAS juga metode perkiraan yang menghasilkan penghitungan-penghitugan yang mendekati penghitungan MLE yang tepat dan parameternya yang benar.

Proc Genmod pada SAS memaksimalkan kemungkinan dengan menentukan titik atau nilai dimana turunannya sama dengan nol, tidak mampu menemukan suatu

penghitungan dalam situasi yang demikian karena prosedur iterasinya tidak konvergen. Masalah yang biasanya muncul dengan satu variabel independen, untuk setiap observasi yang melibatkan kombinasi-kombinasi tertentu dari tingkat variabel independen.

Metode Copy

Suatu metode baru (*copy method*) yang meliputi modifikasi data sehingga MLE dalam data yang sudah dimodifikasi mendekati MLE dari data aslinya. MLE pada pengaturan data yang sudah dimodifikasi selalu berada dalam ruang parameter, sehingga MLE ini bisa ditemukan dengan menggunakan *proc genmod*. Lebih spesifik, kita ajukan bahwa pengaturan data yang baru diciptakan yang berisi *copy-copy* C-1 dari data asli dan 1 *copy* data asli dengan nilai-nilai Y yang saling menggantikan. Misalnya, dapat menggunakan C=1000 dan memiliki 999 *copy* data asli dan 1 data asli dengan semua nilai Y saling menggantikan. Untuk mempertimbangkan ukuran sampel yang terus meningkat, kemudian mengalikan standart error yang sudah dihitung.

Dengan memasukkan data yang sudah dikembangkan tersebut, berharap bisa memperoleh suatu perhitungan MLE baru yang mendekati penghitungan yang sebenarnya tetapi tidak berada pada batas. Tujuannya adalah untuk memilih sejumlah *copy* sehingga penghitungan baru tersebut juga dekat dengan batas sehingga *software* tidak bisa menghitungnya. Pada SAS belum mengalami masalah ini, tetapi kadang-kadang *proc genmod* tidak konvergen ketika jumlah *copy*nya terlalu kecil.

Metode tersebut meliputi perluasan data yang diatur agar bisa menjadi jumlah yang sangat besar dari *copy-copy* data asli yang diatur dengan satu *copy* data asli dari kasus kontrol yang dapat dibalik. Standart error yang dihitung dari rasio prevalensi pada data yang diperluas ini kemudian diatur untuk mendapatkan perhitungan yang tepat pada standart error rasio prevalensi (Deddens, J.A. Petersen, M. R. & Lei, X. 2006).

METODOLOGI

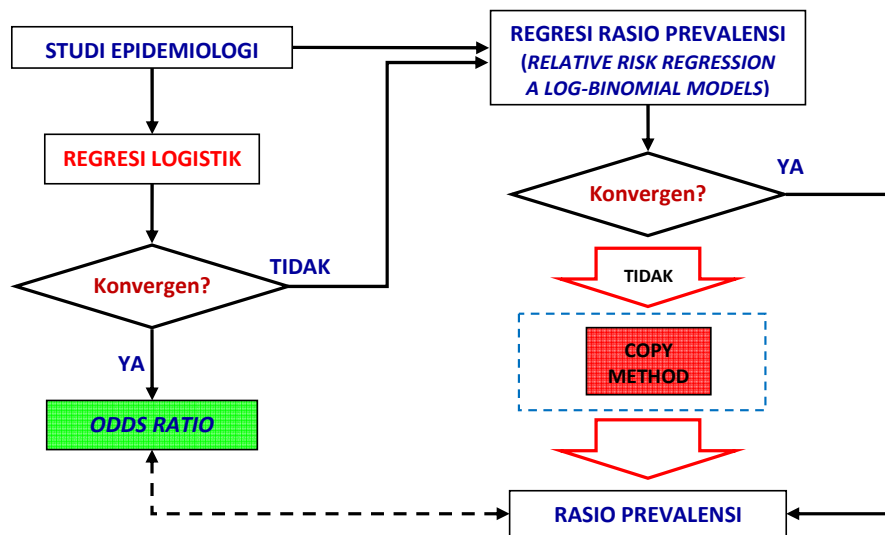
Sebagaimana kerangka logis riset (gambar 1), tahapan penyelesaian masalah ketakkonvergenan pada regresi logistik adalah:

1. Memeriksa kejadian takkonvergen pada model logistik dan log-binomial.
2. Mengatasi ketakkonvergenan, melalui penggunaan model log-binomial dengan algoritma *copy* (*COPY method*) (Deddens, J.A. Petersen, M. R. & Lei, X. 2006)

3. Mendapatkan nilai dugaan rasio prevalensi dari model log-binomial.

Studi kasus epidemiologi yang digunakan dalam penelitian ini adalah data tentang penyakit demam berdarah dengue (DBD) di Rumah Sakit Umum Daerah (RSUD) Dr. Soebandi Jember tahun 2006. Sedangkan variabel yang diteliti adalah umur, jenis kelamin, waktu tahan hidup (hari), status (mati atau sembuh), kondisi pasien (stadium) dan trombosit. Peubah-peubah yang digunakan untuk membuat model regresi risiko relatif adalah sebagai berikut :

1. Peubah respon (Y) adalah kejadian seseorang dengan kasus demam berdarah *dengue* (DBD). Dengan 0 mengindikasikan sembuh dan 1 mengindikasikan meninggal.
2. Peubah penjelas (X) antara lain :
 - a. Jenis kelamin (SEX), yaitu dibagi menjadi 2 kategori antara lain:
 - Wanita, dengan indeks 0; dan
 - Laki-laki, dengan indeks 1.
 - b. Usia (AGE), yaitu dibagi menjadi 2 kategori :
 - ≤ 15 tahun, dengan indeks 0; dan
 - > 15 tahun dengan indeks 1



Gambar 1. Kerangka Logik Riset

- c. Hari (DAY), dalam hal ini hari yang dimaksud adalah waktu seseorang dengan kasus DBD dapat bertahan hidup.
- d. Status, yaitu dibagi menjadi 2 kategori antara lain :

- Sembuh, dengan indeks 0; dan
 - Meninggal, dengan indeks 1.
- e. Stadium, yaitu kondisi seseorang dengan kasus DBD (stadium 1, stadium 2, stadium 3, dan stadium 4).
- Trombosit, yaitu dibagi menjadi 2 kategori yaitu $(1) \leq 100.000 / \mu l$, dengan indeks 0; dan $(2) > 100.000 / \mu l$, dengan indeks 1.

HASIL DAN DISKUSI

Ketakkonvergenan Model Logistik dan Log-Binomial

Dari model yang digunakan pada data studi kasus, ditemui ketakkonvergenan pada model logistik. Kita dapat menggunakan model Log-binomial sebagai alternatif ketika model logistik tidak konvergen. Namun model log-binomial juga sangat mungkin menalami hal yang sama.

Keuntungan menggunakan model log-binomial adalah dapat diperolehnya rasio prevalensi secara langsung dari model. Sedangkan pada model logistik akan menghasilkan odds ratio, yang mana tidak selamanya odds ratio ini dapat mendekati nilai rasio prevalensi yang diinginkan oleh studi epidemiologi.

Pemodelan log-binomial dapat dilakukan menggunakan SAS melalui PROC GENMOD dengan distribusi binomial dan fungsi hubung log. Model log binomial memberikan batasan-batasan pada ruang parameter, dan penghitungan MLE bisa terjadi pada ruang parameter. Metode proc genmod yang memaksimalkan kemungkinan dengan menentukan titik atau nilai dimana turunannya sama dengan nol, tidak mampu menemukan suatu penghitungan, hal ini disebabkan karena proses iterasi yang tidak konvergen (lihat gambar 2).

```

PROC GENMOD did not converge on the original data set.
          Continuing with COPY method ...

-----

Algorithm converged.
PROC GENMOD CONVERGED FOR THE COPY METHOD.
```

Gambar 2. Ketakkonvergenan pada model log-binomial

Mengatasi Ketakonvergenan Model Log-Binomial dengan Metode Copy

Pada kasus dimana terjadi ketakkonvergenan seperti pada kasus ini, dapat dilakukan pendekatan metode copy melalui macro-sas `%pr_copylr.` Metode tersebut meliputi perluasan data yang diatur agar menjadi jumlah yang sangat besar dari copy-copy data asli yang diatur dari satu copy data asli dengan semua nilai Y yang saling menggantikan. . Dalam penelitian ini data yang digunakan adalah data demam berdarah, sebanyak 136, dengan copy = 100.

Standar error yang dihitung dari rasio prevalensi pada data yang diperluas ini kemudian diatur untuk mendapatkan perhitungan yang tepat pada standar error rasio prevalensi.

Model rasio prevalensi yang diperoleh dari data demam berdarah (DBD) dengan model log binomial, dengan model Log binomial sebagai berikut :

$$P(Y = 1|X) = \exp(\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki})$$

Dengan menggunakan data demam berdarah (DBD) yaitu data penelitian yang akan dilihat apakah variabel penjelas akan memperbesar atau memperkecil peluang terjadinya Y(variabel respon), didapat model sebagai berikut :

$$P(Y = 1|X) = \exp \left(\begin{array}{l} -0.2473 + 0.2380X_{11} + 0.2375X_{12} + 0.2364X_{13} + 0.0002X_3 \\ + 0.1423X_{41} - 0.1421X_5 X_6 - 0.1423X_5 X_7 - 0.1421X_5 X_7 \end{array} \right)$$

dimana $X_5 = X_{41}$ dan $X_6 = X_{11}$, $X_7 = X_{12}$, $X_8 = X_{13}$

$X_5 X_6$ adalah kombinasi antara jenis kelamin perempuan dengan stadium 1,

$X_5 X_7$ adalah kombinasi antara jenis kelamin perempuan dengan stadium 2,

$X_5 X_8$ adalah kombinasi antara jenis kelamin perempuan dengan stadium 3,

Dari model yang diperoleh dapat diketahui juga bahwa apabila parameter estimate β yang diperoleh dihitung rasio prevalensi (e^β) yaitu seperti pada tabel.... :

Dari Tabel 1. dapat diketahui bahwa nilai rasio prevalensi (e^β) > 1 , yaitu pada parameter estimate (β) yang bernilai positif. Dengan variabel penjelas stadium 1, stadium 2, stadium 3, hari (*day*), dan jenis kelamin perempuan, yang mempunyai nilai rasio prevalensi 1.268709, 1.268075, 1.266681, 1.000200, dan 1.152920. Hal ini dapat diartikan bahwa pada variabel penjelas yang mempunyai nilai rasio prevalensi (e^β) > 1 , mempunyai peluang sembuh yang lebih besar, yang dipengaruhi

oleh variabel penjelas stadium 1, stadium 2, stadium 3, hari (*day*), dan jenis kelamin perempuan.

Untuk nilai rasio prevalensi (e^β) < 1, yaitu pada parameter estimate (β) yang bernilai negatif. Dengan variabel penjelas kombinasi antara jenis kelamin perempuan dengan stadium 1, dengan stadium 2, dan dengan stadium 3, yang mempunyai nilai rasio prevalensi 0.867534, 0.867361, dan 0.867534. Hal ini dapat diartikan bahwa pada variabel penjelas yang mempunyai nilai rasio prevalensi (e^β) < 1, mempunyai peluang sembuh yang kecil, yang dipengaruhi oleh variabel penjelas kombinasi antara jenis kelamin perempuan dengan stadium 1, dengan stadium 2, dan dengan stadium 3.

Untuk nilai rasio prevalensi (e^β) = 1, yaitu pada parameter estimate (β) yang bernilai 0.0000. Variabel penjelas Stadium 4, Jenis kelamin laki-laki, kombinasi antara jenis kelamin perempuan dengan stadium 4, dan kombinasi jenis kelamin laki-laki dengan stadium 1, 2, 3, 4, mempunyai nilai rasio prevalensi, karena dijadikan kategori acuan. Sedangkan pada variabel penjelas Usia (*age*) dan Trombosit nilai rasio prevalensi (e^β) yang sama dengan 1 dapat diartikan variabel tersebut kurang berpengaruh, karena perbedaan nilai trombosit dan usia tidak memberikan perbedaan peluang sembuh/sehat.

Tabel 2. Hasil analisis parameter estimate model log-binomial dengan metode copy

Variabel Penjelas	Parameter estimate (β)	Rasio Prevalensi (e^β)	Adjusted P_Value
Stadium 1	0.2380	1.268709 ^{**})	0.00100
Stadium 2	0.2375	1.268075 ^{**})	0.00100
Stadium 3	0.2364	1.266681 ^{**})	0.00108
Stadium 4	0.0000	1.000000	-
Usia (<i>age</i>)	0.0000	1.000000 ^{ns})	0.91729
Hari (<i>day</i>)	0.0002	1.000200 ^{ns})	0.92128
JKP	0.1423	1.152920 ^{ns})	0.10283
JKL	0.0000	1.000000	-
Trombosit	0.0000	1.000000 ^{ns})	0.78552
JKP*Stadium1	-0.1421	0.867534 ^{ns})	0.10955
JKP*Stadium 2	-0.1423	0.867361 ^{ns})	0.10890

JKP*Stadium 3	-0.1421	0.867534 ^{ns)}	0.11155
JKP*Stadium 4	0.0000	1.000000 ^{ns)}	-
JKL*Stadium1,2,3,4	0.0000	1.000000 ^{ns)}	-

Dengan melihat nilai adj p-valuenya pada Tabel 2, dapat diinterpretasikan bahwa variabel penjelas lamanya pasien dirawat (hari), umur, jenis kelaminnya, jumlah trombositnya, dan interaksi antara jenis kelamin dengan tingkat stadiumnya tidak berpengaruh nyata terhadap status kesembuhan seorang pasien.

Sedangkan pada variabel penjelas stadium 1,2, dan 3 mempunyai nilai adjusted p-value yang lebih kecil dari taraf nyata 0.05 (adj_p-value < 0.05). Hal ini berarti bahwa variabel penjelas stadium yang dilihat dari tingkatan kritisnya seorang pasien yaitu stadium 1,2, dan 3, mempunyai pengaruh yang nyata terhadap model atau dengan kata lain status kesembuhan pasien yang terkena demam berdarah dipengaruhi oleh stadium 1,2, dan 3.

Estimasi Rasio Prevalensi

Hasil nilai rasio prevalensi yang diperoleh dengan metode copy dapat dilihat pada Tabel 2. Interpretasi hasil rasio prevalensi pada Tabel 2, dapat diketahui bahwa nilai rasio prevalensi pada variabel penjelas stadium yaitu pada stadium 1,2, dan 3 mempunyai nilai rasio prevalensi (RP) > 1, sedangkan stadium 4 mempunyai nilai rasio prevalensi (RP) = 1, hal ini berarti stadium 4 dijadikan sebagai acuan pembading bagi stadium 1,2, dan 3. Pasien dengan stadium 1, 2, dan 3 mempunyai peluang untuk sembuh sekitar 1.26 kali lipat lebih besar daripada stadium 4. Lebih tepatnya, pasien dengan stadium 1 berpeluang sembuh 1.26869 kali lebih besar daripada pasien dengan stadium 4, pasien dengan stadium 2 berpeluang sembuh 1.26807 kali lebih besar daripada pasien dengan stadium 4, sedangkan pasien dengan stadium 3 berpeluang sembuh 1.26667 kali lebih besar daripada pasien dengan stadium 4.

Pada variabel penjelas usia (*age*) mempunyai nilai rasio prevalensi (RP) > 1, hal ini berarti semakin tinggi umurnya maka mempunyai peluang untuk sembuh lebih mudah. Pada variabel penjelas hari (*day*) mempunyai nilai rasio prevalensi (RP) > 1, yang berarti semakin lama pasien dirawat maka peluang untuk sembuh juga semakin

besar. Pada variabel penjas jenis kelamin perempuan mempunyai nilai rasio prevalensi (RP) > 1 , sedangkan jenis kelamin laki-laki mempunyai nilai rasio prevalensi (RP) $= 1$ sebagai acuan pembanding, hal ini berarti pasien yang mempunyai jenis kelamin perempuan mempunyai peluang untuk sembuh 1.15 kali lipat lebih besar daripada pasien yang mempunyai jenis kelamin laki-laki.

Pada variabel penjas trombosit mempunyai nilai rasio prevalensi (RP) < 1 , yang berarti jumlah trombosit pada pasien mempunyai peluang yang kurang berpengaruh pada kesembuhan. Pada variabel penjas kombinasi antara jenis kelamin perempuan dengan stadium 1, dengan stadium 2, dan dengan stadium 3 mempunyai nilai rasio prevalensi (RP) < 1 , hal ini berarti bahwa variabel-variabel penjas tersebut mempunyai peluang untuk sembuh 0.867 kali lipat lebih kecil dibandingkan dengan variabel penjas kombinasi antara jenis kelamin perempuan dengan stadium 4, dan kombinasi antara jenis kelamin laki-laki dengan stadium 1,2,3,4 yang mempunyai nilai rasio prevalensi (RP) $= 1$.

PENUTUP

Dari hasil penelitian yang telah dilakukan, beberapa kesimpulan yang dapat diambil antara lain :

1. Metode Copy sebagai solusi yang dapat digunakan ketika data yang dimodelkan dengan log-binomial gagal menuju konvergen.
2. Kasus kesembuhan demam berdarah (DBD) pada tulisan ini menunjukkan bahwa yang paling berpengaruh adalah stadium pasien saat mendapatkan penanganan medis. Dengan memperhatikan angka prevalen rasionya dapat dikatakan bahwa pasien dengan stadium 1, 2, dan 3 mempunyai peluang untuk sembuh sekitar 1.26 kali lipat lebih besar daripada pasien yang dtangani saat telah mencapai stadium 4 (amat parah). Atau dengan kata lain kegagalan penangan DBD akan naik dengan terlambatnya penanganan.
3. Pada studi epidemiologi, yaitu pada studi kasus demam berdarah (DBD) yang dgunakan pada penelitian ini, dilihat dari nilai rasio prevalensinya maupun nilai adjusted p-valuenya variabel penjas stadium merupakan variabel penjas yang mempunyai pengaruh peluang untuk sembuh paling besar.

UCAPAN TERIMA KASIH

1. Program Hibah Kompetisi A-2 Jurusan Matematika FMIPA Universitas Lampung
2. Seluruh pihak yang telah membantu pengumpulan data.