ISSN 1907-6452

Centre

Civil & Electrical Engineering Journal

Perancangan dan Pembuatan Mesin Jig Test pada Mesin Oven Test untuk Mendeteksi Scrap pada PCB Di PT. Venture Jaya Batam
Dian Tresnawan, Gilang Priyonggi Medansyah

Modifikasi Design Proses Blok Leadframe untuk Menekan Jumlah Defect Unit pada Mesin Die Attach di PT. Unisem Batam
Hadi Puspo Handoyo, Agus Salim

An Experimental Investigation On Comparison Of Heat Transfer Enhancement By Single-Salt Solution And Salts Mixture In The Real Water Used For Spray Cooling
Sabariman

Inprovement Dalam Menekan Pemakaian Daya Mesin Dengan Cara Setup Temperatur Pada Mesin Tamura
Nona Mahditiara Aryuni S., Roni Ardiweri

Algoritma EM STEP Pada Estimasi Model Multilevel
Yayuk Setyaning Astutik

Analisis Kinerja Ruang Parkir di Pelabuhan Ferry International Batam Center
Emil Adly, Prayogo

Volume 10 | Number 1 | June 2015

Centre

Civil & Electrical Engineering Journal
ISSN: 1907-6452
Volume 10, Number 1, June 2015

Editorial Board

Chief Editor	:	Ni'matul Ma'muiyah, M.Eng
Managing Director	:	Nona Mahditiara A. Sumanang, ST., M.Eng
	:	Ir. Mulia Pamadi, MM., M.Hum
Editor	:	Iman Purwoto, ST., MT (Urban Planning Engineering) Universitas Internasional Batam
	:	Emil Adly, ST., M.Eng (Transportation Engineering) Universitas Internasional Batam
	:	Hendra, ST., MT. (Transportation Engineering) Universitas Internasional Batam
	:	Wahyu Setyo Pambudi, ST. (Electronics Engineering) Universitas Internasional Batam
	:	Ni'matul Ma'muiyah, M.Eng (Telecommunication Engineering) Universitas Internasional Batam
	:	Andik Yulianto, ST., MT. (Control System Engineering) Universitas Internasional Batam

Editor and Administration Office	:	Universitas Internasional Batam Jl. Gajah Mada, Baloi Sei Ladi – Batam, Indonesia Telp: +62-778-7437 111 (Hunting) Fax: +62-778-7437 112 E-mail: centre@uib.edu Website: http://www.uib.edu

CENTRE Journal is a scientific journal publised by Universitas Internasional Batam two times year (June and December). Submission of the article on this journal is addressed to editorial office above. All articles will be subjected to editorial review and no charges for submission.

1907 6454

Centre

Civil & Electrical Engineering Journal
ISSN: 1907-6452
Volume 10, Number 1, June 2015

Perancangan dan Pembuatan Mesin Jig Test pada Mesin Oven Test untuk Mendeteksi Scrap pada PCB Di PT. Venture Jaya Batam 1
Dian Tresnawan, Gilang Priyonggi Medansyah

Modifikasi Design Proses Blok Leadframe untuk Menekan Jumlah Defect Unit pada Mesin Die Attach di PT. Unisem Batam 7
Hadi Puspo Handoyo, Agus Salim

An Experimental Investigation On Comparison Of Heat Transfer Enhancement By Single-Salt Solution And Salts Mixture In The Real Water Used For Spray Cooling 13
Sabariman

Inprovement Dalam Menekan Pemakaian Daya Mesin Dengan Cara Setup Temperatur Pada Mesin Tamura 21
Nona Mahditiara Aryuni S., Roni Ardiweri

Algoritma EM STEP Pada Estimasi Model Multilevel 29
Yayuk Setyaning Astutik

Analisis Kinerja Ruang Parkir di Pelabuhan Ferry International Batam Center 35
Emil Adly, Proyogo

Algoritma EM STEP Pada Estimasi Model Multilevel

Yayuk Setyaning Astutik[1)]
Program Studi Teknik Sipil, Universitas Internasional Batam
Email : [1)] *yayuk@uib.ac.id* , *yayuksetyaninga@gmail.com*

Abstrak

EM STEP adalah salah satu metode estimasi parameter yang berbentuk iteras untuk memperoleh titik maksimum pada *fungsi maximum likelihood* yang mempunyai bentuk penyelesaian *close form*. Pada tahap E STEP akan diperoleh ekspektasi bersyarat dan penduga parameternya. Sedangkan pada tahap M STEP parameter yang diperoleh akan digunakan sebagai permulaan E STEP berikutnya. Algoritma EM STEP ini bisa digunakan untuk mencari estimasi pada model multilevel.

Kata Kunci: *Algoritma, EM STEP, Estimasi, Multilevel.*

1. Pendahuluan

Model-model multilevel sering disebut model-model hierarki, efek campuran, kelompok, klaster yang banyak dijumpai dalam ilmu-ilmu sosial, kesehatan dan ilmu pengetahuan lainnnya [1][2]. Analisis multilevel sangat cocok diterapkan untuk data berkelompok atau data dengan klaster serta pada kasus yang diperoleh dari penarikan sampel multi tahap (*multistage sampling*) [3].

Misalkan dalam bidang kesehatan, pasien-pasien yang hipertensi dalam suatu keluarga yang hipertensi mempunyai derajat keparahan yang relatif sama jika dibandingkan dengan pasien-pasien yang hipertensi dalam keluarga yang tidak hipertensi. Hal ini disebabkan adanya perbedaan besarnya paparan dari masing-masing keluarga.

Pada bidang pendidikan adanya perpanjangan dari individu yang diamati diambil dari kelas disuatu sekolah pada wilayanh tertentu. Unit-unit dalam struktur data berhierarki seperti ini mempunyai 4 level yang berbeda, level 1 adalah siswa, level 2 adalah kelas-kelas, level 3 adalah sekolah-sekolah, dan level 4 adalah wilayah. Unit percobaan pada level 1 tersarang (*nested*) dalam unit pada level yang lebih tinggi berikutnya.

Dalam beberapa kasus tertentu dimana data mempunyai struktur bertingkat atau berkelompok yaitu unit-unit pada level yang rendah tersarang dalam unit-unit yang lebih tinggi, analisis menggunakan model regresi logistik dirasa kurang tepat digunakan. Karena model yang dihasilkan kurang mampu mengestimasi model yang diinginkan. Hal ini dikarenakan regresi logistik tidak dapat menganallisis variansi antar kelompok yang terjadi pada data bertingkat di mana observasi-observasi pada data bertingkat tidak sepenuhnya independen sehingga akan dihasilkan tingkat kesalahan yang lebih besar dibandingkan dengan menggunakan model regresi logistik 2 level.

Tujuan dari penelitian ini adalah mencari metode estimasi untuk memperoleh titik maksimum dari *fungsi maximum likelihood* pada regresi multilevel dengan pendekatan iterasi.

2. Kerangka Teoritis

2.1. Pengertian Model Multilevel

Model regresi yang melibatkan satu variabel independen dan fungsi regresinya linier dapat ditulis :

$$y_i = \beta_0 + \beta_1 x_i + e_i \tag{1}$$

di mana β_0 adalah intersep, β_1 adalah koefisien regresi untuk x_i dan e_i adalah *error*.

Model persamaan dengan data observasi yang memiliki struktur bertingkat atau multilevel di mana terdapat lebih dari satu level pada variabel independen yang diduga berpengaruh terhadap variabel dependen, diperoleh persamaan :

$$y_i = \beta_{0j} + \beta_{1j} x_{ij} + e_{ij} \text{ dengan } e_{ij} \sim N\left(0, \sigma^2\right) \tag{2}$$

Di mana $j = 1, 2, \ldots, m$ menunjukkan unit level 2 (missal : sekolah) dan $i = 1, 2, \ldots, n_j$ menunjukkan unit level 1 (misal:siswa). y_{ij} adalah variabel dependen dari siswa ke-i dan sekolah ke-j, β_{0j} dan β_{1j} adalah koefisien regresi sekolah ke-j, x_{ij} adalah variabel independen dari siswa ke-i dan sekolah ke-j, dan e_{ij} adalah *error* dari siswa ke-i dan sekolah ke-j.

2.2. Model Multilevel 2 Level

Perbedaan model regresi 2 level dengan model regresi biasa adalah pada model regresi 2 level diasumsikan koefisien β_{0j} dan β_{1j} berbeda pada masing-masing sekolah dan merupakan variabel random. Sehingga model struktur bertingkat atau multilevel adalah :

$$\begin{aligned} \beta_{0j} &= \beta_0 + u_{0j} \\ \beta_{1j} &= \beta_1 + u_{1j} \\ u_{0j} &\sim N\left(0, \sigma_{u0}^2\right) \\ u_{1j} &\sim N\left(0, \sigma_{u1}^2\right) \end{aligned} \tag{3}$$

Dengan β_0, β_1 adalah koefisien regresi pada masing-masing siswa (level 1), u_{0j}, u_{1j} adalah variabel random *error* level 2 yang berdistribusi normal multivariat dengan parameter :

$$\begin{aligned} &E\left(u_{0j}\right) = E\left(u_{1j}\right) = 0 \\ &Var\left(u_{0j}\right) = \sigma_{u0}^2;\ Var\left(u_{1j}\right) = \sigma_{u1}^2; \\ &Cov\left(u_{0j}, u_{1j}\right) = \sigma_{u_{01}} \end{aligned} \tag{4}$$

Sehingga dapat dibentuk persamaan baru :

$$y_{ij} = \beta_0 + \beta_1 x_{ij} + u_{0j} + u_{1j} x_{ij} + e_{0ij} \tag{5}$$

dengan $\beta_0 + \beta_1 x_{ij}$ merupakan efek tetap dan $u_{0j} + u_{1j} x_{ij} + e_{0ij}$ merupakan efek random dengan $Var\left(e_{0ij}\right) = \sigma_{e0}^2$.

2.3. Estimasi Model Multilevel

Berdasarkan model komponen variansi 2 level, ketika residual mempunyai distribusi normal maka dapat dilakukan estimasi parameter pada model tersebut dengan menggunakan MLE (*Maximum Likelihood Estimation*), sehingga akan diperoleh estimator untuk koefisien tetap :

$$\beta = \left(X^T V^{-1} X\right)^{-1} X^T V^{-1} Y \tag{6}$$

Di mana kasus ini diperoleh :

$$X = \begin{pmatrix} 1 & x_{11} \\ 1 & x_{21} \\ \vdots & \vdots \\ 1 & x_{nj}j \end{pmatrix} \quad Y = \begin{pmatrix} y_{11} \\ y_{21} \\ \vdots \\ y_{nj}j \end{pmatrix}$$

Dengan j unit level 2, $_{nj}j$ unit level 1 dalam unit ke-j level 2 dan V adalah matriks blok diagonal di mana nilai dari variansi dianggap telah dikatahui [3].

2.4 Konsep Umum Model Regresi 2 Level

Model regresi 2 level umumnya mengunakan distribusi kontinu untuk variabel respon. Namun dalam beberapa kasus, diperlukan adanya distribusi diskrit untuk variabel respon. Maka, untuk mengatasi struktur data bertingkat pada variabel respon yang mengandung distribusi diskrit (biner) maka analisis yang dapat digunakan adalah model regresi logistik 2 level. Dalam model regresi logistik 2 level terdapat 2 klasifikasi variabel yang berperan yaitu :

1) Variabel dependen yaitu variabel yang telah ditetapkan dalam model regresi logistik 2 level yaitu biner atau mempunyai 2 nilai (0 jika gagal dan 1 jika sukses).
2) Variabel independen yaitu variabel yang bebas, tidak bergantung degan variabel lain dan biasa disebut variabel prediktor/penjelas.

3. Metodologi Penelitian

Penelitian ini bersifat kajian teori. Terkait dengan estimasi pada metode statistika dengan penguraian metode secara matematis, sehingga diperoleh estimasi yang tepat dari metode tersebut.

3.1. Spesifikasi Model Regresi Logistik 2 Level

Model regresi logistik 2 level bila $y = 1$ menyatakan sukses dan $y = 0$ menyatakan gagal, maka probabilitas akan terjadi sukses pada subyek level ke-i dan *cluster* level ke-j dapat dinyatakan dengan $P(Y=1) = \pi_{ij}$ dan probabilitas akan terjadi gagal pada subyek level ke-i dan *cluster* level ke-j dapat dinyatakan dengan $P(Y=0) = \pi_{ij}\left(1 - \pi_{ij}\right)$.

Dalam model regresi logistik 2 level terdapat 2 level pengamatan yaitu level 1 yang menjelaskan tentang tingkat individu dan level 2 yang menjelaskan tentang grup. Probabilitas hubungan untuk kondisi x_{ij} oleh fungsi logit yang merupakan hasil persamaan gabungan dari pengamatan kevel 1 dan level 2 dapat ditulis :

$$\begin{aligned} \log it\left(\pi_{ij}\right) &= \ln\left(\frac{\pi_{ij}}{1 - \pi_{ij}}\right) \\ &= \beta_0 + \beta_1 x_{ij} + u_{0j} + u_{1j} x_{ij} + e_{ij} \end{aligned} \tag{7}$$

di mana model level 1 :

$$\log it\left(\pi_{ij}\right)=\ln\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right)=\beta_0+\beta_1 x_{ij}+e_{ij} \tag{8}$$

sedangkan model level 2 :

$$\begin{aligned}\beta_{0j}&=\beta_0+u_{0j}\\ \beta_{1j}&=\beta_1+u_{1j}\end{aligned} \tag{9}$$

Dengan :

π_{ij} : probabilitas subyek ke- i dalam *cluster* ke j ;

β_{0j} : intersep dalam model untuk *cluster* ke- j ;

β_{1j} : *slope* atau kemiringan dalam model untuk *cluster* ke- j ;

x_{ij} : variabel prediktor ke- i dalam grup ke- j ;

u_{0j} : parameter yang menyatakan sampel random untuk level ke- j dengan $u_{0j}\sim N\left(0,\sigma_{u0}^2\right)$;

u_{1j} : parameter yang menyatakan sampel random untuk level ke- j dengan $u_{1j}\sim N\left(0,\sigma_{u1}^2\right)$;

e_{ij} : parameter yang menyatakan *error* untuk level ke- i dan level ke- j dengan $e_{ij}\sim N\left(0,\sigma^2\right)$;

$\log it\left(\pi_{ij}\right)$: prediktor linier yang terdiri dari efek tetap $\left(\beta_0+\beta_1 x_{ij}\right)$ dan efek random $\left(u_{0j}+u_{1j}x_{ij}+e_{ij}\right)$.

Dengan menggunakan fungsi logit link, maka :

$$\begin{aligned}\pi_{ij}&=f\left(\beta_0+\beta_1 x_{ij}+u_{0j}+u_{1j}x_{ij}+e_{ij}\right)\\ &=\left[1+\exp\left\{-\left(\beta_0+\beta_1 x_{ij}+u_{0j}+u_{1j}x_{ij}+e_{ij}\right)\right\}\right]^{-1}\end{aligned} \tag{10}$$

3.2. Asumsi Model Regresi Logistik 2 Level

Asumsi-asumsi yang mendasari model regresi logistik 2 level adalah :

1) Koefisien random (residual) level 2 $\left(u_{0j},\ldots,u_{pj}\right)$ independen dengan residual level 1 $\left(e_{ij}\right)$ dan semua residual mempunyai *mean* populasi 0 untuk semua variabel prediktor/independen.
2) Residual level 1 $\left(e_{ij}\right)$ mempunyai distribusi normal.
3) Variansi residual level 1 $\left(\sigma^2 I\right)$ konstan atau bersifat *homoskedastisitas*.
4) Variabel efek random level 2 $\left(u_{0j},\ldots,u_{pj}\right)$ mempunyai distribusi normal multivariat.

4. Analisa dan Pembahasan

4.1. *Maximum Likelihood Estimation* (MLE)

Dibawah asumsi distribusi bersyarat Y hasil dari distribusi *Bernoulli* dengan probabilitas η *fungsi likelihood bersyarat* dengan bentuk *Binomial* diberikan :

$$f\left(Y|U;\beta\right)=\prod_{j=1}^{J}\prod_{i=1}^{n}\left(\frac{\pi_{ij}}{1-\pi_{ij}}\right)^{y_{ij}}\left(1-\pi_{ij}\right)^{n_{ij}} \tag{11}$$

di mana Y adalah vektor yang terdiri dari respon dan β adalah koefisien efek tetap dengan $n_{ij}\equiv 1$ untuk data *biner*. Untuk memperoleh *marginal likelihood*, maka dengan cara memperbanyak *fungsi likelihood bersyarat* oleh densitas U dan menggabungkan efek random diperoleh :

$$L\left(\beta,\Omega\right)=\int f\left(Y|U;\beta\right)f\left(U;\beta\right)dU \tag{12}$$

di mana $f\left(U;\beta\right)$ adalah fungsi distribusi untuk efek random yang bergantung pada parameter yang tidak diketahui Ω sedangkan $f\left(Y|U;\beta\right)$ adalah fungsi distribusi untuk respon bersyarat pada residual U yang hanya bergantung pada parameter yang tidak diketahui β.

Untuk model regresi logistik 2 level di mana efek random diasumikan mempunyai distribusi normal multivariat, maka efek random akan independen dengan unit yang berseberangan sehingga diperoleh:

$$L\left(\beta,\Omega\right)=\prod_{j}\int\prod_{i}\left\{\left(\pi_{ij}\right)^{y_{ij}}\left(1-\pi_{ij}\right)^{n_{ij}-y_{ij}}\right\}\phi\left(u;\Omega\right)du \tag{13}$$

di mana $\phi\left(u;\Omega\right)$ adalah fungsi densitas normal multivariat untuk parameter u sedangkan y_{ij} adalah probabilitas kemungkinan sukses atau

gagal. Karena nilai *maximum likelihood* parameter random dengan menggunakan sampling tidak diketahui maka untuk mendapatkan pendekatan yang baik untuk Ω maka secara umum terdapat $\tilde{\Omega}$ sedemikian sehingga :

$$L(\beta,\Omega)=\int\frac{f(Y|U;\beta)f(U;\Omega)f(U;\tilde{\Omega})}{f(U;\tilde{\Omega})}dU \qquad (14)$$

dan dilakukan estimasi untuk unit ke- j menjadi :

$$\sum_{h=1}^{N}\frac{f(Y|U_j;\beta)f_h(U_j;\Omega)}{f_h(U_j;\tilde{\Omega})}\Big/N \qquad (15)$$

Selanjutnya fungsi *loglikelihood* $\ell(\beta,\Omega)$ diperoleh :

$$L(y,\beta,\Omega)=\prod_{j=1}^{J}\int\prod_{i=1}^{I}g(y_{ij}|u_j,\beta)\phi(u_j;\Omega)du_j$$

$$\cong\prod_{j=1}^{J}\frac{1}{N}\sum_{h=1}^{N}\prod_{i=1}^{I}\frac{g(y_{ij}|u_j,\beta)f_h(U_j;\Omega)}{f_h(U_j;\tilde{\Omega})}$$

$$\log L(y,\beta,\Omega)1\cong\log\prod_{j=1}^{J}\frac{1}{N}\sum_{h=1}^{N}\prod_{i=1}^{I}\frac{g(y_{ij}|u_j,\beta)f_h(U_j;\Omega)}{f_h(U_j;\tilde{\Omega})}$$

$$\cong\log\prod_{j=1}^{J}\frac{1}{N}\sum_{h=1}^{N}\prod_{i=1}^{I}\left\{\left(\pi_{ij}^{(h)}\right)^{y_{ij}}\left(1-\pi_{ij}^{(h)}\right)^{n_{ij}-y_{ij}}\right\}\frac{f_h(U_j;\Omega)}{f_h(U_j;\tilde{\Omega})}$$

$$=\sum_{j=1}^{J}\left(\log\left\{\sum_{h=1}^{N}\left[\prod_{i=1}^{J}\left\{\left(\pi_{ij}^{(h)}\right)^{y_{ij}}\left(1-\pi_{ij}^{(h)}\right)^{n_{ij}-y_{ij}}\right\}\frac{\phi(u_j^{(h)};\Omega)}{\phi(u_j^{(h)};\tilde{\Omega})}\right]\Big/N\right\}\right)$$

$$=\sum_{j=1}^{J}\left(\log\left\{\sum_{h=1}^{N}\exp\sum_{i=1}^{J}\left[y_{ij}\log\left(\pi_{ij}^{(h)}\right)\right.\right.\right.$$

$$\left.\left.\left.+\left(n_{ij}-y_{ij}\right)\log\left(1-\pi_{ij}^{(h)}\right)\right]+\log\left(\frac{\phi(u_j^{(h)};\Omega)}{\phi(u_j^{(h)};\tilde{\Omega})}\right)\right\}-\log(N)\right)$$

$$=\sum_{j=1}^{J}\log\left\{\sum_{h=1}^{N}\exp\left(\sum_{i=1}^{J}\left[y_{ij}\log\left(\pi_{ij}^{(h)}\right)+\left(n_{ij}-y_{ij}\right)\log\left(1-\pi_{ij}^{(h)}\right)\right]\right.\right.$$

$$\left.\left.+\left[\frac{1}{2}u_j^{(h)^T}\left(\tilde{\Omega}^{-1}-\Omega^{-1}\right)u_j^{(h)}+\frac{1}{2}\log\left(\frac{|\tilde{\Omega}|}{|\Omega|}\right)\right]\right)-\log(N)\right\} \qquad (16)$$

Dimana h adalah sekumpulan efek random dan u adalah sebuah vektor *(px1)* untuk setiap unit level 2 dimana p adalah sejumlah koefisien random. Unsur konstanta N dapat dihilangkan karena tidak berpengaruh terhadap pemilihan Ω.

$$\log L(y;\beta,\Omega)=\sum_{j=1}^{J}\log\sum_{h=1}^{N}\exp\left[\sum_{i=1}^{J}\left[y_{ij}\log\left(\pi_{ij}^{(h)}\right)\right.\right.$$

$$\left.\left.+\left(n_{ij}-y_{ij}\right)\log\left(1-\pi_{ij}^{(h)}\right)\right]\right]$$

$$+\sum_{j=1}^{J}\log\sum_{h=1}^{N}\exp\left[\frac{1}{2}u_j^{(h)^T}\left(\tilde{\Omega}^{-1}-\Omega^{-1}\right)u_j^{(h)}+\frac{1}{2}\log\left(\frac{|\tilde{\Omega}|}{|\Omega|}\right)\right]$$

$$=\sum_{j=1}^{J}\log\prod_{h=1}^{N}\sum_{i=1}^{J}\left[y_{ij}\log\left(\pi_{ij}^{(h)}\right)+\left(n_{ij}-y_{ij}\right)\log\left(1-\pi_{ij}^{(h)}\right)\right]$$

$$+\sum_{j=1}^{J}\log\prod_{h=1}^{N}\left[\frac{1}{2}u_j^{(h)^T}\left(\tilde{\Omega}^{-1}-\Omega^{-1}\right)u_j^{(h)}+\frac{1}{2}\log\left(\frac{|\tilde{\Omega}|}{|\Omega|}\right)\right]$$

$$=\sum_{j=1}^{J}\sum_{h=1}^{N}\sum_{i=1}^{J}\left[y_{ij}\log\left(\pi_{ij}^{(h)}\right)+\left(n_{ij}-y_{ij}\right)\log\left(1-\pi_{ij}^{(h)}\right)\right]$$

$$+\sum_{j=1}^{J}\sum_{h=1}^{N}\left[\frac{1}{2}u_j^{(h)^T}\left(\tilde{\Omega}^{-1}-\Omega^{-1}\right)u_j^{(h)}+\frac{1}{2}\log\left(\frac{|\tilde{\Omega}|}{|\Omega|}\right)\right]$$

$$=\sum_{j=1}^{J}NI\left[y_{ij}\log\left(\pi_{ij}^{(h)}\right)+\left(n_{ij}-y_{ij}\right)\log\left(1-\pi_{ij}^{(h)}\right)\right]$$

$$+\sum_{j=1}^{J}N\left[\frac{1}{2}u_j^{(h)^T}\tilde{\Omega}^{-1}u_j^{(h)}-\frac{1}{2}u_j^{(h)^T}\Omega^{-1}u_j^{(h)}+\frac{1}{2}\log\left(\frac{|\tilde{\Omega}|}{|\Omega|}\right)\right] \qquad (17)$$

Jika $\Omega = 0$ maka tahap dalam Ω dapat dihilangkan. Untuk menemukan solusi *maksimum likelihood* yang mendekatai Ω maka dibutuhkan pencarian numeris dengan bergantung pada sekumpulan β dan Ω. Penyelesaian dengan menggunakan *maximum likelihood* terlihat bahwa tidak dapat menghasilkan dalam bentuk *closed form,* maka digunakanlah *prosedur iteratif algoritma EM* untuk menyelesaikan data yang tidak lengkap akibat adanya pengelompokan.

4.2. Iterasi EM STEP

Ada beberapa metode yang dapat digunakan untuk memperoleh titik maksimum dari *fungsi maximum likelihood.* Salah satunya dengan pendekatan EM (*Expected Maximation*) *Step* untuk memperoleh Ω yang memaksimalkan *fungsi maximum likelihood* pada analisis regresi logistik 2 level. Algoritma EM merupakan sebuah metode optimasi iteratif untuk MLE yang berguna dalam permasalahan data tidak lengkap. Ide dasar algoritma ini adalah mengasosiasikan suatu *complete data problem* dengan *incomplete data problem* dengan tujuan agar secara komputasi menjadi lebih mudah. Setiap iterasi pada EM step

terdiri dari 2 tahap yaitu Ekspektasi (*E Step*) dan tahap Maksimasi (*M Step*).

4.3. Tahap E STEP

Bertujuan untuk menemukan ekspektasi bersyarat dari *missing data* dengan syarat data diketahui nilainya dan penduga parameternya, kemudian mensubtitusi nilai ekspektasi yang diperoleh terhadap *missing data*. Dalam hal ini *missing data* yang dimaksud adalah yang muncul dalam *complete data loglikelihood.*

$$Q(\Omega,\Omega_n) = E\left(\log L(y;\beta;\Omega)\middle| y,\Omega_n\right)$$

$$= E\left(\begin{array}{l}\sum_{j=1}^{J} NI\left[y_{ij}\log\left(\pi_{ij}^{(h)}\right)+\left(n_{ij}-y_{ij}\right)\log\left(1-\pi_{ij}^{(h)}\right)\right] \\ +\sum_{j=1}^{J} N\left[\frac{1}{2}u_j^{(h)^T}\tilde{\Omega}^{-1}u_j^{(h)}-\frac{1}{2}u_j^{(h)^T}\Omega^{-1}u_j^{(h)}+\frac{1}{2}\log\left(\frac{|\tilde{\Omega}|}{|\Omega|}\right)\right]\end{array}\middle| y,\Omega_n\right)$$

$$= \sum_{j=1}^{J} NI\left[y_{ij}\log\left(\pi_{ij}^{(h)}\right)+\left(n_{ij}-y_{ij}\right)\log\left(1-\pi_{ij}^{(h)}\right)\right]$$

$$+\sum_{j=1}^{J} N\left[\frac{1}{2}tr\left\{\tilde{\Omega}^{-1}N^{-1}\sum_{h=1}^{N}E\left(u_j^{(h)^T}u_j^{(h)}\middle| y,\Omega\right)\right\}\right]$$

$$\left.\begin{array}{l}-\frac{1}{2}tr\left\{\Omega^{-1}N^{-1}\sum_{h=1}^{N}E\left(u_j^{(h)^T}u_j^{(h)}\middle| y,\Omega\right)\right\} \\ +\frac{1}{2}\log\left(|\tilde{\Omega}|-|\Omega|\right)\end{array}\right]$$

$$= \sum_{j=1}^{J} NI\left[g_j^{(h)}\right]+\sum_{j=1}^{J} N\left[\frac{1}{2}tr\left\{\tilde{\Omega}^{-1}k_j^{(h)}\right\}\right]$$

$$-\frac{1}{2}tr\left\{\Omega^{-1}k_j^{(h)}\right\}+\frac{1}{2}\log\left(|\tilde{\Omega}|-|\Omega|\right)\Bigg] \qquad (18)$$

Dengan :

$$g_j^{(h)} = y_{ij}\log\left(\pi_{ij}^{(h)}\right)+\left(n_{ij}-y_{ij}\right)\log\left(1-\pi_{ij}^{(h)}\right)$$

$$k_j^{(h)} = N^{-1}\sum_{h=1}^{N}E\left(u_j^{(h)^T}u_j^{(h)}\middle| y,\Omega\right)$$

4.4. Tahap M STEP

Langkah untuk memperoleh Ω yang memaksimalkan *ekspektasi likelihood* yang telah didapat pada **E STEP**. Parameter-parameter yang didapat pada **M STEP** akan digunakan sebagai permulaan **E STEP** berikutnya. Proses ini akan terus berulang hingga didapatkan estimasi parameter yang dapat memaksimalkan harga harapan dari *fungsi likelihood* secara stabil (*convergence*). Iterasi ini dilakukan sampai diperoleh $\Omega^{(n+1)} \approx \Omega^{(n)}$, sehingga Ω konsisten.

Persamaan umum iterasi **M STEP** :

$$\Omega_{n+1} = \underset{\Omega}{\arg\max}\; Q(\Omega,\Omega_n)$$

$$= \underset{\Omega}{\arg\max}\; \sum_{j=1}^{J} NI\left[g_j^{(h)}\right]+\sum_{j=1}^{J} N\left[\frac{1}{2}tr\left\{\tilde{\Omega}^{-1}k_j^{(h)}\right\}\right.$$

$$\left.-\frac{1}{2}tr\left\{\Omega^{-1}k_j^{(h)}\right\}+\frac{1}{2}\log\left(|\tilde{\Omega}|-|\Omega|\right)\right] \qquad (19)$$

$$g_j^{(h)} = y_{ij}\log\left(\pi_{ij}^{(h)}\right)+\left(n_{ij}-y_{ij}\right)\log\left(1-\pi_{ij}^{(h)}\right)$$

$$k_j^{(h)} = \left\{N-1\sum_{h=1}^{N}E\left(u_j^{(h)^T}u_j^{(h)}\middle| y,\Omega\right)\right\}$$

5. Kesimpulan

Berdasarkan uraian formula di atas, dapat diambil kesimpulan bahwa EM STEP merupakan salah satu metode yang bisa digunakan untuk memperoleh titik maksimum dari *fungsi maximum likelihood* pada analisis regresi logistik 2 level. Algoritma EM STEP ini merupakan metode optimasi iteratif untuk MLE yang berguna dalam permasalahan data tidak lengkap (*missing data*).

Pada tahap E STEP akan diperoleh ekspektasi bersyarat dan penduga parameternya. Sedangkan pada tahap M STEP parameter yang diperoleh akan digunakan sebagai permulaan E STEP berikutnya.

6. Daftar Pustaka

[1] Rodriguez, G. dan Goldman, N.,(1995),An Assessment of Estimation Procedures for Multilevel Models with Binary Responses, Journal of the Royal Statistical Society (Series A), 151,hal. 73-89.

[2] Rabe-Hesketh, S. and Skrondal, A. (2008). *Classical latent variable models for medical research.* Statistical Methods in Medical Research 17, 5-32.

[3] Goldstein, H. (2011). *Multilevel Statistical Models.* West Sussex : John Wiley & Sons.

[4] Bain, L. J & Engelhardt, M. (1991). *Introduction to Probability and Mathematical Statistics.* Belmont : Duxbury.

[5] Hox, J. (2002). *Multilevel Analysis: Techniques and Applications* . New Jersey : Lawrence Erlbaum Associates, Inc.

[6] Hox, J and Mass, C. J. M (2005). *Multilevel Analysis: Encyclopedia of Social Measurement Vol-2*.

[7] Kusumaningrum, D. riska. (2012). *Model Regresi Logistik 2 Level*. Yogyakarta : UGM.

[8] Nindya, W. (2011). *Pemodelan Multilevel-2 Level pada regresi Linier Ganda*. Yogyakarta : UGM.