Bias Maximum Likelihood Estimator (MLE) dalam Model Multinomial Logit pada Respons Saling Berkorelasi

Oleh
Jaka Nugraha, Suryo Guritno, Sri Haryatmi
Univesitas Islam Indonesia

Abstraks

Model multinomial logit (MNL) didasarkan atas asumsi bahwa komponen errornya berdistribusi *extreme value* tipe I (Gumbel) dan saling independen. Dengan asumsi ini akan menghasilkan persamaan probabilitas masing-masing alternatif dalam bentuk persamaan tertutup dan konsisten dengan sifat *random utility*. Jika terdapat korelasi antar komponen error dalam masing-masing alternatif, maka dengan menggunakan maximum likelihood estimator akan menghasilkan estimator yang bias. Semakin tinggi korelasinya maka makin besar bias yang dihasilkan.
Kata Kunci : *Discrete Choice Model, Random Utiliti, Maximum Likelihood*

1. Pendahuluan

Pada umumnya pemodelan statistik menggunakan asumsi bahwa distribusi errornya adalah normal. Asumsi normalitas error dalam Dicrete Choice Model (DCM) dipakai dalam model probit. Kendala utama dalam model probit ini adalah estimasi parameter dan perhitungan probabilitas masing-masing alternatifnya akan melibatkan integral rangkap sehingga memerlukan teknik simulasi.(Harris ,2000 dan Contoyanis,2001)

Model Multinomial Logit menjadi popular, setelah McFadden (1974) membuktikan bahwa MNL dengan kondisi tertentu dapat diturunkan dari prinsip utiliti maksimum dan probabilitas masing-masing alternatifnya mempunyai bentuk persamaan tertutup. MNL memenuhi sifat DCM, yaitu pembuat keputusan memilih alternatif yang memberikan utiliti terbesar. Estimasi parameter dalam MNL dapat menggunakan metode Maximum Likelihood Estimator (MLE). MLE ini mempunyai sifat-sifat yang baik untuk sampel besar, khususnya *asymptotically efficient* (Horrowitz, 2001). Salah satu sifat penting dalam MLN adalah rasio probabilitas dua buah alternatif tidak tergantung pada alternatif lain selain kedua alternatif tersebut. Sifat ini dinamakan *independence from irrelevant alternatives* (IIA).

Sifat IIA didasarkan pada asumsi bahwa ε_{ij} saling independen untuk semua i dan j. Munizaga, Heydecker dan Ortuzar (1997) telah menunjukkan bahwa dengan mengabaikan korelasi antar alternatif akan menghasilkan estimator bias. Train (2003) menyampaikan bahwa MNL dapat merepresentasikan variasi sistematik dari karakteristik pembuat keputusan tetapi tidak dapat menangani adanya korelasi antar pilihan. Untuk data respon biner, Prentice (1988) telah menyampaikan stategi untuk

Dipresentasikan dalam Seminar Nasional MIPA 2007 dengan tema **"Peningkatan Keprofesionalan Peneliti, Pendidik & Praktisi MIPA"** yang diselenggarakan oleh Fakultas MIPA UNY Yogyakarta pada tanggal 25 Agustus 2007

menangani adanya korelasi antar respon dengan menggunakan generalized estimating equation (GEE) untuk mendapatkan estimator/penaksir parameter yang bersifat konsisten dan asimtotis normal.

Dalam makalah ini akan dibahas bagaimana pengaruh adanya korelasi antar anternatif dalam MLN terhadap sifat estimator parameter menggunakan metode Maksimum Likelihood Estimator. Pengamatan didasarkan pada data simulasi dan estimasi parameter maupun perhitungan probabilitas alternatifnya menggunakan program R. 2.5.0.

2. Model Multinomial Logit

Seorang pembuat keputusan dinotasikan dengan i, yang berhadapan dengan pilihan sebanyak J anternatif. Pembuat keputusan mempunyai tingkat utiliti (keuntungan) untuk setiap alternatif. Misalkan U_{ij} untuk j=1,…,J adalah utiliti pembuat keputusan i jika memilih alternatif j. Nilai U_{ij} yang sesungguhya tidak diketahui oleh pengamat (peneliti). Tentunya pembuat keputusan memilih alternatif yang mempunyai utiliti terbesar, sehingga memilih alternatif k jika dan hanya jika $U_{ik} > U_{ij}\ \forall j \neq k$. Karena nilai utiliti yang sesungguhnya tidak diketahui peneliti maka

$$V_{ij} \neq U_{ij} \text{ dan } U_{ij} = V_{ij} + \varepsilon_{ij} \tag{1}$$

$\varepsilon_i = (\varepsilon_{i1}, \ldots, \varepsilon_{iJ})'$ adalah variabel random yang mempunyai densitas $f(\varepsilon_i)$. V_{ij} merupakan faktor terobservasi yang memuat atribut pembuat keputusan dan atribut yang berkaitan dengan masing-masing alternatif. ε_{ij} merupakan faktor tidak terobservasi dalam utiliti.

Beberapa sifat utiliti yang berkaitan dengan spesifikasi dan estimasi parameter dalam DCM adalah sifat *"Only differences in utility matter"*. Penambahan dengan konstanta tertentu terhadap semua U_{ij}, tidak akan merubah utiliti tertingginya (peringkat utiliti).

$$P_{ij} = P(U_{ik} > U_{ij}) = P(U_{ik} - U_{ij} > 0) \quad \forall k \neq j, \tag{2}$$

hanya tergantung pada selisih dalam utility. Sifat yang lain adalah "*The scale of utility is arbitrary*". Dengan mengalikan setiap U_{ij} dengan bilangan positif λ tidak akan merubah peringkat utilitinya.

Probabilitas pembuat keputusan i memilih alternatif k dapat dinyatakan sebagai

$$\begin{aligned} P_{ik} &= P(U_{ik} > U_{ij}) \quad \forall j \neq k \\ &= P(\varepsilon_{ij} - \varepsilon_{ik} < V_{ik} - V_{ij}) \quad \forall j \neq k \end{aligned}$$

$$= \int_{\varepsilon} I(\varepsilon_{ij} - \varepsilon_{ik} < V_{ik} - V_{ij}) f(\varepsilon_i) d\varepsilon_i \quad \forall j \neq k$$

(3)

Model MNL diturunkan dengan asumsi bahwa ε_{ij} berdistribusi *extreme value* untuk semua j. Fungsi densitas extreme value tipe I (Gumbel) adalah

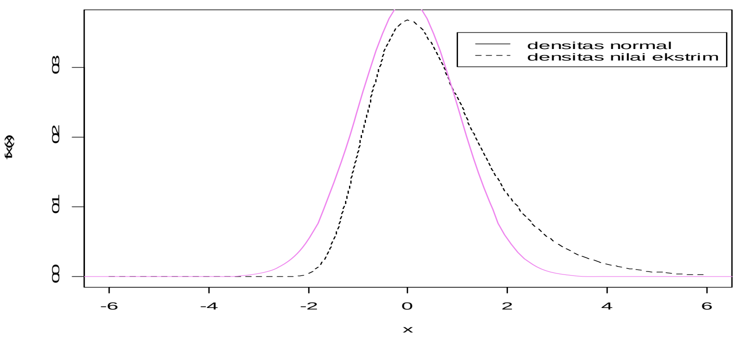
$$f(\varepsilon_{ij}) = \exp(-\varepsilon_{ij}).\exp(-\exp(-\varepsilon_{ij})) \quad (4)$$

Mean dan variansi dari distribusi ini masing-masing adalah 0,5772 dan $\pi^2/6$.

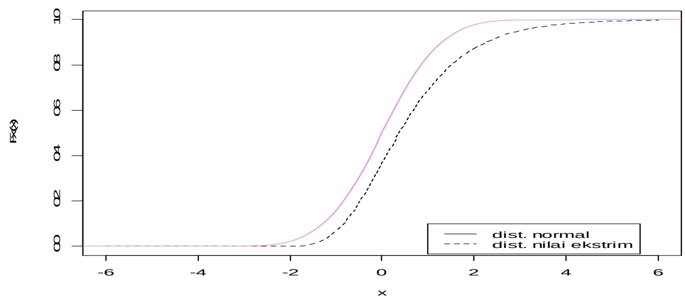
Distribusi kumulatif *extreme value* tipe I adalah

$$F(\varepsilon_{ij}) = \exp(-\exp(-\varepsilon_{ij})) \quad (5)$$

Fungsi densitas extreme value ini hampir simetrik dengan distribusi normal standard, hanya saja distribusi ini mempunyai ekor yang lebih tebal dibanding dengan distribusi normal. Grafik fungsi densitas dan fungsi distribusi extreme value jika dibandingkan dengan distribusi Normal Standart ditampilkan di gambar 1. dan Gambar 2.



Grafik 1. Densitas distribusi normal standar dan distribusi extreme value



Grafik 2. Densitas kumulatif distribusi normal standar dan distribusi extreme value

Probabilitas pembuat keputusan i memilih alternatif k dinyatakan sebagai:

$$P_{ik} = \frac{\exp(V_{ik})}{\sum_{j=1}^{J} \exp(V_{ij})} \quad (6)$$

Formula ini dinamakan probabilitas logit (Train, 2003) . Jika V_{ij} merupakan fungsi linear dari x_{ij} maka dapat dinyatakan menjadi

$$P_{ik} = \frac{\exp(\beta^t x_{ik})}{\sum_{j=1}^{J} \exp(\beta^t x_{ij})} \quad (7)$$

Estimasi parameter β dapat dilakukan dengan prosedur maksimum likelihood. Misalkan n sampel dari individu yang membuat keputusan, probabilitas individu i memilih sebuah alternatif dapat dinyatakan sebagai

$$\prod_{j} (P_{ij})^{y_{ij}}$$

Dengan y_{ij} = 1 jika individu i memilih j dan nol jika memilih yang lainnya. Dengan mengasumsikan bahwa setiap keputusan antar individu saling independen maka probabilitas masing-masing individu dalam sampel memilih sebuah alternatif adalah

$$L(\beta) = \prod_{i=1}^{n} \prod_{j} (P_{ij})^{y_{ij}} \quad (8)$$

Dengan β merupakan vektor parameter dalam model. Fungsi Log likelihood nya menjadi

$$LL(\beta) = \sum_{i=1}^{n} \sum_{j} y_{ij} \ln(P_{ij}) \quad (9)$$

Penaksir β adalah nilai β yang memaksimumkan fungsi LL(β).

Penaksir β dengan menggunakan prosedur maksimum likelihood adalah penyelesaian dari persamaan (Nugraha dkk, 2006)

$$\sum_{i=1}^{n} \sum_{j=1}^{J} (y_{ij} - P_{ij}) x_{ij} = 0 \quad (10)$$

Untuk sebarang dua alternatif j dan k, rasio probabilitas logitnya dapat dinyatakan sebagai

$$\frac{P_{ij}}{P_{ik}} = \exp(V_{ij} - V_{ik}) \quad (11)$$

Rasio ini tidak tergantung pada alternatif lain selain j dan k. Sifat ini dinamakan *independence from irrelevant alternatives* (IIA).

Uji hipotesis dan interval konvidensi untuk parameter (slope) adalah (Koppelman dkk, 2006)

a. Uji untuk masing-masing slope

$H_0 : \beta_j = \beta_j^{(0)}$ didasarkan pada statistik Wald :

$$Z_0 = \frac{B_j - \beta_j^0}{S\hat{E}(B_j)} \tag{12}$$

b. Uji untuk masing-masing slope

$H_0 : \beta_j = ... = \beta_q = 0$ didasarkan pada statistik

$$X^2 = G^2_{\text{model 1}} - G^2_{\text{model 2}} \tag{13}$$

yang berdistribusi chi kuadrat dengan derajad bebasnya sama dengan selisih banyaknya parameter dari kedua model. G^2 adalah deviance yang mempunyai nilai -2log L

Untuk menguji kecocokan model dapat digunakan statistik Pseudo R^2 yang identik dengan nilai R^2 (koefisien deterministik).

$$\text{pseudo } R^2 = 1 - \frac{G_1^2}{G_0^2} \tag{14}$$

Jika model secara sempurna memprediksi nilai Y ($P_i = 1$ maka $y_i = 1$ dan jika P_i=0 maka y_i=0) maka log L = 0 (atau nilai deviancenya nol). Sehingga nilai maksimum dari pseudo R^2 adalah satu. Statistik pseudo R^2 secara luas digunakan untuk menjelaskan kecocokan model dalam DCM secara intuitif. Pemasalahan dalam penggunaan pseudo R^2 ini adalah tidak adanya kaidah untuk menyatakan pada nilai berapa sedemikian hingga model dikatakan baik. Permasalahan kedua adalah peningkatan nilai pseudo R^2 pada penambahan variabel independen tidak dapat menjelaskan seberapa penting variabel tersebut (Koppelman dkk, 2006).

3. Rancangan Percobaan dan Membangkitkan data

Tujuan dari penelitian ini adalah mengamati pengaruh adanya korelasi antar alternatif pada MNL terhadap estimatornya. Data diperoleh dari membangkitkan data dengan nilai korelasi antar alternatifnya ditentukan. Pengamatan dilakukan dengan mengambil model untuk tiga alternatif dengan memasukan variabel atribut pembuat keputusan (X_i) dan variabel atribut masing masing alternatif (Z_{ij}). Variabel X biasa disebut variabel sosio ekonomik/geografi, misalnya penghasilan, jenis kelamin, asal daerah, jumlah anak. Sedangkan variabel Z_{ij} misalkan untuk pilihan penggunaan alat tranportasi (Bus, mobil pribadi, sepeda motor) maka Z_{ij} dapat berupa waktu tempuh, biaya.

Model $U_{ij} = X_iB_j + Z_{ij}\gamma + \varepsilon_{ij}$ untuk i=1,2,...,n dan j=1,2,3.

$U_{i1} = B_{01} + X_iB_1 + Z_{i1}\gamma + \varepsilon_{i1}$

$U_{i2} = B_{02} + X_iB_1 + Z_{i2}\gamma + \varepsilon_{i2}$

$U_{i3} = B_{03} + X_iB_3 + Z_{i3}\gamma + \varepsilon_{i3}$

Dengan mengambil alternatif ke-tiga sebagai *base line*, maka model terestimasinya menjadi

$U^*_{i1} = B^*_{01} + X_iB_{13} + Z_{i1}\gamma + \varepsilon_{i1}$

$U^*_{i2} = B^*_{02} + X_iB_{23} + Z_{i2}\gamma + \varepsilon_{i2}$

$U^*_{i3} = Z_{i3}\gamma + \varepsilon_{i3}$

dimana $B_{13} = B_1 - B_3$, $B_{23} = B_2 - B_3$, $B^*_{01} = B_{01} - B_{03}$ dan $B^*_{02} = B_{02} - B_{03}$Jadi terdapat 5 buah parameter yang akan diestimasi.

Data dibangkitkan pada nilai parameter B01=2 , B02=1, B03=0.2, B1=-3,B2=-2, B3=-1 dan γ=0.8. Diambil 6 struktur kovariansi ε_i , Cov(ε) =Σ yaitu

Model 1 : $\Sigma_A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. Model 2 :$\Sigma_B = \begin{pmatrix} 1 & 0.25 & 0.25 \\ 0.25 & 1 & 0.25 \\ 0.25 & 0.25 & 1 \end{pmatrix}$.

Model 3 : $\Sigma_C = \begin{pmatrix} 1 & 0.5 & 0.5 \\ 0.5 & 1 & 0.5 \\ 0.5 & 0.5 & 1 \end{pmatrix}$. Model 4 : $\Sigma_D = \begin{pmatrix} 1 & 0.9 & 0.9 \\ 0.9 & 1 & 0.9 \\ 0.9 & 0.9 & 1 \end{pmatrix}$.

Model 5 : $\Sigma_E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0.5 \\ 0 & 0.5 & 1 \end{pmatrix}$. Model 6 : $\Sigma_F = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0.9 \\ 0 & 0.9 & 1 \end{pmatrix}$

$X_i \sim NID(0,1)$, $Z_{ij} \sim NID(0,1)$ dan $\varepsilon_i \sim N(\mathbf{0},\boldsymbol{\Sigma})$

Untuk variabel X_i dan Z_{ij} yang berupa data diskrit tidak disampaikan disini karena menghasilkan kesimpulan yang sama dengan data kontinu. Dipilih n=100 dan replikasi masing-masing kovariansi sebanyak 20, 50 dan 100.

Untuk mendapatkan data multivariate normal tersebut digunakan persamaan

$\varepsilon = L\eta$ dan $\Sigma = LL^t$

dimana η~N(0,1) dan L didefinisikan sebagai matrik segi tiga bawah dari faktor Cholesky. Program R menyediakan fasilitas membangkitkan data multivariat normal dalam library ”MASS” dan program estimasi MLE terdapat dalam library ”MicEcon” (Henningsen, 2007).

Program disusun dalam dua tahap, pertama adalah proses membangkitkan data dengan distribusi dan struktuk kovariansi tertentu. Kedua, adalah melakukan estimasi parameter untuk mendapatkan model MLN.

4. Hasil dan Pembahasan

Menyusun nilai utiliti agar parameternya estimable sangat penting untuk dilakukan. sebagaimana dalam contoh, utiliti asli mempunyai 7 parameter, tetapi hanya 5 yang terestimasi. Pengurangan jumlah parameter ini tidak mengubah probabilitas pilihan maupun interprestasinya. Dengan mengambil parameter B01=2 , B02=1, B03=0.2, B1=-3,B2=-2, B3=-1 dan γ=0.8 maka utiliti masing-masing alternatif dan individu adalah

$$U^*_{i1} = 1.8 - 2X_i + 0.8Z_{i1} + \varepsilon_{i1}$$

$$U^*_{i2} = 0.8 - X_i + 0.8Z_{i2} + \varepsilon_{i2}$$

$$U^*_{i3} = 0.8Z_{i3} + \varepsilon_{i3}$$

Model probabilitas masing alternatif dan individu yang akan diestimasi adalah

$$\hat{p}_{1i} = \frac{\exp(b_{10} + b_1 X_i + cZ_{1i})}{\exp(b_{10} + b_1 X_i + cZ_{1i}) + \exp(b_{20} + b_2 X_i + cZ_{2i}) + \exp(cZ_{3i})}$$

$$\hat{p}_{2i} = \frac{\exp(b_{20} + b_2 X_i + cZ_{2i})}{\exp(b_{10} + b_1 X_i + cZ_{1i}) + \exp(b_{20} + b_2 X_i + cZ_{2i}) + \exp(cZ_{3i})}$$

$$\hat{p}_{3i} = \frac{\exp(cZ_{3i})}{\exp(b_{10} + b_1 X_i + cZ_{1i}) + \exp(b_{20} + b_2 X_i + cZ_{2i}) + \exp(cZ_{3i})}$$

dan $p_{1i} + p_{2i} + p_{3i} = 1$. Estimator yang akan dicari adalah b_{10}, b_{20}, b_1, b_2 dan c.

Data yang dibangkitkan untuk masing-masing kovariansi sebanyak 17.000 (dengan n=100, pada replikasi 20, 50 dan 100). Disamping itu juga dibangkitkan data n=500 dengan replikasi 20.

4.1. Bias dan nilai pseudo R^2 pada korelasi tinggi

Diambil tiga nilai replikasi, yaitu 20, 50 dan 100 dimaksudkan untuk melihat kestabilan masing-masing estimator. Diperoleh hasil bahwa rata-rata estimator yang dihasilkan pada ketiga nilai replihasi tidak berbeda secara signifikan. Pengujian dilakukan dengan uji anava satu arah, dengan hipotesis $H_0 : \theta_{R1} = \theta_{R2} = \theta_{R3}$. Pada model satu ($\Sigma_A$) diperoleh hasil

Tabel 1. Estimator pada model 1

	rata-rata θ pada replikasi			Signifikansi
parameter	20	50	100	
B10	2.445245	2.596763	2.558173	0.2892
B20	1.116578	1.24624	1.231416	0.3738
B1	-2.732932	-2.877022	-2.775457	0.06004
B2	-1.372098	-1.550125	-1.429905	0.04827
γ	1.093627	1.0728	1.077740	0.0689

Jada secara umum dapat disimpulkan bahwa replikasi tidak cukup berpengaruh terhadap estimator. Demikian juga pada model yang lain (Σ_B, Σ_C,Σ_D,Σ_E,Σ_F) diperoleh kesimpulan yang sama, yaitu rata-rata estimator pada ketiga nilai replikasi tidak berbeda secara signifikan.

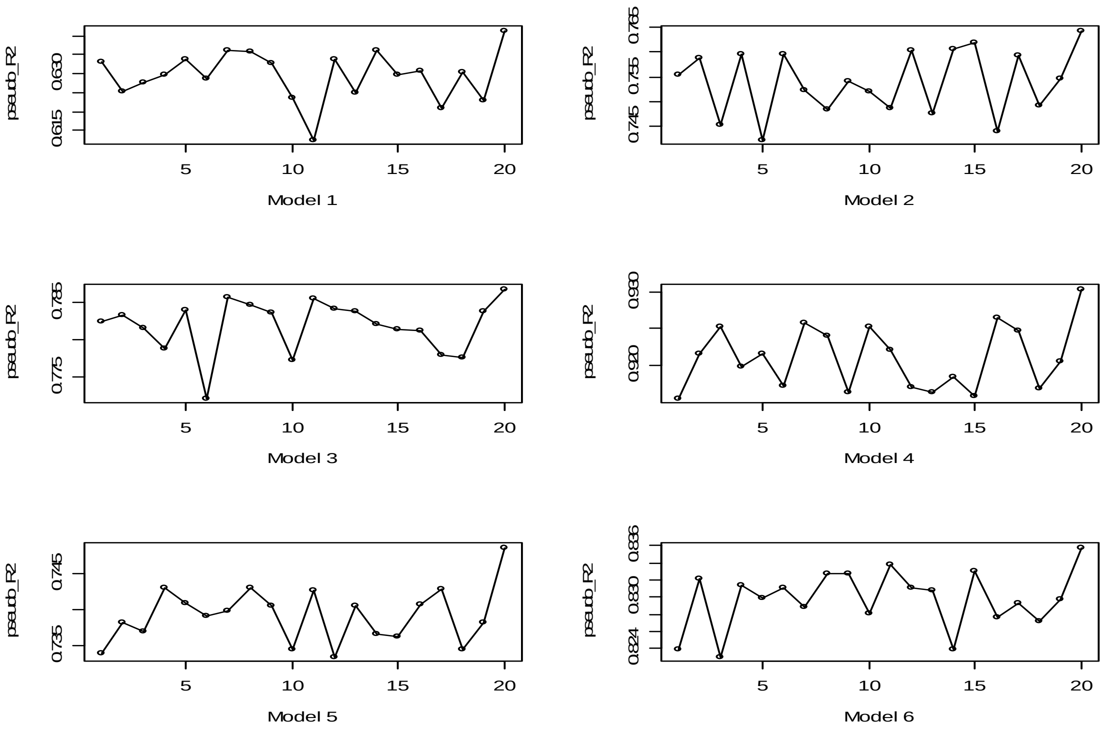
Adanya korelasi antar suku error (ε_j), dari tabel (2) nampak jelas berakibat pada bias estimator terhadap parameternya, Semakin tinggi tinggi tingkat korelasinya, maka bias estimator terhadap parameter semakin besar.

Tabel 2. Nilai estimasi parameter untuk replikasi 20 dengan n=100

prmt	$\sigma_{12}=\sigma_{13}=\sigma_{23}$ =0	$\sigma_{12}=\sigma_{13}=\sigma_{23}$ =0.25	$\sigma_{12}=\sigma_{13}=\sigma_{23}$ =0.5	$\sigma_{12}=\sigma_{13}=\sigma_{23}$ =0.9	$\sigma_{12}=\sigma_{13}=$ 0,σ_{23}=0.5	$\sigma_{12}=\sigma_{13}=$ 0,σ_{23}=0.9
B01	2.445245	2.971488	3.848194	9.361293	2.9512	3.688772
B02	1.116578	1.479699	1.962332	4.151584	1.555506	2.033002
B1	-2.732932	-3.226701	-4.333718	-10.52819	-3.012371	-3.475285
B2	-1.372098	-1.607597	-2.364538	-5.327487	-1.595781	-2.057577
γ	1.093627	1.214585	1.567595	4.140567	1.147604	1.351584

Yang perlu menjadi perhatian adalah pada korelasi tinggi (sekitar 0.9) maka kemungkinan tidak diperoleh estimator karena pada proses iterasi menghasilkan matrik hessian yang tidak positif definit. Pada jumlah sampel 100 dan replikasi sebanyak 100, beberapa diantaranya tidak diperoleh estimator karena matrik hessiannya tidak positif definit. Hal serupa juga terjadi ketika jumlah sampelnya relatif kecil walaupun dengan korelasi rendah. Dari beberapa percobaan, untuk 5 parameter minimum sampel yang dibutuhkan sebesar 500 agar estimatornya dapat diperoleh..

Semakin besar korelasi antar komponen error dari masing –masing alternatif menyebabkan semakin besar nilai bias eatimator. Dari percobaan yang telah dilakukan, semakin besar korelasinya juga semakin besar nilai pseudo R^2. (lihat grafik 3.).

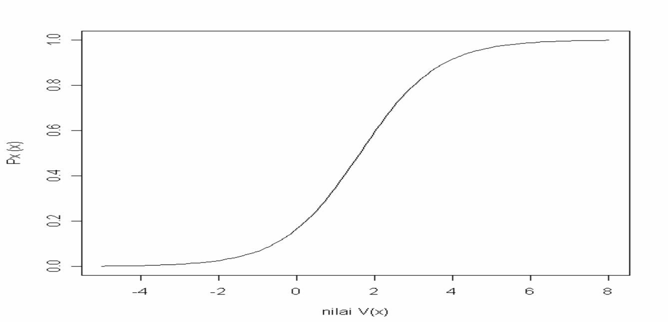


Grafik 3. Nilai R^2 pada n=500 replikasi 20

Peristiwa ini dapat dijelaskan sebagai berikut: persamaan (9) dan (14) menyatakan

$$\max(P_{i1},...P_{iJ}) \rightarrow 1 \Rightarrow LL(\hat{\beta}) \rightarrow 0 \Rightarrow \text{pseudo } R^2 \rightarrow 1$$

Nilai $p_{ij} \rightarrow 1$ untuk i yang memilih j dicapai pada nilai V_{ij} yang cukup besar, seperti terlihat dalam grafik (4) yang merupakan hubungan antara V_{ij} dan p_{ij} pada V_{ik} tetap untuk j≠k. Karena adanya korelasi antar alternatif mengakibatkan terjadinya bias dan dengan adanya bias ini berarti nilai mutlak semua estimator (koefisien) relatif lebih besar dibandingkan estimator pada korelasi rendah. Semakin besar estimator maka nilai V(x) semakin besar, sebaliknya semakin kecil estimator maka nilai V(x) semakin kecil.



Grafik 4. Nilai V(x) terhadap p(x)

Misalkan pada tiga alternatif pilihan, masing-masing mempunyai utiliti dengan komponen deterministiknya

$V_{i1} = b_{01} + X_i b_{13} + Z_{i1} c$

$V_{i2} = b_{02} + X_i b_{23} + Z_{i2} c$

$V_{i3} = cZ_{i3}$

Jika terjadi korelasi antar alternatif menghasilkan bias sebesar k (bernilai positif), maka

$V^*_{i1} = kb_{01} + kX_i b_{13} + kZ_{i1} c = kV_{i1}$

$V^*_{i2} = kb_{02} + kX_i b_{23} + kZ_{i2} c = kV_{i2}$

$V^*_{i3} = kcZ_{i3} = kV_{i3}$

Dengan adanya faktor pengali k, maka berdasarkan persamaan (6) diperoleh rasio

$\exp(V^*_{i1}) : \exp(V^*_{i2}) : \exp(V^*_{i3}) = [\exp(V_{i1})]^k : [\exp(V_{i2})]^k : [\exp(V_{i3})]^k$

Probabilitas masing-masing alternatifnya menjadi

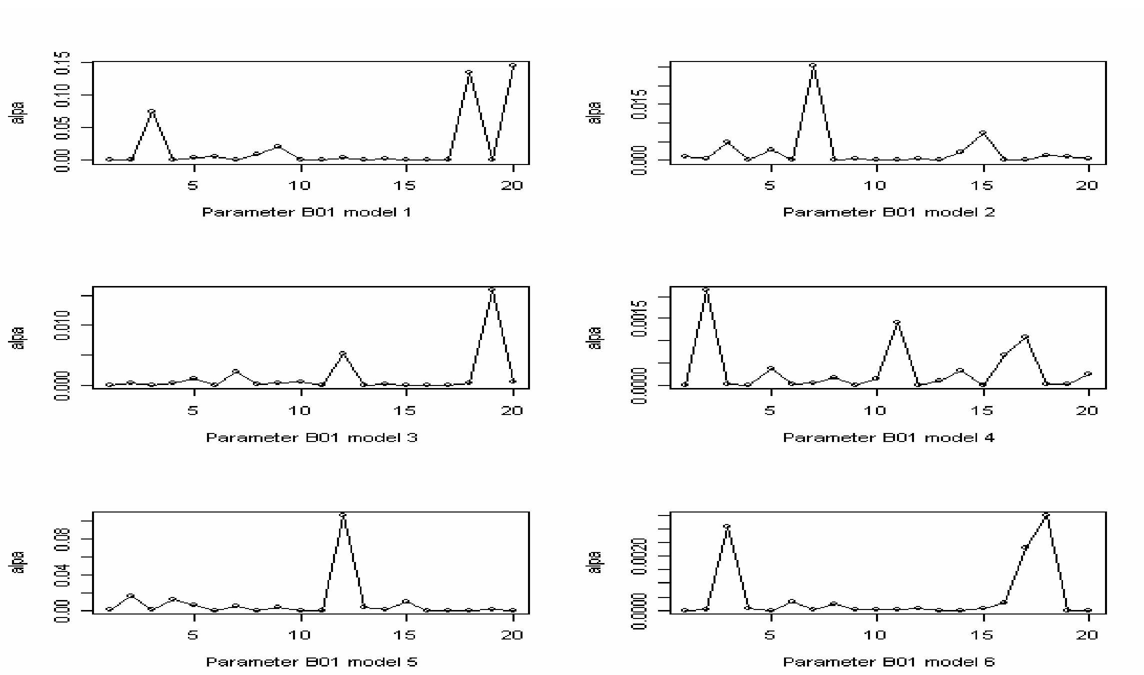
$$\hat{p}_{ij} = \frac{[\exp(V_{ij})]^k}{[\exp(V_{i1})]^k + [\exp(V_{i3})]^k + [\exp(V_{i3})]^k} \quad \text{untuk j=1,2,3.}$$

Dengan adanya bias berakibat nilai $\min(\hat{p}_{i1}, \hat{p}_{i2}, \hat{p}_{i3})\ \forall_i$ akan semakin kecil dan $maks(\hat{p}_{i1}, \hat{p}_{i2}, \hat{p}_{i3})\ \forall_i$ akan semakin besar dibandingkan pada kondisi tidak ada korelasi. Selanjutnya hal ini mengakibatkan nilai psedo R^2 makin besar.

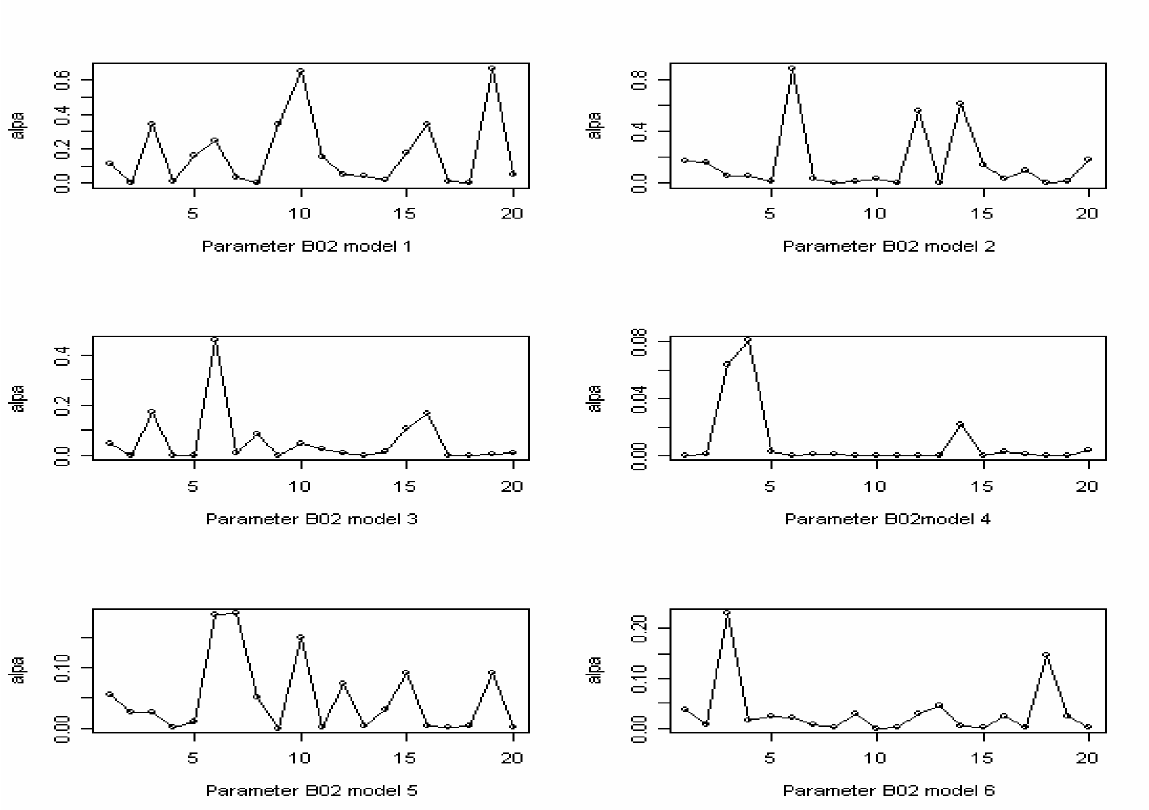
Mengingat konskekuensi adanya korelasi antar alternatif mengakibatkan nilai pseudo R^2 besar (mendekati satu), maka harus hati-hati dalam menggunakan statistik pseudo R^2. Pada kondisi nilai $\exp(V_{i1}), ..., \exp(V_{iJ})$ atau nilai probabilitas masing-masing alternatif relatif sama, statistik pseudo R^2 dari persamaan (9) dan (14) akan bernilai rendah.

4.2. Kesalahan Tipe I

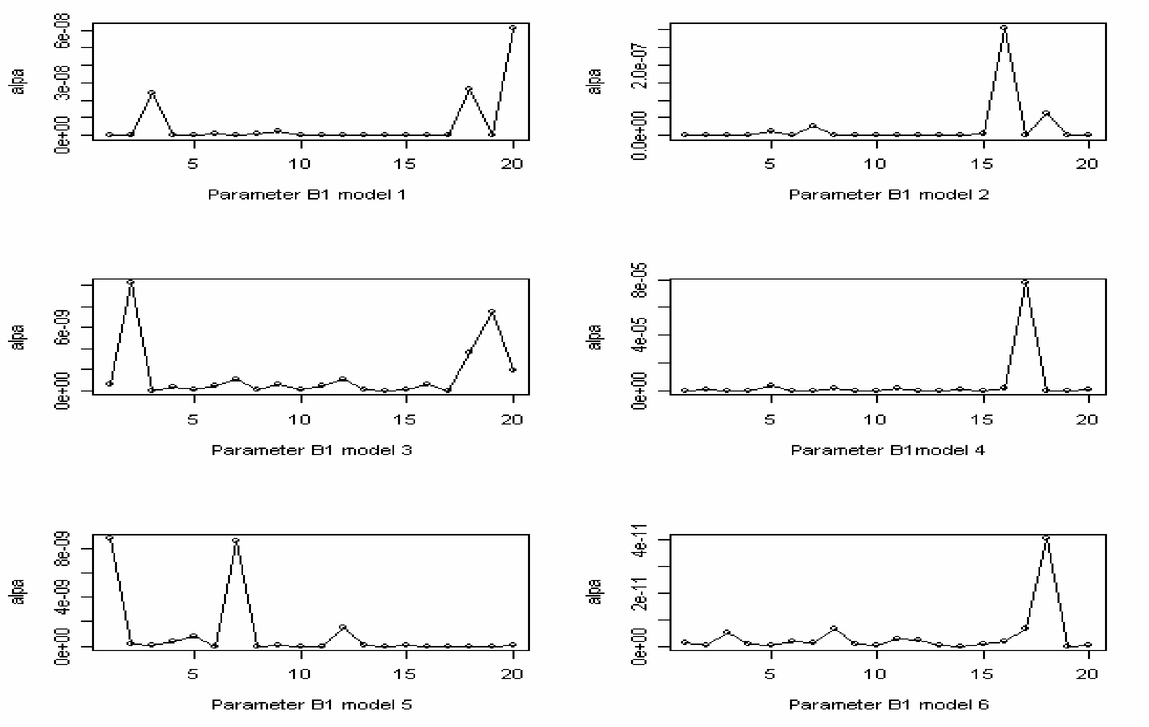
Hasil uji signifikansi parameter secara parsial dapat dilihat di grafik (5) s/d grafik (9). Secara umum dapat disimpulkan bahwa parameter 3, 4 dan 5 pada semua bentuk kovariansi pengaruhnya signifikan. Sementara itu untuk parameter 1 dan 2 (parameter konstanta) pada terdapat nilai alpa yang cukup besar (tidak signifikan) diantara 20 replikasi. Pada analisis berdasarkan data empirik, jika dijumpai uji parameternya hasilnya tidak signifikan sebaiknya dikeluarkan dari model. Sebagaimana pada kasus disini, yaitu parameter kontanta (B01 dan B02) pengaruhnya tidak signifikan, maka parameter ini sebaiknya tidak masuk dalam model. Karena yang dikeluarkan dari model adalah parameter kontanta, maka akan mengakibatkan nilai pseudo R^2 menjadi lebih rendah.



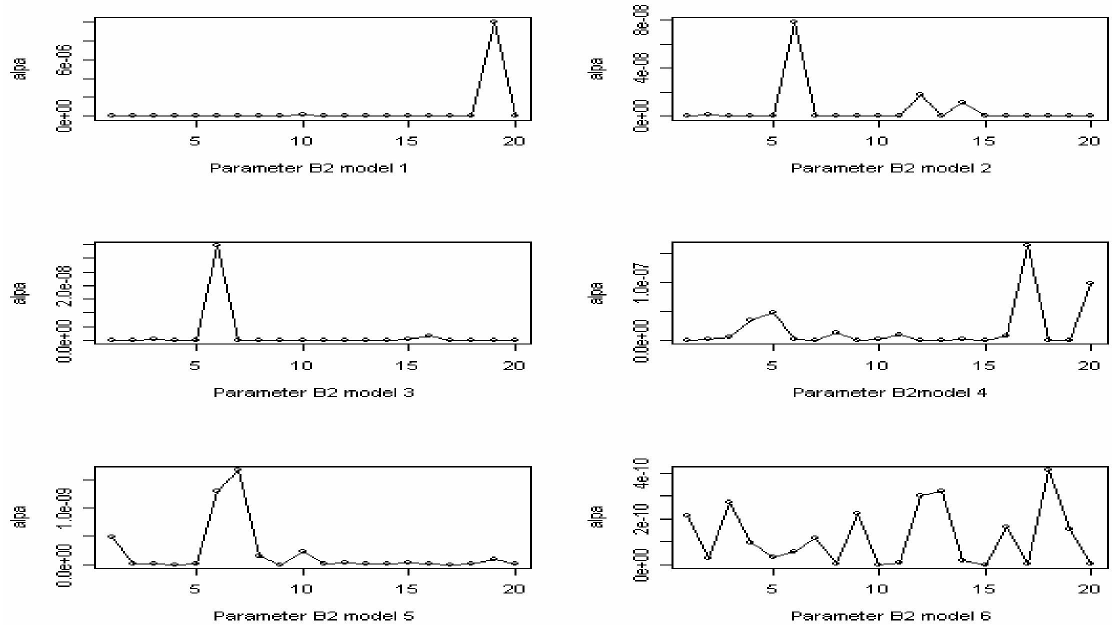
Gambar 5. Kesalahan tipe I (alpa) pada n=500 replikasi 20 untuk parameter B01



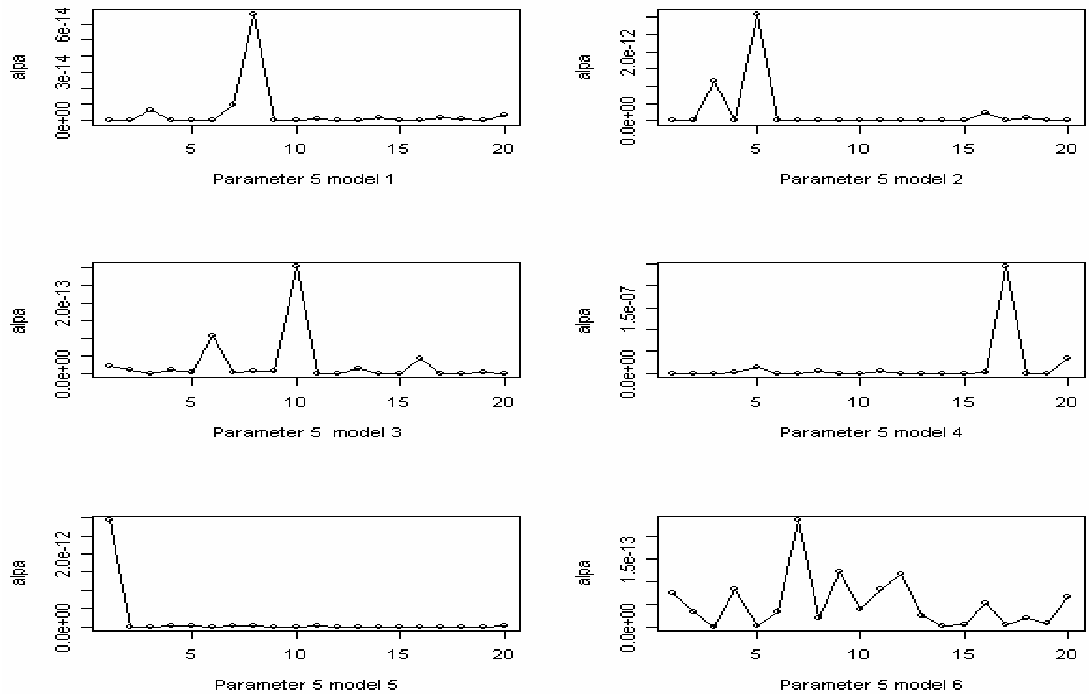
Gambar 6. Kesalahan tipe I (alpa) pada n=500 replikasi 20 untuk parameter B02



Gambar 7. Kesalahan tipe I (alpa) pada n=500 replikasi 20 untuk parameter B1



<u>Gambar 8.</u> Kesalahan tipe I (alpa) pada n=500 replikasi 20 untuk parameter B2



Gambar 9. Kesalahan tipe I (alpa) pada n=500 replikasi 20 untuk parameter γ

Berapa model lain yang telah dikembangkan untuk mengakomudasi adanya korelasi antar alternatif pilihan antara lain Model Probit Multinomial pendekatan klasik maupun bayesian. Ruud dkk (1996) mengembangkan strategi estimasi yang baru untuk

Model Multinomial Probit (MNP) yaitu didasarkan pada metode Simulated Moments. Pendekatan bayesian dilakukan oleh oleh McCullach dkk (2000) dan Imey dkk (2005).

Model lain yang telah dikembangkan dengan memasukkan unsur korelasi adalah model Generalized extreme Value (GEV). Model GEV merupakan pengembangan dari model logit yang disusun berdasarkan adanya korelasi antar alternatif pilihan yang tersedia. Jika semua korelasi diantara alternatif yang ada bernilai nol, maka GEV menjadi model logit standar. Beberapa jenis model yang temasuk dalam GEV antara lain model generalized nested logit (Wen dan Koppelman, 2001) dan model Cross-Nested Logit (Bhat , 2003).

5. Kesimpulan

Adanya korelasi antar komponen error masing-masing alternatif dalam model MNL mengakibatkan terdapat bias estimatornya. Dengan adanya bias ini, nilai pseudo R^2 cenderung besar. Sehingga jika diperoleh nilai pseudo R^2 mendekati satu, perlu dicurigai kemungkinan adanya korelasi. Untuk keperluan uji kecocokan model, harus hati-hati dalam menggunakan statistik pseudo R^2, apalagi pada kondisi nilai probabilitas antar alternatif relatif sama akan mengahsilkan nilai pseudo R^2 yang kecil.

Agar estimator masing-masing parameter dapat diperoleh dan dilakukan pengujian, untuk model MNL dengan 5 parameter disarankan menggunakan sampel sebesar 500 atau lebih.

Jika diduga terdapat korelasi, perlu dipertimbangkan penggunaan model yang lain seperti model MNP maupun Model GEV.

Daftar Pustaka

Contoyannis P, 2001, An Introduction to Simulation-Based Estimation, *Working Paper* Department of Economics and Related Studies, University of York.

Bhat CR., (2003), Econometric Choice Formulation : Alternative Model Structures, Estimation Techniques, and Emerging Direction. *International Conference on Travel Behaviour Research Lucerna.*

Harris, M.N and Macquarie L.R and Siouclis AJ. , (2000) ; Comparison of alternative Estimators for Binary Panel Probit Models, *Melbourne Institute Working Paper* no 3/00

Henningsen, Arne, (2007), The micEcon Package, http://www.r-project.org/.

Horowitz J.L. dan Savin N.E (2001), Binary Response Models : Logits, Probits and Semiparametrics, *Journal of Economic Perspectives, Volume 15, 4 pages 43-56*

Koppelman F.S and Bhat Chandra,(2006) , *A Self Intructing Course in Mode Choice Modeling : Multinomial and Nested Logit Models*, U.S. Department of Transportation Federal Transit Adminitration

Kosuke Imai and David A. van Dyk (2005), A Bayesian analysis of the multinomial probit model using marginal data augmentation, *Journal of Econometrics* 124. 311 – 334

McFadden, D. (1974), Conditional Logit Analysis of Qualitative Choice Behavior, dalam Zarembka, *Frontiers in Econometrics*, Academic Press, New York, pp. 105–142

Munizaga, Heydecker dan Ortuzar (1997) : On error structure of discrete choice models. *Traffic Engineering and Control*, Vo. 38, No. 11, 593-597.

Nugraha J, Haryatmi S. dan Guritno S, (2006), Model Discrete Choice dan Regresi Logistik, *Proseding Seminar Nasional MIPA*, UNY.

Prentice (1988), Correlated Binary Regression with Covariates Specific to Each Binary Observation. *Biometrics 44, 1043-1048.*

Robert E. McCulloch, Nicholas G. Polson and Peter E. Rossi, (2000), A Bayesian analysis of the multinomial probit model with fully identified parameters, *Journal of Econometrics* 99, 173-193

Ruud P.A. (1996), Approximation and Simulation of Multinomial Probit Model: An analysis of Covariance Matrix Estimation, *Working Paper* University of California, Berkeley.

Train, Kenneth (2003), *Discrete Choice Methods with Simulation*, UK Press, Cambridge

Wen, C.-H. and F.Koppelman (2001), The generalized nested logit model, *Transportation Research B* **35**, 627–641.

Lampiran.

```
#PROGRAM R
npara=5 #jumlah parameter yang diestimasi
N=500 #jumlah sampel
X=rnorm(N,3)
E=matrix(,N,3)
Z=matrix(,N,3)
U=matrix(,N,3)
V=matrix(,N,3)
Rep=20 #jumlah replikasi
esti=array(0,dim=c(npara,4,Rep))
s12=0 #korelasi 1 dan 2
s13=0 #korelasi 1 dan 2
s23=0 #korelasi 1 dan 2
sigma=matrix(c(1,s12,s13,s12,1,s23,s13,s23,1),3,3)
for (i in 1:3){Z[,i]=rnorm(N)}
B01=2; B02=1 ;B03=0.2 ;B1=-3 ; B2=-2 ; B3=-1 ;G=0.8
V[,1]=B01-B03 + X*(B1-B3) + Z[,1]*G
V[,2]=B02-B03 + X*(B2-B3) + Z[,2]*G
V[,3]= Z[,3]*G
n=1
EE=array(0,dim=c(N,3,Rep))
for (i in 1:Rep)
{
E=mvrnorm(N,c(0,0,0),sigma)
EE[,,n]=E
        U[,1]=V[,1]+ E[,1]
        U[,2]=V[,2]+ E[,2]
        U[,3]=V[,3]+ E[,3]
        Y=matrix(0,N,3)
for(i in 1:N)(if (U[i,1] > U[i,3] & U[i,1] >U[i,2]) Y[i,1]=1 else (if (U[i,2] >
U[i,1] & U[i,2] > U[i,3] ) Y[i,2]=1 else Y[i,3]=1))
hasil=summary( maxLik(MLE.logit, start=c(0,0,0,0,0)))
hasil2=hasil[6]$estimate
esti[,,n]<-hasil2
n=n+1
}

MLE.logit<-function(a)
{
b01=a[1]
b02=a[2]
b1=a[3]
b2=a[4]
g=a[5]
p1=sum(Y[,1]*(b01 + X*b1 + Z[,1]*g)- Y[,1]*log(exp(b01 + X*b1 + Z[,1]*g
)+exp(b02 + X*b2 + Z[,2]*g )+exp(Z[,3]*g)))
p2=sum(Y[,2]*(b02 + X*b2 + Z[,2]*g)- Y[,2]*log(exp(b01 + X*b1 + Z[,1]*g
)+exp(b02 + X*b2 + Z[,2]*g )+exp(Z[,3]*g)))
p3=sum(Y[,3]* (Z[,3]*g)- Y[,3]*log(exp(b01 + X*b1 + Z[,1]*g )+exp(b02 +
X*b2 + Z[,2]*g )+exp(Z[,3]*g)))
lg= p1+p2+p3
lg
}
```