MATEMATIKA, 2007, Volume 23, Number 1, 51–66

Perbandingan Teori Model Binari
(Comparisons of Theoritical Binary Models)

[1]**Mohd Saifullah Rusiman**, [2]**Zalina Mohd Daud** & [3]**Ismail Mohamad**
[1]Pusat Pengajian Sains, Kolej Universiti Teknologi Tun Hussein Onn
Peti Bersurat 101, 86400 Parit Raja, Batu Pahat, Johor, Malaysia
[2]Bahagian Akademik, Akademi Tentera Malaysia,
Kem Sungai Besi, 57000 Kuala Lumpur, Malaysia
[3]Jabatan Matematik, Fakulti Sains, Universiti Teknologi Malaysia
81310 UTM Skudai, Johor, Malaysia
e-mail: [1]saifulah@uthm.edu.my, [3]im@mel.fs.utm.my

Abstrak Tujuan kertas ini ialah untuk membincangkan teori bagi model logit, model probit dan model kebarangkalian linear. Perbandingan model dari segi penganggaran parameter β_k dilakukan dengan menggunakan kaedah kebolehjadian maksimum dan kaedah kuasadua terkecil berpemberat. Beberapa perbandingan ujian lain ditunjukkan seperti multikolineariti, ujian reja, pengujian hipotesis, kebagusan penyuaian dan diagnosis data berpengaruh serta data terpencil.

Katakunci Model Logit, Model Probit, Model Kebarangkalian Linear, Kaedah Kebolehjadian Maksimum, Kaedah Kuasadua Terkecil Biasa, Kaedah Kuasadua Terkecil Berpemberat.

Abstract *This paper discusses the theory of logit, probit and linear probability models. A comparison of the models' performance in terms of parameter β_k was done using the method of maximum likelihood and weighted least squared method. Several other comparison of tests are shown such as multicollinearity, residual test, hypothesis testing, goodness of fit test dan diagnosis of influential data and outliers.*

Keywords *Logit Model, Probit Model, Linear Probability Model, Maximum Likelihood Method, Ordinary Least Squares Method, Weighted Least Squares Method.*

1 Pengenalan

Dalam kajian analisis regresi, banyak data harian yang diperolehi melibatkan pembolehubah bersandar kualitatif binari seperti keputusan lulus atau gagal, jawapan "ya" atau "tidak" dan sebagainya. Jadi pembolehubah ini perlu dikodkan kepada 0 atau 1. Jika data ini dimodelkan dengan model regresi biasa, anggaran parameter yang didapati adalah kurang sesuai kerana tidak mematuhi andaian bagi nilai reja. Ini disebabkan oleh varians reja yang tidak tetap (heteroskedastik). Lantaran itu kajian menggunakan model kualitatif adalah perlu supaya anggaran parameter yang didapati adalah lebih menghampiri keadaan sebenar kerana mematuhi andaian bagi nilai reja.

Terdapat banyak pemodelan atau ujian yang boleh dilakukan terhadap sekumpulan data yang mengandungi satu pembolehubah bersandar dan sekurang-kurangnya satu pembolehubah tak bersandar. Di antara model atau ujian yang popular dijalankan ialah model log-linear, model logit, model probit, model regresi linear berganda biasa, model regresi linear berganda berpemberat, jadual kontigensi, ujian t, analisis varians dan analisis kovarians Dobson[8]. Bagaimanapun ujian-ujian atau pemodelan-pemodelan tersebut hanya sesuai dijalankan bergantung kepada jenis pembolehubah bersandar dan tak bersandar iaitu sama ada ianya berjenis binari, kategori atau selanjar.

Bagaimanapun Aldrich & Nelson[3] dan Dobson[8] menyatakan analisis data perlu dilakukan menggunakan model regresi, model logit dan model probit bagi mencari model yang terbaik bila pembolehubah bersandar berjenis binari atau kategori dan pembolehubah-pembolehubah tak bersandar berjenis binari, kategori atau selanjar. Ditambah lagi pemodelan logit, probit dan kebarangkalian linear merupakan pemodelan yang agak popular dewasa ini.

Umumnya hasil keputusan akan menjauhi kesimpulan yang sebenar sekiranya sesuatu data kuantitatif dikategorikan kerana terdapat sebahagian maklumat data tepat akan hilang atau terpendam. Kaedah penganggaran kuasa dua terkecil berpemberat dan penganggaran kebolehjadian maksimum juga mempunyai peratus ketepatan ramalan yang agak berbeza bila bilangan data yang dikaji berbeza. Aldrich dan Nelson[3] telah menunjukkan pengiraan bahawa bila bilangan data adalah kecil, kaedah penganggaran kuasa dua terkecil berpemberat menghasilkan keputusan ramalan yang lebih baik berbanding dengan kaedah penganggaran kebolehjadian maksimum.

Bila kajian yang dijalankan melibatkan bilangan data yang besar, kaedah penganggaran kebolehjadian maksimum akan menghasilkan kesimpulan yang lebih baik berbanding dengan kaedah penganggaran kuasa dua terkecil berpemberat. Finney[10] juga menyatakan bahawa penganggar kebolehjadian maksimum adalah lebih efisien bila melibatkan sampel besar.

Jadi di dalam sesuatu kajian, memang perlu perbandingan di antara pemodelan logit, probit dan kebarangkalian linear dijalankan untuk mengenalpasti kaedah mana yang lebih sesuai kerana bilangan data yang berbeza menghasilkan keputusan yang berbeza. Ditambah lagi, kemungkinan data itu akan tertabur menghampiri kepada sama ada taburan logistik, taburan normal piawai atau taburan linear berpemberat. Dalam sesuatu kajian, tidak berlaku keadaan yang mana hanya satu model digunakan. Sekiranya hanya satu model digunakan, mungkin ada model lain lagi yang terbaik yang dapat menggambarkan keadaan sebenar data. Jadi keputusan mengenai sesuatu kajian menjadi kurang tepat.

2 Penganggaran Model

2.1 Penganggaran Model Kebarangkalian Linear

Dalam penganggaran model kebarangkalian linear, terdapat dua pemodelan yang perlu dijalankan. Model regresi linear berganda biasa perlu diperolehi terlebih dahulu. Ini diikuti dengan mendapatkan model regresi linear berganda berpemberat.

2.1.1 Model Regresi Linear Berganda Biasa

Langkah pertama dalam pemodelan ini ialah menggunakan kaedah penganggaran kuasa dua terkecil biasa untuk mendapatkan nilai-nilai pemalar β_j. Pada tahun 1980 dan 1984, Lewis Beck dan Gauss Markov masing-masing menyatakan sebelum anggaran dibuat, model regresi mestilah memenuhi semua andaian-andaian klasik berikut iaitu;

(i) Ralat rawak, ε_i tidak bersandar dengan pembolehubah tak bersandar, $X_{ij}(j = 1, 2, ...k)$.

(ii) Ralat rawak, ε_i mempunyai min sifar atau $\mathrm{E}(\varepsilon_i) = 0$ (untuk setiap i).

(iii) Ralat rawak, ε_i mempunyai varians malar (homoskedastik) dan tidak berkorelasi, iaitu;

$$\mathrm{V}(\varepsilon_i) = \mathrm{E}(\varepsilon_i, \varepsilon_j) = \begin{cases} 0 \ ; & i = j \\ \sigma^2 \ ; & i \neq j \end{cases}$$

(iv) Ralat rawak, ε_i tertabur secara normal iaitu $\varepsilon_i \sim \mathrm{N}(0, \sigma^2)$.

Model regresi linear berganda boleh diungkap sebagai;

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + ...\beta_k X_{ik} + \varepsilon_i$$

$$\text{atau } \mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \text{ (Bentuk matriks)} \tag{1}$$

Bagi model regresi linear berganda biasa, fungsi kaedah kuasa dua terkecil adalah;

$$S(\beta_0, \beta_1, \beta_2, ..., \beta_k) = S(\boldsymbol{\beta}) = \sum_{j=0}^{k} \varepsilon_i^2 \quad \text{atau} \quad \varepsilon^T \varepsilon \tag{2}$$

Daripada persamaan (1), $\varepsilon = \mathbf{Y} - \mathbf{X}\boldsymbol{\beta}$. Maka $S(\boldsymbol{\beta})$ menghasilkan

$$\mathbf{Y^T Y} - 2\boldsymbol{\beta}^T \mathbf{X^T Y} + \boldsymbol{\beta}^T \mathbf{X^T X}\boldsymbol{\beta}.$$

Bagi menentukan penganggar kuasa dua terkecil, SSE (jumlah kuasa dua ralat) perlu diminimumkan iaitu terbitan bagi fungsi kaedah kuasa dua terkecil = 0.

$$\left.\frac{\partial S}{\partial \boldsymbol{\beta}}\right|_{\hat{\beta}} = -2\mathbf{X^T\hat{Y}} + 2\mathbf{X^T X}\boldsymbol{\beta} = 0$$

Oleh itu penganggar kaedah kuasa dua terkecil biasa adalah,

$$\hat{\boldsymbol{\beta}} = (\mathbf{X^T X})^{-\mathbf{1}} \mathbf{X^T \hat{Y}} \tag{3}$$

dan $(\mathbf{X^T X})^{-\mathbf{1}}$ wujud jika pembolehubah-pembolehubah tak bersandar($\mathbf{X}$) adalah linear secara tak bersandar, Montgomery dan Peck[17].

Ralat, ε_i mempunyai min sifar. Jadi penganggar kuasa dua terkecil biasa bagi β_j akan menjadi saksama. Nilai varians bagi Y_i dan ε_i ialah $p_i(1-p_i)$ iaitu ia berbeza secara sistematik dengan nilai-nilai pembolehubah tak bersandar (bebas) iaitu ε_i adalah heteroskedastik (ralat mempunyai nilai varians yang berubah-ubah iaitu varians ralat bersandar pada nilai X_{ij}). Sekiranya nilai p_i berhampiran dengan 0 atau 1, varians ralat adalah rendah secara relatifnya. Sebaliknya jika nilai p_i berhampiran dengan 0.5, varians ralat adalah tinggi.

Oleh itu ralat rawak ε_i tidak memenuhi andaian-andaian klasik yang mana varians ralat yang malar tidak dipenuhi. Kaedah penganggaran kuasa dua terkecil biasa masih saksama tetapi bukan penganggar terbaik. Jadi kaedah penganggaran kuasa dua terkecil biasa yang melibatkan pembolehubah bersandar yang binari kurang sesuai digunakan. Walaubagaimanapun, pengubahsuaian yang sedikit terhadap kaedah penganggaran kuasa dua terkecil biasa boleh dilakukan untuk mendapat penganggaran yang lebih tepat seperti kaedah berpemberat.

2.1.2 Model Regresi Linear Berganda Berpemberat

Goldberger[12] mencadangkan 2 langkah pengiraan bagi penganggar pemberat untuk mengatasi masalah heteroskedastik dalam model kebarangkalian linear. Langkah yang pertama ialah mendapatkan nilai anggaran $\hat{\beta}_j$ yang saksama dengan menggunakan kaedah kuasa dua terkecil biasa. Seterusnya binakan satu set pemberat bagi setiap pemerhatian ke i, iaitu,

$$W_i = \left[\frac{1}{(\sum_{j=0}^{k} \hat{\beta}_j X_{ij})(1 - \sum_{j=0}^{k} \hat{\beta}_j X_{ij})} \right]^{1/2} \quad \text{atau} \quad \frac{1}{\sqrt{p_i(1-p_i)}} \tag{4}$$

Langkah kedua ialah dengan mendarabkan kedua-dua belah persamaan (1) dengan W_i. Jadi kita akan mendapat persamaan baru iaitu,

$$\hat{Y}_i = W_i Y_i = \sum_{j=0}^{k} W_i \beta_j X_{ij} + W_i \varepsilon_i \tag{5}$$

$W_i \varepsilon_i$ mempunyai varians malar untuk semua nilai X_{ij}. Jadi, regresi dengan kaedah kuasa dua terkecil bagi persamaan (5) akan menghasilkan penganggar baru, $\hat{\beta}_j$ yang saksama dan mempunyai varians sampel yang terkecil. Ralat piawai bagi set penganggar yang kedua, $\hat{\beta}_j$ akan digunakan untuk menjalankan ujian hipotesis dan lain-lain ujian yang berkaitan. Walaupun ε_i mempunyai 2 nilai dan tidak tertabur secara normal, tetapi bagi sampel besar, $\hat{\beta}_j$ dianggarkan menghampiri taburan normal. Jadi pengujian hipotesis dan pengujian lain boleh digunakan seperti biasa.

Satu masalah yang wujud di dalam penganggaran kuasa dua terkecil biasa ialah anggaran bagi $\hat{\beta}_j$ mungkin menghasilkan $\sum \beta_j X_{ij}$ di bawah 0 atau melebihi 1 iaitu menyimpang daripada nilai data sebenar(Yi) yang disebabkan oleh ralat pensampelan. Sekiranya ini berlaku, model kebarangkalian linear yang diperkenalkan oleh Goldberger tidak boleh digunakan. Penyelesaian yang praktikal ialah dengan menganggar nilai $\sum \beta_j X_{ij}$ berdekatan dengan 0 atau 1 iaitu contohnya 0.001 atau 0.999. Jika banyak nilai $\sum \beta_j X_{ij}$ berada di luar selang 0 dan 1, kesesuaian model tersebut perlu dilakukan semula.

Aldrich dan Nelson[3] juga mencadangkan bahawa persamaan pemberat (5) boleh diubahsuai lagi. Contohnya kuasa $\frac{1}{2}$ yang terdapat di dalam persamaan (5) ditukar menjadi kuasa 1 iaitu,

$$W_i = \frac{1}{p_i(1-p_i)} \tag{6}$$

Bagi model regresi linear berganda berpemberat, fungsi kaedah kuasa dua terkecil berpemberat, $R(\beta)$ diperolehi dengan mendarabkan pemberat, W_i dengan $S(\boldsymbol{\beta})$ yang terdapat dalam persamaan (1) iaitu; $R(\beta) = W_i S(\boldsymbol{\beta}) = \sum_{j=0}^{k} W_i \varepsilon_i^2$ menghasilkan

$$\mathbf{Y}^T\mathbf{W}\mathbf{Y} - \mathbf{2}\boldsymbol{\beta}^T\mathbf{X}^T\mathbf{W}\mathbf{Y} + \boldsymbol{\beta}^T\mathbf{X}^T\mathbf{W}\mathbf{X}\boldsymbol{\beta}$$

Bagi menentukan penganggar kuasa dua terkecil, SSE (jumlah kuasa dua ralat) perlu diminimumkan iaitu terbitan bagi fungsi kaedah kuasa dua terkecil = 0.

$$\left.\frac{\partial R}{\partial \boldsymbol{\beta}}\right|_{\hat{\beta}} = -2\mathbf{X}^T\mathbf{W}\hat{\mathbf{Y}} + \mathbf{2}\mathbf{X}^T\mathbf{W}\mathbf{X}\boldsymbol{\beta} = 0$$

Oleh itu penganggar kaedah kuasa dua terkecil berpemberat adalah,

$$\hat{\beta}_j = (\mathbf{X}^T\mathbf{W}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{W}\hat{\mathbf{Y}} \tag{7}$$

Nilai $W_i\varepsilon_i$ mempunyai min sifar. Jadi penganggar kuasa dua terkecil berpemberat bagi $\hat{\beta}_j$ akan menjadi saksama. Varians bagi W_iY_i dan $W_i\varepsilon_i$ pula bernilai 1.

2.2 Penganggaran Model Logit

Dalam pemodelan logit, parameter β_j dianggar menggunakan kaedah penganggar kebolehjadian (*likelihood*) maksimum yang mana fungsi kebolehjadian adalah berada pada nilai maksimum. Oleh kerana pemodelan logit adalah tidak linear, maka algoritma berulangan (kaedah lelaran) diperlukan untuk menganggar parameter dalam model. Di samping itu varians bagi Y_i iaitu $p_i(1-p_i)$ adalah tidak malar iaitu bergantung pada nilai p_i dan p_i pula bergantung pada nilai X_{ij}. Disebabkan varians yang tidak malar ini, penganggar kebolehjadian maksimum akan menghasilkan sisihan piawai yang lebih kecil berbanding dengan penganggar kuasa dua terkecil, Agresti[1]. Penganggar kebolehjadian maksimum juga lebih digemari daripada penganggar fungsi pembezalayan di dalam menganggar parameter model logistik, Kleinbaum[14]. Pada tahun 1980, Lewis-Beck menyatakan bahawa sebelum anggaran parameter dibuat, model logit mesti memenuhi semua andaian berikut, iaitu;

(i) $Y_i \in 0, 1; i = 1, 2, ..., n$

(ii)

$$P(Y_i = 1|X_i) = \frac{e^{\sum_{j=0}^{k} \beta_j X_{ij}}}{1 + e^{\sum_{j=0}^{k} \beta_j X_{ij}}} \tag{8}$$

(iii) $Y_1, Y_2, ..., Y_n$ adalah tak bersandar secara statistik.

(iv) Tidak wujudnya saling bersandaran atau tidak wujudnya multikolineariti di antara pembolehubah-pembolehubah $X_j : j = 1, 2, ..., k$.

Katalah $Y_1, Y_2, ..., Y_n$ adalah n pembolehubah rawak dengan fungsi ketumpatan kebarangkalian, $f(Y_1, Y_2, ..., Y_n : \theta_1, \theta_2, ..., \theta_p)$ atau diringkaskan sebagai $f(y;\theta)$ yang mana y adalah pembolehubah rawak dan θ tetap. Fungsi kebolehjadian $L(\theta; y)$ adalah sama dengan $f(y;\theta)$ secara algebra yang mana dalam $L(\theta; y)$, y adalah tetap dan θ adalah pembolehubah rawak sementara. Di dalam fungsi kebolehjadian $L(\theta; y)$, ω ialah perwakilan bagi $(\theta_1, \theta_2, ..., \theta_p)$ yang tidak diketahui dan perlu dianggarkan (Ω ialah ruang parameter). Penganggar kebolehjadian maksimum bagi θ ialah $\hat{\theta}$ yang memaksimumkan fungsi kebolehjadian, $L(\theta; y)$ iaitu,

$$L(\hat{\theta}; y) \geq (L(\theta; y) \text{ bagi semua } \theta \text{ di dalam } \Omega.$$

Memaksimumkan fungsi kebolehjadian, $L(\theta; y)$ adalah sama caranya dengan memaksimumkan fungsi log kebolehjadian, $\ln L(\theta; y)$ dan membezakan $\ln L(\theta; y)$ terhadap θ_j. Seterusnya $\ln L(\theta; y)$ dibezakan kali kedua iaitu bila

$$\frac{\partial^2 \ln L(\theta, y)}{\partial \theta_j \partial \theta_k} < 0$$

maka hanya ada satu penganggar $\theta = \hat{\theta}$ yang merupakan satu nilai maksimum bagi $\theta = \hat{\theta}$, Dobson[8].

Bagi pembolehubah Y_i yang binari iaitu bernilai 0 dan 1, setiap pemerhatiannya dipertimbangkan sebagai suatu percubaan Bernoulli. Merujuk kepada andaian model logit dan persamaan (8), $P(Y_i = 1)$ boleh dinyatakan sebagai p_i atau

$$\frac{e^{\sum_{j=0}^{k} \beta_j X_{ij}}}{1 + e^{\sum_{j=0}^{k} \beta_j X_{ij}}}$$

dan $\mathrm{P}(Y_i = 0|X_i) = 1 - p_i$. Secara amnya $\mathrm{P}(Y_i|X_i) = p_i^{Y_i}(1 - p_i)^{1-Y_i}$.

Jika semua n pemerhatian Y_i tak bersandar, fungsi ketumpatan kebarangkalian atau $\mathrm{P}(Y|X)$ dinyatakan sebagai hasil produk n set kebarangkalian iaitu,

$$P(\mathbf{Y}|\mathbf{X}) = \prod_{i=1}^{n} p_i^{Y_i}(1 - p_i)^{1-Y_i}$$

Diketahui juga p_i dan $P(\mathbf{Y}|\mathbf{X})$ bergantung kepada k parameter bagi $\boldsymbol{\beta}$. Fungsi kebolehjadian juga dinyatakan sebagai, $L(\mathbf{Y}|\mathbf{X}, \boldsymbol{\beta}) \equiv P(\mathbf{Y}|\mathbf{X})$. Dalam hal ini, kita perlu mencari nilai anggaran $\boldsymbol{\beta}$ yang menghasilkan nilai maksimum bagi fungsi kebolehjadian iaitu,

$$L(\mathbf{Y}|\mathbf{X}, \hat{\boldsymbol{\beta}}) = \underset{\beta}{maks}\, L(\mathbf{Y}|\mathbf{X}, \boldsymbol{\beta})$$

Jadi, fungsi kebolehjadian logit boleh dinyatakan sebagai,

$$\begin{aligned} L(\mathbf{Y}|\mathbf{X}, \boldsymbol{\beta}) = P(\mathbf{Y}|\mathbf{X}) &= \prod_{i=1}^{n} p_i^{Y_i}(1 - p_i)^{1-Y_i} \qquad (9) \\ &= \prod_{i=1}^{n} \left(\frac{e^{\sum\limits_{j=0}^{k} \beta_j X_{ij}}}{1 + e^{\sum\limits_{j=0}^{k} \beta_j X_{ij}}} \right)^{Y_i} \left(\frac{1}{1 + e^{\sum\limits_{j=0}^{k} \beta_j X_{ij}}} \right)^{1-Y_i} \end{aligned}$$

Dengan mengambil logaritma asas e bagi fungsi kebolehjadian dalam (9),

$$\begin{aligned}\ln L(\mathbf{Y}|\mathbf{X},\boldsymbol{\beta}) &= \ln\prod_{i=1}^{n} p_i^{Y_i}(1-p_i)^{1-Y_i} \\ &= \sum_{i=0}^{n}\beta_j X_{ij}Y_i - \sum_{i=0}^{n}\ln\left(1+e^{\sum\limits_{j=0}^{k}\beta_j X_{ij}}\right) \quad \text{untuk} \;\; j=0,1,...,k \qquad (10)\end{aligned}$$

Membezakan fungsi $\ln L(\mathbf{Y}|\mathbf{X},\boldsymbol{\beta})$ dalam (10) terhadap β_j, dan mempersamakan dengan sifar, maka diperoleh

$$\frac{\partial \ln L(\mathbf{Y}|\mathbf{X},\boldsymbol{\beta})}{\partial\beta_j} = \sum_{i=0}^{n} X_{ij}Y_i - \sum_{i=0}^{n}\left(\frac{X_{ij}e^{\sum_{j=0}^{k}\beta_j X_{ij}}}{1+e^{\sum_{j=0}^{k}\beta_j X_{ij}}}\right) = 0$$

menghasilkan

$$\sum_{i=0}^{n}\left(Y_i - \frac{1}{1+e^{-\sum_{j=0}^{k}\beta_j X_{ij}}}\right)X_{ij} = 0 \quad \text{untuk} \quad j=0,1,...,k$$

atau

$$\sum_{i=0}^{n}[Y_i - P(Y_i=1|X_i)]X_{ij} = 0 \qquad (11)$$

Persamaan (11) adalah persamaan yang tidak linear. Jadi ia memerlukan lelaran untuk mendapatkan nilai anggaran bagi β_j. Kaedah lelaran yang biasa digunakan ialah kaedah Newton-Raphson atau dikenali sebagai 'method of scoring' yang dinyatakan di bawah,

$$b^{(m)} = b^{(m-1)} - \left[\frac{\partial^2 l}{\partial\beta_j\partial\beta_k}\right]^{-1}_{\beta=b^{(m-1)}} U^{(m-1)} \qquad (12)$$

yang mana $\mathbf{U}$ ialah nilai bagi pembezaan peringkat pertama bagi $\ln L(\mathbf{Y}|\mathbf{X},\boldsymbol{\beta})$ atau kecerunan atau 'gradient' dan nilai

$$\left[\frac{\partial^2 l}{\partial\beta_j\partial\beta_k}\right]^{-1}_{\beta=b^{(m-1)}}$$

adalah songsangan bagi pembezaan peringkat kedua bagi $\ln L(\mathbf{Y}|\mathbf{X},\boldsymbol{\beta})$ yang dinilai pada $\boldsymbol{\beta} = \boldsymbol{b^{(m-1)}}$. Nilai $\left[\frac{\partial^2 l}{\partial\beta_j\partial\beta_k}\right]$ juga dikenali sebagai 'information matrix' atau 'hessian matrix' atau 'matrik varians-kovarians'.

Nilai terbitan peringkat kedua bagi $\ln L(\mathbf{Y}|\mathbf{X},\boldsymbol{\beta})$ perlu dicari iaitu,

$$\begin{aligned}\frac{\partial^2 l}{\partial\beta_j\partial\beta_k} &= \sum_{i=0}^{n} X_{ij}\left(1+e^{-\sum_{j=0}^{k}\beta_j X_{ij}}\right)^{-2}\left(-X_{ik}e^{-\sum_{j=0}^{k}\beta_j X_{ij}}\right) \\ &= -\sum_{i=0}^{n} X_{ij}p_i(1-p_i)X_{ik} \;\; \text{dengan} \quad j=1,2,...,k \qquad (13)\end{aligned}$$

Dalam bentuk matriks, persamaan (11) diolah menjadi $\mathbf{X}^T(y-q)$, sementara persamaan (13) diolah menjadi $-\mathbf{X}^T\mathbf{V}\mathbf{X}$. Dengan mengolah persamaan (11), (12) dan (13), secara ringkasnya persamaan Newton-Raphson yang digunakan ialah,

$$\mathbf{b}^{(m)} = \mathbf{b}^{(m-1)} + (\mathbf{X}^T\mathbf{V}^{(m-1)}\mathbf{X})^{-1}\mathbf{X}^T(\mathbf{y} - \mathbf{q}^{(m-1)}) \tag{14}$$

Sebelum itu tatatanda matriks berikut digunakan iaitu,

$$\mathbf{Z_i} = \sum_{j=0}^{k} \beta_j X_{ij} = \beta_0 + \beta_1 X_{i1} + ... + \beta_k X_{ik} = \mathbf{X_i}\boldsymbol{\beta} \;\text{ atau }\; \mathbf{X}\boldsymbol{\beta} \;\text{ untuk }\; i = 1, 2, ..., n$$

dengan

$$\mathbf{X} = \begin{pmatrix} 1 & X_{11} & X_{12} & ... & X_{1k} \\ 1 & X_{21} & X_{22} & ... & X_{2k} \\ . & . & . & ... & . \\ . & . & . & ... & . \\ . & . & . & ... & . \\ 1 & X_{n1} & X_{n2} & ... & X_{nk} \end{pmatrix}_{n\times p}, \mathbf{X}^T = \begin{pmatrix} 1 & 1 & ... & 1 \\ X_{11} & X_{21} & ... & X_{n1} \\ X_{12} & X_{22} & ... & X_{n2} \\ . & . & ... & . \\ . & . & ... & . \\ . & . & ... & . \\ X_{1k} & X_{2k} & ... & X_{nk} \end{pmatrix}_{p\times n},$$

$$\boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ . \\ . \\ . \\ \beta_k \end{pmatrix}_{p\times 1}, \boldsymbol{q}^{(\boldsymbol{m-1})} = \left(\frac{1}{1+e^{-X\beta}}\right)$$

dan $V^{(m-1)}$ matriks pepenjuru bagi $p_i^{(m-1)}(1-p_i^{(m-1)})$ atau $\left[\frac{e^{-X\beta}}{(1+e^{-X\beta})^2}\right]^{(m-1)}$; iaitu $V^{(m-1)}$ bersamaan dengan

$$\begin{pmatrix} \left[\dfrac{e^{-X\beta}}{(1+e^{-X\beta})^2}\right]^{(m-1)} & 0 & 0 & ... & 0 \\ 0 & \left[\dfrac{e^{-X\beta}}{(1+e^{-X\beta})^2}\right]^{(m-1)} & 0 & ... & 0 \\ . & . & . & ... & . \\ . & . & . & ... & . \\ . & . & . & ... & . \\ 0 & 0 & 0 & ... & \left[\dfrac{e^{-X\beta}}{(1+e^{-X\beta})^2}\right]^{(m-1)} \end{pmatrix}_{p\times p}$$

Seterusnya tatacara bagi kaedah Newton-Raphson digunakan iaitu;

(i) Mulakan dengan nilai $m = 1$ dan dengan nilai $\boldsymbol{b}^{(0)} = 0$, dapatkan nilai $\boldsymbol{b}^{(1)}$ dengan menggunakan persamaan (14).

(ii) Teruskan dengan lelaran berikutnya iaitu bagi $m = 2, 3, ...$ dan dapatkan nilai $\boldsymbol{b}^{(m)}$ yang baru.

(iii) Proses lelaran diteruskan hingga mendapat nilai $\boldsymbol{b}^{(m)} \approx \boldsymbol{b}^{(m-1)}$.

Proses lelaran akan terhenti bila nilai $|\boldsymbol{b}^{(m)} - \boldsymbol{b}^{(m-1)}| < 0.1$ yang mana nilai $\boldsymbol{b}^{(m)}$ akan menumpu ke suatu nilai. Nilai anggaran bagi $\boldsymbol{\beta_j}$ bagi kaedah penganggar kebolehjadian maksimum bagi model logit ialah nilai $\boldsymbol{b}^{(m)}$ yang terakhir. Bagi sampel besar, kaedah kebolehjadian maksimum menghasilkan ciri-ciri penganggar yang saksama, cekap dan tertabur secara normal. (Aldrich & Nelson [3])

2.3 Penganggaran Model Probit

Dalam pemodelan probit, parameter β dianggar dengan menggunakan kaedah penganggar kebolehjadian (*likelihood*) maksimum yang mana fungsi kebolehjadian adalah berada pada nilai maksimum. Oleh kerana pemodelan probit adalah tidak linear, maka algoritma berulangan (kaedah lelaran) diperlukan untuk menganggar parameter dalam model.

Andaian yang perlu dipenuhi oleh model probit adalah:

(i) $Y_i \in 0, 1 \ ; \ i = 1, 2, ..., n$

(ii)

$$P(Y_i = 1|X_i) = \Phi(\sum_{j=0}^{k} \beta_j X_{ij}) \tag{15}$$

(iii) $Y_1, Y_2, ..., Y_n$ adalah tak bersandar secara statistik.

(iv) Tidak wujudnya saling bersandaran atau tidak wujudnya multikolineariti di antara pembolehubah-pembolehubah $X_j; j = 1, 2, ..., k$.

Seperti lazimnya bagi model regresi, ralat rawak, ε_i mesti memenuhi taburan normal piawai, iaitu $\varepsilon_i \sim N(0, 1)$. Kaedah penganggaran bagi model probit mempunyai pendekatan yang serupa dengan kaedah penganggaran bagi model logit. Bagi pembolehubah Y_i yang bernilai 0 dan 1, setiap pemerhatiannya dipertimbangkan sebagai suatu percubaan Bernoulli. Merujuk kepada andaian model probit dan persamaan (15), P($Y_i = 1$) boleh dinyatakan sebagai p_i atau $\Phi(\sum_{j=0}^{k} \beta_j X_{ij})$ dan P($Y_i = 0|X_i$) = 1 - p_i.

Jadi, fungsi logaritma natural kebolehjadian probit boleh dinyatakan sebagai,

$$\begin{aligned} \ln L(\mathbf{Y}|\mathbf{X}, \boldsymbol{\beta}) &= \ln \prod_{i=1}^{n} p_i^{Y_i} (1 - p_i)^{1-Y_i} \\ &= \sum_{i=0}^{n} Y_i \ln \left[\frac{\Phi\left(\sum_{j=0}^{k} \beta_j X_{ij}\right)}{1 - \Phi\left(\sum_{j=0}^{k} \beta_j X_{ij}\right)} \right] + \sum_{i=0}^{n} \ln \left[1 - \Phi\left(\sum_{j=0}^{k} \beta_j X_{ij}\right) \right] \\ &\quad \text{untuk } j = 0, 1, ..., k \end{aligned} \tag{16}$$

Bezakan fungsi $\ln L(\mathbf{Y}|\mathbf{X},\boldsymbol{\beta})$ pada (16) terhadap β_j, dan mempersamakan dengan sifar

$$\frac{\partial \ln L(\mathbf{Y}|\mathbf{X},\boldsymbol{\beta})}{\partial \beta_j} = \sum_{i=0}^{n} Y_i \left[\frac{1-\Phi\left(\sum\limits_{j=0}^{k}\beta_j X_{ij}\right)}{\Phi\left(\sum\limits_{j=0}^{k}\beta_j X_{ij}\right)}\right]\left[\frac{X_{ij}\Phi\left(\sum\limits_{j=0}^{k}\beta_j X_{ij}\right)}{\left(1-\Phi\left(\sum\limits_{j=0}^{k}\beta_j X_{ij}\right)\right)^2}\right] - \sum_{i=0}^{n} X_{ij}\left[\frac{\Phi\left(\sum\limits_{j=0}^{k}\beta_j X_{ij}\right)}{1-\Phi\left(\sum\limits_{j=0}^{k}\beta_j X_{ij}\right)}\right] = 0$$

menghasilkan

$$\sum_{i=0}^{n}\left[Y_i - \Phi\left(\sum_{j=0}^{k}\beta_j X_{ij}\right)\right]\left[\frac{\Phi\left(\sum\limits_{j=0}^{k}\beta_j X_{ij}\right)X_{ij}}{\Phi\left(\sum\limits_{j=0}^{k}\beta_j X_{ij}\right)\left(1-\Phi\left(\sum\limits_{j=0}^{k}\beta_j X_{ij}\right)\right)}\right] = 0$$

atau

$$\sum_{i=0}^{n}[Y_i - P(Y_i=1|X_i)]A_i X_{ij} = 0 \quad \text{untuk} \quad j = 0, 1, ..., k \tag{17}$$

Maka diperoleh

$$A_i = \frac{\Phi(Z_i)}{\Phi(Z_i)[1-\Phi(Z_i)]} \quad \text{dengan} \quad Z_i = \sum_{j=0}^{k}\beta_j X_{ij}$$

Juga diperoleh

$$\frac{\partial^2 \ln L(Y|X,\beta)}{\partial \beta_j \partial \beta_k} = \sum_{i=0}^{n}\left[\frac{X_{ij}X_{ik}\left[\Phi\left(\sum\limits_{j=0}^{k}\beta_j X_{ij}\right)\right](Y_i-1)}{\left[1-\Phi\left(\sum\limits_{j=0}^{k}\beta_j X_{ij}\right)\right]^2}\right]$$

$$= \sum_{i=0}^{n}[X_{ij}\boldsymbol{V}X_{ik}(Y_i-1)] \tag{18}$$

dengan V ialah matriks pepenjuru $p_i/[1-p_i]^2$

Persamaan (17) dan (18) adalah persamaan yang tidak linear. Jadi kaedah lelaran yang biasa digunakan ialah kaedah Newton-Raphson atau dikenali sebagai *'method of scoring'* yang dinyatakan di bawah,

$$\begin{aligned}\boldsymbol{b}^{(m)} &= \boldsymbol{b}^{(m-1)} - \left[\frac{\partial^2 \boldsymbol{l}}{\partial\beta_j\partial\beta_k}\right]^{-1}_{\beta=b^{(m-1)}} U^{(m-1)} \\ &= \boldsymbol{b}^{(m-1)} - \left[\frac{\partial^2 \ln L(\mathbf{Y}|\mathbf{X},\boldsymbol{\beta})}{\partial\beta_j\partial\beta_k}\right]^{-1}_{\beta=b^{(m-1)}} \left[\frac{\partial \ln L(\mathbf{Y}|\mathbf{X},\boldsymbol{\beta})}{\partial\beta_j}\right]_{\beta=b^{(m-1)}}\end{aligned} \tag{19}$$

dengan $\boldsymbol{U}$ ialah nilai bagi pembezaan peringkat pertama bagi $\ln L(\mathbf{Y}|\mathbf{X},\boldsymbol{\beta})$. Nilai

$$\left[\frac{\partial^2 l}{\partial\beta_j\partial\beta_k}\right]^{-1}_{\beta=b^{(m-1)}}$$

adalah sonsangan bagi pembezaan peringkat kedua bagi $\ln L(\mathbf{Y}|\mathbf{X},\boldsymbol{\beta})$. Sebelum nilai $\hat{\boldsymbol{\beta}}_j$ dikira, nilai min ($\boldsymbol{V}$), matrik varians-kovarians ($\boldsymbol{\Sigma}$) dan nilai kebarangkalian, p_i perlu dikira yang mana cara pengiraan bagi model probit adalah agak berbeza dengan model logit . Terdapat 3 cara pengiraan yang terlibat di dalam mencari nilai min ($\boldsymbol{V}$), matrik kovarians ($\boldsymbol{\Sigma}$) dan nilai kebarangkalian, p_i iaitu,

(i) Kaedah pengamiran berangka

(ii) Kaedah simulasi Monte Carlo

(iii) Kaedah penghampiran berangka

Kaedah yang akan dibincangkan di sini ialah kaedah penghampiran berangka. Katakan pembolehubah U atau Y tertabur secara taburan multivariat normal dengan nilai min ($\boldsymbol{V}$) dan matriks varians-kovarians ($\boldsymbol{\Sigma}$). Pada langkah $i-1$, hitung
$\tilde{m}_i = m_i + (\tilde{m}_{i-1} - m_i)\Phi(\alpha_i) + a_i\phi(\alpha_i)$
$\tilde{\sigma}_i^2 = \tilde{\tilde{m}}_i - \tilde{m}_i^2$
$\tilde{\sigma}_{ij}^2 = \sigma_{ij}^2 + (\tilde{\sigma}_{i-1,j}^2 - \sigma_{ij}^2)\Phi(\alpha_i), \tilde{\sigma}_{ij}^2 = \tilde{\sigma}_{ji}^2$
$a_i = \sqrt{\tilde{\sigma}_{i-1}^2 + \sigma_i^2 - 2\tilde{\sigma}_{i-1,i}^2}$ dan $\alpha_i = (\tilde{m}_{i-1} - m_i)/a_i$ untuk j = i+1, i+2, ..., k.

Pada setiap cerapan i, kita perlu mengira hanya satu nilai bagi $\phi(\alpha_i), \Phi(\alpha_i), \tilde{m}_i, \tilde{\tilde{m}}_i$ dan $(k-i)$ varians dan kovarians. Daripada formula ini, kita akan memperolehi min ($\boldsymbol{V}$) dan matriks varians-kovarians ($\boldsymbol{\Sigma}$).

Nilai kebarangkalian boleh dinyatakan sebagai,

$$p_i(\boldsymbol{V},\boldsymbol{\Sigma}_\xi) = P\left\{U_i > \underset{i\neq j}{mak}\,(U_i)|\boldsymbol{V},\boldsymbol{\Sigma}_\xi\right\} = \Phi\left(\frac{V_I - m_{I-1}^*}{\sqrt{\sigma_I^2 + \tilde{\sigma}_{I-1}^{2*} - 2\tilde{\sigma}_{I-1,I}^{2*}}}\right)$$

dengan

$$m_{I-1}^* = E\left[\underset{j<I}{mak}\,(U_j)\right] \tilde{\sigma}_{I-1}^{2*} = Var\left[\underset{j<I}{mak}\,(U_j)\right]$$

dan

$$\tilde{\sigma}_{I-1,I}^{2*} = Kov\left[\underset{j<I}{mak}\,(U_j), U_I\right].$$

Seterusnya nilai $\hat{\boldsymbol{\beta}}_j$ dikira menggunakan kaedah kebolehjadian maksimum. Terdapat 3 kaedah yang terlibat di dalam memaksimumkan fungsi ln kebolehjadian iaitu kaedah 'steepest-ascent', 'Newton-Raphson' dan 'variable-metric'. Walau bagaimanapun hanya kaedah Newton-Raphson diterangkan di sini memandangkan kaedah lelaran ini menumpu dengan cepat berbanding dengan dua kaedah yang lain. Lagipun kebanyakan perisian statistik menggunakan kaedah ini dalam mencari nilai $\hat{\boldsymbol{\beta}}_j$. Tatacara bagi kaedah Newton-Raphson bagi model probit adalah sama seperti model logit.

3 Metodologi

(1) Bagi model kebarangkalian linear, penganggar parameter $\hat{\beta}_k$ diperolehi menggunakan kaedah kuasa dua terkecil biasa dan berpemberat. Bagi model logit dan probit pula, penganggar parameter $\hat{\beta}_k$ diperolehi menggunakan kaedah kebolehjadian maksimum.

(2) Sebelum penganggar parameter $\hat{\beta}_k$ diperolehi, diagnosis multikolineariti perlu dilakukan.

(3) Terdapat 3 cara pemilihan pembolehubah iaitu kaedah langkah demi langkah ke hadapan, kaedah langkah demi langkah penghapusan dari belakang dan kaedah secara manual berasaskan prinsip 'parsimonious model'.

(4) Bagi model kebarangkalian linear, langkah pertama ialah mendapatkan nilai parameter $\hat{\beta}_k$ dengan menggunakan kaedah penganggaran kuasa dua terkecil biasa yang mana kaedah masukan digunakan. Nilai $P(Yi = 1)$, pemberat W_1 dan W_2 dihitung yang mana nilai $W_1 = [1/(P(Y_i)(1 - P(Y_i))]^1$ dan $W_2 = [1/(P(Y_i)(1 - P(Y_i))]^{1/2}$. Sedikit pengubahsuaian dilakukan iaitu dengan menggantikan nilai $P(Y_i = 1) > 1$ dengan nilai 0.999 dan $P(Y_i = 1) < 0$ dengan nilai 0.001. Masalah heteroskedastik diatasi dengan mendarabkan pemberat, W kepada model regresi linear berganda. Seterusnya kaedah penganggaran kuasa dua terkecil berpemberat dilakukan dan pemilihan pembolehubah X yang paling berpengaruh dilakukan.

(5) Bagi model logit, pemilihan pembolehubah yang paling berpengaruh diperolehi dengan menggunakan ketiga-tiga kaedah pemilihan pembolehubah. Bagi model probit pula, pemilihan pembolehubah yang paling berpengaruh diperolehi hanya dengan menggunakan kaedah secara manual berasaskan prinsip *'parsimonious model'*.

(6) Seterusnya analisis reja perlu dilakukan bagi memastikan andaian klasik bagi ralat dipenuhi.

(7) Pengujian hipotesis bagi setiap parameter $\hat{\beta}_k$ diuji dengan menggunakan ujian t (model kebarangkalian linear) dan statistik *Wald* (model logit dan probit) bagi memastikan setiap pembolehubah X yang terpilih bererti.

(8) Ujian kebagusan penyuaian yang dilakukan pula ialah analisis varians, nilai R^2 (model kebarangkalian linear) serta nisbah kebolehjadian dan nilai pseudo R^2 (model logit dan model probit). Bagaimanapun menurut Aldrich dan Nelson[3], nilai R^2 dan pseudo R^2 tidak boleh dijadikan perbandingan ukuran kerana perbezaan formula yang digunakan. Jadi nilai R^2 dan pseudo R^2 kurang sesuai digunakan dan tidak diketengahkan.

(9) Diagnosis data terpencil dan data berpengaruh dilakukan terhadap data bagi ketiga-tiga jenis model. Setelah didiagnosis, pemodelan akan dilakukan sekali lagi dan model dengan peratus ketepatan ramalan yang tertinggi akan diketengahkan sebagai model yang terbaik.

4 Persamaan dan Perbezaan antara Tiga Model

Jadual 1 menunjukkan ringkasan keseluruhan bagi persamaan dan perbezaan bagi ketiga-tiga jenis model.

5 Kesimpulan

Ketiga-tiga kaedah pemodelan perlu dijalankan terhadap sesuatu data yang melibatkan data bersandar binari dan data tak bersandar berjenis binari, kategori atau kuantitatif. Ia bertujuan bagi mencari model yang terbaik dengan peratus ketepatan ramalan tertinggi. Salah satu kaedah iaitu sama ada kaedah penganggaran kuasa dua terkecil berpemberat atau kaedah kebolehjadian maksimum akan menghasilkan kesimpulan yang lebih baik dan lebih efisien bila melibatkan sampel besar. Pembuangan data yang dikesan sebagai data terpencil dan berpengaruh juga perlu dilakukan bagi meningkatkan lagi peratus ketepatan ramalan. Pengurusan atau pengolahan data yang berbeza juga akan menghasilkan kesimpulan yang berbeza kerana pengurusan data yang kurang kemas memungkinkan hasil keputusan menjauhi keputusan sebenar. Perbezaan kesimpulan boleh dibuat bagi ketiga-tiga model dari segi bentuk taburan data, pembolehubah bererti X yang terpilih, model fungsi dan tafsiran model serta peratus ketepatan ramalan.

Rujukan

[1] A. Agresti, *An Introduction to Categorical Data Analysis*, John Wiley & Sons, Inc., New York, 1996.

[2] A. Agresti dan B. Finlay, *Statistical Methods for the Social Sciences*, Prentice-Hall, Inc., USA, 1997.

[3] J.H. Aldrich dan F.D. Nelson, *Linear Probability, Logit and Probit Models*, Sage Publications, Inc., California, 1985.

[4] E.B. Andersen, *Introduction to Statistical Analysis of Categorical Data*, Springer-Verlag, Germany, Berlin, Heidelberg, 1997.

[5] V. Barnett, dan T. Lewis, *Outliers in Statistical Data.* John Wiley & Sons Ltd. England. 1994.

[6] D.A. Belsley, Kuh, E., dan R.E Welsch, *Regression Diagnosis, Identifying Influential Data and Sources of Collinearity.* John Wiley & Sons, Inc. New York, 1980.

[7] C. Daganzo, *Multinomial Probit : The Theory and its Application to Demand Forecasting.* Academic Press, Inc. New York, 1979.

Jadual 1 Persamaan dan Perbezaan antara Tiga Model

Ciri-Ciri Model	Model Kebarangkalian Linear	Model Logit	Model Probit
Sejarah Penemuan	Ditemui oleh Johann Friedrich Carl Gauss pada tahun 1795	Ditemui oleh Joseph Berkson pada tahun 1944	Ditemui oleh C.R. Bliss pada tahun 1934
Model	(1) Regresi linear berganda biasa biasa $\hat{Y}_i = Y_i = \sum_{j=0}^{k} X_{ij} + \varepsilon_i$ (2) Regresi linear berganda berpemberat $\hat{Y}_i = W_i Y_i = \sum_{j=0}^{k} W_i \beta_j X_{ij} + W_i \varepsilon_i$ (Model berbentuk linear)	$\hat{Y}_i = P(Y_i = 1 \| X_i) = p_i = \frac{e^{\sum_{j=0}^{k} \beta_j X_{ij}}}{1 + e^{\sum_{j=0}^{k} \beta_j X_{ij}}}$ Secara amnya, $P(Y_i \| X_i) = p_i^{Y_i}(1-p_i)^{1-Y_i}$ (Model berbentuk tidak linear)	$\hat{Y}_i = P(Y_i = 1 \| X_i) = p_i = \phi\Big(\sum_{j=0}^{k} \beta_j X_{ij}\Big)$ Secara amnya, $P(Y_i \| X_i) = p_i^{Y_i}(1-p_i)^{1-Y_i}$ (Model berbentuk tidak linear)
Andaian Model Bagi Ralat (ε_i)	1. Ralat tak bersandar dengan X_{ij} 2. Ralat mempunyai min 0 dan homoskedastik (varians malar) serta tidak berkorelasi. 3. Ralat rawak, ε_i tertabur secara normal, $\varepsilon_i \sim N(0, \sigma^2)$	Ralat rawak tertabur secara fungsi taburan logistik dengan $E(\varepsilon_i) = 0$ dan $V(\varepsilon_i) = \frac{1}{3}\pi^2$.	Ralat rawak, ε_i tertabur secara taburan normal piawai iaitu $\varepsilon_i \sim N(0,1)$ iaiu, $E(\varepsilon_i) = 0$ dan $V(\varepsilon_i) = 1$.
Kaedah Penganggaran Parameter β	Kaedah kuasa dua terkecil biasa dan berpemberat iaitu dengan meminimumkan fungsi, $\sum_{i=0}^{k} \varepsilon_i^2 = \sum_{i=0}^{k} (Y_i - \hat{Y}_i)^2$ dan $\sum_{i=0}^{k} W_i \varepsilon_i^2 = \sum_{i=0}^{k} W_i (Y_i - \hat{Y}_i)^2$	Kaedah kebolehjadian maksimum dengan lelaran Newton-Raphson iaitu dengan memaksimumkan fungsi, $L(\mathbf{Y} \| \mathbf{X}, \boldsymbol{\beta}) = \prod_{i=1}^{n} \left(\frac{e^{\sum_{j=0}^{k} \beta_j X_{ij}}}{1 + e^{\sum_{j=0}^{k} \beta_j X_{ij}}} \right)^{Y_i} \times \left(\frac{1}{1 + e^{\sum_{j=0}^{k} \beta_j X_{ij}}} \right)^{1-Y_i}$	Kaedah kebolehjadian maksimum dengan lelaran Newton-Raphson iaitu dengan memaksimumkan fungsi, $L(\mathbf{Y} \| \mathbf{X}, \boldsymbol{\beta}) = \prod_{i=1}^{n} \Big[\phi\Big(\sum_{j=0}^{k} \beta_j X_{ij}\Big)\Big]^{Y_i} \times \Big[1 - \phi\Big(\sum_{j=0}^{k} \beta_j X_{ij}\Big)\Big]^{1-Y_i}$
Anggaran $\hat{\beta}$	$\hat{\beta} = (\mathbf{X^T X})^{-1} \mathbf{X^T \hat{Y}}$ $\hat{\beta} = (\mathbf{X^T W X})^{-1} \mathbf{X^T W \hat{Y}}$	$\mathbf{b}^{(m)} = \mathbf{b}^{(m-1)} + (\mathbf{X}^T \mathbf{V}^{(m-1)} \mathbf{X})^{-1} \mathbf{X}^T (\mathbf{y} - \mathbf{q}^{(m-1)})$	$\boldsymbol{b}^{(m)} = \boldsymbol{b}^{(m-1)} - \left[\frac{\partial^2 \ln L(\mathbf{Y} \| \mathbf{X}, \boldsymbol{\beta})}{\partial \beta_j \partial \beta_k} \right]^{-1}_{\beta = b^{(m-1)}} \times \left[\frac{\partial \ln L(\mathbf{Y} \| \mathbf{X}, \boldsymbol{\beta})}{\partial \beta_j} \right]_{\beta = b^{(m-1)}}$

Jadual 1(sambungan) Persamaan dan Perbezaan antara Tiga Model

Ciri-Ciri Model	Model Kebarangkalian Linear	Model Logit	Model Probit
Ciri-ciri Pengganggar	Saksama, kurang efisien dan bersifat normal bila melibatkan sampel besar	Saksama, efisien dan bersifat normal bila melibatkan sampel besar	
Pengujian Hipotesis	Ujian t	Ujian t dan statistik *Wald*	
Kebagusan Penyuaian	Analisis varians (ujian F) dan nilai pekali penentuan, $R^2 = \dfrac{\left[1 - \sum_{i=1}^{n} \left(Y_i - \sum_{j=0}^{k} \hat{\beta} X_{ij}\right)\right]^2}{\sum_{i=1}^{n} (Y_i - \bar{Y})^2}$	Ujian statistik nisbah (c), ujian khi kuasa dua dan kebolehjadian pekali *pseudo* $R^2 = \dfrac{c}{N + c}$	
Peratus Ketepatan Ramalan	Bandingkan Y_i dengan $\hat{Y}_i$ secara manual (dalam *Microsoft Excel*) dengan $\% \text{ ketepatan} = \dfrac{\text{bil. data yang tepat}}{\textit{bilangan data}} \times 100$ Weisberg(1978) : Data dikira tepat bila $Y_i = 1$ (data sebenar) selari dengan nilai anggaran $0.5 \leq P(Y_i = 1) \leq 1$ dan bila jika $Y_i = 0$ (data sebenar) selari dengan nilai anggaran $0 \leq P(Y_i = 1) < 0.5$. Bagi model logit, dinyatakan dalam output komputer.		
Diagnosis Multikolineariti (saling bersandaran di antara pembolehubah X)	Matriks parameter korelasi, ujian VIF (Variance inflation factor), nilai eigen & indeks bersyarat dan kadar huraian varians		
Diagnosis Data Terpencil & Berpengaruh	Reja terpiawai Pearson (mengenalpasti cerapan Y yang terpencil), reja terlaras (mengenalpasti cerapan Y yang terpencil), ujian reja terhapus studentized (mengenalpasti cerapan Y yang terpencil), nilai matriks *leverage hat, H* (mengenalpasti cerapan X yang terpencil), *COVRATIO* (mengenalpasti data terpencil dan berpengaruh), *DFFITS* (mengenalpasti data berpengaruh), jarak *Cook* (mengenalpasti data berpengaruh) dan *DfBeta* (mengenalpasti data berpengaruh).		

[8] A. J. Dobson, *An Introduction to Generalized Linear Models.* Chapman & Hall. London, 1991.

[9] N. R. Draper, dan H. Smith, *Applied Regression Analysis* (2nd Edition), John Wiley, New York, 1981.

[10] D.J. Finney, *Probit Analysis.* Great Britain : Cambridge University Press, Great Britain, 1981.

[11] J. Geweke, dan M. Keane, (1997). *Mixture of Normals Probit Models.* Federal Reserve Bank of Minneapolis Research Department Staf Report 237. http://ideas.repec.org/p/fip/fedmsr/237.html. Tarikh akses : 9 Oct 2002.

[12] A. S. Goldberger, (1964). *Econometric Theory.* John Wiley, New York.

[13] Horowitz, J.L. and Savin, N.E. (2001). *Binary Response Models : Logits, Probits and Semiparametrics.* Journal of Economic Perspectives, Vol 15, Issue 4 : 43-56.

[14] Kleinbaum, D. G. (1994). *Logistic Regression : A Self-Learning Text.* New York : Springer-Verlag, New York, Inc.

[15] Liao, Tim Futing (1994). *Interpreting Probability Models : Logit, Probit and Other Generalized Linear Models.* California : Sage Publications, Inc.

[16] Maddala, G.S. (1983). *Limited Dependent and Qualitative Variables in Econometrics.* Cambridge : Cambridge University Press.

[17] Montgomery, D.C. and Peck, E.A. (1992). *Introduction to Linear Regression Analysis* (2nd Edition). New York : John Wiley & Sons.

[18] Morgan, B.J.T. (1992). *Analysis of Quantal Response Data.* London : Chapman & Hall.

[19] Murray, A., Anderson, D., Francis, B. and Hinde, J. (1990). *Statistical Modelling in GLIM.* Great Britain : Oxford Time Publications.

[20] Neter, J., Wasserman, W. and Kutner, M.H. (1983). *Applied Linear Regression Models.* New York : Richard D. Irwin, Inc.

[21] Seber, G.A.F. (1977). *Linear Regression Analysis.* New York : John Wiley.

[22] Wei, Ching Chang (1976). *On the History of Statistics and Probability.* New York : Marcel Dekker, Inc.

[23] Weeks, M. and Orme, C. (1999). *The Statistical Relationship between Bivariate and Multinomial Choice Models. Journals of Economy.* http://econpapers.repec.org/paper/camcamdae/default6.htm. Tarikh akses : 9 Oct 2002

[24] Weisberg, S. (1985). *Applied Linear Regression* (2nd Edition). New York : John Wiley & Sons.