

PENANGANAN MULTIKOLINEARITAS (KEKOLINEARAN GANDA) DENGAN ANALISIS REGRESI KOMPONEN UTAMA

Tatik Widiharah
Jurusan Matematika FMIPA UNDIP

Abstrak

Multikolinearitas yang tinggi diantara peubah-peubah bebas, mengakibatkan pendugaan dengan metode kuadrat terkecil tidak dapat diandalkan, ditandai dengan koefisien regresi tidak nyata dan adanya multikolinieritas. Pendeteksian multikolinearitas dapat dilakukan secara informal salah satunya dengan koefisien korelasi linear antar peubah bebas maupun dengan cara formal dengan faktor inflasi ragam. Analisis regresi komponen utama digunakan untuk menghilangkan multikolinieritas dan semua peubah bebas masuk dalam model, analisis regresi ini merupakan teknik analisis regresi yang dikombinasikan dengan teknik analisis komponen utama. Analisis komponen utama bertujuan menyederhanakan peubah yang diamati dengan mereduksi dimensinya, hal ini dilakukan dengan jalan menghilangkan korelasi diantara peubah-peubah melalui transformasi. Teknik analisis komponen utama dijadikan sebagai tahap analisis antara untuk memperoleh hasil akhir dalam analisis regresi.

Kata kunci : regresi, kuadrat terkecil, multikolinearitas, komponen utama.

1. PENDAHULUAN

Dalam bidang-bidang ekonomi, ilmu-ilmu sosial dan biologi, peubah-peubah bebas cenderung saling berkorelasi di antara mereka sendiri dan berkorelasi dengan peubah lain yang tidak disertakan kedalam model padahal peubah ini berhubungan dengan peubah bebas (Netter, et al, 1990). Pada kasus seperti ini bila digunakan pendugaan dengan metode kuadrat terkecil hasil yang diperoleh akan menyimpang dengan yang diharapkan. Penyimpangan ini berupa antara lain koefisien regresi bagi peubah-peubah bebas penting memberikan hasil yang tidak nyata secara statistik, tanda koefisien regresi dugaan bertentangan dengan yang diharapkan berdasarkan pertimbangan teoritis ataupun pengalamannya. Apabila dilakukan dengan pemilihan regresi terbaik (best subset), kemungkinan ada beberapa peubah bebas yang tidak muncul dalam model, padahal peneliti tetap menginginkan semua peubah bebas masuk dalam model.

Usaha untuk mengatasi multikolinearitas kadang-kadang ditempuh dengan membangun analisis regresi bertatar (stepwise regression), di mana berdasarkan prosedur bertatar akan dikeluarkan peubah-peubah yang menyebabkan multikolinearitas sehingga dihasilkan persamaan regresi yang terdiri dari peubah-peubah yang nyata secara statistik (Gasperz, 1992). Dengan demikian analisis regresi bertatar hanya mempertimbangkan kriteria statistik dan tidak mempertimbangkan kriteria teoritik dengan masalah yang dihadapi. Pada analisis regresi komponen utama semua peubah bebas masuk kedalam model, tetapi sudah tidak terjadi multikolinieritas karena sudah dihilangkan pada tahap analisis komponen utama.

Analisis komponen utama pada dasarnya bertujuan untuk menyederhanakan peubah yang diamati dengan cara menyusutkan (mereduksi) dimensinya. Hal ini dilakukan dengan jalan menghilangkan korelasi di antara peubah melalui transformasi peubah asal kepeubah baru (merupakan kombinasi linear dari peubah-peubah asal) yang tidak saling berkorelasi. Dari p buah

peubah asal dapat dibentuk p buah komponen utama, dipilih k buah komponen utama saja ($k < p$) yang telah mampu menerangkan keragaman data cukup tinggi (antara 80% sampai dengan 90%). Komponen utama yang dipilih tersebut (k buah) telah dapat mengganti p buah peubah asal tanpa mengurangi informasi .

Analisis regresi komponen utama merupakan analisis regresi dari peubah takbebas terhadap komponen-komponen utama yang tidak saling berkorelasi , dimana setiap komponen utama merupakan kombinasi linear dari semua peubah bebas yang telah dispesifikasikan sejak awal (Hawkins, 1973; Draper and Smith, 1992; Montgomery and Peck, 1982). Pembahasan pertama pada tulisan ini adalah pendeteksian multikolinearitas (kekolinearan ganda) baik secara informal (dengan menggunakan koefisien korelasi antar peubah bebasnya) maupun formal (dengan menggunakan faktor inflasi ragam), membentuk komponen utama, memilih komponen utama yang telah mampu menerangkan keragaman cukup tinggi (80%-90%) sampai dengan menentukan model persamaan regresi komponen utama. Untuk memperjelas pembahasan diberikan contoh penerapannya dimana semua perhitungan yang diperlukan menggunakan paket minitab release 11. 12.

2. DIAGNOSTIK MULTIKOLINEARITAS (KEKOLINEARAN GANDA)

Bila peubah-peubah bebas saling berkorelasi sangat tinggi mengakibatkan koefisien regresi dugaannya cenderung bervariasi sangat besar dari sampel satu kesampel lainnya, sehingga tidak diperoleh informasi yang tepat mengenai koefisien regresi yang sebenarnya (populasi). Bahkan bisa terjadi masing-masing koefisien regresi dugaan secara statistik tidak nyata walaupun antara peubah takbebas dan peubah-peubah bebas terdapat hubungan statistik yang nyata .

Diagnostik bagi kekolinearan ganda dapat dilakukan secara informal maupun diasnotik formal. Indikasi adanya masalah kekolinearan ganda yang serius ditunjukkan oleh diagnostik-diasnotik informal sebagai berikut :

1. Terjadi perubahan besar pada koefisien regresi dugaan bila suatu peubah bebas ditambahkan atau dibuang dari model, atau bila suatu amatan diubah atau dibuang.
2. Uji-uji individu terhadap koefisien regresi bagi peubah-peubah bebas penting memberikan hasil yang tidak nyata.
3. Tanda koefisien regresi dugaan yang diperoleh bertentangan dengan yang diharapkan berdasarkan pertimbangan teoritis atau pengalaman sebelumnya.
4. Koefisien korelasi sederhana yang besar antara pasangan-pasangan peubah bebas dalam matriks korelasi.
5. Selang kepercayaan yang lebar bagi koefisien regresi peubah bebas penting.

Suatu metode formal untuk mendeteksi adanya kekolinearan ganda yang banyak digunakan adalah faktor inflasi ragam (FIR, Variance Inflation Faktor = VIF). Faktor ini mengukur seberapa besar ragam koefisien regresi dugaan membesar dibandingkan seandainya peubah-peubah bebasnya tidak berkorelasi linear. Faktor inflasi ragam (FIR) bagi dugaan koefisien regresi ke k dinotasikan $(FIR)_k$ dan didefinisikan dengan :

$(FIR)_k = (1 - R_k^2)^{-1}$ dengan R_k^2 adalah koefisien determinasi ganda bila peubah bebas ke k (X_k) diregresikan terhadap peubah-peubah bebas X lainnya dalam model. Nilai FIR terbesar diantara semua peubah bebas X digunakan sebagai indikator tingkat keparahan kekolinearan ganda. Nilai FIR terbesar yang lebih besar dari 10 digunakan sebagai indikasi adanya kekolinearan ganda. Rataan nilai-nilai FIR yang cukup jauh lebih besar dari 10 mengindikasikan adanya masalah kekolinearan ganda yang serius. Dalam tulisan ini diagnostik multikolinearitas dilakukan dengan metode formal.

3. ANALISIS REGRESI KOMPONEN UTAMA

Andaikan dipunyai peubah takbebas Y , dan peubah-peubah bebas X_1, X_2, \dots, X_p , dan berdasarkan diasnotik multikolinearitas pada bagian (2) diperoleh bahwa ada kekolinearitas ganda dari peubah-peubah bebas tersebut. Akan ditentukan persamaan regresi dengan menggunakan analisis regresi komponen utama.

Pertama-tama akan dianalisis dengan teknik analisis komponen utama untuk menentukan komponen-komponen utama yang mewakili p buah peubah bebas tersebut. Pada analisis komponen utama, didasarkan bahwa skala pengukuran dari X_1, X_2, \dots, X_p sama, kemudian dibentuk peubah baru W disebut sebagai komponen utama yang merupakan kombinasi linear dari X_1, X_2, \dots, X_p dengan bentuk sebagai berikut (Draper and Smith, 1992; Gasperz, 1992) :

$$\begin{aligned} W_1 &= a_{11} X_1 + a_{21} X_2 + \dots + a_{p1} X_p \\ W_2 &= a_{12} X_1 + a_{22} X_2 + \dots + a_{p2} X_p \\ W_p &= a_{1p} X_1 + a_{2p} X_2 + \dots + a_{pp} X_p \end{aligned} \quad \dots \dots \dots (3.1)$$

Bila ditulis dengan notasi matriks : $W = AX$

Cov (X) = (yang nilainya tidak diketahui sehingga diduga dari sampel yaitu Cov (X) = S

Maka Cov (W) = ASA^T . Sehingga Var (W_i) = $a_i^T S a_i$ dengan $a_i^T = (a_{i1} a_{i2} \dots a_{ip})$.

Komponen utama pertama adalah kombinasi linear dari X_1, X_2, \dots, X_p yang dapat menerangkan keragaman terbesar.

$$W_1 = a_{11} X_1 + a_{21} X_2 + \dots + a_{p1} X_p = a_1^T X \quad \dots \dots \dots (3.2)$$

$$\text{Var} (W_1) = a_1^T S a_1$$

Vektor a_1^T adalah vektor normal dan $a_1^T a_1 = 1$ dipilih sehingga Var (W_1) maksimum. Sehingga dipunyai masalah : maksimumkan Var (W_1) = $a_1^T S a_1$ dengan kendala $a_1^T a_1 = 1$

atau $a_1^T a_1 - 1 = 0$. Dibentuk fungsi lagrange $L = a_1^T S a_1 - (\lambda_1 (a_1^T a_1 - 1))$.

$(L / a_1 = (S - (\lambda_1 I) a_1$. Diambil $(L / a_1 = 0$, maka $(S - (\lambda_1 I) a_1 = 0$ (persamaan ini dikenal dengan persamaan karakteristik) dari matriks S , $(\lambda_1$ adalah akar karakteristik dari matriks S dan a_1 adalah vektor karakteristik yang bersesuaian dengan akar karakteristik $(\lambda_1$. Berdasarkan $(S - (\lambda_1 I) a_1 = 0$ dapat diperoleh $a_1^T S a_1 = (\lambda_1$. Karena $a_1^T S a_1$ harus maksimum maka $(\lambda_1$ dipilih akar karakteristik yang terbesar.

Komponen utama kedua adalah kombinasi linear dari X_1, X_2, \dots, X_p yang tidak berkorelasi dengan komponen utama pertama, serta memaksimalkan sisa keragaman data setelah diterangkan oleh komponen utama pertama.

$$W_2 = a_{12} X_1 + a_{22} X_2 + \dots + a_{p2} X_p = a_2^T X \quad \dots \dots \dots (3.3)$$

$$\text{Var} (W_2) = a_2^T S a_2$$

Vektor a_2^T adalah vektor normal yang dipilih sehingga keragaman komponen utama kedua maksimum serta ortogonal terhadap a_1^T .

Sehingga dipunyai masalah : maksimumkan Var (W_2) = $a_2^T S a_2$ dengan kendala $a_2^T a_2 = 1$

atau $a_2^T a_2 - 1 = 0$ dan $a_1^T a_2 = 0$. Dapat dibentuk fungsi lagrange sebagai berikut :

$L = a_2^T S a_2 - (\lambda_2 (a_2^T a_2 - 1)) - (\mu_1 a_1^T a_2$. Bila fungsi lagrange diturunkan secara parsial terhadap a_2 dan nilai turunan tersebut sama dengan nol akan diperoleh :

$(S - (\lambda_2 I) a_2 - (\mu_1 a_1 = 0$, sehingga diperoleh $a_2^T S a_2 = (\lambda_2$. Karena $a_2^T S a_1$ harus maksimum maka $(\lambda_2$ dipilih akar karakteristik yang terbesar kedua.

Secara sama untuk komponen utama ke j ($j=1,2,\dots,p$) dapat dinyatakan dalam bentuk :

$$W_j = a_{1j} X_1 + a_{2j} X_2 + \dots + a_{pj} X_p = a_j^T X \quad \dots \dots \dots (3.4)$$

$$\text{Var} (W_j) = a_j^T S a_j$$

Vektor a_j^T adalah vektor normal yang dipilih sehingga komponen utama ke j maksimum, serta

ortogonal terhadap \mathbf{a}_i^T dengan $i \neq j$ dan diperoleh $\text{Var}(W_j) = \lambda_j$ yang merupakan akar karakteristik terbesar ke j .

Pada masalah regresi biasanya skala pengukuran untuk peubah-peubah bebas X_1, X_2, \dots, X_p biasanya belum sama, sehingga perlukan disamakan dengan cara transformasi kedalam peubah baku Z sebagai berikut :

$$Z_i = (X_i - E(X_i)) / (\text{Var}(X_i))^{1/2} \dots\dots\dots(3.5)$$

Komponen utama dibentuk sebagai kombinasi linear dari Z_1, Z_2, \dots, Z_p sebagai berikut :

$$\begin{aligned} W_1 &= a_{11} Z_1 + a_{21} Z_2 + \dots + a_{p1} Z_p \\ W_2 &= a_{12} Z_1 + a_{22} Z_2 + \dots + a_{p2} Z_p \\ W_p &= a_{1p} Z_1 + a_{2p} Z_2 + \dots + a_{pp} Z_p \end{aligned} \dots\dots\dots(3.6)$$

Bila ditulis dengan notasi matriks : $\mathbf{W} = \mathbf{AZ}$

$\text{Cov}(\mathbf{Z})$ adalah matriks korelasi \mathbf{R} . Semua formula yang telah diturunkan berdasarkan peubah-peubah X_1, X_2, \dots, X_p dengan matriks \mathbf{S} akan berlaku untuk peubah-peubah Z_1, Z_2, \dots, Z_p dengan matriks \mathbf{R} . Sehingga diperoleh komponen utama pertama :

$$W_1 = a_{11} Z_1 + a_{21} Z_2 + \dots + a_{p1} Z_p = \mathbf{a}_1^T \mathbf{Z} \dots\dots\dots(3.7)$$

$$\text{Var}(W_1) = \mathbf{a}_1^T \mathbf{R} \mathbf{a}_1$$

Harga $\text{Var}(W_1) = \lambda_1$ yang merupakan akar karakteristik terbesar dari \mathbf{R} dan \mathbf{a}_1 merupakan vektor karakteristik yang bersesuaian dengan λ_1 . Demikian juga untuk komponen utama ke dua dan seterusnya, secara umum komponen utama ke j :

$$W_j = a_{1j} Z_1 + a_{2j} Z_2 + \dots + a_{pj} Z_p = \mathbf{a}_j^T \mathbf{Z} \dots\dots\dots(3.8)$$

$$\text{Var}(W_j) = \mathbf{a}_j^T \mathbf{R} \mathbf{a}_j$$

Vektor \mathbf{a}_j^T adalah vektor normal yang dipilih sehingga komponen utama ke j maksimum, serta ortogonal terhadap \mathbf{a}_i^T dengan $i \neq j$ dan diperoleh $\text{Var}(W_j) = \lambda_j$ yang merupakan akar karakteristik terbesar ke j .

Pentingnya suatu komponen utama ke j diukur dengan prosentase keragaman total yang mampu diterangkan oleh komponen utama ke j yaitu sama dengan λ_j / p . Dalam analisis komponen utama dari p buah komponen utama dipilih k ($k < p$) buah komponen utama yang telah mampu menerangkan keragaman cukup tinggi, sekitar 80% - 90%. Setelah dipilih komponen-komponen utama yang akan digunakan (sebanyak k buah) selanjutnya ditentukan persamaan regresi dari peubah takbebas Y dengan komponen utama tersebut. Diperlukan penghitungan skor komponen utama dari setiap pengamatan dengan menggunakan rumus :

$$SK-W_{hi} = \mathbf{a}_i^T \mathbf{Z}_h \dots\dots\dots(3.9)$$

Dengan : $SK-W_{hi}$ adalah skor komponen ke i untuk pengamatan ke h

\mathbf{a}_i^T vektor pembobot komponen utama ke i

\mathbf{Z}_h vektor skor baku dari peubah yang diamati pada pengamatan ke h

Bentuk umum model persamaan regresi komponen utama adalah :

$$Y = c_0 + c_1 W_1 + c_2 W_2 + \dots + c_k W_k + \dots\dots\dots(3.10)$$

Dengan menggunakan metode kuadrat terkecil diperoleh dugaan persamaan regresi :

$$Y^{\wedge} = b_0 + b_1 W_1 + b_2 W_2 + \dots + b_k W_k \dots\dots\dots(3.11)$$

Dengan menggunakan persamaan (3.6) diperoleh persamaan dengan bentuk :

$$Y^{\wedge} = c_0 + c_1 Z_1 + c_2 Z_2 + \dots + c_p Z_p \dots\dots\dots(3.12)$$

Untuk menguji signifikansi koefisien c_i sebagai berikut :

$$H_0 : c_i = 0 \text{ vs } H_1 : c_i \neq 0 \text{ digunakan statistik hitung } t_{hit}(c_i) = c_i / (\text{Var}(c_i))^{1/2}$$

Dengan : $\text{Var}(c_i) = (KTG / JKT) (\lambda_j a_{ij}^2 / \lambda_j)$. Tolak H_0 jika $|t_{hit}(c_i)| > t_{n-k}(\alpha)$.

Akhirnya dengan menggunakan persamaan (3.5) diperoleh persamaan regresi :

$$\hat{Y} = d_0 + d_1 X_1 + d_2 X_2 + \dots + d_p X_p \dots\dots\dots(3.13)$$

Untuk mengetahui tingkat responsif (sensitivitas) dari peubah takbebas Y terhadap perubahan dalam peubah-peubah bebas X_i , dihitung elastisitas rata-rata dari peubah takbebas Y terhadap setiap peubahan bebas X_i , dalam model regresi persamaan (3.13) dengan rumus :



$$; \text{ dengan } i = 1, 2, \dots, p \dots\dots\dots(3.14)$$

Pada dasarnya elastisitas rata-rata Y terhadap setiap peubah bebas X_i mengukur prosentase perubahan-perubahan dalam nilai rata-rata Y apabila terjadi perubahan 1% dalam nilai rata-rata peubah bebas X_i .

4. KASUS PENERAPAN.

Suatu penelitian dilakukan untuk menentukan hubungan antara penjualan suatu produk (Y) dalam ribuan dolar dengan potensi wilayah (X_1 , dikodekan) biaya promosi (X_2) dalam puluhan dolar, banyaknya merek saingan (X_3) dan banyaknya active accounts (X_4) dari 13 wilayah penjualan (Draper and Smith, 1992), diperoleh data sebagai berikut :

X1	X2	X3	X4	Y
7	26	6	60	78.5
1	29	15	52	74.3
11	56	8	20	104.3
11	31	8	47	87.6
7	52	6	33	95.9
11	55	9	22	109.2
3	71	17	6	102.7
1	31	22	44	72.5
2	54	18	22	93.1
21	47	4	26	115.9
1	40	23	34	83.8
11	66	9	12	113.3
10	68	8	12	109.4

Persamaan regresi dari data tersebut dengan metode kuadrat terkecil dengan menggunakan paket program minitab 11.12 adalah :

$$Y = 62.4 + 1.55 X_1 + 0.510 X_2 + 0.102 X_3 - 0.144 X_4$$

Berdasarkan output minitab juga dapat disimpulkan bahwa semua koefisien regresi tersebut tidak nyata, dengan faktor inflasi ragam dari X_1 , X_2 , X_3 dan X_4 sebesar 38.5, 254.4, 46.9 dan 282.5. Harga faktor inflasi ragam (FIR) semuanya jauh lebih besar dari 10 sehingga menandakan adanya kekolinearitas ganda (multikolinearitas) yang sangat serius. Selanjutnya akan diselesaikan dengan regresi komponen utama. Karena skala pengukurannya belum sama digunakan transformasi (3.5) kemudian dianalisis dengan analisis komponen yang didasarkan pada matriks korelasi dan diperoleh hasil sebagai berikut :

Eigenvalue : ($\lambda_1 = 2.2357$, $\lambda_2 = 1.5761$, $\lambda_3 = 0.1866$ dan $\lambda_4 = 0.0016$
Proporsi : 0.559 0.394 0.047 0.000

Kumulatif : 0.559 0.953 1.000 1.000

Komponen utama :

$$W_1 = 0.476 Z_1 + 0.564 Z_2 - 0.394 Z_3 - 0.548 Z_4$$

$$W_2 = 0.509 Z_1 - 0.414 Z_2 - 0.605 Z_3 + 0.451 Z_4$$

$$W_3 = 0.676 Z_1 - 0.314 Z_2 + 0.638 Z_3 - 0.195 Z_4 \dots\dots\dots(3.15)$$

$$W_4 = 0.241 Z_1 + 0.642 Z_2 + 0.268 Z_3 + 0.677 Z_4$$

Komponen utama pertama (W_1) dan komponen utama kedua (W_2) telah mampu menjelaskan 95.3% keragaman dari Y, sehingga cukup dipilih komponen utama pertama dan kedua saja untuk membentuk regresi komponen utama. Diperlukan harga skor komponen utama dengan menggunakan rumus (3.9), kemudian dibentuk persamaan regresi dari Y dengan W_1 dan W_2 melalui skor W_1 dan W_2 menggunakan metode kuadrat terkecil biasa, diperoleh persamaannya sebagai berikut :

$$Y^{\wedge} = 95.4231 + 9.8831 W_1 - 0.1250 W_2 , \text{ dengan menggunakan (3.15) diperoleh :}$$

$$Y^{\wedge} = 95.4231 + 4.6407306 Z_1 + 5.6258184 Z_2 - 3.8183164 Z_3 - 5.4723138 Z_4$$

Untuk pengujian koefisien regresi secara parsial $H_0 : c_i = 0$ vs $H_1 : c_i \neq 0$ digunakan statistik hitung $t_{\text{hit}}(c_i) = c_i / (\text{Var}(c_i))^{1/2}$ Dengan : $\text{Var}(c_i) = (\text{KTG} / \text{JKT}) (j a_{ij}^2 / j)$. Tolak H_0 jika $|t_{\text{hit}}(c_i)| > t_{n-k}(\alpha)$. Diperoleh tabel berikut :

Tabel 3.1 Pengujian Koefisien Regresi Secara Parsial (Berdasarkan Pendekatan Analisis Regresi Komponen Utama)

Peubah (Z _i)	Koef. Reg. (c _i)	Galat baku (s(c _i))	t _{hit}	T _{10;0.01%}
Z ₁	4.6407306	0.03049	152.2050	4.587
Z ₂	5.6258184	0.02963	189.8689	4.587
Z ₃	-3.8183164	0.03248	-117.5590	4.587
Z ₄	-5.4723138	0.03035	-180.3069	4.587

Berdasarkan tabel 3.1 dapat disimpulkan bahwa semua koefisien regresi nyata .

Dengan menggunakan persamaan (3.5) diperoleh :

$$\hat{Y} = 55.3082 + 0.7892X_1 + 0.3616X_2 - 0.5957X_3 - 0.3269X_4$$

Penghitungan elastisitas dengan persamaan (3.14) diperoleh tabel berikut :

Tabel 3.2 Elastisitas Rata-rata Y terhadap X_i

Peubah (X _i)	Koef.Reg. (b _i)	Rataan X _i	Elastisitas (E _i)	Urutan
X ₁	0.7892	7.46	0.06169	4
X ₂	0.3616	48.15	0.18246	1
X ₃	-0.5957	11.77	-0.07348	3
X ₄	-0.3269	30.00	-0.10277	2

Berdasarkan tabel 3.2 terlihat bahwa elastisitas terbesar pada peubah X₂ yaitu 0.18246. Apabila prosentase promosi bertambah 1% akan meningkatkan penjualan secara rata-rata sebesar 0.18246%.

Daftar Pustaka

1. Draper N and Smith H, *Analisis Regresi Terapan*. Edisi 2. Alih Bahasa : Bambang-Sumantri. Penerbit PT Gramedia Pustaka Utama, Jakarta, 1992.
2. Gaspersz V, *Teknik Analisis Dalam Penelitian Percobaan*, Jilid 2, Tarsito, Bandung, 1992.
3. Hawkins D. M, *On the Investigation of Alternative Regression by Principle Component Analysis*, Appl. Statis, 1973, 22 : 75 - 286.
4. Montgomery D. C and Peck E. A, *Introduction to Linear Regression Analysis*, John Wiley & Sons Inc, New York, 1982.
5. *Minitab Reference Manual Release 7*, Data Tech Industries, Inc, Valley Forge, PA Third Printing, 3/90, 1989.
6. Neter J, Wasserman W and Kutner M. H, *Model Linear Terapan*, Buku 1, Analisis Regresi Linear Sederhana, Alih Bahasa Bambang Sumantri, Jurusan Statistika, FMIPA IPB, Bogor, 1997.