BAB II

TEORI PENUNJANG

Seperti yang telah disebutkan dalam bab 1 bahwa teori penunjang berisi tentang model regresi linier berganda , model klasifikasi dua arah , pendekatan model regresi terhadap model klasifikasi dua arah dan regresi minmad .

2.1. MODEL REGRESI LINIER BERGANDA

Dalam kehidupan sehari-hari sering ditemukan persoalan yang melibatkan dua atau lebih variabel yang ada atau yang diduga ada dalam suatu hubungan tertentu . Ada dua jenis variabel dalam model regresi yaitu variabel tak bebas dan variabel bebas.

Dilihat dari banyaknya variabel bebas model regresi terbagi menjadi dua bagian yaitu model regresi linier sederhana dan model regresi linier berganda. Dalam bagian ini akan dibahas mengenai regresi linier berganda . Model regresi linier berganda adalah suatu model regresi yang mencakup lebih dari satu variabel bebas.

Jika Y menyatakan variabel tak bebas dan X_1 ,X_2 ,..., X_k menyatakan variabel-variabel bebas , maka model regresi linier berganda dengan k variabel bebas dapat ditulis sebagai :

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k + \varepsilon \qquad \text{............(2.1)}$$

Dengan :

Y = variabel respon yang merupakan variabel tak bebas

X_j = variabel regresor yang merupakan variabel bebas , $j = 1,2,\dots,k$

β_j = koefisien regresi yang belum diketahui

ε = komponen error yang merupakan variabel random

Parameter β_j menyatakan rata-rata perubahan Y per unit akibat perubahan x_j bila seluruh sisa variabel-variabel bebas x_i $(i \neq j)$ konstan. Parameter β_j sering disebut koefisien regresi parsial, karena koefisien tersebut menggambarkan pengaruh parsial satu variabel bebas bila variabel bebas lainnya dalam model tersebut konstan.

2.2. MODEL KLASIFIKASI DUA ARAH

Seperti yang telah disebutkan dalam Bab I bahwa Model Klasifikasi Dua Arah adalah suatu metode percobaan dimana unit dibagi dalam beberapa kelompok / blok sehingga untuk tiap blok bersifat relatif homogen dan untuk tiap blok diberi suatu perlakuan dengan jumlah yang sama .

Model secara umum dapat dirumuskan seperti pada persamaan (1.1) :

$$Y_{ijk} = \mu + \tau_i + \beta_j + \varepsilon_{ijk} \qquad \text{............(2.2)}$$

$$i = 1,\dots,a \; ; \; j = 1,\dots,b \; ; \; k = 1,\dots,n_{ij}$$

Dalam model klasifikasi dua arah ada dua macam model efek perlakuan yaitu model efek tetap dan model efek random. Model efek tetap merupakan model dimana cara pengambilan sejumlah a perlakuan tersebut dipilih secara khusus oleh pelaku percobaan. Dalam model efek tetap , perlakuan τ_i dan blok β_j didefinisikan sebagai deviasi keseluruhan rata-rata, sehingga

$$\sum_{i=1}^{a} \tau_i = 0 \text{ dan } \sum_{j=1}^{b} \beta_j = 0 \text{ sedangkan } \varepsilon_{ijk} \sim NID\,(0,\sigma^2) \qquad ....(2.3)$$

2.3. REGRESI MINMAD

2.3.1. MODEL REGRESI MINMAD

Regresi minmad didefinisikan sebagai regresi linier yang didapat dengan meminimumkan rata-rata deviasi mutlak antara pengamatan Y_i dan nilai prediksi $\hat{Y}_i$ untuk pengamatan ke-i . Dalam mengestimasi parameter β , meminimumkan rata-rata deviasi mutlak sama artinya dengan meminimumkan $\sum_i |d_i| = \sum_i \left| Y_i - \hat{Y}_i \right|$.

Berdasarkan model regresi linier berganda pada persamaan (2.1) dengan k=1, ... , p-1 variabel bebas , maka :

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_{p-1} X_{p-1} + \varepsilon$$

Parameter $\beta_0, \beta_1, \ldots, \beta_{p-1}$ diperoleh dengan meminimumkan :

$$\frac{1}{n} \sum_i \left| Y_i - \beta_0 - \beta_1 X_1 - \beta_2 X_2 - \ldots - \beta_{p-1} X_{p-1} \right| \qquad ....(2.4).$$

Meminimumkan (2.4) sama artinya dengan meminimumkan

$$\sum_i \left| Y_i - \beta_0 - \beta_1 X_1 - \beta_2 X_2 - \ldots - \beta_{p-1} X_{p-1} \right| \qquad \ldots\ldots\ldots(2.5).$$

Berdasarkan model regresi Y = $X\beta + d$, masalah (2.5) dapat disusun seperti masalah program linier dengan n kendala dan (p+2n) variabel sebagai berikut :

$$\text{Min} \quad \sum_i |d_i|$$

$$\text{Kendala} \quad X\beta + d = Y$$

dimana d dan β tak dibatasi tanda ....(2.6).

Untuk titik data ke-i , d_{1i} dan d_{2i} merupakan bagian positif dan negatif dari d_i . Sehingga untuk $|d_i| = d_{1i} + d_{2i}$ dan $d_i = d_{1i} - d_{2i}$ dengan d_{1i} , d_{2i} tak negatif, maka masalah (2.6) dapat dinyatakan dalam bentuk :

$$\text{Min} \quad \sum_i d_{1i} + \sum_i d_{2i}$$

$$\text{Kendala} \quad X\beta + d_1 - d_2 = Y$$

β tak terbatas tanda ; d_1 , d_2 ≥ 0 ....(2.7).

Dalam bentuk matriks , masalah (2.7) dapat disusun sebagai :

$$\text{Min} \quad e' d_1 + e' d_2$$

$$\text{Kendala} \quad x\beta + Id_1 - id_2 = Y$$

d_1 , d_2 ≥ 0 ; β tak terbatas tanda . ...(2.8)

dengan :

e' = vektor baris ukuran 1 x n yang mempunyai elemen '1'

Y' = (Y_1 , ... , Y_n)

d_1' = (d_{11} , ... , d_{1n})

$d_2' = (d_{21} , \dots , d_{2n})$

$\beta' = (\beta_0 , \dots , \beta_{p-1})$

X matriks n x p dengan n > p , dengan asumsi bahwa rank kolom matriks X (rank X) = p.

2.3.2. METODE SIMPLEX UNTUK REGRESI MINMAD

Metode simpleks ialah suatu metode yang secara sistematis dimulai dari suatu solusi basis fisibel ke solusi basis fisibel lainnya , dilakukan berulang-ulang (dengan jumlah ulangan terbatas) sehingga akhirnya tercapai suatu solusi basis fisibel optimal. Untuk masalah minimum akan didapat solusi optimal yang minimal dengan batas bawah adalah nol .

Dalam bentuk matriks, masalah regresi minmad disajikan seperti masalah program linier :

Min $e' d_1 + e' d_2$

Kendala $x\beta + Id_1 - id_2 = Y$

$d_1 , d_2 \geq 0$; β tak terbatas tanda

Definisi 2.1

Sebarang (β',d_1',d_2') yang memenuhi $x\beta + Id_1 - id_2 = Y$ adalah solusi untuk masalah (2.8) .

Misalkan (X, I , -I) adalah matriks A dengan ordo n x (p+2n) dan (β',d_1',d_2') adalah matriks W . Sebarang kolom dari A dinyatakan sebagai a_j ; j= 1,2,...,p+2n

, sehingga sebarang W yang memenuhi AW = Y merupakan solusi untuk (2.8). Misalkan C′ merupakan vektor (0 ,e′ ,e′) dimana 0 vektor nol ordo 1 x p dan e′= (1, ... ,1) ordo 1 x n . Maka C′ W adalah fungsi tujuan untuk (2.8) .

Definisi 2.2

Himpunan n kolom bebas linier dari matriks A berukuran n x (p+2n) disebut basis dari A .

Definisi 2.3

Sebarang solusi W untuk masalah (2.8) jika memenuhi $W_j > 0$; j = p+1 , ... ,p+2n maka disebut solusi fisibel .

Definisi 2.4

Sebarang solusi fisibel W disebut solusi basis fisibel jika kolom A yang sesuai dengan komponen tak nol dari W membentuk himpunan bebas linier dari vektor-vektor pada R^n .

Langkah-langkah metode simpleks pada regresi minmad :

Langkah 1

Menentukan solusi basis fisibel :

$$W_{p+r} = \begin{cases} Y_r , & Y_r > 0 \\ 0 , & \text{lainnya} , 1 \le r \le n \end{cases}$$

$$W_{p+n+r} = \begin{cases} -Y_r , & Y_r < 0 \\ 0 , & \text{lainnya} , 1 \le r \le n \end{cases}$$

$W_j = 0 \quad ; j = 1, 2, \ldots ,p$

Diberikan α_{ij} yang sesuai dalam bentuk tabel . Basis B yang sesuai untuk solusi ini berisi tepat satu kolom e_r atau $-e_r$ untuk semua r = 1, ..., n dan invers B adalah dirinya sendiri ; $B = B^{-1}$. Sehingga $\alpha_j = B^{-1} a_j$ dengan mengalikan baris ke-r dari A dengan +1 jika W_{p+r} dalam basis dan −1 jika W_{p+n+r} dalam basis dan $B^{-1} Y = W_B$ memberikan solusi basis fisibel yang sesuai . Sedangkan $C_j - Z_j = C_j - \sum_{i=1}^{n} C_{B_i}\alpha_{ij}$ setiap j dan $C_j - Z_j = 0$ untuk a_j yang merupakan basis .

Tabel awal :

C_B	Vektor dalam basis	W_B	α_1	α_2	...	α_j	...	α_{2n+p}
1	a_{p+1} / a_{p+n+1}	$\lvert Y_1 \rvert$	α_{11}	α_{21}	...	α_{j1}	...	$\alpha_{1,2n+p}$
.	.	.	.	.		.		.
.	.	.	.	.		.		.
.	.	.	.	.		.		.
1	a_{p+r} / a_{p+n+r}	$\lvert Y_n \rvert$	α_{n1}	α_{n2}	...	α_{nj}	...	$\alpha_{n,2n+p}$
	$C_k - Z_k$	$Z = \sum_{t=1}^{n} \lvert Y_t \rvert$					$C_j - Z_j$	

Langkah 2

Pilih a_j yang tidak terdapat dalam basis . Untuk mengubah vektor dalam basis dilakukan :

a. misal $C_{j1} - Z_{j1} = \max_{C_k - Z_k > 0} C_k - Z_k$, W_k variabel tak terbatas tanda

b. misal $| C_{j2} - Z_{j2} | = \max_{C_k - Z_k < 0} | C_k - Z_k |$

Pilih j untuk $| C_j - Z_j | = \max [C_{j1} - Z_{j1} , | C_{j2} - Z_{j2} |]$, yang merupakan syarat untuk memilih vektor yang masuk ke basis. Jika j_1 dan j_2 tidak ditemukan maka dilanjutkan ke langkah 5, jika ditemukan dilanjutkan ke langkah 3 .

Langkah 3

Merupakan syarat untuk memilih vektor yang meninggalkan basis .

Jika $C_j - Z_j > 0$ maka pilih r sedemikian sehingga $\dfrac{W_{B_r}}{\alpha_{rj}} = \max\limits_{i \in R_B} \left[\dfrac{W_{B_i}}{\alpha_{ij}}\right] \quad \alpha_{ij} < 0$

Jika $C_j - Z_j < 0$ maka pilih r sedemikian sehingga $\dfrac{W_{B_r}}{\alpha_{rj}} = \min\limits_{i \in R_B} \left[\dfrac{W_{B_i}}{\alpha_{ij}}\right] \quad \alpha_{ij} > 0$

Langkah 4

Membuat tabel baru sesuai dengan basis baru B .

$$\hat{W}_{B_r} = \frac{W_{B_r}}{\alpha_{rj}} \text{ dan } \hat{\alpha}_{rl} = \frac{\alpha_{rl}}{\alpha_{rj}}$$

$$\hat{W}_{B_i} = W_{B_i} - \frac{W_{B_r}}{\alpha_{rj}}.\alpha_{ij} \text{ ; } i \neq r \text{ ; } i = 1, \dots, n$$

$$\hat{\alpha}_{il} = \alpha_{il} - \frac{\alpha_{ij}.\alpha_{rl}}{\alpha_{rj}} \text{ ; } i \neq r \text{ ; } i = 1, \dots, n \text{ ; } l = 1, \dots, p+2n$$

$$\hat{Z} = Z + \left(C_j - Z_j\right)\frac{W_{B_r}}{\alpha_{rj}} \text{ dan } \hat{C}_1 - \hat{Z}_1 = \left(C_1 - Z_1\right) - \left(C_j - Z_j\right)\frac{\alpha_{rl}}{\alpha_{rj}}$$

Kemudian ke langkah 2 dan dilanjutkan dengan B = $\hat{B}$.

Langkah 5

Basis optimal .

Proses perubahan basis optimal jika :

1. $C_j - Z_j \geq 0$ untuk semua variabel nonbasis terbatas secara positif W_j.
2. $C_j - Z_j = 0$ untuk semua variabel nonbasis tak terbatas tanda W_j .

Kondisi optimal pada masalah minmad tidak terpenuhi jika :

1. Variabel nonbasis tak terbatas dengan $C_j - Z_j > 0$ (<0) dan tidak satupun dari $\alpha_{ij} < 0$ (>0) , $i \in R_B$.
2. Variabel nonbasis terbatas dengan $C_j - Z_j < 0$ dan $\alpha_{ij} \leq 0$, $i \in R_B$.

W_1 , ... W_p merupakan parameter $\beta_0,\ldots,\beta_{p-1}$. Dalam model klasifikasi dua arah yang didekati dengan model regresi linier berganda parameter $\beta_0,\ldots,\beta_{p-1}$ sama artinya parameter $(\alpha_1, \alpha_2, \ldots, \alpha_a, \beta_1, \beta_2, \ldots, \beta_b)$. Sehingga tabel akhir didapat :

C_B	Vektor dalam basis	W_B	α_1	α_p	...	α_{p+i}	...	α_{2n+p}
0	a_1	β_0	1	0	...	$\alpha_{1,p+i}$	...	$\alpha_{1,2n+p}$
.	.	.	.	.		.		.
.	.	.	.	.		.		.
.	.	.	.	.		.		.
0	a_p	β_{p-1}	0	1	...	$\alpha_{p,p+i}$	...	$\alpha_{p,2n+p}$
1	a_{p+r} / a_{p+n+r}	d_{1r} / d_{2r}	0	0		$\alpha_{r,p+i}$		$\alpha_{r,p+2n}$
	$C_k - Z_k$	$Z=\sum_{t=1}^{n}\lvert Y_t\rvert$					$C_j - Z_j$ positif	

Berdasar tabel pada iterasi akhir dapat ditunjukkan hal-hal sebagai berikut:

1. Jumlah dari $(C_{p+i} - Z_{p+i})$+ $(C_{p+n+i} - Z_{p+n+i})$ selalu tetap dua untuk i=1,...,n. Diberikan sebarang basis B sedemikian sehingga $B^{-1} a_{p+i} = -B^{-1} a_{p+n+i}$, jika $a_{p+i} = e_i = - a_{p+n+i}$ sehingga $\alpha_{p+i} = -\alpha_{p+n+i}$, oleh karena itu $(C_{p+i} - Z_{p+i}) = 1- C_B\, \alpha_{p+i} = 1 + C_B\, \alpha_{p+n+i}$.

   Selanjutnya $(C_{p+i} - Z_{p+i})+(C_{p+n+i} - Z_{p+n+i})=(1- C_B\alpha_{p+i})+(1 + C_B\alpha_{p+n+i})= 2$.

   Jika $(C_{p+i} - Z_{p+i})$ atau $(C_{p+n+i} - Z_{p+n+i})$ telah diketahui, maka yang lainnya dapat ditentukan tanpa menghitung dengan menggunakan langkah 4 dari algoritma diatas .

2. Jika a_{p+i} merupakan basis maka $B^{-1}a_{p+i} = \alpha_{p+i} = e_i$ oleh karena itu $\alpha_{p+n+i} = -e_i$. Sehingga kedua kolom ini tidak perlu dihitung dalam suatu tabel seperti $\alpha_{p+i}(\alpha_{p+n+i})$ merupakan e_i atau $-e_i$ tergantung apakah a_{p+i} (a_{p+n+i}) dalam basis atau di luar basis. Tetapi biasanya a_{p+i} (a_{p+n+i}) tidak terdapat pada basis secara bersamaan kecuali keduanya bebas linier .
3. Jika a_{p+i} dan a_{p+n+i} merupakan kolom nonbasis maka cukup menghitung α_{p+i} *atau* α_{p+n+i} .