

# Perbandingan Hasil Klasifikasi Analisis Diskriminan dan Regresi Logistik Respon Biner Pada Pengklasifikasian Kejadian Bayi dengan Berat Badan Lahir Rendah (BBLR)

Elvy Kurniati<sup>1</sup>, Sigit Nugroho<sup>2</sup>, dan Baki Swita<sup>2</sup>

<sup>1</sup>Alumni Jurusan Matematika Fakultas MIPA Universitas Bengkulu

<sup>2</sup>Staf Pengajar Jurusan Matematika Fakultas MIPA Universitas Bengkulu

## ABSTRAK

Analisis diskriminan adalah analisis multivariat yang bertujuan untuk mengklasifikasikan suatu pengamatan baru ke dalam salah satu dari kelompok (grup) yang telah diketahui. Dalam penerapannya memerlukan asumsi multivariat normal dan kesamaan struktur matriks varian kovarian antar grup, tetapi kedua asumsi ini tidak selalu dapat dipenuhi. Regresi logistik respon biner adalah bentuk khusus regresi logistik dengan respon biner dimana persamaannya menghasilkan nilai peluang yang dipakai sebagai dasar untuk klasifikasi. Analisis ini tidak memerlukan asumsi seperti halnya analisis diskriminan sehingga dapat diterapkan dalam berbagai skala data. Tujuan penelitian ini adalah membentuk fungsi klasifikasi analisis diskriminan dan regresi logistik respon biner serta membandingkan hasil klasifikasi tersebut dalam kasus kejadian bayi dengan berat badan lahir rendah (BBLR). Hasil yang diperoleh menunjukkan bahwa asumsi multivariat tidak dapat dipenuhi dan matriks varian kovarian dipenuhi. Hasil penelitian ini menunjukkan juga apabila variabel penjelas berskala kategori, regresi logistik respon biner menghasilkan ketepatan klasifikasi yang tinggi dibandingkan dengan analisis diskriminan.

*Kata kunci* : Analisis Diskriminan, Regresi Logistik Respon Biner, Klasifikasi

## PENDAHULUAN

Analisis diskriminan adalah analisis multivariat yang diterapkan untuk memodelkan hubungan antara satu variabel respon yang bersifat dikotomis atau multikotomis maupun kualitatif dengan variabel penjelas yang bersifat kuantitatif (Hair *et al.*, 1998). Analisis diskriminan ini dilakukan berdasarkan perhitungan statistik terhadap kelompok (grup) yang terlebih dahulu diketahui secara jelas. Apabila terdapat dua atau lebih populasi yang telah diukur dalam beberapa karakter  $(x_1, x_2, \dots, x_k)$ , maka dapat dibangun fungsi linear tertentu dari pengukuran itu, dimana fungsi tersebut merupakan fungsi pembeda (pemisah) terbaik bagi populasi-populasi yang diamati. Fungsi linier yang dibangun itu disebut sebagai fungsi diskriminan (*discriminant function*) (Gaspersz, 1992).

Tujuan dari analisis diskriminan ini adalah untuk menghasilkan suatu aturan yang menggambarkan pemisahan beberapa kelompok (grup) pengamatan, serta mengklasifikasikan suatu pengamatan baru ke dalam salah satu dari kelompok (grup) yang telah diketahui. Oleh karena itu, analisis diskriminan ini dapat dipergunakan sebagai metode pengklasifikasian. Dalam penerapannya, fungsi diskriminan ini mengasumsikan bahwa data pada masing-masing kategori/kelompok berdistribusi normal multivariat dan mempunyai struktur matriks varian kovarian yang sama (Hair *et al.*, 1998).

Regresi logistik respon biner adalah bentuk khusus analisis regresi dengan respon biner dan variabel penjelas yang dapat terdiri dari data kontinu, kategori atau campuran antara keduanya. Persamaan regresi logistik respon biner ini tidak menghasilkan nilai pada variabel respon, namun menghasilkan nilai peluang kejadian pada variabel respon. Nilai peluang ini yang dipakai sebagai ukuran untuk mengklasifikasikan pengamatan. Regresi logistik respon biner tidak

memerlukan asumsi distribusi multivariat normal dan kesamaan varian kovarian dalam masing-masing grup (Hosmer & Lemeshow dalam Wibowo, 2002). Sehingga metode ini cukup tahan (*robust*) untuk dapat diterapkan dalam berbagai skala/keadaan data (Hair *et al.*, 1998).

Berdasarkan penjelasan diatas, dapat diketahui bahwa salah satu kegunaan kedua metode tersebut adalah pengklasifikasian, walaupun kedua metode ini berbeda dalam asumsi atau metode estimasi serta inferensi parameternya. Sehingga dengan melihat hasil perbandingan klasifikasi antara analisis diskriminan dan regresi logistik respon biner dapat diketahui pada situasi bagaimana kedua metode tersebut sebaiknya digunakan. Salah satu contoh penerapannya adalah untuk melihat ketepatan klasifikasi kejadian bayi dengan berat badan lahir rendah (BBLR) terhadap faktor-faktor yang mempengaruhinya dengan menggunakan analisis diskriminan dan regresi logistik respon biner.

Berdasarkan uraian diatas maka penulis tertarik untuk mengetahui ketepatan pengklasifikasian kejadian bayi dengan BBLR. Oleh karena variabel respon dalam penelitian ini adalah kejadian bayi dengan BBLR yang berupa data kategori (dua kategori atau biner) dan variabel penjelas adalah umur ibu, parietas, jarak kelahiran, *Antenatal Care* (ANC) serta kenaikan berat badan merupakan data kuantitatif maka analisis ini dapat digunakan analisis diskriminan dan regresi logistik respon biner dalam menganalisis ketepatan pengklasifikasian kejadian bayi dengan BBLR.

## TINJAUAN PUSTAKA

### a. Faktor-faktor yang Mempengaruhi BBLR

Bayi berat badan lahir rendah (BBLR) adalah bayi baru lahir yang berat badan lahirnya pada saat kelahiran kurang dari 2500 gram. Faktor-faktor yang mempengaruhi bayi BBLR ini antara lain (Syafrie dkk., 2004) umur ibu, parietas, jarak kelahiran dengan bayi terdahulu, *antenatal care* (ANC) yang merupakan pelayanan kesehatan bagi ibu hamil dan janinnya oleh tenaga profesional meliputi pemeriksaan kehamilan sesuai dengan standar pelayanan yaitu minimal 4 kali pemeriksaan selama kehamilan, 1 kali pada trimester I, 1 kali pada trimester II dan 2 kali pada trimester III, serta kenaikan berat badan ibu selama hamil.

### b. Analisis Diskriminan

Analisis diskriminan adalah analisis multivariat yang diterapkan untuk memodelkan hubungan antara satu variabel respon yang bersifat dikotomis atau multikotomis maupun kualitatif dengan variabel penjelas yang bersifat kuantitatif (Hair *et al.*, 1998). Tujuan dari analisis diskriminan ini adalah untuk menghasilkan suatu aturan yang menggambarkan pemisahan beberapa kelompok (grup) pengamatan, serta mengklasifikasikan suatu pengamatan baru ke dalam salah satu dari kelompok (grup) yang telah diketahui.

Jika  $\mathbf{x}$  adalah pengamatan baru yang tidak diketahui asalnya, maka dapat dihitung skor diskriminan linier yaitu:

$$W_{ij} = \mathbf{x}'\mathbf{S}_{pl}^{-1}(\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_j) - \frac{1}{2}(\bar{\mathbf{x}}_i + \bar{\mathbf{x}}_j)'\mathbf{S}_{pl}^{-1}(\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_j) \quad (1)$$

Sehingga kriteria klasifikasi berdasarkan statistik  $W$  (*Wald-Anderson*) adalah:

$$\text{golongan } \mathbf{x} \text{ ke grup } i \text{ jika } W_{ij} > 0 \text{ untuk semua } j \neq i \quad (2)$$

dengan catatan  $W_{ij} = -W_{ji}$  (Morrison, 1978).

Keterangan:

$\bar{\mathbf{x}}_i$  = vektor rata-rata sampel grup ke- $i$

$\bar{\mathbf{x}}_j$  = vektor rata-rata sampel grup ke- $j$

$W_{ij}$  = Fungsi diskriminan yang akan menggolongkan individu ke grup  $i$  atau grup  $j$

Menurut Gazpert (1992), jika hanya ada dua grup maka kriteria klasifikasi berdasarkan satatistik  $W$  adalah:

1. Golongkan  $\mathbf{x}$  ke dalam grup 1 jika  $W_{12} > 0$
2. Golongkan  $\mathbf{x}$  ke dalam grup 2 jika  $W_{12} \leq 0$

Salah satu cara untuk menghindari bias adalah membagi sampel menjadi dua bagian yaitu *training sample* yang digunakan untuk membentuk aturan klasifikasi dan *validation sample* yang digunakan untuk mengevaluasi fungsi klasifikasi (Rhencher, 1995). Hair *et al.* (1998) menyatakan, tidak ada acuan pasti dalam membagi sampel menjadi *training sample* dan *validation sample*. Proporsi pembagian sampel ini tidak harus sama besar untuk masing-masing bagian misalnya 60% dan 40%, 75% dan 25% atau yang lainnya. Wibowo (20 02) membuat proporsi pembagian sampel sebesar 80% dan 20% dan Suhaeni (2005) sebesar 60% dan 40%. Proses validasi terhadap fungsi diskriminan yang terbentuk di *training sample* harus dilakukan berkali-kali, yang jelas tidak cukup hanya sekali (Supranto, 200 4).

### c. Analisis Regresi Logistik Respon Biner

Analisis regresi logistik respon biner adalah bentuk khusus analisis regresi dengan respon biner dan variabel penjelas yang dapat terdiri dari data kontinu, kategori atau campuran antara keduanya (Hosmer & Lemeshow dalam Wibowo, 2002). Regresi logistik dengan respon biner ini merupakan regresi logistik dimana hanya terdapat dua kemungkinan dari nilai  $Y$ . Analisis ini digunakan untuk melihat pengaruh sejumlah variabel penjelas  $x_1, x_2, \dots, x_k$  terhadap variabel respon  $Y$  yang berupa variabel respon biner atau juga untuk memprediksi nilai suatu variabel respon  $Y$  (yang berupa variabel biner) berdasarkan nilai variabel-variabel penjelas  $x_1, x_2, \dots, x_k$ .

Model regresi logistik didefinisikan sebagai berikut, misalkan  $Y_1, \dots, Y_n$  adalah variabel respon Bernoulli dan  $\pi_i$  sebagai nilai harapan dari  $Y_i$  yaitu  $\pi_i = E[Y_i] = P[Y_i = 1]$ . Nilai harapan  $\pi_i$  dapat dinyatakan dalam bentuk dari variabel penjelas  $x_{i,1}, x_{i,2}, \dots, x_{i,k}$  yaitu

$$\pi_i = \frac{1}{1 + \exp\left(-\beta_0 - \sum_{j=1}^k \beta_j x_{i,j}\right)} \quad (3)$$

dimana  $\beta = (\beta_1, \dots, \beta_k)$  adalah parameter yang tidak diketahui. Jika persamaan (3) dilakukan transformasi logit, akan didapatkan hubungan linier logit ( $\pi_i$ ).

Sehingga didapat persamaan logit:

$$\text{logit}(\pi_i) = \ln\left(\frac{\pi_i}{1 - \pi_i}\right) = \beta_0 + \sum_{j=1}^k \beta_j x_{i,j} \quad (4)$$

Persamaan (4) dinamakan model logit atau logistik dan merupakan fungsi yang linier dalam parameternya.

Penduga parameter dalam regresi logistik respon biner menggunakan maximum likelihood (Larsen, 2006). Pada dasarnya metode maksimum likelihood memberikan nilai dugaan parameter dengan memaksimumkan suatu fungsi likelihood bagi  $\beta$  dengan iterasi Newton Raphson. Estimasi maksimum likelihood merupakan pendekatan dari estimasi *Weighted Least Square*, dimana matrik pembobotnya berubah setiap putaran. Proses menghitung estimasi maksimum likelihood ini disebut juga sebagai *Iteratif Reweighted Least Square*.

Nilai *odds ratio* merupakan besaran yang digunakan untuk melihat perbandingan masing-masing kategori dari variabel penjelas dalam menerangkan variabel respon. *Odds ratio* biasanya dilambangkan dengan  $\theta$  didefinisikan sebagai ratio dari odds untuk  $x = 1$  terhadap  $x = 0$  atau menunjukkan besarnya perbedaan nilai variabel respon ketika variabel penjelas  $(x + 1)$  dan nilai variabel respon ketika variabel penjelas  $x$ , untuk setiap nilai  $x$ . Ratio antara dua odds didefinisikan sebagai berikut:

$$\theta = \frac{\pi(1)/1 - \pi(1)}{\pi(0)/1 - \pi(0)} \quad (4)$$

Pemilihan antara kedua metode tersebut biasanya tergantung pada asumsi yang diperlukan oleh kedua metode tersebut. Analisis diskriminan mengasumsikan data berdistribusi multivariat normal, sedangkan regresi logistik respon biner tidak mengasumsikan data harus berdistribusi tertentu.

Penelitian dengan dua metode ini telah dilakukan oleh Wibowo (2002), dimana apabila variabel penjelas bersifat campuran antara kontinu dan kategori dalam studi kasus data kerusakan *wall tile*, analisis diskriminan sebaiknya digunakan karena akan menghasilkan ketepatan yang lebih tinggi dibanding regresi logistik respon biner apabila tidak ada pelanggaran asumsi. Apabila variabel penjelas bersifat kategori dalam studi kasus data faktor resiko dan kualitas penderita penyakit stroke, regresi logistik respon biner sebaiknya dipergunakan karena akan menghasilkan ketepatan klasifikasi yang tinggi dibanding analisis diskriminan dan apabila variabel prediktor bersifat kontinu dalam studi kasus data demografi, analisis diskriminan sebaiknya dipergunakan karena akan menghasilkan ketepatan klasifikasi yang lebih tinggi dibanding regresi logistik respon biner meskipun ada pelanggaran asumsi, dengan catatan tidak banyak data yang outlier.

## METODE PENELITIAN

Jenis penelitian ini merupakan penelitian terapan (*applied research*), yaitu penelitian yang dimaksudkan untuk menerapkan suatu teori dalam kaitannya dengan pemanfaatannya dalam bidang kesehatan. Sampel yang digunakan dalam penelitian ini adalah *purposif* sampel yaitu sebanyak 123 responden ibu yang melahirkan bayi dengan BBLR dan tidak BBLR. Data yang digunakan adalah data sekunder, yang diperoleh dari dokumentasi dari klinik bersalin Tiara Sella kota Bengkulu. Gambaran variabel-variabel yang akan diteliti sebagai berikut:

**Tabel 3.1 Gambaran Variabel-variabel yang Akan Diteliti**

Variabel	Label	Skala	Kategori
Respon (Y)	Kejadian BBLR	Nominal	1= ya
			0= tidak
Penjelas (X <sub>1</sub> )	Umur ibu	Ordinal	1= < 20 dan >35
			2= 20-35*
(X <sub>2</sub> )	Parietas	Ordinal	1= <2 dan >3
			2= 2-3*
(X <sub>3</sub> )	Jarak Kelahiran	Ordinal	1= <24 dan >36
			2= 24-36*
(X <sub>4</sub> )	Kenaikan Berat Badan	Ordinal	1= < 10
			2= >=10*
(X <sub>5</sub> )	Frekuensi ANC	Ordinal	1= < 4 kali
			2= ≥ 4 kali*

Keterangan: \* merupakan kategori pembanding

Tahap-tahap analisis data dalam penelitian ini sebagai berikut:

1. Menguji signifikansi secara parsial terhadap variabel-variabel yang ikut dalam analisis secara parsial.
2. Menguji apakah asumsi multivariat normal dan kesamaan matriks varian kovarian terpenuhi oleh data pada masing-masing grup respon.
3. Membagi data menjadi 2 bagian yaitu *training sample* dan *validation sample* dengan proporsi 75% dan 25%, 80% dan 20%, 60 dan 40%. Data yang masuk ke *training sample* maupun *validation sample* dipilih secara acak.
4. Membentuk model klasifikasi dengan analisis diskriminan dan regresi logistik menggunakan data pada *training sample*. Model klasifikasi yang telah diperoleh ini selanjutnya divalidasi dengan data pada *validation sample*. Jika ketepatan klasifikasi dari kedua sampel hampir sama besar, dikatakan fungsi diskriminan dan logistik respon biner dari *training sample* sudah valid.

## HASIL DAN PEMBAHASAN

Hasil uji signifikansi secara parsial, dari lima variabel ternyata ada empat variabel yang memberikan pengaruh secara signifikan (Sig. < 0,05) terhadap kejadian bayi dengan BBLR yaitu umur ibu, jarak kelahiran, kenaikan berat badan serta frekuensi ANC. Variabel parietas ternyata tidak berbeda secara signifikan (Sig. > 0,05).

Untuk selanjutnya, data yang melibatkan semua variabel dalam analisis termasuk variabel parietas disebut sebagai data 1 dan data yang melibatkan semua variabel yang berpengaruh secara signifikan (tanpa variabel parietas) disebut sebagai data 2.

### Analisis Data 1

Berdasarkan hasil pengujian asumsi untuk data 1 ini, asumsi multivariat normal tidak dapat dipenuhi sedangkan matriks varian kovarian dapat dipenuhi. Penelitian ini akan dilihat hasil klasifikasi analisis diskriminan dan regresi logistik respon biner jika variabel berskala kategori dan asumsi tidak mutlak dipenuhi dalam kasus kejadian bayi dengan BBLR.

Hasil penerapan analisis diskriminan dan regresi logistik respon biner dari data 1, dengan proporsi 80% dan 20% untuk *training* dan *validation sample* menghasilkan fungsi diskriminan dan regresi logistik respon biner seperti tertera pada tabel 4.1. Dari tabel tersebut dapat dilihat, pada data *training sample* regresi logistik respon biner memiliki ketepatan klasifikasi yang tinggi (85,7%) dibandingkan analisis diskriminan (83,7%). Sedangkan pada *validation sample* regresi logistik juga menghasilkan ketepatan klasifikasi yang tinggi (80%) dibanding analisis diskriminan (72%). Penilaian valid atau tidaknya fungsi diskriminan maupun regresi logistik respon biner adalah jika ketepatan dari *training* dan *validation sample* sudah sama besar (Santoso, 2004), artinya selisih dari keduanya seminimal mungkin. Jika dilihat dari hasil kedua metode tersebut, selisih yang paling minimum adalah metode analisis regresi logistik respon biner. Jadi, metode analisis regresi logistik respon biner dapat memberikan hasil klasifikasi yang lebih baik daripada analisis diskriminan.

Tabel 4.1 Koefisien Fungsi Diskriminan dan Regresi Logistik Respon Biner 1a

Variabel	Analisis Diskriminan	Regresi Logistik Respon Biner
Umur Ibu	0,9924	1,8987
Parietas	0,5077	0,9570
Jarak Kelahiran	0,5919	1,8241
Kenaikkan Berat Badan	1,9893	2,9315
frekuensi ANC	1,3926	2,8564
Konstanta	-8,6852	-7,0504

Tabel 4.2 Hasil Klasifikasi Analisis Diskriminan dan Regresi Logistik Respon Biner 1a

			Analisis Diskriminan Predicted Group		Ketepatan Klasifikasi
			Tidak BBLR	BBLR	
Training	Actual	Tidak BBLR	65	12	83,70%
Sample	Group	BBLR	4	17	
			Tidak BBLR	BBLR	
Validation	Group	Tidak BBLR	15	5	72%
Sample	Actual	BBLR	2	3	
			Regresi Logistik Respon Biner Predicted Group		Ketepatan Klasifikasi
			Tidak BBLR	BBLR	
Training	Actual	Tidak BBLR	71	6	85,70%
Sample	Group	BBLR	8	13	
			Tidak BBLR	BBLR	
Validation	Group	Tidak BBLR	19	1	80%
Sample	Actual	BBLR	4	1	

Kemudian, hasil penerapan analisis diskriminan dan regresi logistik respon biner dari data 1, dengan proporsi 75% dan 25% akan menghasilkan fungsi diskriminan dengan ketepatan klasifikasi sebesar 84,8% dan regresi logistik respon biner dengan ketepatan klasifikasi 87% pada *training sample* seperti tertera pada tabel 4.4. Jika dilihat dari hasil kedua metode tersebut, metode analisis regresi logistik respon biner masih memberikan hasil klasifikasi yang lebih tinggi daripada analisis diskriminan.

Tabel 4.3 Koefisien Fungsi Diskriminan dan regresi Logistik Respon Biner 1b

Variabel	Analisis	
	Diskriminan	Regresi Logistik Respon Biner
Umur Ibu	1,013	1,993
Parietas	0,424	0,771
Jarak Kelahiran	0,531	1,737
Kenaikkan Berat Badan	2,099	3,172
frekuensi ANC	1,333	2,882
Konstanta	-8,640	-7,049

Tabel 4.4 Hasil Klasifikasi Analisis Diskriminan dan Regresi Logistik Respon Biner 1b

			Analisis Diskriminan Predicted Group		Ketepatan Klasifikasi
			Tidak BBLR	BBLR	
Training	Actual	Tidak BBLR	62	11	84,8%
Sample	Group	BBLR	3	16	
			Tidak BBLR	BBLR	
Validation	Group	Tidak BBLR	17	7	71%
Sample	Actual	BBLR	2	5	

			Regresi Logistik Respon Biner Predicted Group		Ketepatan Klasifikasi
			Tidak BBLR	BBLR	
Training Sample	Actual	Tidak BBLR	68	5	87%
	Group	BBLR	7	12	
			Tidak BBLR	BBLR	
Validation Sample	Group	Tidak BBLR	22	2	77,4%
	Actual	BBLR	5	2	

Jika data 1 dengan proporsi 60% dan 40%, maka menghasilkan fungsi diskriminan dan regresi logistik respon biner seperti tertera pada tabel 4.5. Dari tabel 4.9 dapat dilihat, pada data *training sample* regresi logistik respon biner memiliki ketepatan klasifikasi yang tinggi (86,5%) dibandingkan analisis diskriminan (83,8%). Sedangkan pada *validation sample* regresi logistik respon biner juga menghasilkan ketepatan klasifikasi yang tinggi (79,6%) dibanding analisis diskriminan (71,4%).

Tabel 4.5 Koefisien Fungsi Diskriminan dan Regresi Logistik Respon Biner 1c

Variabel	Analisis	
	Diskriminan	Regresi Logistik Respon Biner
Umur Ibu	1,113	2,114
Parietas	0,321	0,382
Jarak Kelahiran	0,528	1,558
Kenaikkan Berat Badan	2,166	3,153
frekuensi ANC	1,116	2,395
Konstanta	-8,424	-6,318

Tabel 4.6 Hasil Klasifikasi Analisis Diskriminan dan Regresi Logistik Respon Biner 1c

			Analisis Diskriminan Predicted Group		Ketepatan Klasifikasi
			Tidak BBLR	BBLR	
Training Sample	Actual	Tidak BBLR	49	9	83,8%
	Group	BBLR	3	13	
			Tidak BBLR	BBLR	
Validation Sample	Group	Tidak BBLR	27	12	71,4%
	Actual	BBLR	2	8	
			Regresi Logistik Respon Biner Predicted Group		Ketepatan Klasifikasi
			Tidak BBLR	BBLR	
Training Sample	Actual	Tidak BBLR	53	5	86,6%
	Group	BBLR	5	11	
			Tidak BBLR	BBLR	
Validation Sample	Group	Tidak BBLR	36	3	79,6%
	Actual	BBLR	7	3	

Jika dilihat dari hasil kedua metode tersebut, metode analisis regresi logistik respon biner dapat memberikan tingkat klasifikasi yang tinggi daripada analisis diskriminan. Ini berarti bahwa data kategori dengan metode analisis regresi logistik respon biner dapat memberikan hasil klasifikasi yang lebih baik daripada analisis diskriminan.

## Analisis Data 2

Berdasarkan hasil pengujian asumsi untuk data 2 ini, asumsi multivariat normal tidak dapat dipenuhi sedangkan matriks varian kovarian dapat dipenuhi. Penelitian ini akan dilihat hasil klasifikasi analisis diskriminan dan regresi logistik respon biner jika variabel berskala kategori dan asumsi tidak mutlak dipenuhi dalam kasus kejadian bayi dengan BBLR

Hasil penerapan analisis diskriminan dan regresi logistik respon biner dari data 2, dengan proporsi 80% dan 20% untuk *training* dan *validation sample* akan menghasilkan fungsi diskriminan dan regresi logistik respon biner seperti tertera pada tabel 4.7. Dari tabel tersebut dapat dilihat, pada data *training sample* regresi logistik respon biner memiliki ketepatan klasifikasi yang tinggi (85,7%) dibandingkan analisis diskriminan (81,6%). Sedangkan pada *validation sample* regresi logistik juga menghasilkan ketepatan klasifikasi yang tinggi (80%) dibanding analisis diskriminan (64%). Regresi logistik respon biner juga memberikan selisih yang paling minimum antara *training* dan *validation sample*. Jadi, metode analisis regresi logistik respon biner dapat memberikan hasil klasifikasi yang lebih baik daripada analisis diskriminan.

Tabel 4.7 Hasil Klasifikasi Analisis Diskriminan dan Regresi Logistik Respon Biner

Analisis Diskriminan			
Ketepatan Klasifikasi			
	80%-20%(2a)	75%-25%(2b)	60%-40%(2c)
Training Sample	83,80%	82,60%	81,60%
Validation Sample	75,50%	74,20%	64%
Analisis Regresi Logistik Respon Biner			
Ketepatan Klasifikasi			
	80%-20%(2a)	75%-25%(2b)	60%-40%(2c)
Training Sample	86,5%	87%	85,7%
Validation Sample	81,6%	80,6%	80%

Kemudian, hasil penerapan analisis diskriminan dan regresi logistik respon biner dari data 2, dengan proporsi 75% dan 25% akan menghasilkan fungsi diskriminan dengan ketepatan klasifikasi sebesar 82,6% dan regresi logistik respon biner dengan ketepatan klasifikasi 87%. Pada *validation*



*sample* analisis diskriminan menghasilkan ketepatan klasifikasi 74,2% sedangkan regresi logistik respon biner 80,6%. Jika dilihat dari hasil kedua metode tersebut, metode analisis regresi logistik respon biner masih memberikan hasil klasifikasi yang lebih baik daripada analisis diskriminan.

Jika data 2 dengan proporsi 60% dan 40%, pada *training sample* regresi logistik respon biner memiliki ketepatan klasifikasi yang tinggi (86,5%) dibandingkan analisis diskriminan (83,8%). Sedangkan pada *validation sample* regresi logistik juga menghasilkan ketepatan klasifikasi yang tinggi (81,6%) dibanding analisis diskriminan (75,5%). Jika dilihat dari hasil kedua metode tersebut, metode analisis regresi logistik respon biner dapat memberikan tingkat klasifikasi yang tinggi daripada analisis diskriminan.

Dari berbagai cara analisis seperti tersebut diatas, dapat dikatakan bahwa data 1 dan 2 yang mana variabelnya berskala kategori, metode analisis regresi logistik respon biner akan memberikan hasil ketepatan klasifikasi yang tinggi bila dibandingkan dengan diskriminan. Hal ini bisa terjadi dimungkinkan karena semua variabel penjelasnya berskala kategori sehingga penyebaran datanya sangat menyimpang dari bentuk multivariat normal. Menurut Wibowo (2002) pelanggaran asumsi multivariat normal pada analisis diskriminan biasanya menghasilkan tingkat ketepatan klasifikasi yang rendah.

Dari analisis data 1, ternyata ketepatan klasifikasi data 1b (84,8%-87%) lebih tinggi bila dibanding dengan data 1a (83,7%-85,7%) dan 1c (83,8%-86,5%), namun data 1a memberikan hasil klasifikasi yang lebih baik karena selisih antara ketepatan klasifikasi data pada *training* dan *validation sample* lebih minimum dibanding data 1b dan 1c. Sedangkan dari analisis data 2, ternyata ketepatan klasifikasi data 2b (82,6%-87%) juga lebih tinggi bila dibanding dengan data 2a (81,6%-85,7%) dan 2c (83,8%-86,5%). Hal ini berarti bahwa ternyata cara pembagian sampel dapat memberikan pengaruh terhadap hasil analisis diskriminan dan regresi logistik respon biner. Hasil analisis ini menunjukkan regresi logistik respon biner yang menghasilkan ketepatan klasifikasi yang tinggi dengan proporsi 75%-25%.

Dari berbagai cara analisis yang telah dilakukan, maka yang memberikan ketepatan klasifikasi yang tinggi adalah analisis regresi respon biner dengan data 2 (tanpa variabel parietas) yang mempunyai proporsi 75%-25%, yaitu sebesar 87% dari *training sample* dan 80,6% dari *validation sample*.

Hasil penelitian ini menunjukkan bahwa kejadian bayi dengan BBLR dipengaruhi oleh umur ibu, jarak kelahiran, kenaikan berat badan serta frekuensi ANC. Untuk melihat seberapa besar kecenderungan kejadian BBLR berdasarkan variabel-variabel yang mempengaruhinya digunakan *odds ratio* ( $\theta$ ).

Tabel 4.8 Output *Odds Ratio* dalam Analisis Regresi Logistik Respon Biner Data 2b

Variabel	Exp(B)
Umur	7,036
Jarak Kelahiran	6,118
Kenaikan Berat Badan	22,268
Frekuensi ANC	18,341
Konstan	0,002

Nilai rasio kecenderungan masing-masing variabel yang berpengaruh terhadap kejadian bayi dengan BBLR yang tertera pada tabel 4.13 adalah umur ibu sebesar 7,036 berarti bahwa peluang bagi responden yang berusia kurang dari 20 tahun dan besar dari 35 tahun berisiko melahirkan bayi dengan BBLR 7,036 kali lebih tinggi dibandingkan responden yang berumur 20 -35 tahun. Nilai rasio untuk jarak kelahiran sebesar 6,118 yang mana memperlihatkan bahwa peluang bagi responden yang jarak kelahiran kurang dari 24 bulan dan besar dari 36 bulan berpeluang melahirkan BBLR 6,118 kali lebih tinggi dibandingkan responden dengan jarak kelahiran 24 -36 bulan. Nilai rasio untuk kenaikan berat

badan sebesar 22,268 artinya responden dengan kenaikan berat badan kurang dari 10 kg memiliki kecenderungan memperoleh BBLR sebesar 22,268 kali dibandingkan responden yang kenaikan berat badannya lebih dari 10 kg. Nilai rasio untuk frekuensi ANC sebesar 18,341 bermakna bahwa responden dengan frekuensi ANC kurang dari 4 kali memiliki kecenderungan memperoleh BBLR sebesar 18,341 kali dibandingkan responden dengan frekuensi ANC lebih dari 4 kali.

## Kesimpulan

Berdasarkan analisis dan pembahasan pada bab sebelumnya, maka dapat disimpulkan yaitu apabila variabel penjelas berskala kategori dalam kasus kejadian bayi dengan BBLR ini, regresi logistik respon biner menghasilkan ketepatan klasifikasi yang tinggi dibandingkan dengan analisis diskriminan. Variabel-variabel yang mempengaruhi kejadian bayi dengan BBLR serta nilai rasionya adalah umur ibu dengan nilai ratio sebesar 7,036, jarak kelahiran (6,118), kenaikan berat badan (22,268) serta frekuensi ANC (18,341).

## DAFTAR PUSTAKA

- [1]. Anonim. 2002. *Logistic Regression*. <http://online.sfsu.edu/~efc/classes/biol710/logistic/logisticreg.htm>
- [2]. Anonim. 2004. *Estimation and Hypothesis Testing for Logistic Regression*. <http://courses.washington.edu/b515/113.pdf>
- [3]. Anonim. 2005. *Discriminant Function Analysis*. <http://www.statsoft.com/textbook/stdiscan.html>
- [4]. Anonim. 2005. *Indikator Kesejahteraan Rakyat Propinsi Bengkulu*. Badan Pusat Statistik. Bengkulu.
- [5]. Fox, J. 2005. *Maximum Likelihood Estimation of the Logistic Regression Model*. [www.socserv.mcmaster.ca/jfox/Courses/UCLA/logistic-regression-notes.pdf](http://www.socserv.mcmaster.ca/jfox/Courses/UCLA/logistic-regression-notes.pdf).
- [6]. Garson, G.D. 2006. *Logistic Regression*. <http://www2.chass.ncsu.edu/garson/pa765/logistic.htm>.
- [7]. Gaspersz, V. 1992. *Teknik Analisis Dalam Penelitian Percobaan*. Tarsito. Bandung.
- [8]. Hair, J.F. et al. 1998. *Multivariate Data Analysis*. Prentice Hall. New Jersey.
- [9]. Larsen, P.V. 2006. *Logistic Regression*. <http://statmaster.sdu.dk/courses/st111/module14/module.pdf>.
- [10]. Morrison, D.F. 1978. *Multivariate Statistical Methods*. International Student Edition, McGrawhill.
- [11]. Permata, A. 2006. *Analisis Regresi Logistik Status Kemiskinan Penduduk Kota Bengkulu Tahun 2004*. Skripsi FMIPA Universitas Bengkulu.
- [12]. Rhencher, A.C. 1995. *Method of Multivariate Analysis*. John Wiley and Sons. New York.
- [13]. Ridgeway, G. 2004. *Maximum Likelihood and Logistic Regression*. [www.i-pensieri.com/gregr/ModernPrediction/L2logistic.pdf](http://www.i-pensieri.com/gregr/ModernPrediction/L2logistic.pdf).
- [14]. Santoso, S. 2002. *SPSS Statistik Multivariat*. Gramedia. Jakarta.
- [15]. Seber, G. A. F. 1984. *Multivariate Observation*. John Wiley and Sons. New York.
- [16]. Sitohang, N.A. 2004. *Asuhan Keperawatan Pada Bayi Berat Badan Lahir Rendah*. <http://library.usu.ac.id/download/fk/04006076.pdf>.
- [17]. Suhaeni, C. 2005. *Analisis Diskriminan Untuk Prediksi Indeks Prestasi Mahasiswa Pada Semester Pertama*. Skripsi FMIPA Universitas Bengkulu.
- [18]. Supranto, J. 2004. *Analisis Multivariat: Arti dan Interpretasi*. Rineka Cipta. Jakarta.
- [19]. Syafrie, dkk.. 2004. *Hubungan Pelayanan Antenatal Terhadap Kejadian Bayi Berat badan Lahir Rendah di Kabupaten Purworejo*. [www.akatiga.or.id/black\\_paper/van\\_bem\\_grijns.pdf](http://www.akatiga.or.id/black_paper/van_bem_grijns.pdf)

- [20]. Wibowo, W. 2002. *Perbandingan Hasil Klasifikasi Analisis Diskriminan dan Regresi Logistik Pada Pengklasifikasian Data Respon Biner*. Kappa, Vol 3, No 1:36-45. [www.fmipa.its.ac.id/isi%20mipajurnal/jurnal/KAPPA%20\(2002\)%20Vol.3,%20No.1,%2036-45.pdf](http://www.fmipa.its.ac.id/isi%20mipajurnal/jurnal/KAPPA%20(2002)%20Vol.3,%20No.1,%2036-45.pdf)