

ANALISIS REGRESI LOGISTIK MULTINOMIAL PADA FAKTOR-FAKTOR YANG MEMPENGARUHI KADAR GULA DARAH PENDERITA DIABETES MELLITUS

MULTINOMIAL LOGISTIK REGRESSION ANALYSIS OF FACTORS THAT AFFECTING THE BLOOD SUGAR LEVELS DIABETES MELLITUS PATIENTS

Oleh: Leila Anggi Gusniar,

Jurusan Pendidikan, FMIPA, UNY.

e-mail: leilaagusniar@gmail.com dan listyani@uny.ac.id

Abstrak

Penelitian ini bertujuan untuk melakukan analisis regresi logistik multinomial dan mengetahui faktor yang secara signifikan mempengaruhi kadar gula darah penderita diabetes mellitus serta mencari model regresi logistik terbaik. Metode yang digunakan dalam penelitian ini adalah metode maksimum likelihood dan uji parameter menggunakan uji simultan dan uji parsial. Variabel terikat yaitu kadar gula darah yang dibedakan menjadi kadar gula darah rendah ($Y=1$), kadar gula darah normal ($Y=2$) dan kadar gula darah tinggi ($Y=3$). Variabel bebas yaitu usia, jenis kelamin, indeks massa tubuh, tekanan darah dan 5 jenis serum darah yaitu kadar kolesterol, *low density lipoprotein* (LDL), *high density lipoprotein* (HDL), *thyrocalcitonin hormone* (TCH) dan *loss trigliserida* (LTG). Hasil dari penelitian ini menunjukkan faktor-faktor yang secara signifikan mempengaruhi kadar gula darah penderita diabetes mellitus adalah usia, indeks massa tubuh, LDL, dan TCH.

Kata kunci: regresi logistik multinomial, Metode maksimum likelihood, diabetes mellitus.

Abstract

The purpose of this study is to analyzing multinomial logistic regression and to know the factors that significantly affect blood sugar levels of people with diabetes mellitus and look for the best logistic regression model. The method used in this research is the maximum likelihood method and parameter test using simultaneous test and partial test. The dependent variable is blood sugar level which is differentiated into low blood sugar level ($Y = 1$), normal blood sugar level ($Y = 2$) and high blood sugar level ($Y = 3$). The independent variables are age, sex, body mass index, blood pressure and 5 types of blood serum ie cholesterol, low density lipoprotein (LDL), high density lipoprotein (HDL), thyrocalcitonin hormone (TCH) and triglyceride loss (LTG). The results of this study shows the factors that significantly affect blood sugar levels of patients with diabetes mellitus are age, body mass index, LDL, and TCH.

Keywords: multinomial logistic regression, maximum likelihood method, diabetes mellitus.

PENDAHULUAN

Statistika merupakan salah satu disiplin ilmu yang penerapannya hampir di semua aspek kehidupan. Hal ini

menunjukkan bahwa peranan statistika sangat diperlukan dalam kehidupan sehari-hari. Begitu pula dalam penelitian ilmiah, statistika merupakan alat yang

berguna bagi perencanaan dan evaluasi hasil penelitian, sehingga dapat dilakukan perbaikan dan penyempurnaan terhadap hasil penemuan (Gunardi, 1999). Salah satu analisis pada statistika adalah analisis regresi. Analisis regresi merupakan suatu metode yang digunakan untuk menganalisis hubungan antara dua atau lebih variabel. Berdasarkan pola hubungannya analisis regresi dibagi menjadi dua yaitu analisis regresi linear dan analisis regresi non-linear. Salah satu model regresi nonlinear adalah regresi logistik.

Regresi logistik merupakan sebuah metode analisis statistik untuk menggambarkan hubungan antara variabel terikat dengan variabel bebas yang mempunyai dua atau lebih kategori dengan variabel terikat yang menggunakan skala kategorik maupun interval (Hosmer dan Lemeshow, 1989). Regresi logistik terbagi menjadi dua yaitu regresi logistik biner dan regresi logistik multinomial. Regresi Logistik biner adalah suatu analisis regresi yang digunakan untuk menggambarkan hubungan antara variabel bebas dengan sekumpulan variabel terikat, dimana variabel terikat bersifat biner atau dikotomus. Variabel kategorik yang

tidak memiliki urutan disebut sebagai variabel nominal sedangkan yang memiliki urutan disebut variabel ordinal. Kedua jenis variabel ini, baik nominal maupun ordinal sering disebut juga sebagai variabel multinomial.

Regresi logistik multinomial, yang tidak mempertimbangkan sifat ordinal data, juga dapat diterapkan untuk meneliti sebuah variabel ordinal namun memanfaatkan sifat ordinal data dapat meningkatkan kesederhanaan dan kekuatan model (Agresti, 2002). Salah satu metode yang lebih umum dan digunakan pada sebagian besar paket program komputer adalah *maximum likelihood methods* atau metode maksimum likelihood. Metode maksimum likelihood merupakan metode pendugaan parameter yang digunakan pada model regresi logistik.

Pada penelitian ini, ingin diketahui model terbaik pada model logistik multinomial dengan variabel terikat data ordinal pada kasus kadar gula darah penderita diabetes mellitus. Penelitian ini bertujuan untuk menganalisis regresi logistik multinomial pada kasus diabetes mellitus, mengetahui faktor apa saja yang mempengaruhi kadar gula darah penderita diabetes mellitus serta mencari model terbaiknya.

Diabetes mellitus atau penyakit kencing manis merupakan suatu penyakit menahun yang ditandai dengan kadar glukosa darah (gula darah) melebihi nilai normal (Misnadiarly, 2006). Penderita diabetes mellitus terus meningkat setiap tahunnya. Organisasi Kesehatan Dunia (WHO) memprediksi bahwa di Indonesia pada tahun 2030 diperkirakan pasien penyakit diabetes mellitus akan mencapai angka 21,3 juta jiwa dari 8,4 juta jiwa pada tahun 2010. Awalnya diabetes mellitus hanya diderita sebagian kecil orang, namun dengan meningkatnya gaya hidup sekarang ini jumlah pasien pun bertambah.

Beberapa faktor yang dianggap mempengaruhi pada penderita diabetes mellitus adalah usia, indeks masa tubuh, tekanan darah, kadar glukosa dan kadar kolesterol (Misnadiarly, 2006). Menurut Graha (2010), umumnya kadar kolesterol diukur melalui total kolesterol, *low-density lipoprotein*, *high density lipoprotein*, *thyrocalcitonin hormone*, serta *loss trigliserida*. Pada penulisan skripsi ini faktor-faktor yang berpengaruh tersebut digunakan sebagai variabel bebas dan kadar gula darah penyakit diabetes mellitus sebagai variabel terikat akan di analisis dengan

menggunakan analisis regresi logistik multinomial.

METODE PENELITIAN

Data yang digunakan dalam penelitian ini adalah data sekunder dari hasil penelitian Universitas Stanford tahun 2004 tentang pasien penderita diabetes mellitus. Analisis data dilakukan menggunakan regresi logistik multinomial dengan langkah-langkah sebagai berikut:

1. Variabel terikat dikategorikan menjadi 1, 2 dan 3, yaitu kadar gula darah rendah, kadar gula darah normal dan kadar gula darah tinggi.
2. Melakukan *likelihood ratio test* atau uji simultan untuk menguji keseluruhan model dengan menggunakan seluruh variabel bebas. Uji simultan ini bertujuan untuk mengetahui apakah variabel bebas berpengaruh signifikan terhadap variabel terikat secara keseluruhan. Dari persamaan $g(x_i) = \beta_0 + \beta_1 X_{i1} + \dots + \beta_k X_{ik}$ diperoleh hipotesis yang akan diuji sebagai berikut:

$H_0: \beta_1 = \beta_2 = \dots = \beta_j = 0$, artinya tidak ada pengaruh signifikan antara variabel bebas dengan variabel terikat secara simultan.

$H_1: \exists \beta_j \neq 0$, artinya minimal ada satu variabel bebas yang berpengaruh signifikan terhadap model.

Jika H_0 ditolak, maka minimal ada satu variabel bebas yang berpengaruh signifikan terhadap variabel terikat.

- Melakukan uji parameter model dengan menggunakan uji parsial (uji *Wald*) untuk menguji tiap variabel bebas terhadap variabel terikat. Uji parsial ini bertujuan untuk mengetahui peran setiap variabel bebas dalam model secara individu.

Hipotesis yang digunakan adalah:

$H_0: \beta_j = 0$, artinya tidak ada pengaruh antara variabel bebas ke- j terhadap variabel terikat.

$H_1: \beta_j \neq 0$, artinya ada pengaruh antara variabel bebas ke- j terhadap variabel terikat.

Jika H_0 ditolak, maka ada pengaruh signifikan antara variabel bebas ke- j terhadap variabel terikat.

- Uji kebaikan model regresi logistik dilakukan untuk menguji layak atau tidaknya model yang dihasilkan.

Hipotesis yang digunakan adalah:

H_0 : tidak terdapat perbedaan yang signifikan antara hasil pengamatan dengan kemungkinan hasil atau prediksi model (model sesuai)

H_1 : terdapat perbedaan yang signifikan antara hasil pengamatan dengan kemungkinan hasil atau prediksi model (model tidak sesuai).

Jika H_0 diterima maka model sesuai.

- Odd rasio merupakan ukuran untuk mengetahui risiko kecenderungan suatu kategori terhadap kategori lainnya.
- Model peluang persamaan regresi logistik

$$\pi(x) = \frac{\exp(\beta_0 + \beta_{1i}x_i + \dots + \beta_{ki}x_k)}{1 + \exp(\beta_0 + \beta_{1i}x_i + \dots + \beta_{ki}x_k)}$$

Dengan model transformasi logit untuk model menjadi:

$$g(x) = \frac{[\ln \pi(x)]}{[1 - \pi(x)]} = (\beta_0 + \beta_{1i}x_i + \dots + \beta_{ki}x_k)$$

HASIL DAN PEMBAHASAN

Uji Parameter

Sebelum dilakukan uji parameter akan dilihat pendugan parameter menggunakan metode *maximum likelihood*. Hasil pendugaan parameter disajikan dalam Tabel 1 berikut ini:

Tabel 1. Pendugaan Parameter

	B	Wald	Sig
Rendah	0,369	0,011	0,917
Normal	4,049	1,268	0,260
Usia	0,045	7,050	0,008
Gender	0,091	0,032	0,857
IMT	2,243	19,964	0,000
BP	0,478	0,134	0,715
TC	-0,001	0,820	0,365
LDL	-0,024	5,890	0,015
HDL	-0,012	0,497	0,481
TCH	0,202	2,890	0,027
LTG	0,603	1,207	0,089

Menggunakan variabel terikat kategori ke tiga yaitu kategori kadar gula darah tinggi sebagai pembanding di dapatkan model regresi logistik sebagai berikut:

$$g_1(x) = 0,369 + 0,045X_1 + 2,243X_3 - 0,024X_6 + 0,202X_8$$

$$g_2(x) = 4,049 + 0,045X_1 + 2,243X_3 - 0,024X_6 + 0,202X_8$$

Selanjutnya dilakukan uji signifikansi model secara simultan menggunakan uji G atau *likelihood ratio test* dengan hipotesis sebagai berikut:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_9 = 0$$

$$H_1 : \exists \beta_j \neq 0$$

Hasil pengujian diperoleh nilai signifikansi atau p-value sebesar 0,000 dengan selang kepercayaan 95% didapatkan keputusan tolak H_0 . Tabel *likelihood ratio test* tersaji dalam Tabel 2 berikut:

Tabel 2. Uji Simultan

Model Fitting Information				
Model	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	183,119			
Final	142,483	40,636	10	,000

Link function: Logit.

Berikutnya untuk mengetahui variabel bebas yang signifikan secara individu dan layak masuk ke dalam model digunakan uji Parsial menggunakan uji *Wald*. Hipotesisnya yaitu:

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0, j = 1, 2, 3, \dots, 9$$

Tabel 3. Uji Parsial

Variabel	Chi-Square	Sig	Keterangan
Intersep	7,551	0,023	Tolak H_0
Usia	10,889	0,004	Tolak H_0
Gender	5,757	0,056	Terima H_0
IMT	21,270	0,000	Tolak H_0
BP	4,252	0,119	Terima H_0
TC	2,471	0,291	Terima H_0
LDL	0,346	0,041	Tolak H_0
HDL	5,138	0,077	Terima H_0
TCH	2,009	0,035	Tolak H_0
LTG	4,382	0,112	Terima H_0

Berdasarkan Tabel 3 diketahui bahwa dengan $\alpha = 0,05$ dari sembilan faktor didapatkan empat faktor yang berpengaruh secara signifikan terhadap kadar gula darah penderita diabetes mellitus yaitu usia, IMT, LDL dan TCH.

Uji Keباikan Model

Selanjutnya akan dilakukan uji kebaikan model menggunakan uji *Pearson* dengan hipotesis sebagai berikut:

H_0 : model layak digunakan atau model sesuai

H_1 : model tidak layak digunakan atau model tidak sesuai

Tabel 4. Uji Keباikan Model

	Chi-Square	Df	Sig
Pearson	175,386	188	0,736
Deviance	142,483	188	0,994

Berdasarkan Tabel 4 diketahui bahwa nilai signifikansi baik *Pearson* maupun *Deviance* lebih besar dari $\alpha = 0,05$ artinya H_0 diterima dan model sesuai.

Koefisien Determinasi

Nilai koefisien determinasi dapat diketahui dari nilai *Mc Fadden*, *Cox and Snell* dan *Nagelkerke* seperti pada Tabel 5 berikut ini.

Tabel 5. Koefisien Determinasi

Cox and Snell	0,330
Nagelkerke	0,393
McFadden	0,219

Berdasarkan Tabel 5 diketahui bahwa nilai Cox and Snell sebesar 0,33, nilai Nagelkerke sebesar 0,393 dan nilai Mc Fadden sebesar 0,219. Artinya variabel bebas nya mampu menjelaskan variabel terikat sebesar 31,4%.

Odds Ratio

Apabila model telah diuji dan hasil modelnya baik serta signifikansinya nyata maka data tersebut dapat diinterpretasikan dengan menggunakan uji *odds ratio* seperti pada tabel 6 berikut ini:

Tabel 6. Uji *Odds Ratio*

Kadar glukosa	Variabel	Sig	Odds Ratio
Rendah	Usia	0,007	0,869
	IMT	0,029	9,585
	LDL	0,038	1,014
	TCH	0,018	4,420
Normal	Usia	0,019	0,969
	IMT	0,000	9,226
	LDL	0,048	0,999
	TCH	0,039	1,599

Berdasarkan Tabel 6 diketahui bahwa semakin bertambahnya usia pada penderita kadar gula darah rendah maka peluang seseorang menderita diabetes mellitus lebih kecil 0,869 kali dibanding penderita kadar gula darah tinggi. Sedangkan pada penderita kadar gula darah normal maka peluang seseorang menderita diabetes mellitus lebih kecil 0,969 kali dibanding penderita kadar gula darah tinggi. Uji *odd ratio* untuk IMT adalah semakin bertambahnya berat badan pada penderita kadar gula darah rendah maka peluang seseorang menderita diabetes mellitus lebih besar 9,585 kali dibanding penderita kadar gula darah tinggi. Sedangkan pada penderita kadar gula darah normal maka peluang seseorang menderita diabetes mellitus lebih besar 9,226 kali dibanding penderita kadar gula darah tinggi. Uji *odd ratio* untuk LDL adalah semakin tinggi kadar LDL pada penderita kadar gula darah rendah maka peluang seseorang menderita diabetes mellitus

lebih besar 1,014 kali dibanding penderita kadar gula darah tinggi. Sedangkan pada penderita kadar gula darah normal maka peluang seseorang menderita diabetes mellitus lebih kecil 0,999 kali dibanding penderita kadar gula darah tinggi. Uji *odd ratio* untuk TCH adalah semakin tinggi kadar TCH pada penderita kadar gula darah rendah maka peluang seseorang menderita diabetes mellitus lebih besar 4,420 kali dibanding penderita kadar gula darah tinggi. Sedangkan pada penderita kadar gula darah normal maka peluang seseorang menderita diabetes mellitus lebih besar 1,599 kali dibanding penderita kadar gula darah tinggi.

Pemodelan Regresi Logistik Multinomial

Setelah melakukan uji-uji tersebut maka akan dibentuk model logit terbaik regresi logistik multinomial pada kasus diabetes mellitus yaitu:

Logit 1

$$g_1(x) = \ln\left[\frac{\pi_1(x)}{\pi_3(x)}\right]$$

$$g_1(x) = 0,369 + 0,045X_1 + 2,243X_3 - 0,024X_6 + 0,202X_8$$

Logit 2

$$g_2(x) = \ln\left[\frac{\pi_2(x)}{\pi_3(x)}\right]$$

$$g_2(x) = 4,049 + 0,045X_1 + 2,243X_3 - 0,024X_6 + 0,202X_8$$

KESIMPULAN DAN SARAN

Kesimpulan

dari hasil penelitian didapatkan faktor-faktor yang secara signifikan mempengaruhi kadar gula darah penderita diabetes mellitus adalah usia, jenis kelamin, *low density lipoprotein* (LDL) dan *thyrocalcitonin hormone* (TCH) dengan model regresi logistik multinomial terbaik adalah $g_1(x) = 0,369 + 0,045X_1 + 2,243X_3 - 0,024X_6 + 0,202X_8$ dan $g_2(x) = 4,049 + 0,045X_1 + 2,243X_3 - 0,024X_6 + 0,202X_8$

Saran

Penelitian ini terbatas pada cara mengestimasi parameter menggunakan metode maksimum likelihood, padahal masih ada beberapa metode seperti *noniterative weighted least squares* dan *discriminant function analysis*. Untuk penelitian selanjutnya dapat digunakan metode lain seperti NN (*Neural Networks*), Fungsi Radial Basis, model CART dan yang lainnya yang memungkinkan memberikan hasil yang lebih baik.

Pencegahan faktor risiko yang mempengaruhi kadar gula darah dapat dilakukan dengan melakukan aktivitas fisik secara teratur, penerapan pola makan gizi seimbang dengan banyak

mengonsumsi sayuran dan buah-buahan, menghindari rokok serta minuman keras.

DAFTAR PUSTAKA

Agresti, Alan. (2002). *Categorical Data Analysis*. New York: John Willey & Sons

Graha, Chairinniza. (2010). *100 Questions and Answers: Kolesterol*. Jakarta: Elex Media Komputido

Gunardi. (1999). *Metode Statistics*. Yogyakarta: FMIPA Universitas Gadjah Mada.

Hosmer, D.W., dan Lemeshow, S. (1989). *Applied Logistic Regression*. John Wiley & Sons Inc., New York.

Misnadiarly. (2006). *Diabetes Mellitus Gangren, Ulcer, Infeksi*. Jakarta: Pustaka Populer Obor