

Analisis Regresi Logistik Status Kemiskinan Penduduk di Kota Bengkulu Tahun 2004

Atika Permata¹, Sigit Nugroho², dan Fachri Faisal²

¹Alumni Jurusan Matematika Fakultas MIPA Universitas Bengkulu

²Staf Pengajar Jurusan Matematika Fakultas MIPA Universitas Bengkulu

ABSTRAK

Penelitian ini bertujuan untuk mengetahui variabel-variabel apa saja yang mempengaruhi kecenderungan status kemiskinan penduduk di Kota Bengkulu tahun 2004. Data yang digunakan adalah data mentah yang diperoleh dari Badan Pusat Statistik (BPS). Analisis regresi logistik dengan metode enter, stepwise forward dan stepwise backward digunakan. Hasil penelitian menunjukkan bahwa variabel-variabel yang mempengaruhi status kemiskinan penduduk di Kota Bengkulu tahun 2004 adalah umur kepala rumah tangga, jumlah anggota rumah tangga, pendidikan yang ditamatkan kepala rumah tangga, morbiditas, cara berobat, fasilitas air bersih, fasilitas sanitasi, status pekerjaan utama kepala rumah tangga dan lapangan usaha utama kepala rumah tangga.

Kata Kunci : analisis regresi logistik, kemiskinan, Bengkulu

PENDAHULUAN

Informasi mengenai jumlah penduduk miskin yang dikeluarkan BPS Provinsi Bengkulu hanya memberikan perkiraan secara makro artinya hanya berdasarkan garis kemiskinan yang ditentukan atas dua kriteria yaitu pengeluaran konsumsi per kapita per bulan yang setara dengan 2100 kalori per kapita per hari dan nilai kebutuhan minimum komoditi bukan makanan. Perkiraan tersebut secara operasional tidak dapat menjelaskan untuk menentukan penduduk atau rumah tangga yang dikategorikan miskin atau tidak miskin.

Penelitian ini bertujuan untuk mengetahui variabel-variabel apa saja yang mempengaruhi kecenderungan status kemiskinan penduduk di Kota Bengkulu tahun 2006 serta besar rasio kecenderungan variabel-variabel tersebut terhadap status kemiskinan penduduk berdasarkan analisis regresi logistik.

TINJAUAN PUSTAKA

Analisis regresi logistik adalah analisis yang digunakan untuk melihat hubungan fungsional antara variabel respon yang bersifat kategori (dikotomis dan polykotomis) dan variabel-variabel penjelas berupa data kualitatif (nominal atau ordinal) maupun kuantitatif (interval atau rasio) (Anonim, 2004a).

Variabel penjelas dalam regresi logistik adalah variabel yang bertipe kualitatif maupun kuantitatif. Untuk variabel penjelas bertipe kualitatif digunakan variabel dummy (Agung, 2002) sedangkan variabel kuantitatif didefinisikan langsung tanpa pengkodean. Variabel respon dalam model regresi logistik berbentuk dikotomis (biner atau dua kategori) maupun pol ykotomis (banyak kategori).

Model regresi logistik dengan X variabel penjelas dan Y variabel respon satu-nol adalah sebagai berikut:

$$\pi(x) = \frac{\exp(\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k)}{1 + \exp(\beta_0 + \beta_1 x_1 + \dots + \beta_k x_k)}$$

dimana $\beta = (\beta_0, \dots, \beta_k)$ adalah parameter yang tidak diketahui.

Sehingga bentuk fungsi linier dari fungsi regresi logistik adalah sebagai berikut

$$\ln \left[\frac{\pi(\underline{x})}{1 - \pi(\underline{x})} \right] = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k = \underline{x}'\underline{\beta}$$

Pendugaan parameter dalam regresi logistik menggunakan metode maksimum likelihood (Hosmer dan Lemeshow, 1989) Penduga dengan variabel respon dikotomus berdistribusi Bernoulli dapat dituliskan sebagai berikut :

$$P(Y_i = y_i) = \pi(x_i)^{y_i} [1 - \pi(x_i)]^{1-y_i},$$

Untuk $Y_i = 0$ atau 1 ; $i = 1, 2, \dots, n$

maka fungsi logaritma likelihood bagi $\underline{\beta}$:

$$\ln L(\underline{\beta}) = l(\underline{\beta})$$

$$= \sum_{i=1}^n \left\{ y_i \underline{x}'_i \underline{\beta} + \ln \left[1 - \frac{1}{1 + \exp(-\underline{x}'_i \underline{\beta})} \right] \right\}$$

$$= \sum_{i=1}^n \left\{ y_i \underline{x}'_i \underline{\beta} + \ln [1 + \exp(\underline{x}'_i \underline{\beta})] \right\}$$

Nilai $\underline{\beta}$ diperoleh dengan memaksimalkan fungsi logaritma likelihood $l(\underline{\beta})$ dengan mencari nilai stasioner fungsi logaritma likelihood yaitu turunan pertama fungsi logaritma likelihood terhadap $\underline{\beta}$ sama dengan nol

$$\begin{aligned} \frac{\partial l(\underline{\beta})}{\partial \underline{\beta}} &= \sum_{i=1}^n \left\{ y_i \underline{x}'_i - [1 + \exp(\underline{x}'_i \underline{\beta})]^{-1} \exp(\underline{x}'_i \underline{\beta}) \underline{x}'_i \right\} \\ &= X' [y - \pi(\underline{x})] \end{aligned}$$

Karena turunan pertama dari fungsi tersebut merupakan persamaan non linier maka pendugaan parameter $\underline{\beta}$ diperoleh dengan menggunakan metode iterasi Newton-Raphson.

Pengujian Parameter

Pada penelitian ini, pengujian parameter dilakukan untuk memeriksa apakah variabel penjelas mempunyai peranan yang nyata di dalam model. Uji parameter yang digunakan adalah uji signifikan model (Likelihood Ratio Test) dan Uji parameter Parsial (Wald Test).

Uji Signifikan Model (Likelihood Ratio Test)

Hipotesis:

H_0 : $\beta_1 = \beta_2 = \beta_3 = \dots = \beta_k = 0$ (tidak ada pengaruh variabel penjelas pada variabel respon)

H_1 : Minimal ada satu $\beta_j \neq 0$ (minimal ada satu variabel penjelas yang berpengaruh terhadap variabel respon); $j = 1, 2, \dots, k$

Statistik Uji :

$$G^2 = -2 \ln \left[\frac{L_0}{L_p} \right]$$

Kesimpulan:

H_0 ditolak jika $G^2 > \chi^2_{0,05; k}$ atau $p\text{-value} < 0,05$, yang berarti variabel penjelas secara bersama-sama mempengaruhi variabel respon.

Uji Parameter Secara Parsial (Wald Test)

Hipotesis:

$H_0 : \beta_j = 0$ (tidak ada pengaruh variabel penjelas ke-j pada variabel respon)

$H_1 : \beta_j \neq 0$ (ada pengaruh variabel penjelas ke-j pada variabel respon);
 $j = 1, 2, \dots, k$

Statistik Uji :

$$W = \left[\frac{\hat{\beta}_j}{se(\hat{\beta}_j)} \right]^2$$

Kesimpulan:

H_0 ditolak jika $W > \chi^2_{0,05; 1}$ atau $p\text{-value} < 0,05$, yang berarti variabel penjelas secara parsial mempengaruhi variabel respon.

Rasio Kecenderungan

Rasio kecenderungan didefinisikan sebagai rasio 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 x .

Nilai rasio kecenderungan dirumuskan:

$$= e$$

METODE PENELITIAN

Penelitian ini merupakan penelitian terapan dengan sampel sebanyak 416 rumah tangga dan variabel-variabel yang digunakan adalah sebagai berikut :

Variabel	Label	Kategori
Respon (Y)	Status Kemiskinan Penduduk	1. Penduduk Miskin 2. Penduduk Tidak Miskin
	Penjelas X1	umur kepala rumah tangga
X2	Status perkawinan krt	1. Belum Kawin 2. Kawin 3. Duda 4. Janda
X3	Jumlah art	1. > 4 2. <= 4
X4	Pendidikan tertinggi yang ditamatkan krt	1. SMP Ke bawah 2. SMA Ke Atas
X5	kemampuan membaca dan menulis	1. Tidak Dapat 2. Dapat
X6	morbiditas	1. Ada 2. Tidak Ada
X7	cara berobat	1. Mengobati Sendiri 2. Puskesmas 3. Rumah Sakit/dokter
X8	fasilitas air bersih	1. Tidak Ada/Umum 2. Ada
X9	fasilitas sanitasi	1. Tidak Ada/Umum 2. Ada
X10	luas lantai per kapita	1. <10m ² 2. >=1 m ²

X11	status pekerjaan utama	1. Bekerja sendiri 2. Pengusaha 3. buruh/Karyawan
X12	lapangan usaha	1. Jasa-jasa 2. lainnya
X13	sektor pekerjaan	1. Sektor informal 2. Formal

Data diperoleh dari BPS, teknik analisis dalam penelitian ini adalah

1. Mengubah data nominal/ordinal ke dalam bentuk variabel dummy
2. Analisis regresi logistik
3. Interpretasi hasil analisis regresi logistik

HASIL DAN PEMBAHASAN

Pada penelitian ini pembentukan variabel untuk variabel-variabel penjelas berupa data nominal atau ordinal yang berkategori tidak dilakukan secara manual karena program komputer (SPSS versi 11.5) yang digunakan untuk analisis ini secara otomatis akan membentuk variabel dummy. Hasil uji asumsi memperlihatkan adanya multikolinieritas antara status pekerjaan utama dengan sektor pekerjaan yaitu korelasi sebesar 0,641. Oleh karena itu salah satu variabel dapat dihilangkan untuk analisis selanjutnya.

Hasil analisis regresi logistik terbaik (metode stepwise backward) dari beberapa metode yang digunakan diperoleh bahwa nilai $G^2 = 149,201$ dengan nilai $p\text{-value} = 0,000$ lebih kecil dari $\alpha = 0,05$. Jadi secara bersama-sama variabel-variabel penjelas berikut: umur kepala rumah tangga, jumlah anggota rumah tangga, pendidikan yang ditamatkan kepala rumah tangga, morbiditas, cara berobat, fasilitas air bersih, fasilitas sanitasi, status pekerjaan utama kepala rumah tangga, dan lapangan usaha utama kepala rumah tangga berpengaruh signifikan di dalam model regresi logistik. Hasil uji Wald untuk masing-masing variabel umur kepala rumah tangga, jumlah anggota rumah tangga, pendidikan yang ditamatkan kepala rumah tangga, morbiditas, cara berobat, fasilitas air bersih, fasilitas sanitasi, status pekerjaan utama kepala rumah tangga, dan lapangan usaha utama kepala rumah tangga memiliki nilai $\text{sig} < 0,05$ artinya semua variabel berpengaruh signifikan secara parsial.

Model regresi logistik berdasarkan nilai koefisien (e) dari masing-masing variabel penjelas menunjukkan besarnya rasio variabel penjelas terhadap status kemiskinan

Sehingga diperoleh penjelasan bahwa nilai 0,954 menunjukkan kecenderungan penduduk berumur x tahun berstatus miskin sebesar 0,954 kali dibanding umur kepala rumah tangga $(x+1)$ tahun. Angka 3,560 menunjukkan bahwa anggota rumah tangga lebih dari 4 untuk berstatus miskin adalah sebesar 3,560 kali dibanding anggota rumah tangga kurang dari 4. Kepala rumah tangga tamatan SMP ke bawah cenderung berstatus miskin sebesar 1,908 kali dibanding tamatan SMP ke atas.

Variabel Penjelas	(e)
umur kepala rumah tangga	0,954
Jumlah art (Art>4)	3,560
Pendidikan tertinggi yang ditamatkan krt (SMP Ke bawah)	1,908
Morbiditas (ada)	4,394
cara berobat (Mengobati sendiri)	8,858
fasilitas air bersih (Umum/Tidak ada)	8,147
fasilitas sanitasi (Umum/Tidak ada)	3,195
status pekerjaan utama (Bekerja sendiri)	1,963
Konstan	0,087

Penduduk memiliki morbiditas cenderung berstatus miskin 4,394 kali dibanding dengan tidak ada. Mengobati sendiri cenderung berstatus miskin sebesar 8,858 kali dibanding dengan penduduk yang berobat ke rumah sakit atau ke dokter. Penduduk yang tidak memiliki fasilitas air bersih/ fasilitas air bersih umum cenderung berstatus miskin sebesar 8,174 dibanding penduduk yang memiliki fasilitas air bersih. Penduduk yang tidak memiliki fasilitas sanitasi/ fasilitas sanitasi umum cenderung berstatus miskin sebesar 3,195 dibanding penduduk yang memiliki fasilitas sanitasi.. Penduduk dengan kepala rumah tangga bekerja sendiri cenderung berstatus miskin sebesar 1,963 bila dibandingkan dengan penduduk yang bekerja sebagai buruh/karyawan.

KESIMPULAN

1. Variabel-variabel yang mempengaruhi status kemiskinan penduduk di Kota Bengkulu tahun 2004 adalah umur kepala rumah tangga, jumlah anggota rumah tangga, pendidikan yang ditamatkan kepala rumah tangga, morbiditas, cara berobat, fasilitas air bersih, fasilitas sanitasi, status pekerjaan utama kepala rumah tangga dan lapangan usaha utama kepala rumah tangga.
2. Hasil analisis regresi logistik yang paling baik diantara yang lain yang dipakai dalam analisis ini adalah dengan metode stepwise backward.

DAFTAR PUSTAKA

- [1]. Agung, I.G.N. 2002. *Statistika: Penerapan Metode Analisis untuk Tabulasi Sempurna dan Tak Sempurna*. Jakarta: PT RajaGrafindo Persada.
- [2]. Anonim, 2004a. *Pembentukan Model Regresi Logistik*.
- [3]. Hosmer, W.D & S. Lemeshow. 1989. *Applied Logistic Regression*. New York: Jhon Wiley and Sons. <http://noc.its.ac.id/sokam/wiwiek-stat/tugas-logistik.doc>
- [4]. Santosa,P.B. & Ashari. 2005 *Analisis Statistik dengan Microsoft Excel & SPSS*. Yogyakarta: Andi.