ISBN 978-602-14413-1-2

SEMINAR NASIONAL MATEMATIKA DAN APLIKASINYA

**Peranan Matematika dan Sistem Informasi di Era Big
Data untuk Menunjang Perkembangan IPTEK di Indonesia**

Surabaya, 21 Oktober 2017

Prosiding

SNMA 2017

DEPARTEMEN MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS AIRLANGGA

REGRESI LOGISTIK UNTUK PEMODELAN INDEKS PEMBANGUNAN KESEHATAN MASYARAKAT KABUPATEN/KOTA DI PULAU KALIMANTAN

M. Fathurahman

*Jurusan Matematika, Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Mulawarman*

Jl. Barong Tongkok No. 4 Kampus Gunung Kelua, Samarinda, Kalimantan Timur

fathur@fmipa.unmul.ac.id

Abstract— Regresi logistik merupakan model regresi yang paling sering digunakan untuk pemodelan data kategorik. Pada penelitian ini dilakukan pemodelan regresi logistik dan penerapannya pada Indeks Pembangunan Kesehatan Masyarakat (IPKM) kabupaten/kota di Pulau Kalimantan tahun 2013. Metode *Maximum Likelihood Estimation* (MLE) digunakan untuk penaksiran parameter. Metode *Likelihood Ratio Test* (LRT) dan uji Wald digunakan untuk pengujian parameter. Hasil penelitian menunjukkan bahwa penaksir parameter dengan metode MLE berbentuk fungsi yang tidak eksplisit. Sehingga digunakan pendekatan numerik dengan metode *Fisher Scoring*. Berdasarkan metode LRT dan uji Wald, statistik uji untuk pengujian parameter mendekati distribusi *chi-square* dan distribusi normal standar. Berdasarkan model regresi logistik terbaik, faktor-faktor yang berpengaruh terhadap IPKM kabupaten/kota di Pulau Kalimantan tahun 2013 adalah Indeks Pembangunan Manusia (IPM), tingkat kepadatan penduduk dan persentase penduduk miskin.

Keywords— Regresi logistik, MLE, LRT, uji Wald, IPKM

I. PENDAHULUAN

Metode regresi merupakan bagian dari analisis data yang menjelaskan hubungan antara variabel respon dan satu atau lebih variabel bebas. Jika variabel respon merupakan data kategorik, baik dua kategori atau lebih, maka dapat digunakan regresi logistik. Model regresi logistik terdiri atas regresi logistik biner, regresi logistik multinomial, dan regresi logistik ordinal (Hosmer, Lemeshow, dan Sturdivant 2013).

Beberapa penelitian yang mengkaji regresi logistik diantaranya adalah Pregibon (1981) mengkaji model regresi logistik, penaksiran parameter dengan metode MLE, pendekatan numerik untuk penaksiran parameter, dan diagnosis terhadap model regresi logistik. Hosmer, Jovanovic, dan Lemeshow (1989) mengkaji model regresi logistik, penaksiran parameter model dengan metode MLE, pengujian hipotesis parameter model dengan

metode LRT dan uji Wald serta pemilihan model regresi logistik terbaik dengan metode *Akaike Information Criterion* (AIC) dan *Cp Mallows*. Yanagihara, Sekiguchi, dan Fujikoshi (2003) mengkaji model regresi logistik, pemilihan model regresi logistik terbaik dengan metode AIC, dan koreksi bias pada metode AIC. Bianco dan Martinez (2009) mengkaji model regresi logistik *robust*. Hosseini dan Morgenthaler (2011) mengkaji penaksiran parameter model regresi logistik dengan pendekatan *Least Absolute Deviation* (LAD). Pingel (2014) mengkaji matriks kovarians pada regresi logistik. Ozkale dan Arican (2015) mengkaji masalah multikolinieritas pada regresi logistik. Zhang, Jeske, dan Wong (2016) mengkaji model regresi logistik dan pengklasifikasian dengan regresi logistik untuk data longitudinal.

Penerapan regresi logistik dalam penelitian ini adalah pemodelan Indeks Pembangunan Kesehatan Masyarakat (IPKM) kabupaten/kota di pulau Kalimantan tahun 2013. IPKM merupakan indikator komposit yang menggambarkan kemajuan pembangunan kesehatan dan bermanfaat untuk menentukan peringkat provinsi dan kabupaten/kota dalam mencapai keberhasilan pembangunan kesehatan masyarakat. Selain itu IPKM dapat juga digunakan untuk menentukan prioritas daerah yang memerlukan bantuan dalam peningkatan pembangunan kesehatan (Badan Penelitian dan Pengembangan Kesehatan, 2010).

Berdasarkan hasil publikasi IPKM oleh Badan Penelitian dan Pengembangan Kesehatan Kementerian Kesehatan tahun 2013, dari 55 kabupaten/kota yang ada di Pulau Kalimantan tahun 2013, Kota Balikpapan menempati peringkat ke-1 diikuti oleh Kota Palangkaraya dan Kota Pontianak. Kota Samarinda dan Kota Banjarmasin masing-masing menempati peringkat ke-6 dan ke-7. Peringkat IPKM Kota Samarinda dan Kota Banjarmasin berada di bawah Kota Bontang dan Kabupaten Penajam Paser Utara. Kabupaten Banjar menempati peringkat ke-55 (Badan Penelitian dan Pengembangan Kesehatan, 2014). Hasil

peringkat IPKM kabupaten/kota di pulau Kalimantan tahun 2013 tersebut sangat beragam dan perlu dikaji faktor-faktor yang menyebabkan keragamannya.

II. TINJAUAN PUSTAKA

2.1 Regresi logistik

Regresi logistik merupakan metode regresi yang dapat menjelaskan hubungan antara variabel respon kategorik dan satu atau lebih variabel bebas. Hasil pengamatan variabel random respon (y) mempunyai dua kategori yaitu 0 dan 1, sehingga mengikuti distribusi Bernoulli dengan fungsi probabilitas (Hosmer dkk., 2013).

$P(Y = y) = \pi^y(1 - \pi)^{1-y}, y = 0, 1$ (1)
dimana $P(Y = 1) = \pi$ menyatakan probabilitas untuk variabel respon dengan kategori $y = 1$ dan $P(Y = 0) = 1 - \pi$ menyatakan probabilitas untuk variabel respon dengan kategori $y = 0$.

Model regresi logistik dengan k variabel bebas dapat ditulis sebagai berikut (Hosmer dkk., 2013).

$$\pi(x) = \frac{\exp(\beta^T x)}{1 + \exp(\beta^T x)} \quad (2)$$

dimana $\beta^T = [\beta_0 \beta_1 \beta_2 \dots \beta_k]$ adalah vektor parameter dan $x = [x_0 \ x_1 \ x_2 \dots x_k]^T$ adalah vektor variabel bebas, dengan $x_0 = 1$. Untuk mendapatkan model regresi logistik yang linier dalam parameter dilakukan transformasi logit terhadap persamaan (2) dan diperoleh model regresi logistik seperti persamaan (3).

$$g(x) = \ln \left[\frac{\pi(x)}{1 - \pi(x)} \right] = \beta^T x \quad (3)$$

Variabel bebas pada persamaan (3) tidak saling berkorelasi atau bebas multikolinieritas. Untuk mengetahui adanya multikolinieritas digunakan nilai *Variance Inflation Factor* (VIF). Jika nilai VIF antar variabel bebas kurang dari 10, maka tidak ada multikolinieritas (Kutner, Nachtsheim dan Neter, 2004).

Untuk mendapatkan model regresi logistik terbaik digunakan kriteria nilai *Akaike Information Criterion* (AIC) seperti persamaan (4).

$$AIC = -2\ell + 2q \quad (4)$$

dimana ℓ adalah fungsi *log-likelihood* dan q adalah banyaknya parameter. Model terbaik adalah model yang mempunyai nilai AIC terkecil (Hosmer dkk., 2013).

2.2 Indeks Pembangunan Kesehatan Masyarakat

Pengembangan model Indeks Pembangunan Kesehatan Masyarakat (IPKM) tahun 2013 bertujuan untuk memperkaya informasi indikator kesehatan sehingga dapat menghasilkan penajaman program yang harus diintervensi di kabupaten/kota. Pengembangan model IPKM 2013 juga mencakup perubahan metode penghitungan indeks. Metode ini mengelompokkan 30 indikator terpilih menjadi

tujuh kelompok indikator kesehatan yang meliputi kesehatan balita (6 indikator), kesehatan reproduksi (3 indikator), pelayanan kesehatan (5 indikator), perilaku kesehatan (5 indikator), penyakit tidak menular (6 indikator), penyakit menular (3 indikator) dan kesehatan lingkungan (2 indikator). Kemudian dihitung nilai sub indeks masing-masing. Hasil nilai sub indeks ini dapat memberikan gambaran baik buruknya kondisi kesehatan di tiap kabupaten/kota menurut masing-masing kelompok. Nilai indeks mendekati satu menunjukkan kondisi yang baik.

III. METODE PENELITIAN

3.1 Data

Data yang digunakan pada penelitian ini adalah data sekunder yang diperoleh dari Badan Pusat Statistik (BPS) Pusat, BPS Provinsi Kalimantan Barat, BPS Provinsi Kalimantan Tengah, BPS Provinsi Kalimantan Selatan, BPS Provinsi Kalimantan Timur, dan Badan Penelitian dan Pengembangan Kesehatan (Balitbangkes) Kementerian Kesehatan. Data yang diperlukan meliputi IPKM kabupaten/kota di pulau Kalimantan tahun 2013 sebagai variabel respon (y) dengan kategori 0 = IPKM < IPKM Indonesia, 1 = IPKM $\geq$ IPKM Indonesia. Selanjutnya tujuh variabel bebas yang diduga berpengaruh terhadap IPKM, yaitu Indeks Pembangunan Manusia (IPM) (x_1), tingkat kepadatan penduduk (x_2), persentase penduduk yang tamat perguruan tinggi (x_3), persentase penduduk miskin (x_4), pertumbuhan ekonomi (x_5), kontribusi terhadap jumlah Produk Domestik Regional Bruto (PDRB) (x_6) dan Tingkat Pengangguran Terbuka (TPT) (x_7).

Unit pengamatan yang digunakan dalam penelitian ini adalah kabupaten/kota di pulau Kalimantan pada tahun 2013 yang terdiri dari 46 kabupaten dan 9 kota. Sehingga jumlah keseluruhan dari unit pengamatan adalah sebanyak 55 kabupaten/kota.

3.2 Metode

Metode analisis data untuk pemodelan IPKM kabupaten/kota di pulau Kalimantan tahun 2013 dengan regresi logistik adalah sebagai berikut :

1. Melakukan analisis statistika deskriptif terhadap variabel respon dan variabel bebas.
2. Melakukan pengujian multikolinieritas pada variabel bebas.
3. Melakukan pemodelan regresi logistik dengan tahapan sebagai berikut :
 - a. Menghitung penaksiran parameter dengan metode MLE dan *Fisher Scoring*.
 - b. Melakukan pengujian serentak terhadap parameter.
 - c. Melakukan pengujian parsial terhadap masing-masing parameter.

- d. Melakukan pemilihan model terbaik dengan nilai AIC terkecil.

IV. HASIL DAN PEMBAHASAN

4.1 Penaksiran Parameter

Penaksiran terhadap parameter model regresi logistik dilakukan dengan metode *Maximum Likelihood Estimation* (MLE). Langkah awal adalah membentuk fungsi *likelihood* dan fungsi *log-likelihood*,

$$L(\beta) = \prod_{i=1}^n \pi(x_i)^{y_i} (1 - \pi(x_i))^{1-y_i} \quad (5)$$

$$\ell = \ln L(\beta) = \sum_{i=1}^n \{y_i \ln[\pi(x_i)] + (1 - y_i) \ln[1 - \pi(x_i)]\}. \quad (6)$$

Selanjutnya memaksimumkan fungsi *log-likelihood* dengan cara menentukan turunan parsial pertama fungsi *log-likelihood* terhadap parameter yang ditaksir kemudian disamakan dengan nol.

$$\frac{\partial \ell}{\partial \beta_j} = \sum_{i=1}^n y_i x_{ij} - \sum_{i=1}^n x_{ij} \pi(x_i) \quad (7)$$

Berdasarkan persamaan (7), terlihat bahwa hasil turunan parsial pertama fungsi *log-likelihood* terhadap parameter yang ditaksir menghasilkan fungsi yang tidak eksplisit. Oleh karena itu diperlukan pendekatan numerik untuk mendapatkan penaksir parameter. Salah satu pendekatan numerik yang dapat digunakan adalah metode iterasi *Fisher Scoring*. Metode ini membutuhkan turunan parsial kedua fungsi *log-likelihood* terhadap parameter yang ditaksir seperti pada persamaan (8) dan (9).

$$\frac{\partial^2 \ell}{\partial \beta_j^2} = - \sum_{i=1}^n y_i x_{ij}^2 \pi(x_i) (1 - \pi(x_i)) \quad (8)$$

$$\frac{\partial^2 \ell}{\partial \beta_j \partial \beta_l} = - \sum_{i=1}^n x_{ij} x_{il} \pi(x_i) (1 - \pi(x_i)), \quad \text{untuk } j, l = 1, 2, \dots, k. \quad (9)$$

Persamaan yang digunakan untuk mendapatkan penaksir parameter dengan metode iterasi *Fisher Scoring* adalah :

$$\hat{\beta}_{t+1} = \hat{\beta}_t + \mathbf{I}^{-1}(\hat{\beta}_t) \mathbf{D}(\hat{\beta}_t) \quad (10)$$

dimana

$\hat{\beta} = [\hat{\beta}_0 \hat{\beta}_1 \hat{\beta}_2 \dots \hat{\beta}_k]^T$ adalah parameter model regresi logistik yang ditaksir, $\mathbf{I}^{-1}(\hat{\beta})$ adalah matriks informasi Fisher,

$$\mathbf{I}^{-1}(\hat{\beta}) = -E \begin{bmatrix} \frac{\partial^2 \ell}{\partial \beta_0^2} & \frac{\partial^2 \ell}{\partial \beta_0 \partial \beta_1} & \dots & \frac{\partial^2 \ell}{\partial \beta_0 \partial \beta_k} \\ \frac{\partial^2 \ell}{\partial \beta_1^2} & \dots & \dots & \frac{\partial^2 \ell}{\partial \beta_1 \partial \beta_k} \\ \vdots & \ddots & \ddots & \vdots \\ \text{Simetris} & & & \frac{\partial^2 \ell}{\partial \beta_k^2} \end{bmatrix}$$

dan $\mathbf{D}(\hat{\beta})$ adalah vektor gradien,

$$\mathbf{D}(\hat{\beta}) = \left[\frac{\partial \ell}{\partial \beta_0} \frac{\partial \ell}{\partial \beta_1} \frac{\partial \ell}{\partial \beta_2} \dots \frac{\partial \ell}{\partial \beta_k} \right]^T.$$

Proses iterasi *Fisher Scoring* untuk parameter β akan berhenti bila terpenuhi kondisi konvergen, yaitu $\|\beta_{t+1} - \beta_t\| \leq \varepsilon$, dimana ε adalah bilangan yang sangat kecil. Hasil penaksiran yang diperoleh adalah $\hat{\beta}_{t+1}$ pada saat iterasi terakhir. Selanjutnya dari matriks Informasi Fisher dapat diperoleh penaksir untuk matriks varians kovarians untuk parameter β , yaitu

$$\text{Cov}(\hat{\beta}) = \mathbf{I}^{-1}(\hat{\beta}) \quad (11)$$

4.2 Pengujian Parameter

Pengujian parameter digunakan untuk mengetahui adanya pengaruh variabel bebas terhadap variabel respon, baik secara serentak maupun parsial. Hipotesis yang digunakan untuk uji serentak adalah :

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

$$H_1 : \text{paling tidak ada satu } \beta_j \neq 0, j = 1, 2, \dots, k.$$

Berdasarkan metode LRT, statistik uji untuk uji ini adalah (Hosmer dkk., 2013) :

$$G = 2 \left\{ \sum_{i=1}^n [y_i \ln(\hat{\pi}(x_i)) + (1 - y_i) \ln(1 - \hat{\pi}(x_i))] - [n_1 \ln(n_1) + n_0 \ln(n_0) - n \ln(n)] \right\} \quad (12)$$

dimana $n_1 = \sum_{i=1}^n y_i$ dan $n_0 = \sum_{i=1}^n (1 - y_i)$.

Daerah penolakan :

Tolak H_0 bila $G > \chi^2_{\alpha, v}$.

Setelah dilakukan uji serentak, pengujian hipotesis parameter selanjutnya adalah uji parsial. Hipotesis yang digunakan adalah :

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0$$

hipotesis uji ini untuk $j = 1, 2, \dots, k$.

Statistik uji yang digunakan untuk uji ini adalah statistik Wald (Hosmer dkk., 2013)

$$W_j = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \quad (13)$$

dimana $W_j \sim N(0,1)$ untuk ukuran sampel n yang besar dan $SE(\hat{\beta}_j)$ adalah standar *error* dari $\hat{\beta}_j$.

Daerah penolakan :

Tolak H_0 bila $|W| > Z_{\alpha/2}$.

4.3 Pemodelan IPKM

Pemodelan IPKM kabupaten/kota di pulau Kalimantan tahun 2013 diawali dengan deskripsi dari variabel penelitian. Deskripsi dari variabel

respon dan variabel bebas disajikan berturut-turut pada Tabel 1 dan Tabel 2.

TABEL 1. DESKRIPSI VARIABEL RESPON

IPKM	Frekuensi	Persentase
IPKM < IPKM Indonesia	39	70,9
IPKM $\geq$ IPKM Indonesia	16	29,1
Total	55	100

TABEL 2. DESKRIPSI VARIABEL BEBAS

Variabel	Min	Max	Mean	Std. Dev.
x_1	66,81	79,87	73,19	3,08
x_2	1,39	9037,81	393,19	1410,77
x_3	1,60	14,47	5,23	3,10
x_4	2,48	14,18	7,14	2,67
x_5	0,70	9,03	6,10	1,32
x_6	1,34	32,39	9,10	6,97
x_7	0,78	11,19	4,64	2,69

Berdasarkan Tabel 1, terlihat bahwa terdapat 39 kabupaten/kota yang mempunyai IPKM < IPKM Indonesia. Sedangkan kabupaten/kota yang mempunyai IPKM $\geq$ IPKM Indonesia sebanyak 16 kabupaten/kota (29,1%).

Berdasarkan Tabel 2, terlihat bahwa IPM terendah sebesar 66,81 dan tertinggi sebesar 79,87. Rata-rata IPM kabupaten/kota di pulau Kalimantan tahun 2013 sebesar 73,19 dengan standar deviasi sebesar 3,08. Tingkat kepadatan penduduk terendah sebesar 1,39 jiwa/km² dan tertinggi sebesar 9037,81 jiwa/km². Rata-rata tingkat kepadatan penduduk pada kabupaten/kota di pulau Kalimantan tahun 2013 sebesar 393,19 jiwa/km² dengan standar deviasi sebesar 1410,77 jiwa/km². Persentase penduduk yang tamat perguruan tinggi terendah sebesar 1,6 persen dan tertinggi sebesar 14,47 persen. Rata-rata persentase penduduk yang tamat perguruan tinggi pada kabupaten/kota di pulau Kalimantan tahun 2013 sebesar 5,23 persen dengan standar deviasi sebesar 3,10 persen. Persentase penduduk miskin terendah sebesar 2,48 persen dan tertinggi sebesar 14,18 persen. Rata-rata persentase penduduk miskin pada kabupaten/kota di pulau Kalimantan tahun 2013 sebesar 7,14 persen dengan standar deviasi sebesar 2,67 persen. Pertumbuhan ekonomi terendah sebesar 0,70 persen dan tertinggi sebesar 9,03 persen. Rata-rata pertumbuhan ekonomi kabupaten/kota di pulau Kalimantan tahun 2013 sebesar 6,10 persen dengan standar deviasi sebesar 1,32 persen. Kontribusi kabupaten/kota terhadap jumlah PDRB terendah sebesar 1,34 persen dan tertinggi sebesar 32,39 persen. Rata-rata kontribusi terhadap jumlah PDRB untuk kabupaten/kota di pulau Kalimantan tahun 2013

sebesar 9,10 persen dengan standar deviasi sebesar 6,97 persen. Tingkat Pengangguran Terbuka (TPT) terendah sebesar 0,78 persen dan tertinggi sebesar 11,19 persen. Rata-rata TPT kabupaten/kota di pulau Kalimantan tahun 2013 sebesar 4,64 persen dengan standar deviasi sebesar 2,69 persen.

Untuk selanjutnya sebagai prasyarat analisis data untuk model regresi logistik dilakukan uji multikolinieritas. Untuk mengetahui adanya kasus multikolinieritas adalah dengan mempertimbangkan nilai VIF. Hasil nilai VIF untuk masing-masing variabel bebas disajikan pada Tabel 3.

TABEL 3. NILAI VIF VARIABEL BEBAS

Variabel	Nilai VIF
x_1	3,322
x_2	1,669
x_3	3,205
x_4	1,310
x_5	1,273
x_6	1,694
x_7	1,441

Terlihat pada Tabel 3, bahwa nilai VIF terbesar ditunjukkan oleh x_1 sebesar 3,322, sedangkan nilai VIF terkecil ditunjukkan oleh x_5 sebesar 1,273. Secara keseluruhan nilai VIF untuk masing-masing variabel bebas kurang dari 10. Hal ini mengindikasikan bahwa tidak ada kasus multikolinieritas antara variabel bebas, sehingga semua variabel bebas bisa dimasukkan dalam model.

Pemodelan IPKM kabupaten/kota di pulau Kalimantan, dimulai dengan menentukan taksiran parameter model regresi logistik yang dilakukan dengan metode MLE dan iterasi *Fisher Scoring* sebagaimana disajikan pada persamaan (10). Hasil penaksiran parameter, standar error dan nilai statistik uji Z yang diperoleh berdasarkan perhitungan melalui *software R*, disajikan pada Tabel 4.

TABEL 4. TAKSIRAN PARAMETER DAN NILAI STATISTIK UJI PARSIAL MODEL REGRESI LOGISTIK

Parameter	Est	Se	Z
β_0	-55,914	33,179	-1,685
β_1	0,758	0,463	1,637
β_2	0,012	0,007	1,664
β_3	-0,072	0,335	-0,217
β_4	0,431	0,247	1,743
β_5	-0,609	0,352	-1,730
β_6	-0,063	0,086	-0,733

Parameter	Est	Se	Z
β_7	-0,140	0,238	-0,591

Pengujian parameter secara serentak pada model regresi logistik dilakukan dengan metode LRT yang diformulasikan sebagai berikut :

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_7 = 0$$

$$H_1 : \text{paling tidak ada satu } \beta_j \neq 0, j = 1, 2, \dots, 7.$$

Nilai statistik uji serentak yang diperoleh berdasarkan perhitungan dengan *software* R adalah $G = 32,662$ yang lebih besar dari $\chi^2_{(0,1;7)} = 12,017$, sehingga tolak H_0 . Hal itu berarti, paling sedikit ada satu variabel bebas yang berpengaruh signifikan terhadap variabel respon.

Untuk mengetahui variabel bebas mana yang berpengaruh signifikan terhadap variabel respon, dilakukan pengujian parameter secara parsial. Untuk menguji secara parsial parameter β_j ($j = 1, 2, 3, \dots, 7$) diformulasikan sebagai berikut :

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0$$

Nilai penaksiran parameter dan statistik uji Z yang disajikan pada Tabel 4, terlihat bahwa beberapa variabel bebas mempunyai nilai $|W_j| > Z_{\alpha/2} = 1,645$ yang menunjukkan bahwa variabel tersebut berpengaruh signifikan terhadap variabel respon. Variabel-variabel tersebut adalah x_2, x_4 , dan x_5 .

Selanjutnya dilakukan pemilihan model regresi logistik terbaik, sehingga model yang terbentuk memuat variabel bebas yang minimal tetapi memberikan kontribusi yang optimal terhadap variabel respon. Untuk pemilihan model regresi logistik terbaik digunakan metode *Best Subsets Regression* dengan mempertimbangkan nilai AIC. Dua nilai AIC terkecil untuk semua kombinasi variabel bebas yang diperoleh berdasarkan perhitungan dengan *software* R, disajikan pada Tabel 5.

TABEL 5. NILAI AIC DUA TERKECIL UNTUK KOMBINASI VARIABEL BEBAS MODEL REGRESI LOGISTIK

Banyak Var Bebas	Variabel Bebas dalam Model	AIC
7	$x_1, x_2, x_3, x_4, x_5, x_6, x_7$	48,662
6	$x_1, x_2, x_3, x_4, x_5, x_6$	47,022
	$x_1, x_2, x_3, x_4, x_5, x_7$	47,238
5	x_1, x_2, x_3, x_4, x_5	45,875
	x_1, x_2, x_3, x_4, x_7	48,332
4	x_1, x_2, x_3, x_4	46,456
	x_1, x_2, x_3, x_5	46,378
3	x_1, x_2, x_4	44,558 ^{*)}
	x_2, x_3, x_5	46,692
2	x_1, x_2	45,666
	x_2, x_3	48,085
1	x_2	50,439

Banyak Var Bebas	Variabel Bebas dalam Model	AIC
	x_3	52,879

^{*)} Nilai AIC terkecil

Berdasarkan Tabel 5, terlihat bahwa model regresi logistik terbaik dengan nilai AIC terkecil adalah model dengan 3 variabel bebas, yaitu x_1, x_2 dan x_5 dengan nilai AIC sebesar 44,539.

Untuk selanjutnya model regresi logistik terbaik yang telah diperoleh, ditentukan kembali nilai taksiran parameter. Hasil penaksiran parameter, standar *error*, nilai statistik uji Z dan *odds ratio* (OR) yang diperoleh berdasarkan perhitungan melalui *software* R, disajikan pada Tabel 6.

TABEL 6. TAKSIRAN PARAMETER DAN NILAI STATISTIK UJI PARSIAL MODEL REGRESI LOGISTIK TERBAIK

Parameter	Est	Se	Z	OR
β_0	-39,228	16,495	-2,378 ^{*)}	$9,191 \times 10^{-18}$
β_1	0,474	0,215	2,208 ^{*)}	1,607
β_2	0,010	0,006	1,922 ^{*)}	1,011
β_4	0,315	0,184	1,714 ^{*)}	1,371

^{*)} Signifikan pada $\alpha = 0,1$.

Berdasarkan Tabel 6, diperoleh model regresi logistik terbaik untuk pemodelan IPKM kabupaten/kota di pulau Kalimantan tahun 2013 seperti persamaan (14).

$$g(x) = -39,228 + 0,474 x_1 + 0,010 x_2 + 0,315 x_5 \quad (14)$$

Dari Tabel 6, terlihat bahwa faktor-faktor yang signifikan berpengaruh terhadap IPKM kabupaten/kota di pulau Kalimantan tahun 2013 adalah IPM, tingkat kepadatan penduduk dan persentase penduduk miskin. Interpretasi dari model regresi logistik terbaik pada persamaan (14) adalah untuk setiap peningkatan IPM pada kabupaten/kota sebesar 1 poin, maka IPKM kabupaten/kota lebih dari atau sama dengan IPKM Indonesia akan meningkat menjadi 1,607. Untuk setiap peningkatan tingkat kepadatan penduduk pada kabupaten/kota sebesar 1 jiwa/km², maka IPKM kabupaten/kota lebih dari atau sama dengan IPKM Indonesia akan meningkat menjadi 1,011. Sedangkan untuk setiap peningkatan persentase penduduk miskin sebesar 1 persen pada kabupaten/kota, maka IPKM kabupaten/kota lebih dari atau sama dengan IPKM Indonesia akan meningkat menjadi 0,729.

Nilai statistik uji serentak model regresi logistik terbaik pada persamaan (14) yang diperoleh berdasarkan perhitungan dengan *software* R adalah $G = 36,539$ lebih dari $\chi^2_{(0,1;3)} = 6,251$, sehingga tolak H_0 . Hal ini menunjukkan bahwa paling sedikit ada satu

variabel bebas yang berpengaruh signifikan terhadap variabel respon.

V. KESIMPULAN DAN SARAN

Berdasarkan hasil dan pembahasan, diperoleh kesimpulan sebagai berikut :

1. Penaksir parameter model regresi logistik yang diperoleh dengan metode MLE berbentuk fungsi yang tidak eksplisit. Pendekatan numerik dengan metode iterasi *Fisher Scoring* dapat digunakan untuk mendapatkan penaksir parameter.
2. Statistik uji untuk pengujian hipotesis parameter model regresi logistik secara serentak dan parsial berdasarkan metode LRT dan uji Wald berturut-turut mendekati distribusi *chi-square* dan distribusi normal standar.
3. Berdasarkan model regresi logistik terbaik, faktor-faktor yang signifikan berpengaruh terhadap IPKM kabupaten/kota di pulau Kalimantan tahun 2013 adalah Indeks Pembangunan Manusia (IPM), tingkat kepadatan penduduk dan persentase penduduk miskin.

Saran yang diajukan dari penelitian ini diantaranya adalah untuk penelitian selanjutnya dapat menggunakan pendekatan numerik yang lain untuk mendapatkan penaksir MLE parameter model regresi logistik, yaitu dengan metode Quasi-Newton menggunakan algoritma Broyden-Fletcher-Goldfarb-Shanno (BFGS) dan metode Nelder-Mead. Untuk mendapatkan faktor-faktor lain yang berpengaruh terhadap IPKM, penelitian selanjutnya dapat menambahkan variabel-variabel bebas lain yang diduga berpengaruh terhadap IPKM.

DAFTAR PUSTAKA

- Badan Penelitian dan Pengembangan Kesehatan Kementerian Kesehatan. 2010, *Pedoman Umum Penanggulangan Daerah Bermasalah Kesehatan*, Jakarta.
- Badan Penelitian dan Pengembangan Kesehatan Kementerian Kesehatan. 2014, *Indeks Pembangunan Kesehatan Masyarakat*, Jakarta.
- Bianco, A. M., dan Martinez, E., 2009, "*Robust Testing in the Logistic Regression Model*", Computational Statistics and Data Analysis, ISSN 0167-9473 Vol. 53, p.4095-4105.
- Hosmer, D. W., Jovanovic, B. dan Lemeshow, S., 1989, "*Best Subsets Logistic Regression*", Biometrics, ISSN 1541-0420 Vol. 45, No. 4, p.1265-1270.
- Hosmer, D. W., Lemeshow, S. dan Sturdivant, R. X. 2013, *Applied Logistic Regression*, Third Editon, John Wiley & Sons, Inc., Hoboken, New Jersey.
- Hosseinian, S. dan Morgenthaler, S. 2011, "*Robust Binary Regression*", Journal of Statistical Planning and Inference, ISSN 0378-3758 Vol. 141, p.1497-1509.
- Kutner, M. H., Nachtsheim, C. J. dan Neter, J. 2004, *Applied Linear Regression Models*, Fourth Editon, McGraw-Hill/Irwin, New York.
- Ozkale, M. R. dan Arican, E., 2016, "*First-order r-d Class Estimator in Binary Logistic Regression Model*", Statistics & Probability Letters, ISSN 0167-7152 Vol. 106, p.19-29.
- Pingel, R., 2014, "*Some Approximations of the Logistic Distribution with Application to the Covariance Matrix of Logistic Regression*", Statistics & Probability Letters, ISSN 0167-7152 Vol. 85, p.63-68.
- Pregibon, D., 1981, "*Logistic Regression Diagnostics*", The Annals of Statistics, ISSN 2168-8966 Vol. 9, No. 4, p.705-724.
- Zhang, X., Jeske, D. R., Li, J. dan Wong, V., 2016, "*A Sequential Logistic Regression Classifier Based on Mixed Effects with Applications to Longitudinal Data*", Computational Statistics and Data Analysis, ISSN 0167-9473 Vol. 94, p.238-249.