

ESTIMASI PARAMETER MODEL REGRESI LOGISTIK BINER MENGUNAKAN METODE JACKKNIFE

Rizal Ariefaidzin Asikin¹, Yudi Setyawan^{*2}

^{1,2}Jurusan Statistika, Fakultas Sains Terapan, Institut Sains dan Teknologi AKPRIND, Yogyakarta

Email : rariefaidzin@gmail.com¹, setyawan@akprind.ac.id²

Abstract. Logistic regression is one of the statistical methods used to identify a causal relationship (causality) between one dependent variable and one or more independent variables, where the dependent variable has 2 or more categories. To find a good logistic regression model, sufficient amount of data is needed, if the amount of data is insufficient, the modeling can be done using the resampling method, namely using Bootstrap or Jackknife. The working principle of the Jackknife method use computers in generating data from the original samples to generate artificial samples. An artificial sample is generated by removing an observation from the original sample which can then be used to calculate estimator value. In this case, the Jackknife Method is used to resampling the data of beras miskin admission in Yogyakarta. Based on the results of the analysis, Jackknife method can minimize the standard error in calculating the estimation of parameters until Jackknife Method erases two datas. Based on data obtained from SUSENAS (Sensus Sosial Ekonomi Nasional) of Yogyakarta in 2016, there are 641 households being sample. The dependent variable which is used was the admission of beras miskin. Beras miskin consists of two categories. That is receiving beras miskin and not-receiving beras miskin. From the number of households affected by sample, it shows that the number of households receiving beras miskin is 15% (98 households) and the number of households which isn't receiving beras miskin is 85% (543 households). The significant variables after logistic regression test were the main floor of house (X_2) and the source of drinking water (X_4). Those variables affect admission of beras miskin with the odds ratio $e^{-1,3959X_2}=0,2476$ and $e^{-1,6095X_4}=0,1999$. The best logistic regression model for case studies discussed in the previous chapter is a logistic regression model with the two datas erased by jackknife method, in which the probability model of a household receiving raskin is:

$$P(Y = 1|X) = \pi(x) = \frac{e^{-0,1494-1,3959X_2(1)-1,6095X_4(1)}}{1+e^{-0,1494-1,3959X_2(1)-1,6095X_4(1)}}$$

With the level of accuracy of the model in predicting at 84.2%

Keywords: Logistic Regression, Jackknife Method, Parameter Estimation, Raskin.

Abstrak. Regresi logistik merupakan salah satu metode statistika yang digunakan untuk mengetahui suatu hubungan sebab akibat (kausalitas) antara satu variabel dependen dengan satu atau lebih variabel independen, dimana variabel dependen memiliki 2 kategori atau lebih. Untuk mengetahui model regresi logistik yang baik diperlukan jumlah data yang mencukupi, apabila jumlah data tidak mencukupi penyusunan model dapat dilakukan dengan metode resampling yaitu menggunakan Bootstrap atau Jackknife. Prinsip kerja metode Jackknife adalah menggunakan komputer dalam membangkitkan data dari sampel asli untuk menghasilkan sampel tiruan yang dihasilkan dengan cara menghapus satu atau lebih observasi dari sampel asli yang selanjutnya dapat digunakan untuk menghitung nilai estimator. Metode Jackknife digunakan untuk melakukan *resampling* data penerimaan beras miskin. Berdasarkan hasil analisis, metode Jackknife dapat memperkecil standar error dalam menghitung estimasi parameter sampai dengan Jackknife terhapus-2. Berdasarkan data yang diperoleh dari SUSENAS (Sensus Sosial Ekonomi Nasional) Yogyakarta pada tahun 2016, ada 641 rumah tangga yang menjadi sampel. Variabel dependen yang digunakan adalah penerimaan beras miskin. Beras miskin terdiri dari dua kategori. Yaitu menerima beras miskin dan tidak menerima beras miskin. Dari jumlah rumah tangga yang terkena sampel, terlihat bahwa jumlah rumah tangga yang menerima beras miskin adalah 15% (98 rumah tangga) dan jumlah rumah tangga yang tidak menerima beras miskin adalah 85% (543 rumah tangga). Variabel yang signifikan setelah uji regresi logistik adalah lantai utama rumah (X_2) dan sumber air minum (X_4). Variabel-variabel tersebut mempengaruhi penerimaan beras miskin dengan odds ratio masing-masing $e^{-1,3959} = 0,2476$ dan $e^{-1,6095} = 0,1999$. Model regresi logistik terbaik untuk studi kasus yang dibahas dalam bab sebelumnya adalah model regresi logistik dengan dua data yang dihapus dengan metode Jackknife, dimana model probabilitas dari rumah tangga yang menerima raskin adalah:

*Corresponding author's email: : setyawan@akprind.ac.id

$$P(Y = 1|X) = \pi(x) = \frac{e^{-0,1494-1,3959X_2(1)-1,6095X_4(1)}}{1+e^{-0,1494-1,3959X_2(1)-1,6095X_4(1)}}$$

Dengan tingkat ketelitian model dalam memprediksi sebesar 84,2%

Kata kunci: Regresi Logistik, Metode Jackknife, Estimasi Parameter, Raskin.

1. PENDAHULUAN

Metode statistika saat ini mengalami perkembangan yang pesat, diikuti dengan arus berbagai permasalahan dunia yang selalu bermunculan dan menuntut manusia untuk melakukan penelitian agar permasalahan yang ada dapat dipecahkan. Hal tersebut memberikan motivasi yang berarti dalam mengembangkan ilmu pengetahuan, guna memberikan solusi-solusi nyata terhadap masalah-masalah yang ada dalam kehidupan sehari-hari. Banyak permasalahan real yang ada dalam kehidupan yang bisa dianalisis dan didekati dengan permodelan statistika. Karena itu statistika sebagai salah satu disiplin ilmu yang telah diterapkan didalam berbagai bidang mulai dari bidang kesehatan, sosial, ekonomi, kemasyarakatan dan berbagai bidang lainnya. Dalam bidang sosial, aplikasi statistika banyak digunakan untuk menyelesaikan permasalahan yang ada didalamnya, salah satunya pada kasus penerimaan beras miskin (raskin).

Metode jackknife merupakan salah satu metode untuk melakukan estimasi parameter pada regresi logistik biner dengan cara menghapus suatu observasi dari sampel asli, dan sampel yang didapat digunakan untuk menghitung nilai estimator. Model Jackknife juga cukup populer dalam menyelesaikan masalah estimasi parameter dengan tingkat akurasi yang baik. Selain itu metode jackknife dapat dijadikan salah satu metode *resampling* yang digunakan untuk memecahkan masalah pada suatu data (Rianto, 2015). Banyak sekali metode *resampling* yang telah ditemukan oleh para peneliti terkemuka, yang populer dan mudah dalam penerapannya adalah metode Jackknife. Menurut Shao dan Tu dalam Fitianingrum (2013) Metode Jackknife merupakan teknik *resampling* nonparametrik yang bertujuan untuk menentukan estimasi bias, standar error dan interval konfidensi dari parameter populasi seperti mean, rasio, median, proporsi, koefisien korelasi/koefisien regresi tanpa menggunakan asumsi distribusi. Prosedur dari metode jackknife, yaitu dengan cara mengestimasi nilai variansi metode jackknife yang kemudian dapat digunakan untuk menghitung standar error. Nilai standar error tersebut merupakan akar dari nilai variansinya.

2. METODE

Pada penelitian ini menggunakan data sekunder, yaitu data yang diperoleh dari pihak lain dengan cara mengambil data Survei Sosial Ekonomi Nasional (SUSENAS) tahun 2016 dari BPS (Badan Pusat Statistik) Provinsi D.I Yogyakarta. Didapat data dengan jumlah 641 rumah tangga. Susenas merupakan survey yang dirancang untuk mengumpulkan data sosial kependudukan yang relatif sangat luas dan dilaksanakan setiap tahun. Data yang dikumpulkan antara lain menyangkut bidang-bidang pendidikan, kesehatan/gizi, perumahan, sosial ekonomi lainnya, kegiatan sosial budaya, konsumsi/pengeluaran dan pendapat masyarakat mengenai kesejahteraan rumah tangganya. Variabel dependen merupakan variabel yang dipengaruhi oleh variabel lain yaitu variabel independen. Variabel dependen yang dimaksud dalam penelitian ini adalah penerimaan raskin dengan pengkategorian Y untuk menerima dan T untuk tidak menerima. Variabel independen merupakan variabel yang mempengaruhi atau yang menjadi penyebab bagi variabel dependen. Variabel yang digunakan dalam penelitian ini adalah: Jenis dinding tempat tinggal (X_1) Bahan utama lantai rumah (X_2) Fasilitas buang air besar (X_3) dan Sumber air minum (X_4).

3. HASIL DAN PEMBAHASAN

Berdasarkan data telah didapat dari SUSENAS (Sensus Sosial Ekonomi Nasional) Kota Yogyakarta pada tahun 2016 (Lampiran 1), dengan jumlah data keseluruhan 641 rumah tangga

yang terkena sampel. Variabel terikat yang digunakan adalah penerimaan raskin yang terdiri dari dua kategori yaitu menerima raskin dan tidak menerima raskin.



Gambar 1 Presentase Penerimaan Raskin di Kota Yogyakarta Tahun 2016

Pada Gambar 1 dari keseluruhan rumah tangga yang terkena sampel terlihat bahwa jumlah rumah tangga yang menerima raskin sebanyak 15% (98 rumah tangga). Sedangkan rumah tangga yang tidak menerima raskin sebanyak 85% (543 rumah tangga). Dari data tersebut dapat dilihat bahwa pendistribusian raskin kepada rumah tangga masih minim dari sampel yang ada, maka perlu dilakukan *resampling*.

Tabel 1 Tabulasi Silang Jenis Dinding Tempat Tinggal dengan Penerimaan Raskin

Jenis Dinding Tempat Tinggal (X_1)	Penerimaan (Y)		Jumlah
	Tidak menerima	Menerima	
Tembok	257	51	308
Bukan Tembok	286	47	333
Jumlah	543	98	641

Sumber :data sekunder yang diolah, 2017

Pada Tabel 1 di atas menunjukkan bahwa mayoritas rumah tangga memiliki jenis dinding tempat tinggal adalah bukan tembok. Dari 333 rumah tangga yang memiliki jenis dinding tempat tinggal bukan tembok sebanyak 286 rumah tangga tidak menerima raskin atau sebesar 85,9%, sedangkan sebanyak 47 rumah tangga yang memiliki jenis dinding tempat tinggal bukan tembok dan menerima raskin atau sebesar 14,1 %.

Tabel 2 Tabulasi Silang Bahan Utama Lantai Rumah dengan Penerimaan Raskin

Bahan Utama Lantai Rumah (X_2)	Penerimaan (Y)		Jumlah
	Tidak menerima	Menerima	
Keramik	447	50	497
Bukan Keramik	96	48	144
Jumlah	543	98	641

Sumber :data sekunder yang diolah, 2017

Pada Tabel 2 di atas menunjukkan bahwa rumah tangga memiliki bahan utama lantai rumah bukan keramik sebanyak 144 rumah tangga. Dengan 96 (66,7%) rumah tangga tidak menerima raskin dan memiliki bahan utama lantai rumah bukan keramik, sedangkan rumah tangga yang memiliki bahan utama lantai rumah bukan keramik dan menerima raskin sebanyak 48 (33,3%) rumah tangga.

Tabel 3 Tabulasi Silang Fasilitas Buang Air Besar dengan Penerimaan Raskin

Fasilitas Buang Air Besar (X ₃)	Penerimaan (Y)		Jumlah
	Tidak menerima	Menerima	
Sendiri	339	57	396
Bersama	204	41	245
Jumlah	543	98	641

Sumber :data sekunder yang diolah, 2017

Pada Tabel 3 di atas menunjukkan bahwa rumah tangga memiliki fasilitas buang air besar bersama sebanyak 245 rumah tangga. Dengan 204 (83,3%) rumah tangga tidak menerima raskin dan memiliki fasilitas buang air besar bersama, sedangkan rumah tangga yang memiliki fasilitas buang air besar bersama dan menerima raskin sebanyak 41 (16,7%) rumah tangga.

Tabel 4 Tabulasi Silang Sumber Air Minum dengan Penerimaan Raskin

Sumber Air Minum (X ₄)	Penerimaan (Y)		Jumlah
	Tidak menerima	Menerima	
Berbayar rutin	316	20	336
Tidak membayar rutin	227	78	305
Jumlah	543	98	641

Sumber: data sekunder yang diolah, 2017

Pada Tabel 4 di atas menunjukkan bahwa rumah tangga memiliki sumber air minum tidak membayar rutin sebanyak 305 rumah tangga. Dengan 227 (74,4%) rumah tangga tidak menerima raskin dan memiliki sumber air minum tidak membayar rutin, sedangkan rumah tangga yang memiliki sumber air minum tidak membayar rutin dan menerima raskin sebanyak 78 (25,6%) rumah tangga.

Setelah dilakukan analisis deskriptif ,maka dilakukan analisis regresi logistik biner. Analisis ini digunakan untuk mengetahui pengaruh variabel bebas terhadap variabel respon. Berikut adalah hasil analisis regresi logistik biner berdasarkan sampel asli dengan jumlah sampel 641 menggunakan *software R*. Maka diperoleh:

Tabel 5 Hasil Analisis Regresi Logistik Biner

Prediktor	$\hat{\beta}$	SE	P-value
Konstan	0,0366	0,2598	0,8878
X ₁ (1)	0,1955	0,2739	0,4753
X ₂ (1)	-1,3816	0,2451	0,0000
X ₃ (1)	-0,4619	0,2747	0,0927
X ₄ (1)	-1,6363	0,2800	0,0000

Sumber :data sekunder yang diolah, 2017

Model regresi logistik yang terbentuk adalah :

$$\pi(x) = \frac{e^{0,0366+0,1955X_{1(1)}-1,3816X_{2(1)}-0,4619X_{3(1)}-1,6363X_{4(1)}}}{1 + e^{0,0366+0,1955X_{1(1)}-1,3816X_{2(1)}-0,4619X_{3(1)}-1,6363X_{4(1)}}}$$

Dimana:

- X_1 = Jenis dinding tempat tinggal
 X_2 = Bahan utama lantai rumah
 X_3 = Fasilitas buang air besar
 X_4 = Sumber air minum

• Uji Simultan

Dilihat dari hasil analisis di atas dilakukan uji signifikansi model, dengan hipotesis :

$$H_0 = \beta_0 = \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$$

(Semua parameter secara simultan sama dengan nol)

$$H_1 = \beta_k \neq 0$$

(Minimal ada satu parameter secara simultan sama dengan nol)

Tingkat signifikansi (α) = 5% (0,05)

Adapun kriteria pengambilan keputusannya yaitu :

Jika G atau $-2 \text{ Log Likelihood} > \chi^2$ tabel atau $p\text{-value} < \alpha$, maka H_0 ditolak

maka H_0 ditolak Jika G atau $-2 \text{ Log Likelihood} < \chi^2$ tabel atau $p\text{-value} > \alpha$ maka H_0 tidak ditolak

Pada Lampiran 3 nilai $-2 \text{ Log Likelihood} = 462,98$ dan $\chi^2_{0,05}$ tabel = 7,815 sehingga $-2 \text{ Log Likelihood} < \chi^2_{0,05}$, maka H_0 ditolak, bisa dikatakan uji signifikansi model minimal ada satu parameter secara simultan sama dengan nol.

• Uji Parsial

Hipotesis :

$H_0 : \beta_j = 0$ (tidak ada pengaruh antara variabel independen ke- j dengan variabel dependen).

$H_1 : \beta_j \neq 0$ (ada pengaruh antara variabel independen ke- j dengan variabel dependen).

Tingkat signifikansi (α) = 5% (0,05)

Adapun kriteria pengambilan keputusannya yaitu :

Jika $W > \chi^2_{(df;\alpha)}$ atau $p\text{-value} < \alpha$, maka H_0 ditolak

Jika $W < \chi^2_{(df;\alpha)}$ atau $p\text{-value} > \alpha$, maka H_0 tidak ditolak

Kemudian berdasarkan Tabel 5, Dapat dilihat variabel mana yang signifikan dan variabel mana yang tidak signifikan, pada kasus ini ada 2 variabel yang signifikan dan 2 variabel yang tidak signifikan. Untuk variabel yang signifikan, yaitu variabel bahan utama lantai rumah (X_2) dan variabel sumber air minum (X_4) terhadap penerimaan raskin dengan nilai koefisien pengaruh sebesar 1,3816 dan 1,6363, sedangkan nilai $p\text{-value}$ signifikansi variabel yang tidak signifikan yaitu variabel jenis dinding tempat tinggal (X_1) dan variabel fasilitas buang air besar (X_3) $> 0,05$ maka tidak ditolak H_0 dengan kata lain tidak terdapat pengaruh yang signifikan terhadap penerimaan raskin.

Dari data tersebut di atas maka dilakukan uji kembali dengan mengeliminasi variabel yang tidak signifikan dan memiliki nilai besar yaitu variabel jenis dinding tempat tinggal (X_1), maka diuji kembali antara variabel penerimaan raskin (Y) dengan variabel bahan utama lantai rumah (X_2), variabel fasilitas buang air besar (X_3) dan variabel sumber air minum (X_4). dengan hasil sebagai berikut :

Tabel 6 Hasil Analisis Regresi Logistik Biner

Prediktor	$\hat{\beta}$	SE	$P\text{-value}$
Konstan	0,0837	0,2516	0,739
$X_2(1)$	-1,3633	0,2433	0,0000
$X_3(1)$	-0,3776	0,2472	0,127
$X_4(1)$	-1,6723	0,2758	0,0000

Sumber : data sekunder yang diolah, 2017

Berdasarkan Tabel 6 di atas tolak hipotesis nol (H_0) jika nilai p-value signifikansi < 0.05 dapat dilihat variabel yang signifikan adalah variabel bahan lantai utama rumah (X_2) dan variabel sumber air minum (X_4) terhadap penerimaan raskin dengan nilai koefisien pengaruh sebesar 1,3633 dan 1,6723, sedangkan variabel yang tidak signifikan yaitu variabel dan variabel fasilitas buang air besar (X_3) > 0.05 tidak ditolak H_0 dengan kata lain tidak terdapat pengaruh yang signifikan terhadap penerimaan raskin. Kemudian dilakukan uji kembali dengan mengeliminasi variabel yang tidak signifikan yaitu variabel fasilitas buang air besar (X_3), dengan hasil sebagai berikut :

Tabel 7 Hasil Analisis Regresi Logistik Biner

Prediktor	$\hat{\beta}$	SE	P-value
Konstan	-0,1494	0,2000	0,455
$X_2(1)$	-1,3959	0,2418	0,0000
$X_4(1)$	-1,6094	0,2713	0,0000

Sumber :data sekunder yang diolah, 2017

Berdasarkan Tabel 7 di atas dapat dilihat bahwa semua variabel yang diuji signifikan, maka dapat model yang terbentuk yaitu :

$$\pi(x) = \frac{e^{-0.1494-1.3959X_{2(1)}-1.6094X_{4(1)}}}{1 + e^{-0.1494-1.3959X_{2(1)}-1.6094X_{4(1)}}}$$

Setelah mengetahui hasil regresi logistik biner berdasarkan sampel asli selanjutnya estimasi parameter berdasarkan metode Jackknife terhapus-1. Untuk metode Jackknife terhapus-1 (dengan ukuran sampel $n-1 = 640$ dan $N = \binom{641}{1} = 641$ kali *resampling*) berikut:

Tabel 8 Hasil untuk metode Jackknife terhapus-1

r	Variabel	1	2	...	639	640	$\hat{\beta}_0^{Jk}$	$\hat{\beta}_2^{Jk}$	$\hat{\beta}_4^{Jk}$
1	Y		T	...	Y	Y	-0,149	-1,392	-1,613
	X_2		Y	...	Y	Y			
	X_4		T	...	T	T			
2	Y	T		...	Y	Y	-0,149	-1,392	-1,613
	X_2	Y		...	Y	Y			
	X_4	T		...	T	T			
.
640	Y	T	T	...	Y		-0,153	-1,415	-1,594
	X_2	Y	Y	...	Y				
	X_4	T	T	...	T				
641	Y	T		...	Y	Y	-0,135	-1,419	-1,663
	X_2	Y		...	Y	Y			
	X_4	T		...	T	T			
	$\hat{\beta}^{(J)}$						-0,149	-1,396	-1,609

Berikut adalah hasil analisis regresi logistik biner berdasarkan metode Jackknife terhapus-1 menggunakan *software R* :

Tabel 9 Estimasi menggunakan Jackknife terhapus-1

Prediktor	$\hat{\beta}$	SE
Konstan	-0,1494	3,156 x 10 ⁻⁴
X ₂ (1)	-1,3959	3,785 x 10 ⁻⁴
X ₄ (1)	-1,6095	4,350 x 10 ⁻⁴

Sumber: data sekunder yang diolah, 2017

Berdasarkan Tabel 9 maka model regresi logistik yang terbentuk yaitu:

$$\pi(x) = \frac{e^{-0,1494-1,3959X_2(1)-1,6095X_4(1)}}{1 + e^{-0,1494-1,3959X_2(1)-1,6095X_4(1)}}$$

Dimana:

X₂ = Bahan utama lantai rumah

X₄ = Sumber air minum

Setelah mengetahui hasil regresi logistik biner berdasarkan metode Jackknife-1 selanjutnya estimasi parameter berdasarkan metode Jackknife terhapus-2. Untuk metode Jackknife terhapus-2 (dengan ukuran sampel $n-2 = 639$ dan $\binom{641}{2} = 205,120$ kali *resampling*) sebagai berikut:

Tabel 10 Hasil untuk Metode Jackknife terhapus-2

R	Variabel	1	2	3	...	638	639	$\hat{\beta}_0^{Jk}$	$\hat{\beta}_2^{Jk}$	$\hat{\beta}_4^{Jk}$
1	Y			T	...	Y	Y	-0,148	-1,388	-1,616
	X ₂			Y	...	Y	Y			
	X ₄			T	...	T	T			
2	Y		T		...	Y	Y	-0,148	-1,388	-1,616
	X ₂		Y		...	Y	Y			
	X ₄		T		...	T	T			
.
205119	Y	T		T	...	Y		-0,139	-1,439	-1,648
	X ₂	Y		Y	...	Y				
	X ₄	T		T	...	T				
205120	Y	T	T	T	...			-0,139	-1,439	-1,648
	X ₂	Y	Y	Y	...					
	X ₄	T	T	T	...					
	$\hat{\beta}^{(j)}$							-0,149	-1,396	-1,609

Berikut adalah hasil analisis regresi logistik biner berdasarkan metode Jackknife terhapus-2 menggunakan *software R* :

Tabel 11 Estimasi menggunakan Jackknife terhapus-2

Prediktor	$\hat{\beta}$	SE
Konstan	-0,1494	2,4949x10 ⁻⁵
X ₂ (1)	-1,3959	2,9923x10 ⁻⁵
X ₄ (1)	-1,6095	3,4396x10 ⁻⁵

Sumber: data sekunder yang diolah, 2017

Berdasarkan Tabel 11 maka model regresi logistik yang terbentuk yaitu:

$$\pi(x) = \frac{e^{-0,1494-1,3959X_2(1)-1,6095X_4(1)}}{1 + e^{-0,1494-1,3959X_2(1)-1,6095X_4(1)}}$$

Dimana:

X_2 = Bahan utama lantai rumah

X_4 = Sumber air minum

Setelah melakukan analisis estimasi menggunakan MLE dengan menggunakan sampel data asli dan menggunakan Metode Jackknife, diperoleh nilai perbandingan standar error untuk masing-masing metode adalah sebagai berikut:

Tabel 12 Perbandingan estimasi standar error

Metode	konstan	$X_2(1)$	$X_4(1)$
Sampel asli	0,2000	0,2418	0,2713
Jackknife terhapus-1	$3,156 \times 10^{-4}$	$3,785 \times 10^{-4}$	$4,350 \times 10^{-4}$
Jackknife terhapus-2	$2,4949 \times 10^{-5}$	$2,9923 \times 10^{-5}$	$3,4396 \times 10^{-5}$

Sumber: data sekunder yang diolah, 2017

Berdasarkan Tabel 12 nilai estimasi standar error di atas, terlihat bahwa metode Jackknife terhapus-2 memiliki nilai standar error yang paling kecil, sehingga metode Jackknife terhapus-2 paling baik digunakan. Dengan model yang terbentuk adalah :

$$\pi(x) = \frac{e^{-0,1494-1,3959X_2(1)-1,6095X_4(1)}}{1 + e^{-0,1494-1,3959X_2(1)-1,6095X_4(1)}}$$

Dengan menggunakan model Jackknife terhapus-2 dapat diuji

$$\hat{\pi}(x) = P(Y = 1|X = x)$$

$$\pi(x_2, x_4) = P(Y = 1|X_2 = x_2, X_4 = x_4)$$

Sebagai contoh, untuk data pertama : $X_2 = 1 = Y$, $X_4 = 0 = T$, diperoleh :

$$\begin{aligned} \hat{\pi}(1,0) &= \frac{e^{-0,1494-1,3959(1) -1,6095(0)}}{1 + e^{-0,1494-1,3959(1) -1,6095(0)}} \\ &= 0,16586 \end{aligned}$$

Karena $\hat{\pi}(0,1) = 0,16586 < 0,5$, maka $\hat{Y}_1 = 0$

Nilai $Y_1 = 0$ dan $\hat{Y}_1 = 0$, sehingga pendugaan tersebut benar.

Kemudian, untuk data kedelapan : $X_2 = 0 = T$, $X_4 = 0 = T$, diperoleh :

$$\begin{aligned} \hat{\pi}(0,0) &= \frac{e^{-0,1494-1,3959(0) -1,6095(0)}}{1 + e^{-0,1494-1,3959(0) -1,6095(0)}} \\ &= 0,55920 \end{aligned}$$

Karena $\hat{\pi}(0,0) = 0,55920 > 0,5$ maka $\hat{Y}_8 = 1$.

Nilai $Y_1 = 0 \neq \hat{Y}_1$, sehingga pendugaan tersebut salah.

Demikian seterusnya sesuai dengan Tabel 4.13.

Berdasarkan data hasil pendugaan model untuk prediksi, dapat diketahui bahwa persentase pendugaan benar yang didapat besar (>50%), dapat dikatakan model cukup baik.

Tabel 13 Hasil Prediksi Model

No.	Penerimaan	$\hat{\pi}$	Hasil
1	T	0.17841	T
2	T	0.17841	T
3	T	0.17841	T
4	T	0.03881	T
5	T	0.03881	T
6	T	0.17841	T
7	T	0.03881	T
8	T	0.5	Y
9	T	0.15678	T
10	T	0.17841	T
.			
.			
.			
632	Y	0.15678	T
633	Y	0.5	Y
634	Y	0.5	Y
635	Y	0.5	Y
636	Y	0.5	Y
637	Y	0.17841	T
638	Y	0.17841	T
639	Y	0.17841	T
640	Y	0.17841	T
641	Y	0.03881	T

Tabel 14 Persentase penerapan model untuk prediksi

	T	Y	Persentase (%)
T	499	44	91.9
Y	57	41	41.8
Total Persentase			84.2

Diketahui bahwa :

Persentase estimasi yang benar sebanyak 540 buah (84,2%)

Persentase estimasi yang salah sebanyak 101 buah (15,8%)

4. KESIMPULAN

Berdasarkan pembahasan pada bab-bab sebelumnya maka dapat ditarik beberapa kesimpulan, diantaranya sebagai berikut:

1. Dari keseluruhan rumah tangga miskin yang terkena sampel di Kota Yogyakarta, mayoritas tidak menerima raskin. Berdasarkan jenis dinding tempat tinggal, rumah tangga miskin yang tidak menerima raskin memiliki jenis dinding tempat tinggal bukan tembok. Berdasarkan bahan utama lantai rumah, rumah tangga miskin yang tidak menerima raskin memiliki bahan utama lantai rumah keramik. Berdasarkan fasilitas buang air besar, rumah tangga miskin yang tidak menerima raskin memiliki fasilitas buang sendiri. Berdasarkan

sumber air minum, rumah tangga miskin yang tidak menerima raskin memiliki sumber air minum berbayar rutin.

2. Pada studi kasus dengan data tersebut metode Jackknife mampu memberikan hasil estimasi yang lebih tepat dibandingkan dengan menggunakan sampel asli secara langsung. Hal ini dibuktikan dengan metode Jackknife yang mampu memperkecil nilai standar error sampai dengan metode Jackknife terhapus-2.
3. Model regresi logistik terbaik untuk studi kasus yang telah dibahas pada bab sebelumnya adalah model regresi logistik dengan metode jackknife terhapus-2, dimana model probabilitas sebuah rumah tangga akan menerima raskin adalah :

$$P(Y = 1|X) = \pi(x) = \frac{e^{-0,1494-1,3959X_2(1)-1,6095X_4(1)}}{1+e^{-0,1494-1,3959X_2(1)-1,6095X_4(1)}}$$

Dengan odds ratio masing-masing variabel X_2 dan X_4 adalah $e^{-1,3959} = 0,2476$ dan $e^{-1,6095} = 0,1999$ kali.

4. Dengan tingkat ketelitian model dalam memprediksi sebesar 84,2%

Saran

1. Mengadakan penelitian lebih lanjut dengan menggunakan variabel-variabel lain yang sekiranya berpengaruh terhadap penerimaan raskin, misalnya pendapatan atau pendidikan.
2. Aplikasi metode *resampling* yang lain seperti Bootstrap dan lainnya merupakan topik yang menarik untuk dibahas dalam karya tulis lainnya.
3. Dalam penghapusan data set observasi sebaiknya peneliti memperhatikan jumlah replikasi, karena semakin banyak replikasi belum tentu baik dan juga pada *software R* terdapat keterbatasan memunculkan data replikasi yang berjumlah sampai ratusan ribu.
4. Untuk pemerintah dan instansi terkait dalam hal penerimaan raskin, setelah dilakukannya *resampling* ulang perlu juga diperhatikan variabel jenis dinding tempat tinggal (X_1) dan fasilitas buang air besar (X_3) di dalam penelitian ini tidak terdapat hubungan atau pengaruh yang signifikan terhadap penerimaan raskin.

Ucapan Terima Kasih

Dalam penyusunan tulisan ini, banyak pihak yang telah memberikan dukungan kepada penulis. Oleh karena itu, pada kesempatan ini penulis ingin menyampaikan terima kasih kepada seluruh dosen dan pimpinan Jurusan Statistika Institut Sains & Teknologi AKPRIND Yogyakarta.

Daftar Pustaka

- Agresti A, 2002, *An Introduction to Categorical Data Analysis*, Wiley Interscience, New York.
- Dede C, 2011, *Analisis Partisipasi Perempuan Dalam Kegiatan Ekonomi di Kabupaten Banjar*, Universitas Negeri Yogyakarta, Yogyakarta.
- Efron B dan Tibshirani, R, 2013, *An Introduction to The Bootstrap*, Chapman and Hall, New York.
- Fitianingrum H, 2013, *Estimasi Parameter Model Regresi Menggunakan Metode Jackknife*, Universitas Gadjah Mada, Yogyakarta.
- Hosmer DW, Lemeshow S, 1989, *Applied Logistic Regression*, Wiley Interscience, New York.
- Jamhari, 2012, *Efektivitas Distribusi Raskin di Pedesaan dan Perkotaan Indonesia*, Yogyakarta.
- Rianto H, Wahono RS, 2015, *Resampling Logistic Regression untuk Penanganan Ketidakseimbangan Class pada Prediksi Cacat Software*, Semarang.
- Sayogya, 1988, *Kemiskinan Absolut dan Relatif*, Andi Offset, Yogyakarta.
- Sujarweni, 2014, *Statistik Inferensi untuk Ekonomi dan Bisnis*, Andi Offset, Yogyakarta.
- Ustyannie W, 2014, *Perbandingan Metode Bootstrap dan Jackknife untuk Estimasi Parameter Model Regresi Logistik Biner*, Institut Sains dan Teknologi AKPRIND, Yogyakarta.