

PEMILIHAN MODEL TERBAIK PADA ANALISIS REGRESI LINIER MULTIVARIAT DENGAN KRITERIA AIC

Best Model Election on Multivariate Linear Regression Analysis with AIC Criterion

M. S. Noya Van Delsen¹, S. N. Aulele^{2*}, H. W. M. Patty³, Natalia Kelbulan⁴

^{1,2,3,4} Jurusan Matematika Fakultas MIPA Universitas Pattimura

Jln. Ir. M. Putuhena, Kampus Unpatti, Poka-Ambon, 97233, Provinsi Maluku, Indonesia

e-mail: ¹ marlonnvd@gmail.com ; ^{2*} salmon.aulele@yahoo.com ; ⁴ nataliakelbulan21@gmail.com

Corresponding author*

Abstrak

Kesejahteraan merupakan tujuan utama pembangunan sebuah negara. Salah satu aspek penting yang dapat digunakan untuk mengukur tingkat kesejahteraan adalah kualitas fisik penduduk itu sendiri, dua diantaranya adalah angka kematian bayi dan status gizi buruk. Model regresi multivariat adalah model regresi dengan lebih dari satu variabel respon yang saling berkorelasi dan satu atau lebih variabel prediktor. Tujuan dari penelitian ini adalah untuk mengetahui faktor-faktor apa saja yang mempengaruhi angka kematian bayi dan status gizi buruk di Provinsi Maluku dengan menggunakan regresi linier multivariat. Hasil penelitian menunjukkan bahwa persentase angka kematian bayi tertinggi yaitu di Kabupaten Maluku Tenggara Barat dan persentase status gizi buruk tertinggi di Kabupaten Kepulauan Aru. Faktor-faktor yang signifikan mempengaruhi angka kematian bayi dan status gizi buruk di Provinsi Maluku dengan menggunakan regresi linier multivariat yaitu persentase ASI Eksklusif (X_1), persentase kunjungan ibu hamil ke Puskesmas untuk memeriksa kandungan (X_3), persentase jumlah sarana kesehatan (X_4) dan persentase penduduk miskin (X_5). Model terbaik untuk Y_1 dan Y_2 diperoleh dengan nilai AIC sebesar 8,440420 dan 0,6999986.

Kata Kunci : Model Regresi Linier Multivariat, Angka Kematian Bayi, Status Gizi Buruk, AIC

Abstract

Welfare is the main objective of developing a country. One important aspect that can be used to measure the level of welfare is the physical quality of the population itself, two of which are infant mortality and poor nutritional status. Infant mortality and poor nutritional status in an area can be influenced by more than one variable. The multivariate regression model is a regression model with more than one correlated response variable and one or more predictor variables. The purpose of this study was to find out what factors influence infant mortality and poor nutritional status in Maluku Province by using multivariate linear regression. The results showed that the highest percentage of infant mortality was in West Southeast Maluku Regency and the highest percentage of malnutrition status was in the Aru Islands Regency. Factors that significantly influence infant mortality and poor nutritional status in Maluku Province using multivariate linear regression, namely the percentage of exclusive breastfeeding (X_1), the percentage of visits of pregnant women to the health center to check the content (X_3), percentage of health facilities (X_4) and percentage of poor people (X_5). The best models for Y_1 and Y_2 are obtained with AIC values of 8,440420 and 0,6999986.

Keywords: Multivariate Regression Model, Infant Mortality, Poor Nutritional Status, AIC

Diterima : 02 November 2018

Direvisi: 24 Januari 2019

Disetujui: 27 Februari 2019

Copyright © Jurusan Matematika FMIPA Unpatti 2019

1. PENDAHULUAN

Menurut PBB dan WHO, kematian adalah hilangnya semua tanda-tanda kehidupan secara permanen yang bisa terjadi setiap saat setelah kelahiran hidup. Perubahan jumlah kematian (naik turunnya) di setiap daerah tidaklah sama, tergantung pada berbagai macam keadaan [1]. Besar kecilnya tingkat kematian ini dapat merupakan petunjuk atau indikator bagi tingkat kesehatan dan tingkat kehidupan penduduk di suatu wilayah. Angka Kematian Bayi (AKB) adalah jumlah penduduk yang meninggal sebelum mencapai usia satu tahun yang dinyatakan dalam 1000 kelahiran hidup pada tahun yang sama [8]. Angka kematian bayi juga merupakan indikator yang lazim digunakan untuk menentukan derajat kesehatan masyarakat, baik pada tingkat provinsi maupun nasional. Selain itu program-program kesehatan di Indonesia banyak yang menitikberatkan pada upaya penurunan Angka Kematian Bayi [4].

Berdasarkan data yang diperoleh, terlihat bahwa rata-rata jumlah kematian bayi di Provinsi Maluku Tahun 2015 adalah 35 dimana Kabupaten Maluku Tenggara Barat (MTB) dan Kabupaten Maluku Tengah (Malteng) memiliki jumlah kematian bayi tertinggi yaitu 66 bayi [2]. Jumlah tersebut dianggap terlampau tinggi, karena jika dibandingkan dengan target MDG's maka angka kematian bayi pertahun hanya 17 bayi. Angka-angka kematian tersebut seharusnya dapat di perhatikan oleh pihak-pihak yang berkompeten, sehingga dapat menindak lanjuti guna meminimalisir dan menyelamatkan manusia (bayi), sebagai aset bangsa khususnya daerah Maluku [3]. Selain AKB, indikator lain yang digunakan untuk menentukan derajat kesehatan masyarakat adalah Status Gizi Buruk (SGB).

Dalam perhitungan kedua indikator ini, ternyata mempunyai kesamaan yaitu dipengaruhi oleh faktor-faktor yang sama [4]. Namun dalam penerapannya hal ini tidak dapat diberlakukan secara nasional. Ini dikarenakan letak geografis dan kondisi di setiap daerah yang ada di Indonesia memiliki perbedaan. Sehingga dalam penelitian ini akan dilakukan pemilihan model terbaik untuk faktor-faktor yang berpengaruh terhadap angka kematian bayi dan status gizi buruk di Provinsi Maluku pada regresi linier multivariat dengan menggunakan nilai AIC.

2. METODE PENELITIAN

2.1 Data dan Variabel Penelitian

Data yang digunakan adalah data sekunder yang diperoleh dari Badan Pusat Statistik (BPS) Provinsi Maluku, Profil Kesehatan Provinsi Maluku serta Survei Sosial Ekonomi Nasional (SUSENAS) tahun 2015. Sampel yang digunakan dalam penelitian ini sebanyak 11, yang merupakan kabupaten/kota yang ada di Provinsi Maluku. Dalam penelitian ini variabel yang digunakan yaitu variabel respon dan variabel prediktor.

Variabel respon terdiri dari:

- Y_1 = Persentase Angka Kematian Bayi (AKB)
- Y_2 = Persentase Status Gizi Buruk (SGB)

Sedangkan Variabel Prediktor terdiri dari:

- X_1 = Persentase Pemberian Asi Eksklusif
- X_2 = Persentase Bayi Mendapat Imunisasi
- X_3 = Persentase Kunjungan Ibu Hamil ke Puskesmas untuk memeriksa Kandungan
- X_4 = Persentase Jumlah Sarana Kesehatan (Puskesmas)
- X_5 = Persentase Penduduk Miskin

2.2. Tahapan Analisis

Tahapan analisis data adalah (1) Mendeskripsikan variabel respon dan prediktor. (2) Menguji korelasi antar variabel respon, apabila tidak terdapat korelasi antar variabel respon maka dilanjutkan dengan analisis regresi univariat. (3) Menguji normal multivariat variabel respon. (4) Menduga parameter model regresi multivariat. (5) Melakukan pemilihan model terbaik dengan memilih satu persatu variabel prediktor dengan kriteria AIC. (6) Menguji signifikansi model regresi multivariat secara serentak dan parsial. (7) Menguji asumsi residual yaitu identik, independent dan berdistribusi normal [11].

3. HASIL DAN PEMBAHASAN

3.1. Statistik Deskriptif

Statistik Deskriptif Variabel Respon

Statistik deskriptif variabel respon disajikan dalam Tabel 1, berikut:

Tabel 1. Statistik Deskriptif Variabel Respon

Variabel Respon	N	Minimum	Maksimum	Mean
Angka Kematian Bayi (AKB)	11	0,00	56,00	28,6364
Status Gizi Buruk (SGB)	11	0,00	0,93	0,6318

Berdasarkan Tabel 1, terlihat bahwa nilai $N = 11$ yang menunjukkan jumlah kabupaten/kota yang ada di Provinsi Maluku. Selain itu dapat dilihat juga persentase Angka Kematian Bayi terendah yaitu 0 (tidak terjadi kematian bayi pada tahun itu) di Kabupaten Kepulauan Aru. Sedangkan angka yang tertinggi yaitu sebesar 56 yang terjadi di Kabupaten Maluku Tenggara Barat. Untuk persentase Status Gizi Buruk terendah yaitu sebesar 0 yang terjadi di Kabupaten Seram Bagian Timur, sedangkan angka tertinggi sebesar 0,93 yang terjadi di Kabupaten Kepulauan Aru.

Statistik Deskriptif Variabel Prediktor

Statistik deskriptif variabel prediktor disajikan dalam Tabel 2, berikut:

Tabel 2. Statistik Deskriptif Variabel Prediktor

Variabel Prediktor	N	Minimum	Maksimum	Mean
Persentase Pemberian ASI Eksklusif (X_1)	11	83,55	98,87	93,7427
Persentase Bayi mendapat Imunisasi (X_2)	11	11,87	73,46	37,5045
Persentase kunjungan ibu hamil ke Puskesmas (X_3)	11	9,90	88,50	60,5091
Persentase Jumlah Sarana Kesehatan/Puskesmas (X_4)	11	9,00	33,00	18,0909
Persentase Penduduk Miskin (X_5)	11	4,64	31,01	22,7100

Berdasarkan Tabel 2, terlihat bahwa persentase Pemberian ASI Eksklusif terendah yaitu sebesar 83,55% di Kabupaten Kepulauan Aru sedangkan yang tertinggi sebesar 98,87% di Kabupaten Buru Selatan. Untuk persentase Bayi mendapat Imunisasi terendah yaitu sebesar 11,87% di Kabupaten Buru Selatan sedangkan yang tertinggi yaitu sebesar 73,46% di Kabupaten Maluku Tenggara. Untuk persentase Kunjungan Ibu Hamil ke Puskesmas terendah yaitu sebesar 9,90% di Kabupaten Kepulauan Aru sedangkan yang tertinggi sebesar 88,50% di Kabupaten Seram Bagian Timur. Untuk persentase Jumlah Sarana Kesehatan terendah yaitu sebesar 9,00% di Kabupaten Buru sedangkan yang tertinggi yaitu sebesar 33,00% di Kabupaten Maluku Tengah. Untuk persentase Penduduk Miskin terendah yaitu sebesar % di Kota Ambon sedangkan yang tertinggi sebesar 31,01% di Kabupaten Maluku Barat Daya.

3.2. Regresi Multivariat

Sebelum dibentuk suatu model regresi multivariat, terlebih dahulu dilakukan pengujian yaitu pengujian korelasi antar variabel respon dan pengujian normal multivariate variabel respon. Jika variabel respon tidak saling berkorelasi maka analisis dilanjutkan dengan menggunakan regresi univariat, tetapi sebaliknya jika terdapat korelasi antar variabel respon maka akan dilanjutkan analisis dengan menggunakan regresi multivariat [7].

Pengujian Korelasi Antar Variabel Respon

Untuk menguji korelasi antar variabel respon maka akan digunakan uji *Bartlett Sphericity*. Hipotesis yang digunakan yaitu [9]:

H_0 : Antar variabel respon bersifat *independent*

H_1 : Antar variabel respon bersifat *dependent*

Statistik uji yang diperoleh yaitu:

$$\begin{aligned} x_{hit}^2 &= -\left\{n - 1 - \frac{2q+5}{6}\right\} \ln|R| \\ &= -\left\{11 - 1 - \frac{2 \times 2 + 5}{6}\right\} \ln \left| \begin{bmatrix} 1 & -0,726 \\ -0,726 & 1 \end{bmatrix} \right| \\ &= -(8,5) \ln(0,473) \\ &= -(8,5) (-0,748) = 6,365 \end{aligned}$$

Dengan menggunakan $\alpha = 0,05$ maka diperoleh $x_{0,05;2}^2 = 5,991$. Karena nilai $x_{hit}^2 = 6,365$ lebih besar dari nilai $x_{0,05;2}^2 = 5,991$ maka Tolak H_0 sehingga antar variabel respon bersifat *dependent* atau antar variabel respon saling berkorelasi.

Distribusi Normal Multivariat Variabel Respon

Asumsi selanjutnya yang harus dipenuhi dalam analisis regresi multivariat adalah variabel respon berdistribusi normal multivariat. Analisis yang digunakan adalah dengan menggunakan nilai *square distance* (d_i^2). Hipotesis yang digunakan yaitu:

H_0 : Variabel respon berdistribusi normal multivariat

H_1 : Variabel respon tidak berdistribusi normal multivariat

Hasil yang diperoleh adalah sebagai berikut:

Tabel 3. Distribusi Normal Multivariat

Kabupaten/ Kota	d_i^2	Kabupaten/ Kota	d_i^2
Maluku Tenggara Barat	2,73224	Seram Bagian Timur	1,70413
Maluku Tenggara	5,48220	Maluku Barat Daya	1,97191
Maluku Tengah	6,59034	Buru Selatan	6,71591
Buru	3,75725	Kota Ambon	7,87845
Kepulauan Aru	7,81502	Kota Tual	1,85552
Seram Bagian Barat	3,49702		

Berdasarkan Tabel 3, terlihat bahwa dari 11 kabupaten/kota, terdapat 7 kabupaten/kota atau sebesar 64% menunjukkan nilai $d_i^2 < x_{tabel}^2 = 5,991$. Karena sudah lebih dari 50% maka terima H_0 sehingga variabel respon berdistribusi normal multivariat. Karena pengujian korelasi antar variabel respon dan pengujian normal multivariate variabel respon telah terpenuhi maka dapat dilanjutkan analisis regresi multivariat.

3.3. Estimasi Parameter

Selanjutnya dilakukan estimasi parameter model dengan melibatkan semua variabel prediktor. Hasil yang diperoleh adalah sebagai berikut:

Tabel 4. Hasil Estimasi Parameter

Variabel Respon	Parameter	Nilai Duga
Angka Kematian Bayi	β_0	-519,4103
	β_1	2,8730
	β_2	-0,2187
	β_3	4,3503
	β_4	1,7501
	β_5	0,2593
Status Gizi Buruk	β_0	5,5322
	β_1	-0,0215
	β_2	0,0024
	β_3	-0,0438
	β_4	-0,0088
	β_5	-0,0133

3.4. Pemilihan Model Terbaik Dengan Metode AIC

Langkah awal untuk pemilihan model dengan menggunakan kriteria AIC adalah membentuk model dengan memasukan satu persatu variabel prediktor [10]. Dari pemodelan regresi multivariat dengan menggunakan kriteria AIC diperoleh model terbaik untuk AKB (Y_1) dan SGB (Y_2) dengan nilai AIC sebesar 8,440420 dan 0,6999986. Model tersebut dengan nilai AIC terkecil yaitu:

$$Y_1 = -614,55 + 1,14X_1 + 2,68X_3 + 0,09X_4 + 0,25X_5$$

$$Y_2 = 0,16 - 0,02X_1 - 0,01X_3 - 0,005X_4 - 0,028X_5$$

3.5. Pengujian Signifikansi Model

Selanjutnya dilakukan pengujian signifikansi secara serentak maupun parsial

Pengujian Signifikansi Model Secara Serentak

Pengujian secara serentak dilakukan dengan tujuan untuk mengetahui signifikansi parameter secara keseluruhan dalam model dengan menggunakan uji *Wilk's Lambda*. Hipotesis yang digunakan yaitu [12]:

$$H_0 : \hat{\beta}_{11} = \hat{\beta}_{12} = \dots = \hat{\beta}_{p1} = \dots \hat{\beta}_{pq} = 0 \text{ (model tidak signifikan)}$$

$$H_1 : \text{paling sedikit ada satu } \hat{\beta}_{pq} \neq 0 \text{ (model signifikan)}$$

Berdasarkan hasil analisis diperoleh nilai signifikansi *Wilk's Lambda* sebesar 0,021. Karena nilai signifikansi kurang dari nilai α (0,05) maka tolak H_0 sehingga dapat dikatakan bahwa paling sedikit ada satu parameter yang berpengaruh terhadap angka kematian bayi dan status gizi buruk.

Pengujian Signifikansi Model Secara Parsial

Untuk pengujian signifikansi model secara parsial digunakan uji multivariat *Wilk's Lambda*. Hasil yang diperoleh disajikan dalam Tabel 5, sebagai berikut:

Tabel 5. Hasil Uji *Wilk's Lambda*

Parameter	Sig.
β_1	0,002
β_3	0,007
β_4	0,007
β_5	0,002

Berdasarkan Tabel 5, terlihat bahwa nilai signifikan dari ke-4 variabel kurang dari 0,05 sehingga ke-4 variabel prediktor signifikan mempengaruhi variabel respon. Maka faktor-faktor yang berpengaruh terhadap angka kematian bayi (Y_1), dan status gizi buruk (Y_2) di Provinsi Maluku yaitu persentase Pemberian ASI Eksklusif (X_1), persentase kunjungan ibu hamil ke Puskesmas untuk memeriksa kandungan (X_3), persentase jumlah sarana kesehatan/puskesmas (X_4) dan persentase penduduk miskin (X_5).

3.6. Uji Asumsi Residual IIDN

Uji Asumsi Residual Identik

Asumsi selanjutnya yang harus dipenuhi dalam pemodelan secara multivariat adalah matriks varian-kovarian residual homogen. Pengujian dilakukan terhadap nilai dari residual yang terlihat. Hipotesis yang digunakan yaitu [5]:

$$H_0 : \Sigma_1 = \Sigma_2 = \Sigma$$

$$H_1 : \text{Minimal ada satu } \Sigma_i \neq \Sigma_j \text{ untuk } i \neq j$$

Berdasarkan hasil analisis diperoleh nilai statistic uji Box's-M adalah 3,529 dan nilai *p-value* sebesar 0,446. Karena nilai *p-value* lebih dari nilai α (0,05) maka terima H_0 yang artinya matriks varian-kovarian residual homogen sehingga dapat disimpulkan bahwa residual identik.

Uji Residual Saling Bebas (*Independent*)

Selain asumsi matriks varian-kovarian residual homogen, asumsi residual saling bebas juga harus dipenuhi [6]. Pengujian asumsi ini dilakukan dengan uji *Bartlett Sphercity*. Hipotesis yang digunakan yaitu:

H_0 : Residual bersifat *independent*

H_1 : Residual bersifat *dependent*

Berdasarkan hasil analisis diperoleh nilai $\chi_{hit}^2 = 1,792$. Dengan menggunakan $\alpha = 0,05$ maka diperoleh $\chi_{0,05;2}^2 = 5,991$. Karena $\chi_{hit}^2 = 1,792$ lebih kecil dari nilai $\chi_{0,05;2}^2 = 5,991$ maka terima H_0 sehingga dapat dikatakan bahwa residual bersifat *independent* (saling bebas).

Uji Asumsi Residual Berdistribusi Normal

Selanjutnya asumsi yang harus dipenuhi adalah residual berdistribusi normal multivariat. Hipotesis yang digunakan yaitu :

H_0 : Residual berdistribusi normal multivariat

H_1 : Residual berdistribusi tidak normal multivariat

Hasil yang diperoleh adalah sebagai berikut:

Tabel 6. Distribusi Normal Multivariat untuk Residual

d_i^2	
4,30177	5,00055
0,17170	0,65205
2,32594	0,92778
1,46945	0,71162
0,45590	0,36916
3,61407	

Berdasarkan Tabel 6, terlihat bahwa 11 pengamatan (100%) menunjukkan nilai $d_i^2 < \chi_{tabel}^2 = 5,991$. Karena sudah lebih dari 50% maka terima H_0 sehingga residual berdistribusi normal multivariat.

4. KESIMPULAN

Berdasarkan hasil analisis data dan pembahasan maka dapat dibuat kesimpulan, sebagai berikut:

1. Berdasarkan pemilihan model dengan menggunakan kriteria AIC, maka, didapatkan model terbaik yaitu:
 $Y_1 = -614,55 + 1,14X_1 + 2,68X_3 + 0,09X_4 + 0,25X_5$ dengan nilai AIC sebesar 8,440420
 $Y_2 = 0,16 - 0,02X_1 - 0,01X_3 - 0,005X_4 - 0,028X_5$ dengan nilai AIC sebesar 0,6999986
2. Faktor-faktor yang mempengaruhi angka kematian bayi dan status gizi buruk di Provinsi Maluku dengan menggunakan regresi linier multivariat adalah persentase Asi Eksklusif (X_1), persentase kunjungan ibu hamil ke puskesmas untuk memeriksa kandungan (X_3), persentase jumlah sarana kesehatan/Puskesmas (X_4) dan persentase penduduk miskin (X_5)

DAFTAR PUSTAKA

- [1] Apriadji, Wied Hary, Gizi Keluarga, Jakarta: Penebar Swadaya, 1986.
- [2] Badan Pusat Statistik Maluku , “ Analisa Penyusunan Kinerja Makro Ekonomi dan Sosial Maluku Tahun 2015,” Badan Pusat Statistik Provinsi Maluku, 2015.
- [3] Badan Pusat Statistik Maluku, “ Badan Pusat Statistik Provinsi Maluku,” dalam *Laporan Eksekutif Kesehatan Provinsi Maluku*, 2015.
- [4] Deparet emen Kesehatan RI, “ Indikator Sehat Indonesia 2010 dan Pedoman Penetapan Indikator Provinsi Sehat dan Kabupaten/Kota Sehat, Keputusan Menteri Kesehatan No. 1202/Menkes/SK/VIII/2003,” Departemen Kesehatan RI, , Jakarta, 2003.
- [5] Draper, N, & Smith, H, Analisis Regresi Terapan, Jakarta: Gramedia, 1992.

- [6] Gujarati, Damodar, *Dasar- Dasar Ekonometrika*, Jakarta: Erlangga, 2006.
- [7] Johnson, R.A, & Wichern, D, *Applied Multivariat Statistical Analysis*, New Jersey: Prentice Hall, 2007.
- [8] Melvie Matulesy, SE. MA, “Kematian Bayi di Maluku masih Tinggi,” www.tribun-maluku.com, 2017.
- [9] Morrison, D.F, “*Multivariat Statistical Methods, Fourth Edition,*,” The Wharton School University of Pennsylvania, 2005.
- [10] Priyatno, Duwi, *Analisis KORELASI, REGRESI dan MULTIVARIAT dengan SPSS*, Yogyakarta: iv+126 hlm: Gava Media, 2013.
- [11] Rencher, A.R, *Methods of Multivariat Analysis Second Edition*, New York.: John Wiley & Sons, 2002.
- [12] Sawyer, S. (2010), “*Multivariate Linier Models,*” <http://www.math.wustl.edu/>, Diakses pada tanggal Juni 2014.

