

ANALISIS DISKRIMINAN UNTUK KLASIFIKASI KABUPATEN/KOTA TERTINGGAL DI PROVINSI BENGKULU

Siti Annisa Cholifah¹, Sigit Nugroho², Pepi Novianti²

¹Mahasiswa Jurusan Matematika, FMIPA UNIB

²Dosen Jurusan Matematika, FMIPA UNIB

Jln. WR Supratman Kandang Limun, Bengkulu 38371

ABSTRACT. *Purpose of this research is the classification the left regency/city in 2012 and to know the factors influenced the different of left district in Bengkulu Province by using discriminant analysis. Discriminant analysis can separate the regency/city in Bengkulu Province into two group left district and unleft district with calculation the function score of discriminant compare to cut off score. The result of this research is 5 regency are classified as left district, 5 regency are classified as unleft district, and the factors influenced the different of left district is the number of life expectancy and the numbers of midwife, with the percentage of classification of decision is 100%.*

Key words: Left District, Discriminant Analysis, The Function Score of Discriminant.

1. PENDAHULUAN

Berdasarkan Kementerian Negara Pembangunan Daerah Tertinggal (KNPDT) 2010-2014 yang ditetapkan dengan Perpres No. 5 tahun 2010, pengertian kabupaten/kota tertinggal adalah kabupaten/kota yang masyarakat serta wilayahnya relatif kurang berkembang dibandingkan daerah lain dalam skala nasional. Suatu daerah dikategorikan sebagai daerah tertinggal, disebabkan oleh beberapa faktor, yaitu perekonomian masyarakat, Sumber Daya Manusia, prasarana (infrastruktur), kemampuan keuangan lokal, aksesibilitas, dan karakteristik daerah.

KNPDT 2010-2014 menyatakan bahwa masih tingginya tingkat kesenjangan pembangunan antar wilayah belum optimal, perkembangan kawasan pertumbuhan yang diharapkan menjadi penggerak daerah tertinggal dan kawasan perbatasan. Ini menunjukkan belum adanya keterkaitan dan integrasi ekonomi wilayah dalam sistem pengembangan wilayah.

Provinsi Bengkulu memiliki sepuluh kabupaten/kota dan diantaranya masuk sebagai kabupaten tertinggal. Seperti dijelaskan oleh Staf Khusus Kementerian

Pembangunan Daerah Tertinggal, Jhon Ramadhan, bahwa masih ada enam kabupaten tertinggal yang masuk dalam Kepres tentang Rencana Pembangunan Jangka Menengah Nasional 2010-2014. Meskipun menyatakan ada enam kabupaten tertinggal, namun hanya ada lima kabupaten tertinggal yang disebutkan. Kabupaten tertinggal tersebut adalah Kabupaten Mukomuko, Kabupaten Lebong, Kabupaten Seluma, Kabupaten Kaur dan Kabupaten Kepahiang.

Terdapat berbagai macam metode statistika untuk menganalisis berbagai macam permasalahan, termasuk dalam menentukan status kabupaten/kota tertinggal dan tidak tertinggal. Salah satu metode yang dapat digunakan untuk mengklasifikasikan suatu objek ke dalam kelompok adalah analisis diskriminan. Analisis diskriminan merupakan teknik multivariat yang berkaitan dengan pemisahan objek dalam kelompok yang berbeda dan mengalokasikan objek ke dalam suatu kelompok yang telah ditetapkan sebelumnya.

2. TINJAUAN PUSTAKA

2.1 Daerah Tertinggal

Daerah tertinggal adalah daerah kabupaten/kota yang relatif kurang berkembang dibandingkan daerah lain dalam skala nasional dan berpenduduk relatif tertinggal. Strategi Nasional Pembangunan Daerah Tertinggal menyatakan bahwa ruang lingkup daerah tertinggal adalah wilayah daerah otonom kabupaten/kota. Penetapan status daerah tertinggal mempertimbangkan hal-hal berikut, yaitu: kondisi sosial, ekonomi, budaya, dan wilayah (fungsi inter dan intra spasial baik pada aspek alam, aspek manusia, maupun prasarana pendukung) dari masing-masing daerah. Kondisi tersebut dijadikan dasar dalam menentukan tingkat ketertinggalan suatu daerah dengan menggunakan perhitungan atas kriteria umum, sub kriteria dan indikator yang telah ditentukan bobotnya.

2.2 Analisis Diskriminan

Analisis diskriminan adalah salah satu metode yang dapat digunakan untuk mengetahui variabel mana yang membedakan suatu kelompok dengan kelompok lain dalam suatu populasi. Analisis diskriminan bertujuan mengklasifikasikan suatu objek ke dalam kelompok yang saling lepas (*mutually exclusive/disjoint*) dan menyeluruh (*exhaustive*) berdasarkan sejumlah variabel bebas dan mengelompokkan objek baru ke dalam kelompok-kelompok yang saling bebas tersebut.

Dalam analisis diskriminan ada dua asumsi yang harus dipenuhi (Nugroho, 2008) yaitu:

- a. Memiliki distribusi multivariat normal.
- b. Memiliki kesamaan matriks varians kovarians antar grup.

2.3 Model Analisis Diskriminan

Analisis diskriminan termasuk dalam *multivariate dependence method*, dengan model:

$$Y = b_0 + b_1X_1 + b_2X_2 + \dots + b_pX_p$$

X_p = variabel independen, yaitu data berjenis interval atau rasio, seperti usia seseorang, tinggi sebuah pohon, kandungan zat besi dalam tubuh dan sebagainya.

b_p = koefisien fungsi diskriminan yang telah dihasilkan.

Y = variabel dependen, yaitu data kategorikal atau nominal, seperti golongan tidak tertinggal (kode 0), golongan tertinggal (kode 1), dan sebagainya. Jika data kategorikal tersebut hanya terdiri atas dua kode saja, maka model bisa disebut *Two-Group Discriminant Analysis*. Sedangkan jika kode lebih dari dua kategori, disebut dengan *Multiple Discriminant Analysis*.

2.4 Analisis Diskriminan Fisher

Prinsip utama dari fungsi diskriminan Fisher adalah pemisahan sebuah populasi. Fungsi diskriminan yang terbentuk dapat digunakan untuk pengelompokkan suatu observasi berdasarkan kelompok-kelompok tertentu. Metode Fisher ini tidak mengasumsikan data harus berdistribusi normal, tetapi dalam perhitungan salah satu syarat yang harus diperhatikan adalah data yang digunakan harus memiliki matriks kovarians yang sama untuk setiap kelompok populasi yang diberikan.

Pada analisis diskriminan dua kelompok, jika variabel bebas itu X_1, X_2, \dots, X_p , dan kombinasi linier yang dicari adalah $Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + \dots + \beta_pX_p$ maka vektor $\beta = (\beta_1, \beta_2, \dots, \beta_p)$ perlu dipilih sedemikian hingga hasil bagi $\frac{\beta' \mu_1 - \beta' \mu_2}{\beta' \Sigma \beta}$ mencapai maksimum dimana $\beta = c \Sigma^{-1}(\mu_1 - \mu_2)$ untuk sembarang konstan $c \neq 0$.

Dimana matriks varians kovarians sampel gabungannya adalah S dengan rumus:

$$S = \frac{1}{n_1 + n_2 - 2} [(X_1'X_1 + X_2'X_2)]$$

2.5 Pengujian Signifikasi

Analisis diskriminan bertujuan juga untuk menghasilkan ukuran tingkat perbedaan antara kelompok-kelompok yang terlibat. Jika fungsi diskriminan untuk dua kelompok K_1 dan K_2 adalah $Y = b'X$, maka selisih antara rata-rata skor diskriminan adalah:

$$\begin{aligned}\bar{Y}_1 - \bar{Y}_2 &= b'\bar{X}_1 - b'\bar{X}_2 = b'(\bar{X}_1 - \bar{X}_2) \\ &= (\bar{X}_1 - \bar{X}_2)'S^{-1}(\bar{X}_1 - \bar{X}_2)\end{aligned}$$

Besaran ruas kanan disebut jarak *Mahalanobis* antara \bar{X}_1 dan \bar{X}_2 , dan dinyatakan dengan tanda $D^2(\bar{X}_1, \bar{X}_2)$ atau D^2 . Ukuran atau statistik D^2 dapat digunakan untuk menyelidiki apakah antara dua vektor rata-rata ada perbedaan yang signifikan.

2.6 Klasifikasi

Aturan klasifikasi atau penempatan individu (objek) kedalam kedua kelompok digunakan untuk memprediksi variabel masuk dalam kelompok mana yang kemudian akan didapat matriks klasifikasi. Proses klasifikasi dapat juga digunakan untuk menilai validasi analisis diskriminan.

Aturan klasifikasi menggunakan rumus:

$$Z_{CU} = \frac{n_1 Z_1 + n_2 Z_2}{n_1 + n_2}$$

di mana:

Z_{CU} = angka kritis, yang berfungsi sebagai *cut off score*.

n_1 = jumlah sampel di grup 1, yang dalam kasus ini adalah grup tidak tertinggal.

n_2 = jumlah sampel di grup 2, yang dalam kasus ini adalah grup tertinggal.

Z_1 = angka centroid pada grup 1.

Z_2 = angka centroid pada grup 2.

3. METODE PENELITIAN

3.1 Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yaitu data hasil Survei Sosial Ekonomi Nasional (SUSENAS) tahun 2012 yang dilaksanakan oleh Badan Pusat Statistik (BPS). Data yang digunakan berupa data faktor penyebab daerah tertinggal untuk wilayah kabupaten/kota di Provinsi Bengkulu yang terdiri dari 10 kabupaten/kota. Data tersebut disesuaikan dengan ketersediaan data yang ada.

3.2 Variabel Penelitian

Variabel-variabel yang digunakan pada penelitian ini yaitu persentase penduduk miskin (X_1), angka harapan hidup (X_2), jumlah puskesmas (X_3), jumlah klinik keluarga berencana (X_4), jumlah bidan (X_5), angka melek huruf (X_6), jumlah sekolah SMP (X_7), persentase rumah tangga pengguna listrik (X_8).

3.3 Analisis Data

Prosedur analisis data adalah sebagai berikut:

1. Tabulasi pada Data View.
2. Uji asumsi dan signifikan.
 - a. *Multivariate normal* dilakukan dengan menguji masing-masing variabel secara individu dengan menggunakan uji *Chi-Square*.
 - b. Kesamaan matriks varian kovarian. Kriteria pengujian untuk kesamaan matriks varian kovarian dengan menggunakan *Box's M*.
 - c. Hasil uji vektor rata-rata dengan menggunakan uji Wilk's Lambda.
3. Pembentukan fungsi diskriminan.
4. Interpretasi hasil dengan cara membandingkan nilai fungsi diskriminan dan Z_{CU} .
5. Mengukur ketepatan klasifikasi menggunakan rumus *Hit Ratio*.

4. HASIL DAN PEMBAHASAN

4.1 Uji Asumsi Analisis Diskriminan

1. Uji Normalitas

Uji normalitas pada multivariat bisa dilakukan pada setiap variabel dengan logika bahwa jika secara individual masing-masing variabel memenuhi asumsi normalitas, maka secara bersama-sama (multivariat) variabel-variabel tersebut juga bisa dianggap memenuhi asumsi normalitas.

Tabel 1 Uji Kenormalan

Tests of Normality						
	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
X1	.190	10	.200*	.918	10	.344
X2	.194	10	.200*	.947	10	.630
X3	.224	10	.169	.854	10	.065
X4	.188	10	.200*	.890	10	.169
X5	.199	10	.200*	.873	10	.109
X6	.137	10	.200*	.979	10	.957
X7	.230	10	.142	.898	10	.209
X8	.221	10	.182	.955	10	.733

a. Lilliefors Significance Correction
*. This is a lower bound of the true significance.

Karena nilai sig. semua variabel pada tabel *kolmogorov-smirnov* $> 0,05$ maka H_0 diterima. Artinya variabel-variabel independen tersebut berdistribusi normal. Jadi variabel-variabel independen tersebut memenuhi asumsi dan dapat digunakan untuk analisis diskriminan.

2. Uji Matriks Varians Kovarians

Pengujian dilakukan untuk menguji varians dari setiap variabel dengan menggunakan uji *Box's M*. Asumsi yang harus terpenuhi adalah *group covariance matrices* setiap variabel relatif sama.

Tabel 2 Uji Kesamaan Matriks Varians Kovarians

Test Results		
Box's M		2.030
F	Approx.	.493
	df1	3
	df2	1.152E4
	Sig.	.687

Tests null hypothesis of equal population covariance matrices.

Karena nilai sig. pada tabel *Test Results* $> 0,05$ maka H_0 diterima. Artinya grup kovarians matrik relatif sama (kedua variabel dependen tertinggal dan tidak tertinggal mempunyai matrik varians-kovarians yang sama pada grup-grup yang ada).

3. Uji Vektor Rata-Rata

Untuk melihat kemampuan variabel *independen* dalam membedakan grup secara *multivariate* digunakan uji *Wilks' Lambda*.

Tabel 3 Uji Vektor Rata-Rata

Wilks' Lambda				
Test of ...	Wilks' Lambda	Chi-square	df	Sig.
1	.241	9.972	2	.007

Karena nilai sig pada tabel $0,007 < 0,05$ maka H_0 ditolak. Artinya bahwa ke delapan variabel *independen* mampu membedakan kategori secara signifikan (terdapat perbedaan rata-rata antara dua kategori).

4.2 Interpretasi Fungsi Diskriminan

1. Pembentukan Fungsi Diskriminan

Dalam membuat fungsi diskriminan, digunakan metode *Stepwise Discriminant Analysis*. Berdasarkan hasil analisis dengan program SPSS diperoleh dua variabel yang membentuk fungsi diskriminan yaitu X_2 dan X_5 , sedangkan variabel X_1 , X_3 , X_4 , X_6 , X_7 , dan X_8 tidak masuk dalam fungsi diskriminan.

2. Fungsi Diskriminan

Menurut Santoso (2002), kegunaan fungsi diskriminan adalah untuk mengetahui *case* (dalam hal ini suatu kabupaten) masuk pada grup yang satu atau pada grup lainnya. Dalam menentukan fungsi diskriminan dapat dilihat pada output SPSS berikut.

Tabel 4 Koefisien Fungsi Diskriminan

Canonical Discriminant Function Coefficients	
	Function
	1
x2	.583
x5	.011
(Constant)	-42.503

Unstandardized coefficients

Fungsi diskriminan tersebut dapat ditulis sebagai berikut:

$$Y = -42,503 + 0,583(X_2) + 0,011(X_5) \quad (4.1)$$

Selain fungsi di atas, dengan dipilihnya *Fisher Function Coefficient* pada proses analisis akan terbentuk fungsi diskriminan dari Fisher yang dapat dilihat pada output SPSS.

Tabel 5 Pengklasifikasian Koefisien Fungsi Diskriminan

Classification Function Coefficients		
	KATEGORI	
	Tidak Tertinggal	Tertinggal
X2	40.724	38.873
X5	.245	.212
(Constant)	-1455.585	-1320.514

Fisher's linear discriminant functions

Fungsi diskriminan dari Fisher (4.2) hampir sama dengan fungsi diskriminan sebelumnya (4.1), hanya saja pembagiannya berdasarkan kode grup:

- Untuk kabupaten yang termasuk dalam kategori Tidak Tertinggal
 $Tidak\ Tertinggal = -1455,585 + 40,724(X_2) + 0,245(X_5)$
- Untuk kabupaten yang termasuk dalam kategori Tertinggal
 $Tertinggal = -1320,514 + 38,873(X_2) + 0,212(X_5)$

Selisih antara grup tidak tertinggal dan tertinggal adalah:

$$Y = -135,075 + 1,851(X_2) + 0,033(X_5) \quad (4.2)$$

Fungsi yang akan digunakan dalam aplikasi pada perhitungan *score* menggunakan fungsi *Unstandardized* (4.1). Sedangkan fungsi Fisher persamaan (4.2) sebenarnya bersifat proporsional dengan fungsi *Unstandardized*, yang untuk kasus ini dengan mengalikan setiap koefisien dari Fisher dengan angka 0,31 maka akan diperoleh fungsi *Unstandardized*.

3. Interpretasi Fungsi Diskriminan

Tabel 6 Function of Group Centroids

Functions at Group Centroids	
kategori	Function
tidak tertinggal	1.589
tertinggal	-1.589

Unstandardized canonical discriminant functions evaluated at group means

Dari tabel di atas akan dilakukan perhitungan *cut off score* pada persamaan (5) sebagai berikut:

$$Z_{CU} = \frac{n_1 Z_1 + n_2 Z_2}{n_1 + n_2} = \frac{(5 \times 1,589) + (5 \times (-1,589))}{5 + 5} = 0$$

Penggunaan angka Z_{CU} berdasarkan keputusan:

- Angka skor kasus di atas Z_{CU} , masuk dalam kategori tidak tertinggal (kode 0).
- Angka skor kasus di bawah Z_{CU} , masuk dalam kategori tertinggal (kode 1)

Dengan menggunakan fungsi diskriminan *Unstandardized* pada persamaan (4.1) diperoleh hasil sebagai berikut.

Misalkan untuk Kabupaten Bengkulu Selatan, skor fungsi diskriminannya adalah:

$$Y = -42,503 + 0,583(67,77) + 0,011(361) = 0,978$$

Karena *score* diskriminan Kabupaten Bengkulu Selatan adalah $0,978 > 0$, maka masuk kedalam kategori tidak tertinggal.

Misalkan untuk Kabupaten Kaur, skor fungsi diskriminannya adalah:

$$Y = -42,503 + 0,583(67,85) + 0,011(92) = -1,934$$

Karena *score* diskriminan Kabupaten Kaur adalah $-1,934 < 0$, maka masuk kedalam kategori tertinggal.

4. Ketepatan Hasil Pengklasifikasian

Setelah fungsi diskriminan diperoleh, kemudian melakukan pengujian klasifikasi fungsi diskriminan, selanjutnya akan dilihat seberapa besar ketepatan klasifikasi kategori tersebut. Hal ini dapat dilihat pada output SPSS pada tabel 7 sebagai berikut.

Tabel 7 Menilai Kelayakan Fungsi Diskriminan

Classification Results ^{a,c}					
			Predicted Group Membership		Total
			tidak tertinggal	tertinggal	
Original	Count	tidak tertinggal	5	0	5
		tertinggal	0	5	5
	%	tidak tertinggal	100.0	.0	100.0
		tertinggal	.0	100.0	100.0
Cross-validated ^a	Count	tidak tertinggal	5	0	5
		tertinggal	1	4	5
	%	tidak tertinggal	100.0	.0	100.0
		tertinggal	20.0	80.0	100.0

a. Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.
b. 100,0% of original grouped cases correctly classified.
c. 90,0% of cross-validated grouped cases correctly classified.

Terlihat pada bagian original, kabupaten yang pada data awal berada pada kategori tidak tertinggal dan setelah klasifikasi tetap pada kategori tidak tertinggal sebanyak 5 kabupaten. Sedangkan kabupaten yang pada data awal berada pada kategori tertinggal dan setelah klasifikasi tetap pada kategori tertinggal sebanyak 5 kabupaten.

Hal ini berarti 100% dari 10 kabupaten/kota yang diolah telah dimasukkan pada grup yang sesuai dengan data semula. Jika dilihat dari validasi silang (*cross validated*), pada kode c angka tersebut 90%. Semakin tinggi nilai validasi, termasuk *cross validated groups* tentu semakin bagus, karena semakin tepat fungsi diskriminan membedakan kedua grup.

5. KESIMPULAN DAN SARAN

1. Kesimpulan

Berdasarkan hasil pengolahan data faktor penyebab daerah tertinggal di Provinsi Bengkulu tahun 2012, dapat diperoleh kesimpulan sebagai berikut:

1. Model fungsi diskriminan yang terbentuk berdasarkan hasil pengolahan data adalah:

$$Y = -42,503 + 0,583(X_2) + 0,011(X_5)$$

2. Faktor yang mempengaruhi perbedaan kabupaten/kota tertinggal di Provinsi Bengkulu yaitu angka harapan hidup dan jumlah bidan.
3. Model diskriminan yang terbentuk ini mempunyai ketepatan pengklasifikasian sebesar 100%. Sehingga model diatas dapat digunakan untuk mengklasifikasikan suatu kabupaten/kota termasuk dalam kategori tertinggal atau tidak tertinggal.

2. Saran

Berdasarkan hasil penelitian di atas, saran yang dapat penulis sampaikan untuk pemerintah Provinsi Bengkulu, agar dapat memprioritaskan pembangunan pada daerah yang masih relatif tertinggal dengan memperhatikan faktor sumber daya manusia pada bidang kesehatan yaitu angka harapan hidup dan jumlah bidan. Hal ini dimaksudkan agar masyarakat menjadi sejahtera.

Untuk melakukan analisis diskriminan sebaiknya variabel bebas yang akan dianalisis memiliki pengaruh besar terhadap variabel terikat, sehingga variabel yang layak diuji tidak sedikit. Kemudian, memastikan pengujian asumsi terpenuhi semua agar diperoleh model diskriminan dengan tingkat ketepatan yang lebih tinggi.

6. DAFTAR PUSTAKA

Anonim. 2010. *Kementrian Pembangunan Daerah Tertinggal-Penetapan Daerah Tertinggal dalam RPJM 2010-2014*. Jakarta: Kementrian Pembangunan Daerah Tertinggal.

Anonim. 2010. *Peraturan Presiden Republik Indonesia No. 5 Tahun 2010-RPJM 2010-2014*. Jakarta: Kementrian Pembangunan Daerah Tertinggal.

Anonim. 2012. *Enam kabupaten di Bengkulu Masih Tertinggal*. ANTARA Bengkulu: Bengkulu. Publikasi: Kamis, 21 Juni 2012.[Online]. Tersedia dalam <http://www.antarabengkulu.com/berita/4113/enam-kabupaten-di-bengkulu-masih-tertinggal>

Anonim. 2013. *Provinsi Bengkulu dalam Angka 2013*. Bengkulu: Badan Pusat Statistik.

- Johnson, R. A. dan D. W. Wichern. 2002. *Applied Multivariate Statistical Analysis*. Fifth Edition. New York: Prentice-Hall. Inc.
- Maria, T. S. N., D, Dasari., dan N, Herrhyanto. 2012. *Analisis Diskriminan Terhadap Perilaku Mahasiswa dalam Mengkonsumsi Makanan Cepat Saji*. [Online]. Tersedia dalam <http://journal.fpmipa.upi.edu/index.php/eurekamatika/article/view/35> [20 Agustus 2015]
- Nugroho, S. 2008. *Statistika Multivariat Terapan*. Unib Press. Bengkulu.
- Santoso, S. 2002. *Buku Latihan SPSS Statistik Multivariat*. Jakarta: PT Elex Media Komtindo.
- Sulistiyorini, I. 2013. *Analisis Faktor-Faktor yang Mempengaruhi Kepuasan Pelanggan Internet Speedy Reguler Menggunakan Analisis Diskriminan di Semarang*. [Skripsi]. Jurusan Matematika. Fakultas Matematika dan Ilmu Pengetahuan Alam. Universitas Negeri Semarang.
- Suryanto. 1988. *Metode Statistika Multivariat*. Jakarta: Departemen Pendidikan dan Kebudayaan Direktorat Jendral Pendidikan Tinggi, Proyek Pengembangan Lembaga Tinggi Tenaga Kependidikan.