

PENERAPAN ANALISIS DISKRIMINAN KUADRATIK KLASIK UNTUK MENDUGA KATEGORI INDEKS PEMBANGUNAN MANUSIA (IPM) PROVINSI DI INDONESIA TAHUN 2012

RATIH FEBI RAMADHANI, HAZMIRA YOZZA, IZZATI RAHMI HG

*Program Studi Matematika,
Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Andalas,
Kampus UNAND Limau Manis Padang, Indonesia.
email : ratihfebi@gmail.com*

Abstrak. Indikator Indeks Pembangunan Manusia (IPM) yang digunakan Badan Pusat Statistika (BPS) dalam penghitungan IPM adalah angka harapan hidup, angka melek huruf, rata-rata lama sekolah, dan daya beli. Untuk mendapatkan IPM dengan indikator lain sebagai pembanding digunakan empat variabel pembeda yaitu angka kematian bayi, persentase penduduk berumur 10 tahun ke atas yang belum/tidak sekolah, persentase penduduk berumur 10 tahun ke atas yang buta huruf dan persentase penduduk miskin. Tujuan dari penelitian ini adalah untuk menduga kategori Indeks Pembangunan Manusia (IPM) provinsi di Indonesia tahun 2012 berdasarkan empat variabel pembeda yang digunakan. Metode yang digunakan dalam penelitian ini adalah analisis diskriminan kuadratik klasik. Data yang digunakan diperoleh dari situs BPS. Variabel pembeda yang digunakan signifikan dalam menduga kategori IPM. Dari 33 provinsi di Indonesia dengan analisis diskriminan kuadratik diduga terdapat dua kelompok kategori IPM yaitu kategori sedang ($70 \leq \text{IPM} \leq 75$) dan kategori tinggi ($\text{IPM} > 75$) dimana 25 provinsi masuk ke dalam kategori sedang dan 8 provinsi masuk ke dalam kategori tinggi. Kesalahan pendugaan kategori IPM menghasilkan nilai APER sebesar 27,27%.

Kata Kunci: Indeks Pembangunan Manusia (IPM), Analisis Diskriminan, Analisis Diskriminan Kuadratik Klasik

1. Pendahuluan

Indeks Pembangunan Manusia (IPM) merupakan ukuran standar pembangunan manusia yang dibentuk dari empat indikator yaitu angka harapan hidup, angka melek huruf, rata-rata lama tahun sekolah dan kemampuan daya beli [1]. Berdasarkan skala internasional capaian IPM dapat dikategorikan menjadi 4 yaitu kategori tinggi, kategori menengah ke atas, kategori menengah ke bawah dan kategori rendah. Berdasarkan nilai IPM, Indonesia termasuk ke dalam kategori menengah ke atas, demikian juga halnya dengan seluruh provinsi di Indonesia, kecuali provinsi Papua. Selain keempat faktor yang digunakan dalam menentukan nilai IPM diduga terdapat faktor lain yang membedakan pengkategorian IPM provinsi di Indonesia. Sangat menarik untuk diketahui, apakah faktor-faktor tersebut berbeda antar provinsi-provinsi pada kelompok IPM yang berbeda. Jika memang berbeda apakah pengelompokan yang terjadi tersebut memang berbeda dengan kelompok kategori berdasarkan penetapan BPS. Oleh karena itu perlu pengelompokan

berdasarkan variabel pembeda yang didasarkan pada pengelompokan sebelumnya dengan analisis diskriminan.

Analisis diskriminan merupakan teknik multivariat yang berkaitan dengan pemisahan objek dalam kelompok yang berbeda dan mengalokasikan objek tersebut ke dalam suatu kelompok yang telah ditetapkan sebelumnya. Terdapat 2 jenis analisis diskriminan yaitu analisis diskriminan linier klasik dan analisis diskriminan kuadrat klasik. Analisis diskriminan linier klasik digunakan jika data menyebar normal multivariat dan matriks ragam peragam antar kelompok sama sedangkan analisis diskriminan kuadrat klasik digunakan jika data menyebar normal multivariat namun matriks ragam peragam antar kelompok berbeda. Jadi dengan menggunakan analisis diskriminan kuadrat klasik dapat ditentukan suatu fungsi diskriminan yang berguna untuk menduga kategori IPM serta dapat menjadi alat pembandingan bagi suatu provinsi dengan provinsi lainnya dalam mengukur tingkat IPM.

2. Analisis Diskriminan Kuadrat Klasik

Analisis diskriminan merupakan teknik multivariat yang berkaitan dengan pemisahan objek ke dalam kelompok yang berbeda dan mengelompokkan objek tersebut ke dalam kelompok yang telah ditentukan sebelumnya [3]. Analisis diskriminan digunakan pada kasus dengan variabel respon berupa data kualitatif dan variabel bebas berupa data kuantitatif. Analisis diskriminan kuadrat klasik dilakukan apabila data memenuhi asumsi normal multivariat dan matriks ragam peragam antar kelompok berbeda [4].

Skor diskriminan kuadrat klasik dengan nilai $x = (x_1, x_2, \dots, x_p)$ dapat didefinisikan sebagai berikut:

$$d_k^{CQ}(x) = \ln(p_k) - \frac{1}{2} \ln(|\sum_k|) - \frac{1}{2}(x - \mu_k)^T \sum_k^{-1}(x - \mu_k)$$

dengan d_k^{CQ} adalah fungsi diskriminan kuadrat klasik untuk kelompok- k , dimana $k = 1, 2, \dots, g$.

Untuk mengestimasi μ_k dan \sum_k , dapat digunakan rata-rata kelompok \bar{x}_k dan matriks ragam peragam S_k , sehingga taksiran skor diskriminan kuadrat klasik menjadi

$$d_k^{CQ}(x) = \ln(p_k) - \frac{1}{2} \ln(|S_k|) - \frac{1}{2}(x - \bar{x}_k)^T S_k^{-1}(x - \bar{x}_k)$$

dengan

\bar{x}_k : vektor mean sampel kelompok ke- k

S_k : matriks ragam peragam sampel kelompok ke- k

Sebuah observasi $x^* = (x_1^*, x_2^*, \dots, x_p^*)$ akan termasuk kedalam kelompok k jika skor diskriminan kuadrat:

$$d_k^{CQ} = \max\{d_k^{CQ}(x); k = 1, 2, \dots, g\}$$

Untuk menghitung ketepatan klasifikasi pada hasil pengelompokan digunakan *Apparent Error Rate* (APER). Dalam menghitung nilai APER digunakan tabel

prediksi pengelompokan sebagai berikut: dengan

Tabel 1. Tabel Prediksi Pengelompokan

Pengelompokkan awal	Pengelompokkan prediksi			
	π_1	π_2	...	π_g
π_1	n_{11}	n_{12}	...	n_{1g}
π_2	n_{21}	n_{22}	...	n_{2g}
\vdots	\vdots	\vdots	\ddots	\vdots
π_g	n_{g1}	n_{g2}	...	n_{gg}

π_{gg} : jumlah pengamatan dari kelompok π_g yang dikelompokkan secara benar ke kelompok π_g

$\pi_{gg'}$: jumlah pengamatan dari kelompok π_g yang dikelompokkan secara benar ke kelompok $\pi_{g'}$

Kemudian *Apparent Error Rate* dihitung dengan menggunakan rumus :

$$APER = \frac{n_{12} + n_{13} + n_{21} + n_{23} + \dots + n_{gg'}}{n_1 + n_2 + \dots + n_g}$$

dimana

n_1 : ukuran contoh kelompok 1

n_2 : ukuran contoh kelompok 2

n_3 : ukuran contoh kelompok 3

\vdots

n_g : ukuran contoh kelompok g

3. Metode Penelitian

Data yang digunakan dalam penelitian ini adalah data indikator Indeks Pembangunan Manusia, angka kematian bayi, persentase penduduk berumur 10 tahun ke atas yang belum/tidak sekolah, persentase penduduk berumur 10 tahun ke atas yang buta huruf dan persentase penduduk miskin dari 33 provinsi di Indonesia yang merupakan data sekunder yang diperoleh dari situs Badan Pusat Statistika (BPS) Indonesia, <http://www.bps.go.id>. Variabel respon (Y) pada penelitian ini adalah kategori awal Indeks Pembangunan Manusia (IPM) yang nilai IPM diperoleh dari perhitungan BPS yang dikelompokkan ke dalam 3 kelompok kategori yaitu kategori

rendah, kategori sedang, dan kategori tinggi dengan variabel bebas (X) adalah

X_1 : Angka kematian bayi

X_2 : Persentase penduduk berumur 10 tahun ke atas yang belum/tidak sekolah

X_3 : Persentase penduduk berumur 10 tahun ke atas yang buta huruf

X_4 : Persentase penduduk miskin

Adapun langkah-langkah dalam penelitian ini adalah :

1. Pembentukan kelompok awal IPM berdasarkan nilai IPM yang diperoleh dari BPS yang dikelompokkan kedalam tiga kategori kelompok yaitu kategori rendah ($IPM < 70$), kategori sedang ($70 \leq IPM \leq 75$) dan kategori tinggi ($IPM > 75$).
2. Analisis deskriptif data.
3. Pemeriksaan asumsi normalitas menggunakan plot antara jarak mahalanobis dan khi-kuadrat.
4. Uji asumsi kesamaan matriks ragam peragam menggunakan statistik uji bartlett.
5. Uji kesamaan vektor nilai tengah menggunakan statistik uji T^2 -Hotelling.
6. Pemilihan variabel pembeda dalam fungsi diskriminan kuadrat klasik menggunakan statistik uji- F .
7. Membentuk fungsi diskriminan kuadrat klasik dan menentukan skor diskriminan kuadrat klasik.
8. Menentukan kesalahan pengelompokan IPM dengan menggunakan nilai APER.

4. Analisis dan Pembahasan

Berikut akan dipaparkan hasil dari pembahasan dalam penelitian ini.

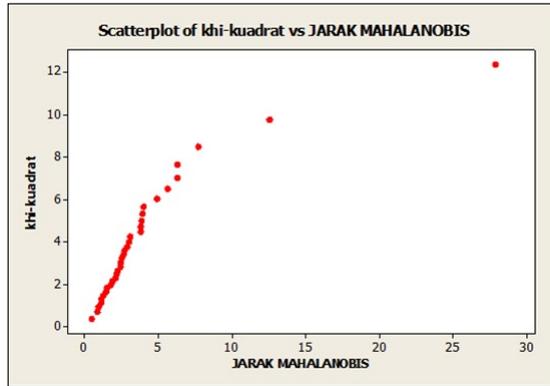
4.1. *Pemeriksaan Asumsi Normalitas*

Berdasarkan hasil plot antara jarak mahalanobis dan nilai khi-kuadrat seperti pada Gambar 2 terlihat bahwa titik-titik cenderung membentuk garis lurus dengan lebih dari 50% data dengan nilai $d_i^2 \leq (X_{p;0,05}^2)$ yang berarti data peubah bebas yang digunakan memenuhi asumsi normal multivariat.

4.2. *Uji Kesamaan Vektor Nilai Tengah Antar Kelompok dan Uji Kesamaan Matriks Ragam Peragam*

Uji kesamaan vektor nilai tengah antar kelompok menggunakan uji T^2 -Hotelling menghasilkan nilai $T^2 > X_{4;0,05}^2$ yang berarti tolak H_0 pada taraf nyata 5% sehingga dapat disimpulkan vektor nilai tengah antar kelompok berbeda. Karena vektor nilai tengah antar kelompok berbeda, maka metode analisis diskriminan merupakan metode yang tepat digunakan

Uji kesamaan matriks ragam peragam digunakan untuk mengetahui metode diskriminan yang tepat digunakan untuk data. Uji ini dilakukan untuk melihat



Gambar 1. Plot Kuantil Khi-kuadrat

apakah matriks ragam peragam dari ketiga kelompok adalah sama. Statistik uji yang digunakan adalah uji Bartlett. Dengan uji Bartlett diperoleh nilai $MC^{-1} = 84,50 > 31,4$ sehingga tolak H_0 pada taraf nyata 5% sehingga dapat disimpulkan ada kelompok dengan matriks ragam peragam berbeda. Jadi perbedaan matriks ragam peragam antar kelompok dapat dijadikan alasan bahwa metode yang digunakan adalah analisis diskriminan kuadratik klasik.

4.3. Pemilihan Variabel Pembeda

Pemilihan variabel pembeda dilakukan untuk mengetahui variabel apa saja yang signifikan dalam fungsi diskriminan kuadratik klasik. Pemilihan variabel pembeda dilakukan dengan menggunakan statistik uji- F . Dengan menggunakan software SPSS 20 disajikan Tabel ANOVA F sebagai berikut:

Tabel 2. Tabel ANOVA

ANOVA						
	Sumber Keragaman	Jumlah kuadrat	Derajat Bebas	Kuadrat Tengah	F Hitung	Nilai p
X_1	Perlakuan	1457,069	2	728,534	4,369	0,022
	Galat	5002,446	30	166,748		
	Total	6459,515	32			
X_2	Perlakuan	329,997	2	164,998	8,689	0,001
	Galat	569,687	30			
	Total	899,648	32			
X_3	Perlakuan	406,350	2	203,175	9,882	0,001
	Galat	616,813	30	20,560		
	Total	1023,162	32			
X_4	Perlakuan	335,265	2	167,632	4,842	0,015
	Galat	1038,564	30	34,619		
	Total	1373,829	32			

Berdasarkan Tabel ANOVA F terlihat bahwa nilai p kurang dari $\alpha(5\%)$ sehingga

variabel pembeda yang digunakan yaitu angka kematian bayi, persentase penduduk berumur 10 tahun ke atas yang belum/tidak sekolah, persentase penduduk berumur 10 tahun ke atas yang buta huruf, dan persentase penduduk miskin berpengaruh dalam membedakan pengelompokan kategori IPM rendah, sedang, dan tinggi.

4.4. Analisis Diskriminan Kuadratik Klasik

Langkah pertama dalam menentukan fungsi diskriminan kuadratik klasik adalah menentukan nilai vektor rata dan matriks ragam-peragam setiap kelompok. Setelah vektor rata-rata dan matriks ragam peragam setiap kelompok diperoleh, selanjutnya dapat dibentuk fungsi diskriminan kuadratik klasik dari ketiga kelompok. Fungsi diskriminan kuadratik klasik untuk IPM kategori rendah dilambangkan dengan d_1^{CQ} , untuk IPM kategori sedang dilambangkan dengan d_2^{CQ} , dan untuk IPM kategori tinggi dilambangkan dengan d_3^{CQ} . Fungsi diskriminan kuadratik klasik yang terbentuk adalah sebagai berikut.

$$\begin{aligned}d_1^{CQ}(x) &= -1,4895632 - \frac{1}{2}(x - \bar{x}_1)^T S_1^{-1}(x - \bar{x}_1) \\d_2^{CQ}(x) &= -5,116485 - \frac{1}{2}(x - \bar{x}_2)^T S_2^{-1}(x - \bar{x}_2) \\d_3^{CQ}(x) &= -3,397166 - \frac{1}{2}(x - \bar{x}_3)^T S_3^{-1}(x - \bar{x}_3).\end{aligned}$$

Dari ketiga fungsi diskriminan kuadratik klasik di atas dapat dihitung skor diskriminan kuadratik klasik masing-masing provinsi. Skor diskriminan yang diperoleh digunakan dalam pengelompokan kembali IPM berdasarkan variabel pembeda yang digunakan. Berdasarkan fungsi diskriminan yang diperoleh, memberikan hasil bahwa 25 provinsi masuk ke dalam IPM kategori sedang, delapan provinsi masuk ke dalam IPM kategori tinggi dan tidak satupun provinsi yang masuk ke dalam IPM kategori rendah.

Pada Tabel 3 disajikan nama-nama provinsi berdasarkan kelompok kategori IPM, dimana provinsi dengan tulisan merah merupakan provinsi yang pengelompokannya tidak sesuai berdasarkan pengelompokan sebelumnya.

Untuk melihat keakuratan pengelompokan kategori IPM di 33 provinsi di Indonesia adalah dengan menghitung tingkat kesalahan pengelompokan dengan menggunakan metode APER. Nilai APER untuk fungsi diskriminan kuadratik klasik dalam pengelompokan IPM adalah sebagai berikut.

$$APER = \frac{n_{12} + n_{13} + n_{21} + n_{23} + n_{31} + n_{32}}{n_1 + n_2 + n_3} = \frac{3 + 1 + 0 + 2 + 0 + 3}{4 + 21 + 8} = 0,2727$$

Berdasarkan dari nilai APER yang diperoleh dapat dijelaskan bahwa fungsi diskriminan kuadratik klasik menghasilkan proporsi kesalahan dalam pengelompokan Indeks Pembangunan Manusia sebesar 27,2%. Jadi kemampuan analisis diskriminan kuadratik klasik dalam mengelompokkan IPM kedalam tiga kategori berdasarkan variabel angka kematian bayi, persentase penduduk 10 tahun keatas yang tidak/belum pernah sekolah, persentase penduduk berumur 10 tahun keatas yang buta huruf, dan persentase penduduk miskin adalah sebesar 72.72%.

Tabel 3. Tabel Pengelompokan Provinsi Perkelompok (Kelompok Hasil dengan Analisis Diskriminan Kuadratik Klasik)

Kelompok 2	Kelompok 3
Nangroe Aceh	Sumatera Barat
Darussalam	Sumatera Utara
Riau	Kepulauan Riau
Jambi	DKI Jakarta
Sumatera Selatan	Kalimantan Tengah
Bengkulu	Kalimantan Timur
Lampung	Kalimantan Selatan
Kepulauan Bangka	Sulawesi Utara
Belitung	Maluku Utara
Jawa Barat	
Jawa Tengah	
Jawa Timur	
DI Yogyakarta	
Banten	
Bali	
Nusa Tenggara Barat	
Nusa Tenggara Timur	
Kalimantan Barat	
Sulawesi Tengah	
Sulawesi Selatan	
Sulawesi Tenggara	
Gorontalo	
Sulawesi Barat	
Maluku	
Papua Barat	
Papua	

5. Kesimpulan

Berdasarkan hasil dan pembahasan dalam penelitian ini, maka dapat disimpulkan bahwa berdasarkan keempat variabel pembeda yang signifikan dalam fungsi diskriminan kuadratik klasik menghasilkan dua kelompok kategori IPM yaitu kategori sedang dan kategori tinggi, dengan 25 provinsi masuk ke dalam kategori sedang dan delapan provinsi masuk ke dalam kategori tinggi.

Keakuratan pengelompokan IPM dengan menggunakan metode analisis diskriminan kuadratik klasik menghasilkan persentase kesalahan pengelompokan (APER) sebesar 27,27%. Ini berarti dengan analisis diskriminan kuadratik klasik pendugaan pengelompokan kategori IPM yang digunakan menghasilkan ketepatan pengelompokan IPM cukup baik.

6. Ucapan Terima Kasih

Penulis mengucapkan terima kasih kepada Bapak Dr. Dodi Devianto, Ibu Dr. Maiyastri, dan Bapak Dr. Ahmad Iqbal Baqi yang telah memberikan masukan dan saran dalam penyempurnaan penulisan artikel ini.

Daftar Pustaka

- [1] Badan Pusat Statistika (BPS). 2008. *Indeks Pembangunan Manusia 2006-2007*. BPS, Jakarta.
- [2] Badan Pusat Statistika (BPS). 2014. *Indeks Pembangunan Manusia 2013*. BPS, Jakarta.
- [3] Johnson, R.A., dan D.W. Winchern. 1998. *Applied Multivariate Statistical Analysis*. Edisi ke-3. Prentice-Hall Inc, New Jersey.
- [4] Sharma, S. 1996. *Applied Multivariate Techniques*, John Wiley & Sons Inc, Canada.