

BAB II

LANDASAN TEORI

2.1 Indeks Pembangunan Manusia

Pembangunan merupakan topik hangat yang selalu diperbincangkan. Keberhasilan pembangunan sering diukur dengan tingginya pertumbuhan ekonomi. Ini merupakan konsep pembangunan sebelum tahun 1970, dimana pertumbuhan ekonomi hanya diukur melalui pertumbuhan Produk Nasional Bruto (PNB). Namun, fakta menunjukkan bahwa banyak negara memiliki pertumbuhan ekonomi tinggi tetapi gagal dalam meningkatkan kualitas penduduknya.

United Nation Development Program (UNDP) pada tahun 1990 mulai memperkenalkan Indeks Pembangunan Manusia (IPM) / *Human Development Indeks (HDI)* yang dimuat dalam laporan tahunan *Human Development Report (HDR)*. UNDP mendefinisikan pembangunan sebagai proses perluasan pilihan bagi penduduk. Peran manusia sebagai tujuan akhir, bukan sebagai alat pembangunan.

IPM merupakan indeks yang mengukur pembangunan manusia yang disusun dari tiga aspek yaitu, kesehatan dengan dimensi umur panjang dan hidup sehat, pendidikan dengan dimensi pengetahuan serta ekonomi dengan dimensi standar hidup layak. Ketiga dimensi ini mencerminkan tingkat keberhasilan pembangunan di suatu wilayah.

Dari tiga dimensi tersebut kemudian diturunkan menjadi empat indikator dalam penghitungan IPM, yaitu Angka Harapan Hidup saat lahir (AHH), Angka Melek Huruf (AMH), gabungan Angka Partisipasi Kasar (APK), dan Produk Domestik Bruto (PDB) per kapita.

Pada tahun 2010, UNDP melakukan penyempurnaan dalam penghitungan IPM. Indikator Angka Melek Huruf (AMH) dan gabungan Angka Partisipasi Kasar (APK) diganti dengan Harapan Lama Sekolah (HLS) dan Rata-rata Lama Sekolah (RLS). Indikator Produk Domestik Bruto (PDB) per kapita diganti dengan Produk Nasional Bruto (PNB) per kapita serta penghitungan rata-rata

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

indeks diubah dari aritmatik menjadi geometrik. Indonesia baru menggunakan metode ini pada tahun 2014.

Kemiskinan adalah kondisi dimana seseorang atau sekelompok orang tidak mampu memenuhi hak-hak dasarnya untuk mempertahankan dan mengembangkan kehidupan yang bermartabat (Bappenas, 2004). Kemiskinan adalah keadaan dimana terjadi ketidakmampuan untuk memenuhi kebutuhan dasar seperti makanan, pakaian, tempat berlindung, pendidikan, dan kesehatan.

Kemiskinan dapat disebabkan oleh kelangkaan alat pemenuh kebutuhan dasar, ataupun sulitnya akses terhadap pendidikan dan pekerjaan (Bappeda, 2011) oleh karena itu tingkat kemiskinan dapat mempengaruhi nilai IPM.

Pengangguran terbuka merupakan bagian dari angkatan kerja yang tidak bekerja atau sedang mencari pekerjaan (baik bagi mereka yang belum pernah bekerja sama sekali maupun yang sudah pernah bekerja sama sekali maupun yang sudah pernah bekerja), atau sedang mempersiapkan suatu usaha, mereka yang tidak mencari pekerjaan karena merasa tidak mungkin untuk mendapatkan pekerjaan dan mereka yang sudah memiliki pekerjaan dan mereka yang sudah memiliki pekerjaan tetapi belum mulai bekerja (Bappeda, 2011). Tingkat pengangguran adalah persentase jumlah pengangguran terbuka terhadap jumlah angkatan kerja. Tingkat pengangguran ini dapat mempengaruhi nilai IPM.

Masalah tenaga kerja tidak terlepas dari upah minimum regional (UMR). Upah minimum ini merupakan salah satu pertimbangan bagi investor yang ingin menanamkan modalnya disuatu daerah terutama investor yang ingin mendirikan pabrik atau industri yang banyak menyerap tenaga kerja. Semakin tinggi upah minimum regional suatu daerah menunjukkan semakin tinggi tingkat ekonominya (Bappeda, 2010).

Pertumbuhan ekonomi menunjukkan sejauh mana aktivitas perekonomian akan menghasilkan tambahan pendapatan masyarakat pada suatu periode tertentu. Perekonomian dianggap mengalami pertumbuhan bila seluruh balas jasa riil terhadap penggunaan faktor produksi pada tahun tertentu lebih besar daripada tahun sebelumnya. Indikator yang digunakan untuk mengukur pertumbuhan

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Adapun dalam regresi berganda variable dependentnya harus metrik, dan jika variabelnya independen, bisa metrik maupun nonmetrik. Sama seperti regresi berganda, dalam analisis diskriminan variabel independen hanya satu, sedangkan variabel independen banyak (*multiple*).

Analisis diskriminan adalah metode statistik untuk mengelompokkan atau mengklasifikasi sejumlah objek ke dalam beberapa kelompok, berdasarkan beberapa variabel, sedemikian hingga setiap obyek yang menjadi anggota lebih dari pada satu kelompok. Pada prinsipnya analisis diskriminan bertujuan untuk mengelompokkan setiap objek ke dalam dua atau lebih kelompok berdasarkan pada kriteria sejumlah variabel bebas.

Analisis diskriminan adalah salah satu teknik analisa Statistika dependensi yang memiliki kegunaan untuk mengklasifikasikan objek beberapa kelompok. Pengelompokan dengan analisis diskriminan ini terjadi karena ada pengaruh satu atau lebih variabel lain yang merupakan variabel independen. Kombinasi linier dari variabel-variabel ini akan membentuk suatu fungsi diskriminan .

Analisis diskriminan adalah teknik multivariate yang termasuk *dependence method*, yakni adanya variabel dependen dan variabel independen. Dengan demikian ada variabel yang hasilnya tergantung dari data variabel independen. Analisis diskriminan mirip regresi linier berganda (*multivariable regression*). Analisis diskriminan merupakan tehnik yang akurat untuk memprediksi seseorang termasuk dalam kategori apa, dengan catatan data-data yang dilibatkan terjamin akurasi.

2.4 Konsep Dasar Analisis Diskriminan

Analisis diskriminan merupakan teknik menganalisis data, kalau variable tak bebas disebut criterion merupakan kategori merupakan nonmetrik ordinal atau nominal bersifat kualitatif sedangkan variable bebas sebagai predictor merupakan metric interval atau rasio bersifat kuantitatif. Misalnya variabel tak bebas merupakan pilihan 3 merek computer pribadi yaitu merk A,B,C dan variable bebasnya merupakan penilaian atribut dari PC skala likert butir, misalnya dari 1 = sangat tidak bermutu sampai 5= sangat bermutu.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Ilustrasi di atas sebenarnya telah menggambarkan penggunaan dari Analisis Diskriminan, yaitu sebagai sebuah teknik analisis yang digunakan untuk mendapatkan beberapa karakteristik yang membedakan objek ke dalam dua atau lebih kelompok.

Kegunaan Analisis Diskriminan bukan hanya untuk mengelompokkan kota di Indonesia, tetapi untuk sangat banyak keperluan. Misalnya :

- a. Sedikit atau banyaknya konsumen membeli produk setiap bulan.
- b. Pengelompokkan *sales* berdasarkan jumlah produk yang terjual dengan kategori banyak, sedang, sedikit.
- c. Menganalisis karakter perusahaan yang berpotensi bangkrut atau tidak. Pemilihan warna produk, misalnya ungu, merah, biru, hitam, dan putih.

Masih banyak lagi kegunaan Analisis Diskriminan yang tidak mungkin disebutkan satu-persatu. Tapi yang paling penting selama kita melakukan ingin melakukan pengelompokkan, maka Analisis Diskriminan dapat digunakan.

Sehubungan dengan variabel atau data, ada beberapa ketentuan yang harus dipatuhi ketika menggunakan Analisis Diskriminan. antara lain :

1. Variabel terikat mempunyai skala pengukuran nonmetriks (nominal atau ordinal). Hal ini sangat beralasan karena karakter inilah yang digunakan untuk membedakan objek menjadi dua kelompok atau lebih.
2. Variabel bebas harus mempunyai skala metric (interval atau rasio).

Adapun hubungan diskriminan dengan regresi dan anova akan di gambarkan oleh Tabel 2.1 sebagai berikut:

Tabel 2.1 Persamaan dan Perbedaan antara Analisis Varians (Anova), Analisis Regresi dan Analisis Diskriminan

Persamaan	Anova	Regresi	Diskriminan
Banyaknya variabel tak bebas(Y)	Satu	Satu	Satu
Banyaknya variabel bebas(X)	Banyak (>1)	Banyak (>1)	Banyak (>1)
Perbedaan sifat variabel tak	Metric kuantitatif	Metric kuantitatif	Kategori kualitatif

bebas(Y)			
Perbedaan sifat variable bebas(X)	Kategori kualitatif	Kategori kualitatif	Metric kuantitatif

Tabel 2.1 menunjukkan hubungan antara analisis diskriminan dengan regresi berganda dan analisis varians. Di dalam analisis varian dan regresi berganda variable tak bebas (Y) merupakan data matriks atau interval sedangkan dalam analisis diskriminan merupakan data kategorikal.

Variabel bebas berupa kategori nonmetrik dalam analisis varian dan berupa data metric dalam hal analisis regresi dan diskriminan. Di dalam regresi berganda pada Tabel 2.1 meramalkan nilai Y kuantitatif berupa angka, di dalam analisis diskriminan meramalkan nilai Y suatu objek, elemen, kasus, atau responden masuk kelompok kategori yang mana.

2.5 Statistik yang Relevan dengan Analisis diskriminan

Korelasi kanonikal mengukur seberapa kuat asosiasi antara skor diskriminan dan kelompok, merupakan ukuran antara fungsi diskriminan tunggal dan set variable yang membentuk anggota kelompok.

Centroid ialah rata-rata nilai skor fungsi diskriminan untuk suatu kelompok tertentu, banyaknya centroid sebanyak kelompok yang ada sebab setiap kelompok mempunyai satu centroid. Rata-rata untuk suatu kelompok pada semua fungsi disebut centroid kelompok (*group centroids*).

Matriks klasifikasi sering disebut matriks kebingungan atau prediksi memuat jumlah objek atau kasus yang secara benar terklasifikasi dan yang salah terklasifikasi. Objek atau kasus yang terklasifikasi secara benar tampak pada diagonal, sebab kelompok yang diramalkan dan yang sebenarnya sama, sedangkan yang berda diluar diagonal mewakili objek yang terklasifikasi secara salah. Banyaknya objek yang berada di diagonal dibagi dengan banyaknya objek atau kasus yang diteliti disebut *hit ratio*.

Koefesien fungsi diskriminan yang tidak dilakukan merupakan pengganda variabel, ketika variabel masih dalam satu ukuran yang asli. Kalau variabel

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

dibakukan disebut koefisien beta. Skor diskriminan koefisien yang tidak dibakukan dikalikan dengan nilai variabel, dijumlahkan kemudian ditambah dengan konstan akan diperoleh nilai atau skor diskriminan sama dengan nilai fungsi diskriminan.

2.6 Fungsi Analisis Diskriminan

Fungsi diskriminan merupakan fungsi atau kombinasi linier peubah-peubah asal yang akan menghasilkan cara terbaik dalam pemisahan kelompok-kelompok. Fungsi ini akan memberikan nilai nilai yang sedekat mungkin dalam kelompok dan sejauh mungkin antar kelompok. Banyaknya fungsi diskriminan yang terbentuk secara umum tergantung dari $\min(p, k - 1)$, dengan p adalah banyaknya peubah pembeda dan k adalah banyaknya kelompok yang telah ditetapkan.

Fungsi diskriminan ini diartikan sebagai keragaman peubah yang terpilih sebagai kekuatan pembeda. Apabila fungsi diskriminan yang terbentuk sebanyak lebih dari satu fungsi, maka dapat dikatakan bahwa fungsi diskriminan pertama akan menjadi kekuatan pembeda yang paling besar, demikian berturut-turut untuk fungsi berikutnya.

2.7 Tujuan Analisis Diskriminan

1. Membuat suatu fungsi diskriminan atau kombinal linear, dari prediktor atau variable yang bisa mendiskriminasi atau membedakan kategori variable tak bebas *criterion* atau kelompok, artinya dapat membedakan suatu objek (responden) masuk kelompok kategori yang mana.
2. Menguji apakah perbedaan signifikan antara kategori/kelompok dengan variabel bebas atau prediktor.
3. Menentukan variabel bebas yang mana yang memberikan pengaruh terbesar terhadap terjadinya perbedaan antar kelompok
4. Mengklasifikasikan objek atau kasus ke dalam suatu kelompok didasarkan pada nilai variabel bebas.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

2.8 Asumsi dan Sampel dalam Analisis Diskriminan

Sejumlah p variabel independen harus berdistribusi normal tidak ada korelasi antar variabel independen. Tidak terdapat data yang *outlier* pada variabel independen.

Menurut Hair et al. (1987 : 76), analisis diskriminan tidak terlalu sensitif dengan pelanggaran asumsi ini, kecuali pelanggarannya bersifat ekstrim. Dan Johnson and Wichern (1988: 472) mengatakan hal yang sama bahwa asumsi ini (kesamaan ragam-peragam) di dalam praktiknya sering dilanggar.

Tidak ada jumlah sampel yang ideal secara pasti pada analisis diskriminan. Selain itu, pada analisis diskriminan sebaiknya digunakan dua jenis sampel, yakni analisis sampel yang digunakan untuk membuat fungsi diskriminan, serta *holdout* sampel (split sampel) yang digunakan untuk menguji hasil diskriminan.

2.9 Model Analisis Diskriminan

Model dasar analisis diskriminan sangat mirip dengan analisis faktor dan regresi berganda. Perbedaannya adalah kalau variabel dependen regresi berganda dilambangkan dengan Y , maka dalam analisis diskriminan dilambangkan dengan D . Model analisis diskriminan adalah sebuah persamaan yang menunjukkan suatu kombinasi linier dari berbagai variabel independen yaitu :

$$D = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_3 + \dots + b_k X_k$$

dengan :

D : skor diskriminan

B : koefisien diskriminasi atau bobot

X : prediktor atau variabel independen

Yang diestimasi adalah koefisien ' b ', sehingga nilai ' D ' setiap grup sedapat mungkin berbeda. Ini terjadi pada saat rasio jumlah kuadrat antargrup (*betweengroup sum of squares*) terhadap jumlah kuadrat dalam grup (*within-group sum of square*) untuk skor diskriminan mencapai maksimum. Berdasarkan nilai D itulah keanggotaan seseorang diprediksi.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Elemen sebagai objek penelitian kemudian diputuskan untuk dimasukkan kedalam kelompok tertentu berdasarkan skor diskriminan. Sebagai contoh di dalam analisis diskriminan dua kelompok, suatu objek akan dimasukkan kedalam kelompok tertentu yang centroidnya terdekat.

Hit ratio atau kasus yang secara benar atau tepat diklasifikasikan, kalau objek memang masuk dikelompok A dimasukkan ke A, kalau objek memang masuk dikelompok B dimasukkan ke B. *hit ratio* dihitung dengan jalan menjumlahkan seluruh elemen yang berada pada diagonal utama klasifikasi dibagi dengan banyaknya objek/elemen.

Kebanyakan program analisis diskriminan juga mengestimasi suatu matriks klasifikasi berdasarkan *analysis sample*, yang pada umumnya memberikan hasil klasifikasi yang lebih apabila dibandingkan dengan matriks klasifikasi yang didasarkan pada *holdout sample*.

2.11 Pengujian Hipotesis

Hipotesis yang akan diuji adalah H_0 yang menyatakan bahwa rata-rata semua variabel dalam semua grup adalah sama. Dalam SPSS, uji dilakukan dengan menggunakan Wilks Lamda. Jika dilakukan pengujian sekaligus beberapa fungsi sebagaimana dilakukan pada analisis diskriminan, statistik Wilk's adalah hasil univariat untuk setiap fungsi. Kemudian, tingkat signifikansi diestimasi berdasarkan *chi-square* yang telah ditransformasi secara statistik. dimana hipotesis yang digunakan adalah:

H_0 : Rata-rata variabel pada semua klasifikasi sama

H_1 : Rata-rata variabel pada semua klasifikasi sama

Setelah hasil analisis diketahui, kemudian dilihat apakah Wilks berasosiasi dengan fungsi diskriminan. Selanjutnya, angka ini ditransformasi menjadi *chi-square* dengan derajat kebebasan (*df*) yang akan digunakan dalam pengambilan kesimpulan dengan uji kriteria hipotesis berikut:

Jika F hitung $> F$ tabel maka H_0 ditolak dan H_1 diterima

Jika F hitung $\leq F$ tabel maka H_0 diterima dan H_1 ditolak

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

the *highest adjusted of partial F ratio* dengan memperhiungkan prediktor yang telah dipilih sebelumnya dan seterusnya.

Setiap prediktor yang telah terpilih diuji untuk retensi berdasarkan pada hubungannya dengan prediktor lain yang telah dipilih. Proses pemilihan dan retensi dilanjutkan sampai semua prediktor memenuhi kriteria signifikansi untuk dimasukkan dalam fungsi diskriminan. Beberapa statistik dihitung pada setiap langkah sebagai tambahan. Pada kesimpulan, suatu ringkasan mengenai prediktor yang masuk dan digeser dari fungsi diskriminan disediakan. Output yang berkaitan dengan metode langsung juga tersedia dari prosedur tahap demi tahap.

Pemilihan prosedur tahap demi tahap didasarkan pada criteria optimisasi yang dipergunakan, prosedur didasarkan pada maksimasi suatu ukuran umum mengenai jarak dua kelompok yang paling dekat. Prosedur ini memungkinkan peneliti pemasaran memanfaatkan secara maksimal data/informasi yang tersedia.

Urutan untuk mana variabel dipilih juga menunjukkan pentingnya variabel-variabel tersebut didalam mendiskriminasi antar-kelompok. Ini lebih lanjut dibenarkan dengan suatu kajian koefesien fungsi diskriminan yang dibakukan dan koefesien korelasi struktur.

2.13 Prosedur Analisis Diskriminan

Langkah pertama yang dilakukan pada analisis kelompok adalah uji kenormalan multivariat. Langkah kedua adalah uji multikolinieritas. Multikolinieritas adalah suatu keadaan dimana terdapat dua variabel atau lebih, saling berkorelasi. Dan langkah ketiga adalah Dalam analisis kelompok, pengelompokan data atau permasalahan dibutuhkan suatu ukuran yang dapat menerangkan kedekatan antara data. Dimana ukuran jarak yang biasa digunakan adalah ukuran jarak *Euclidean*. Setelah itu mulai melakukan analisis kelompok.

Analisis kelompok atau biasa dikenal sebagai *cluster analysis* adalah salah satu teknik statistik yang bertujuan untuk mengelompokkan objek ke dalam suatu kelompok sedemikian sehingga objek yang berada dalam satu kelompok akan memiliki kesamaan yang tinggi dibandingkan dengan objek yang berada di

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

kelompok lain (Sharma, 1996:185). Langkah pengelompokkan dalam analisis kelompok terdiri dari 3 langkah yaitu Mengukur kesamaan jarak, membentuk kelompok secara *K-Means*, menentukan jumlah kelompok.

Dengan melakukan uji normal multivariat, uji homogen, dan uji multikolinieritas, kita dapat melakukan analisis diskriminan. Heteroskedastisitas akan mengakibatkan penaksiran koefisien-koefisien regresi menjadi tidak efisien. Hasil penaksiran akan menjadi kurang dari semestinya. Asumsi homoekedastisitas berarti sama (homo) dan sebaran memiliki varian yang sama (Imam Ghozali, 2009:35). Jika nilai $T_{hitung} < T_{\alpha/2}(n - k - 1)$ atau $T_{hitung} > T_{\alpha/2}(n - k - 1)$ dan nilai signifikan > 0.05 maka mengindikasikan bahwa sisaaan tersebut tidak mengalami heteroskedastisitas (nilai tersebut homogen).

Analisis diskriminn merupakan salah satu metode analisis multivariat yang digunakan untuk mengetahui variabel-variabel ciri yang membedakan tiap-tiap kelompok yang terbentuk dan bertujuan untuk mengklasifikasikan beberapa kelompok data yang sudah terkelompokkan dengan cara membentuk kombinasi linier fungsi diskriminan.

Menurut Dillan (1984) dalam analisis diskriminan dengan p variabel yang diukur terdapat asumsi-asumsi yang harus dipenuhi untuk mendapatkan kombinasi linier fungsi diskriminan yang optimal dengan kesalahan klasifikasi terkecil adalah sebagai berikut:

- a. Data berdistribusi normal multivariat.
- b. Matriks varian kovarians antar kelompok homogen.

Setelah Fungsi diskriminan terbentuk, langkah selanjutnya adalah menilai validitas analisis diskriminan. Hasil klasifikasi dapat didasarkan pada analisis sampel ataupun validitas sampel dapat dilakukan dengan cara sebagai berikut.

- a. Menghitung diskriminan *score*

Setelah dibentuk fungsi liniernya, maka dapat dihitung skor diskriminan untuk tiap observasi dengan memasukkan nilai-nilai variabel penjelasnya.

- b. Menghitung *cutting score*

Cutting score (m) dapat dihitung dengan rumus sebagai berikut :

$$Z_{ce} = \frac{N_a Z_b + N_b Z_a}{N_a N_b}$$

Hak Cipta Dilindungi Undang-Undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Menguji hipotesis komparatif dua sampel independen berarti menguji signifikansi perbedaan nilai dua sampel yang tidak berpasangan. Sampel independen biasanya digunakan dalam penelitian yang menggunakan pendekatan penelitian survey, sedangkan sampel berpasangan banyak digunakan dalam penelitian *experiment*.

Contoh dua sampel independent :

- a. Sampel pengusaha ekonomi kuat dan ekonomi lemah.
- b. Sampel partai status quo dan partai reformis.
- c. Sampel pria dan wanita.

Contoh dua sampel dependent :

- a. Sampel pegawai sebelum dilatih dan setelah dilatih.
- b. Sampel konsumen yang dikenai iklan dan tidak.
- c. Sampel mahasiswa yang mendapat beasiswa dan tidak

2.15 Uji Asumsi Klasik

Sebelum melakukan teknik analisis diskriminan metode *fisher*, ada tiga asumsi yang harus dipenuhi. Pertama, yaitu uji normal multivariate pada masing-masing kelompok. Jika data tidak berdistribusi normal multivariat, maka dapat dilakukan transformasi box cox. Kedua, uji multikolinieritas pada seluruh data.

Uji multikolinieritas bertujuan untuk menguji apakah dalam model regresi ditemukan adanya korelasi antar variabel bebas. Jika terjadi multikolinieritas, maka dapat diatasi dengan mengganti atau mengeluarkan variabel yang memiliki korelasi tinggi.

Uji asumsi klasik yang ketiga yang harus dipenuhi adalah kesamaan matriks varians kovarians pada seluruh data. Jika matriks varians kovarians variabel bebas untuk semua kelompok tidak sama, maka dapat diatasi dengan transformasi box cox. Setelah ketiga uji asumsi terpenuhi, langkah selanjutnya adalah membentuk fungsi diskriminan dengan metode *fisher*.

2.15.1 Uji Normal Multivariat

Uji pertama yang dilakukan pada data pengamatan ini adalah uji kenormala multivariat terhadap masing-masing kelompok.

Menurut Karson (1982), untuk menguji kenormalan multivariat digunakan prosedur yang dikembangkan oleh Mardia (1970) dengan cara menghitung dua macam ukuran statistik yaitu ukuran *skewness* ($b_{1,p}$) dan *kurtosis* ($b_{2,p}$) yaitu:

$$b_{1,p} = (1/n^2) \sum \sum [(X_i - X)' S^{-1} (\bar{X}_i - X)]^3$$

$$b_{2,p} = (1/n) \sum [(X_i - X)' S^{-1} (\bar{X}_i - X)]^2$$

Hipotesis yang digunakan adalah:

H_0 : multivariat mengikuti sebaran normal

H_1 : multivariat tidak mengikuti sebaran normal

Bila ($b_{1,p} n/6 \leq \chi^2_{p(p+1)(p+2)/6}$, dan $[b_{2,p} - p(p+2)] / \sqrt{(8p(p+2)/n)} \leq Z_\alpha$ (tabel normal), maka H_0 diterima, berarti multivariat mengikuti sebaran normal.

Menurut Johnson (1982), untuk menguji kenormalan ganda adalah dengan mencari nilai jarak kuadrat untuk setiap pengamatan yaitu $d_i^2 = (X_i - X)' S^{-1}(X_i - X)$, di mana X_i adalah pengamatan yang ke $-i$ dan S^{-1} adalah kebalikan (*inverse*) matrik kovarians S .

Kemudian d_i^2 diurutkan dari kecil ke besar, selanjutnya dibuat *plot* d_i^2 dengan nilai Chi-Kuadrat $\chi^2_{(i-1/2)/n}$, dimana $i = 1, 2, \dots, n$, dan $p =$ banyaknya peubah. Bila hasil *plot* dapat didekati dengan garis lurus, maka dapat disimpulkan bahwa data menyebar secara normal ganda.

Menurut Norusis (1986), berdasarkan teori Wahl & Kronmal (1977), seringkali kenormalan ganda sulit diperoleh terutama bila sampel yang diambil relatif kecil. Bila hal ini terjadi, uji vektor rata-rata tetap bisa dilakukan selama asumsi kedua (kesamaan kovarians) dipenuhi.

2.15.2 Uji Kesamaan Vektor Rata-Rata

Pengujian multikolinieritas pada variabel bebas diperlukan untuk memeriksa apakah terjadi korelasi antar variabel bebas. Dalam penelitian ini

Kriteria Pengujian : Tolak H_0 jika sign. $< 0,05$ atau sebaliknya. Diharapkan dari uji ini adalah H_0 ditolak.

2.15.3 Uji Kesamaan Matriks Kovarians

Pengujian kehomogenan matriks kovarian gabungan kedua kelompok diperlukan untuk dapat membentuk fungsi diskriminan linier Fisher. Dengan menggunakan statistik uji Box'M.

Untuk menguji kesamaan matrik kovarians (Σ) antar kelompok, digunakan hipotesis:

$$H_0: \Sigma_1 = \Sigma_2$$

$$H_1: \Sigma_1 \neq \Sigma_2$$

Statistik uji yang digunakan untuk perhitungan ,yaitu:

$$M = \frac{|S_1|^{v_1/2} |S_2|^{v_2/2} \dots |S_k|^{v_k/2}}{|S_{pl}|^{\sum_i v_{pl}/2}}$$

dengan

$$S_{pl} = \frac{\sum_{i=1}^k v_i S_i}{\sum_{i=1}^k v_i} \text{ dan } v_i = n_i - 1$$

dan M dihitung melalui pendekatan distribusi χ^2 dan F. Pendekatan distribusi χ^2 melalui persamaan $u = -2(1 - c_1) \ln M$.

Berdistribusi

$$\chi^2_{\left[\frac{1}{2}(k-1)p(p+1)\right]} \text{ dengan } c_1 = \left[\sum_{i=1}^k \frac{1}{v_i} - \frac{1}{\sum_{i=1}^k v_i} \right] \left[\frac{2p^2+3p-1}{6(p+1)(k-1)} \right]$$

Pendekatan distribusi F dihitung bergantung pada nilai c_1 dan c_2 dengan

$$c_2 = \frac{(p-1)(p+2)}{6(k-1)} \left[\sum_{i=1}^k \frac{1}{v_i^2} - \frac{1}{(\sum_{i=1}^k v_i)^2} \right]$$

Juga dengan menghitung

$$a_1 = \frac{1}{2}(k-1)p(p+1)$$

$$a_2 = \frac{a_1+2}{[c_2-c_1^2]}$$

$$b_1 = \frac{1-c_1-a_1 a_2}{a_1}$$

$$b_2 = \frac{1-c_1+2/a_2}{a_2}$$

Jika $c_2 > c_1^2$ maka digunakan $F = -2b_1 \ln M$ yang didekati oleh F_{a_1, a_2}

Jika $c_2 < c_1^2$ maka digunakan $F = \frac{2a_2 b_2 \ln M}{a_1(1+2b_2 \ln M)}$ yang didekati oleh F_{a_1, a_2}

Kriteria pengujian : Tolak H_0 jika sign. $< 0,05$ atau Terima H_0 jika sign. $> 0,05$

2.16 Analisis Diskriminan Metode Fisher

Prinsip utama dari fungsi diskriminan fisher adalah pemisahan sebuah populasi. Fungsi diskriminan yang terbentuk dapat digunakan untuk pengelompokan suatu observasi berdasarkan kelompok-kelompok tertentu.

Metode fisher ini tidak mengasumsikan data harus berdistribusi normal, tapi dalam perhitungan dalam suatu syarat yang harus diperhatikan adalah data yang digunakan harus memiliki matriks kovarians yang sama untuk setiap kelompok populasi yang diberikan.

Metode fisher adalah suatu metode yang bertujuan untuk membentuk fungsi diskriminan dengan pemilihan koefesien-koefesien yang menghasilkan hasil bagi maksimum antara variasi antar kelompok dan variasi dalam kelompok.

Analisis diskriminan linier (LDA) Bergantung pada mean sampel dan matriks kovarians dihitung dari kelompok yang berbeda dari pelatihan Sampel.

Analisis diskriminan linear Fisher (FLDA) adalah kombinasi linear dari pengamatan atau Variabel terukur yang paling menggambarkan perpisahan antara kelompok pengamatan yang diketahui. Tujuan dasarnya Adalah untuk mengklasifikasikan atau memprediksi masalah dimana tergantung variabel muncul dalam bentuk kualitatif . Fisher linier Analisis diskriminan adalah multivariat konvensional Teknik untuk pengurangan dimensi dan klasifikasi.

2.17 Pembentukan Fungsi Linear Fisher

Fisher mengelompokkan suatu observasi berdasarkan nilai skor yang dihitung dari suatu fungsi linier $Y = \lambda' X$ dimana λ' menyatakan vektor yang berisi

koefisien-koefisien variabel penjelas yang membentuk persamaan linier terhadap variabel respon, $\lambda' = [\lambda_1, \lambda_2, \dots, \lambda_p]$.

$$\mathbf{X} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$$

X_k menyatakan matriks data pada kelompok ke- k

$$X_k = \begin{bmatrix} X_{11k} & X_{12k} & \cdots & X_{1pk} \\ X_{21k} & X_{22k} & \cdots & X_{2pk} \\ \vdots & \vdots & \vdots & \vdots \\ X_{n1k} & X_{n2k} & \cdots & X_{npk} \end{bmatrix}$$

$$i = 1, 2, \dots, n$$

$$j = 1, 2, \dots, p$$

$$k = 1, 2$$

X_{ijk} menyatakan observasi ke- i variabel ke- j pada kelompok ke- k .

di bawah asumsi $X_k \sim N(\mu_k, \Sigma_k)$ maka

$$\mu = \begin{bmatrix} E(X_1) \\ E(X_2) \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \text{ dan } \Sigma_k = E(X_k - \mu_k)(X_k - \mu_k)'; \Sigma_1 = \Sigma_2 = \Sigma$$

$$\mu_k = \begin{bmatrix} \mu_{1k} \\ \vdots \\ \mu_{pk} \end{bmatrix}; \mu_k \text{ adalah vector rata-rata tiap variabel } X \text{ pada kelompok ke-}k$$

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \cdot & \sigma_{22} & \cdots & \sigma_{2p} \\ & * & & \vdots \\ & & & \sigma_{pp} \end{bmatrix}$$

$$\sigma_{j_1 j_2} = \begin{cases} \text{varians variabel } j \text{ apabila } j_1 = j_2 \\ \text{ko varians variabel } j_1 \text{ dan } j_2 \text{ apabila } j_1 \neq j_2 \end{cases}$$

Fisher mentransformasikan observasi-observasi x yang multivariate menjadi observasi y yang univariate. Dari persamaan $\mathbf{Y} = \lambda \mathbf{X}$ diperoleh

$$\mu_{ky} = E(Y_k) = E(\lambda' \mathbf{X}) = \lambda' \mu_k;$$

$$\sigma_Y^2 = \text{var}(\lambda' \mathbf{X}) = \lambda' \Sigma \lambda$$

μ_{ky} adalah rata-rata Y yang diperoleh dari X yang termasuk dalam kelompok ke- k
 σ_Y^2 adalah varians Y dan diasumsikan sama untuk kedua kelompok.

Kombinasi linier yang terbaik menurut Fisher adalah yang dapat memaksimalkan rasio antara jarak kuadrat rata-rata Y yang diperoleh dari X dari kelompok 1 dan 2 dengan varians Y , atau dirumuskan sebagai berikut :

$$\frac{(\mu_{1y} - \mu_{2y})^2}{\sigma_Y^2} = \frac{\lambda' (\mu_1 \mu_2) (\mu_1 \mu_2)' \lambda}{\lambda' \Sigma \lambda}$$

Jika $(\mu_1 - \mu_2) = \delta$ maka persamaan di atas menjadi $\frac{(\lambda' \delta)^2}{\lambda' \Sigma \lambda}$. Karena Σ adalah matriks definit positif, maka menurut teori pertidaksamaan *Cauchy-Schwartz*, rasio $\frac{(\lambda' \delta)^2}{\lambda' \Sigma \lambda}$ dapat dimaksimumkan jika

$$\lambda' = c \Sigma^{-1} \delta = c \Sigma^{-1} (\mu_1 - \mu_2)$$

dengan memilih $c=1$, menghasilkan kombinasi linier yang disebut kombinasi linier Fisher sebagai berikut:

$$Y = \lambda' X = (\mu_1 - \mu_2) \Sigma^{-1} X$$

2.18 Analisis Diskriminan untuk Dua Grup atau Kelompok

Fisher menyarankan untuk mengubah multivariate pengamatan terhadap observasi univariat sehingga pengamatan univariat yang berasal dari masing-masing populasi adalah maksimal dipisahkan pemisahan ini univariat pengamatan dapat diperiksa dengan mean mereka perbedaan Aturan klasifikasi Fisher memaksimalkan variasi antara variabilitas sampel sampai dalam sampel variabilitas Analisis diskriminan linier Fisher untuk dua kelompok.

Misalnya kita memiliki dua kelompok populasi yang bebas. Dari populasi 1 diambil secara acak contoh berukuran n_1 dan mempelajari p buah sifat dari contoh itu, demikian pula ditarik contoh acak berukuran n_2 dari populasi 2 serta mempelajari p buah sifat dari contoh itu. dengan demikian ukuran contoh secara keseluruhan dari populasi 1 dan populasi 2 adalah $n = n_1 + n_2$.

Misalkan p adalah sebuah sifat dipelajari itu dinyatakan dalam variable acak berdimensi ganda melalui vektor $p \times X$, X_1, \dots, X_p dalam bentuk catatan matriks dapat dinyatakan sebagai berikut :

$$\bar{X}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} X_{ij}$$

$$S_i = \frac{1}{n_i} \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)(X_{ij} - \bar{X}_i)'$$

$$S_{pooled} = \frac{\sum_{i=1}^{g-2} (n_i - 1) S_i}{\sum_{i=1}^{g-2} (n_i - g)}$$

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengemukakan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Karena diasumsikan bahwa populasi induk memiliki peragam yang sama yaitu Σ , maka matriks peragam contoh S_1 dan S_2 dapat digabung untuk memperoleh matriks gabungan sebagai penduga bagi σ^2 melalui rata-rata terbobot berikut : $S_{pooled} = \frac{(n_1-1)S_1 + (n_2-1)S_2}{n_1+n_2-2}$

Pengujian perbedaan vektor nilai rata-rata di antara dua populasi dilakukan dengan jalan merumuskan hipotesis berikut :

$H_0 : \underline{U}_1 = \underline{U}_2$ artinya vektor nilai rata-rata dari populasi 1 sama dengan dari populasi 2.

$H_1 : \underline{U}_1 \neq \underline{U}_2$ artinya kedua vektor nilai rata-rata berbeda.

Pengujian terhadap hipotesis dilakukan menggunakan uji statistic T^2 - Hotelling yang dirumuskan sbb :

$$T^2 = \frac{n_1 n_2}{n_1 + n_2} (\bar{X}_1 - \bar{X}_2)' S_G^{-1} (\bar{X}_1 - \bar{X}_2)$$

Selanjutnya besaran : $F = \frac{n_1+n_2-p-1}{(n_1+n_2-2)p} T^2$

Jika $T^2 \leq \frac{(n_1+n_2-2)p}{n_1+n_2-p-1}$ akan berdistribusi F dengan derajat bebas $v_1 = p$ dan $v_2 = n_1 + n_2 - p - 1$

Kriteria Uji

Terima $H_0 : \underline{U}_1 = \underline{U}_2$; Jika $T^2 \leq \frac{(n_1+n_2-2)p}{n_1+n_2-p-1} F_{\alpha;v_1v_2}$ selain itu tolak H_0 .

Alternatif lain **kriteria Uji** :

Terima $H_0 : \underline{U}_1 = \underline{U}_2$; Jika $F \leq F_{\alpha;v_1v_2}$

Tolak H_0 : Jika $F_{\alpha;v_1v_2}$

Jika terima H_0 artinya vektor nilai rata-rata dari populasi 1 sama dengan dari populasi 2.

Jika tolak H_0 artinya kedua vektor nilai rata-rata berbeda.