ANALISIS DISKRIMINAN UNTUK MELIHAT MISKLASIFIKASI MAHASISWA FMIPA ANGKATAN 1998/1999 UNIVERSITAS JEMBER

SKRIPSI

Diajukan Guna Memenuhi Persyaratan Memperoleh Gelar Sarjana Sains
Jurusan Matematika Fakultas Matematika Dan Ilmu Pengetahuan Alam
Universitas Jember

Oleh :

Ratna Kusumawati

NIM. 971810101012

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS JEMBER
Oktober, 2002

MOTTO

- Matematika adalah ratu dari ilmu dan ilmu hitung (aritmetika) adalah ratu dari matematika. Ia sering berkenan merendahkan diri menyumbang kepada astronomi dan ilmu alam lainnya, tetapi dalam semua hubungan ia berhak mendapat peringkat pertama.

C. F. Gauss

- Sasaran utama dari semua penyelidikan tentang dunia luar seharusnya menemukan urutan rasional dan keserasian yang telah ditekankan padanya oleh Tuhan dan yang Ia ungkapkan kepada kita dalam bahasa matematika.

J. Keppler

KUPERSEMBAHKAN SKRIPSI INI

KEPADA

- Ayah dan Ibuku yang sangat ananda cintai dan hormati, yang tiada sunyi memberikan doa, sebagai tanda bakti dan terima kasih atas segala ketulusan, kesabaran dan pengorbanannya.
- Kakak dan adikku yang tercinta, Om dan Tanteku yang aku hormati, yang aku sayangi dan seluruh saudara – saudaraku, terima kasih atas segala dukungan yang diberikan,
- Rental Techno Computer dan Bintang Computer, terimakasih atas pelayanannya.
- Teman-teman seperjuangan.
- Almamater dan Tanah Airku Tercinta.

DEKLARASI

Skripsi ini berisi hasil kerja/penelitian mulai 28 Januari - 5 Februari 2002 di bagian akademik Fakultas MIPA Universitas Jember. Bersama ini saya menyatakan bahwa isi skripsi ini adalah hasil pekerjaan saya sendiri kecuali jika disebutkan sumbernya dan skripsi ini belum pernah diajukan pada institusi lain.

Jember, Oktober 2002

Ratna Kusumawati

ABSTRAKSI

Ratna Kusumawati, Oktober, 2002, " **Analisis Diskriminan Untuk Melihat Misklasifikasi Mahasiswa FMIPA Angkatan 1998/1999 Universitas Jember** ". Skripsi, Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Jember.

Pembimbing I : Drs. I Made Tirta, Dip.Sc, M.Sc, PhD.
Pembimbing II : Yuliani S. Dewi, S.Si, M.Si.

Kata Kunci : Fungsi diskriminan, Koefisien yang dibakukan, Misklasifikasi.

Penelitian ini bertujuan untuk membentuk sejumlah fungsi melalui kombinasi linier peubah-peubah asal, yang dapat digunakan sebagai cara terbaik untuk memisahkan kelompok-kelompok mahasiswa, fungsi ini disebut fungsi diskriminan. Fungsi diskriminan dapat digunakan untuk menerangkan perbedaan antar kelompok atau menterjemahkan ciri kelompok, juga dapat digunakan dalam masalah klasifikasi (misklasifikasi).

Ada perbedaan yang signifikan (kurang dari 1%) untuk ketiga fungsi diskriminan yang menerangkan perbedaan antara kelompok matematika, fisika, kimia dan biologi, sehingga analisis diskriminan layak untuk digunakan. Variabel yang membedakan misklasifikasi awal pengelompokan adalah nilai rata-rata matematika (y_2), nilai rata-rata fisika (y_4), nilai NEM kimia (y_5), nilai rata-rata kimia (y_6) dan nilai rata-rata biologi (y_8). Ada tiga fungsi diskriminan yang terbentuk, yakni:

$$Z_{score}\_1 = -0{,}863*y_2 + 0{,}471*y_4 + 0{,}522*y_5 + 0{,}464*y_6 - 0{,}426*y_8$$

$$Z_{score}\_2 = -0{,}080*y_2 - 0{,}656*y_4 + 0{,}330*y_5 + 0{,}391*y_6 + 0{,}704*y_8$$

$$Z_{score}\_3 = -0{,}307*y_2 + 0{,}284*y_4 + 0{,}679*y_5 - 1{,}045*y_6 + 0{,}479*y_8$$

Dari hasil analisis terjadi misklasifikasi mahasiswa FMIPA angkatan 1998/1999 Universitas Jember sebesar 46,4%. Angka yang relatif tinggi mungkin disebabkan tanpa memperhatikan profil kemampuan akademik dari mata kuliah yang terkait (baik NEM, mata kuliah dari kemipaan).

Melihat adanya misklasifikasi yang relatif tinggi, sebaiknya pihak FMIPA Universitas Jember memberikan informasi dan wawasan yang luas mengenai prospeks dan prasyarat dari masing-masing jurusan di masa yang akan datang untuk pelajar SMU yang akan mengikuti UMPTN serta memberikan motivasi cara belajar yang efektif untuk menghasilkan lulusan yang berkualitas.

Sistem penjurusan mahasiswa FMIPA Universitas Jember ada baiknya mempertimbangkan prestasi akademik awal (tahun I) setelah di Perguruan Tinggi selain minat dan profil NEM.

PENGESAHAN

Skripsi ini telah dipertahankan didepan Tim Penguji dan diterima oleh Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Jember pada :

Hari : Selasa

Tanggal : 19 NOV 2002

Tempat : Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Jember.

Tim Penguji

Ketua

Drs. I Made Tirta, Dip.Sc, M.Sc, PhD
NIP. 131 474 500

Sekretaris

Yuliani Setia Dewi, S.Si, M.Si.
NIP. 132 258 183

Anggota :

Rita Ratih T., S.Si, M.Si
NIP. 132 243 343

Agustina Pradjaningsih, S.Si, M.Si
NIP. 132 257 933

Mengesahkan
Dekan FMIPA UNEJ

Ir. Sumadi,
NIP. 130 368 784

KATA PENGANTAR

Puji syukur penulis panjatkan kehadirat Allah SWT, yang telah melimpahkan rahmat dan hidayah-Nya, sehingga penulis dapat menyelesaikan penyusunan skripsi ini. Sholawat serta salam semoga tetap tercurahkan kepada Nabi Muhammad SAW, dan orang – orang yang selalu berada dijalannya.

Selanjutnya penulis sampaikan terima kasih dan penghargaan yang setulusnya atas bantuan yang tidak ternilai kepada :

1. Bapak Ir. Sumadi, MS. Selaku Dekan Fakultas MIPA Universitas Jember.
2. Bapak Drs. Kusno, DEA, PhD. Selaku Ketua Jurusan Matematika FMIPA Universitas Jember.
3. Bapak Drs. I Made Tirta, Dip.Sc, M Sc, Ph.D. Selaku Pembimbing Utama yang dengan penuh kesabaran dan ketulusan membimbing penulis dalam penyusunan skripsi ini.
4. Ibu Yuliani S. Dewi, S.Si, M.Si. Selaku Pembimbing Anggota, yang telah banyak memberikan saran dan petunjuk bagi penulisan skripsi ini.
5. Ibu Rita Ratih T, S.Si, M.Si dan Ibu Agustina Pradjaningsih, S.Si, M.Si. yang telah bersedia menjadi Dosen Penguji.
6. Bapak Drs. Adi Supriono. Selaku Kasubbag. Pendidikan FMIPA Universitas Jember yang telah memberikan ijin untuk mengambil data.
7. Ayahanda dan Ibunda tercinta, Om dan Tanteku, Kakakku, Adikku yang selalu menemani penulis, dan saudara – saudaraku yang selalu memberikan dorongan moril pada penulis.

8. Teman – temanku Mahasiswa Math '97, teman – teman seperjuangan dalam penyusunan skripsi, semoga sukses senantiasa bersama kita.
9. Rekan – rekan dan semua pihak yang telah membantu dan memotivasi dalam penyusunan skripsi ini.

Semoga segala bantuan dan kebaikan yang telah diberikan kepada penulis akan mendapatkan imbalan yang setimpal dari Allah SWT. Amin ...

Akhir kata, penulis berharap semoga apa yang penulis tuangkan dalam skripsi yang sederhana ini bermanfaat bagi pembaca.

Jember, Oktober 2002

Penulis

DAFTAR ISI

DAFTAR TABEL

DAFTAR TABEL

DAFTAR GAMBAR

BAB I
PENDAHULUAN

1.1 Latar Belakang

Analisis diskriminan pertama kali diperkenalkan olah R. A. Fisher (1936) sebagai suatu teknik statistik yang berguna dalam bidang taksonomi dan setelah itu banyak dikembangkan pada riset-riset statistik (Kleinbum, 1987) dalam buku Siswandi dan Budi Suharjo (1999:11).

Menurut Tatsuoka (1971) dalam buku Siswandi dan Budi Suharjo (1999:11) masalah yang dihadapi dalam analisis statistika peubah ganda secara umum adalah bagaimana mempelajari faktor penentu yang membedakan populasi (kelompok) atau mendapatkan kombinasi linear dari peubah-peubah (atribut-atribut) yang diukur untuk menunjukkan ukuran perbedaan dalam nilai tengah populasi tersebut. Analisis diskriminan merupakan suatu metode yang dapat digunakan untuk mendapatkan setiap kombinasi linear dari peubah-peubah tersebut yang kemudian digunakan untuk pemisahan yang terbaik antar kelompok individu.

Menurut Carl.J.Huberty ada empat permasalahan khusus yang berhubungan dengan analisis diskriminan adalah:

1. data hilang (*missing data*);
2. pencilan dan unit yang berpengaruh (*outlier and influential units*);
3. misklasifikasi awal pengelompokan (*initial grup misclassifications*);
4. pendugaan klinis dibandingkan menurut statistik (*statistical versus clinical predictions*) (1994:270-276).

Salah satu yang dipelajari dalam analisis diskriminan adalah misklasifikasi awal pengelompokan. Misalnya misklasifikasi mahasiswa Fakultas MIPA Universitas Jember dilatar belakangi oleh nilai nem yang berbeda, nilai IPK dan jurusan yang dimasukinya tidak sesuai dengan kondisi pada waktu UMPTN diumumkan. Pengumuman hasil UMPTN diasumsikan benar sebelum dianalisis.

Berdasarkan dari uraian diatas penulis mencoba mengadakan penelitian dengan judul **"Analisis Diskriminan Untuk Melihat *Misklasifikasi* Mahasiswa FMIPA Angkatan 1998/1999 Universitas Jember"**.

1.2 Perumusan Masalah

Beberapa permasalahan yang dibahas dalam penelitian ini dibatasi misklasifikasi awal pengelompokan yang dirumuskan sebagai berikut ini.

1. Apakah ada faktor-faktor yang mendominasi terjadinya misklasifikasi mahasiswa FMIPA angkatan 1998/1999 Universitas Jember ?
2. Bagaimana memprediksikan klasifikasi mahasiswa FMIPA angkatan 1998/1999 Universitas Jember ?

1.3 Tujuan Penelitian

Tujuan yang ingin dicapai dalam penelitian ini adalah:

1. mengetahui ada faktor-faktor yang mendominasi terjadinya misklasifikasi mahasiswa FMIPA angkatan 1998/1999 Universitas Jember;
2. memprediksikan klasifikasi mahasiswa FMIPA angkatan 1998/1999 Universitas Jember .

1.4 Manfaat Penelitian

Hasil penelitian ini memberikan informasi tentang penerapan ilmu statistika, khususnya mengenai misklasifikasi awal pengelompokan mahasiswa pada setiap jurusan di FMIPA Universitas Jember dengan menggunakan analisis diskriminan.

Berdasarkan dari uraian diatas penulis mencoba mengadakan penelitian dengan judul **"Analisis Diskriminan Untuk Melihat *Misklasifikasi* Mahasiswa FMIPA Angkatan 1998/1999 Universitas Jember"**.

1.2 Perumusan Masalah

Beberapa permasalahan yang dibahas dalam penelitian ini dibatasi misklasifikasi awal pengelompokan yang dirumuskan sebagai berikut ini.

1. Apakah ada faktor-faktor yang mendominasi terjadinya misklasifikasi mahasiswa FMIPA angkatan 1998/1999 Universitas Jember ?
2. Bagaimana memprediksikan klasifikasi mahasiswa FMIPA angkatan 1998/1999 Universitas Jember ?

1.3 Tujuan Penelitian

Tujuan yang ingin dicapai dalam penelitian ini adalah:

1. mengetahui ada faktor-faktor yang mendominasi terjadinya misklasifikasi mahasiswa FMIPA angkatan 1998/1999 Universitas Jember;
2. memprediksikan klasifikasi mahasiswa FMIPA angkatan 1998/1999 Universitas Jember .

1.4 Manfaat Penelitian

Hasil penelitian ini memberikan informasi tentang penerapan ilmu statistika, khususnya mengenai misklasifikasi awal pengelompokan mahasiswa pada setiap jurusan di FMIPA Universitas Jember dengan menggunakan analisis diskriminan.

BAB II

TINJAUAN PUSTAKA

2.1 Definisi Analisis Diskriminan

Analisis diskriminan merupakan suatu analisis dengan tujuan membentuk sejumlah fungsi melalui kombinasi linear peubah-peubah asal, yang dapat digunakan sebagai cara terbaik untuk memisahkan kelompok-kelompok individu. Fungsi yang terbentuk melalui analisis ini kemudian disebut sebagai fungsi diskriminan.

Menurut Donald.R.Coopeer (1999:151) analisis diskriminan yaitu menghubungkan satu kriteria atau variabel dependen berskala nominal dengan satu atau beberapa variabel independen yang berskala interval atau rasio.

Sedangkan menurut Alvin C Rencher (1996:296) fungsi diskriminan adalah kombinasi linier dari pemisahan variabel-variabel yang digunakan menggambarkan atau menjelaskan perbedaan antara 2 atau lebih pengelompokan.

Dalam prosesnya fungsi diskriminan akan memberikan nilai-nilai yang sedekat mungkin dalam kelompok dan sejauh mungkin antar kelompok. Manfaat lain dari fungsi diskriminan ini disamping dapat digunakan untuk menerangkan perbedaan antar kelompok juga dapat digunakan dalam masalah klasifikasi.

Bila diketahui dengan jelas adanya kelompok-kelompok individu dengan keanggotaannya, masalah yang akan ditelusuri dalam analisis diskriminan adalah:

1. mencari cara terbaik untuk menyatakan perbedaan antar kelompok tersebut (masalah diskriminasi);
2. cara untuk mengalokasikan suatu individu baru kedalam salah satu kelompok tersebut (masalah klasifikasi).

Asumsi yang harus dipenuhi dalam analisis diskriminan:

1. Multivariat normal, atau variabel independen seharusnya berdistribusi normal. Jika data tidak berdistribusi normal, hal ini akan menyebabkan masalah pada ketepatan fungsi diskriminan.
2. Matrik kovarian dari semua variabel independen seharusnya sama.

Jadi analisis fungsi diskriminan digunakan untuk menentukan variabel mana sebagai penduga terbaik dari beberapa kumpulan karakteristik dengan berbagai jenis tipe variabel.

2.2 Bentuk / Jenis Analisis Diskriminan

Ada dua bentuk / jenis analisis fungsi diskriminan yaitu :

a. Analisis Fungsi Diskriminan Untuk 2 Pengelompokan

Kita asumsikan bahwa 2 populasi dapat disamakan mempunyai matrik peragam (*kovarian*) sama Σ, tetapi vektor nilai tengah (*mean*) μ_1 dan μ_2. Kita menggunakan sampel dari populasi pertama $\mathbf{y}_{11}, \mathbf{y}_{22}, \ldots, \mathbf{y}_{1n_1}$ dan populasi kedua adalah $\mathbf{y}_{11}, \mathbf{y}_{22}, \ldots, \mathbf{y}_{1n_2}$. Masing-masing vektor $\mathbf{y}_{ij}$ terdiri dari p variabel. Misalnya $\mathbf{y}_{11} = (y_{111}, y_{112}, \ldots, y_{11p})$

Fungsi diskriminan merupakan kombinasi linier dari p variabel dengan jarak vektor nilai tengah maksimum untuk dua pengelompokan. Kombinasi linier $z = \mathbf{a}'\mathbf{y}$ mentransformasi masing-masing pengamatan dari bentuk vektor menjadi skalar.

$$z_{1j} = \mathbf{a}'\mathbf{y}_{1j} = a_1 y_{1j1} + a_2 y_{1j2} + \ldots + a_p y_{1jp}, \qquad j = 1, 2, \ldots, n_1$$

$$z_{2j} = \mathbf{a}'\mathbf{y}_{2j} = a_1 y_{2j1} + a_2 y_{2j2} + \ldots + a_p y_{2jp}, \qquad j = 1, 2, \ldots, n_2.$$

Jika $n_1 + n_2$ merupakan vektor observasi dalam 2 sampel,

$$\mathrm{n}_1 = \begin{bmatrix} \mathbf{y}_{11} \\ \mathbf{y}_{12} \\ \vdots \\ \mathbf{y}_{1n_1} \end{bmatrix}_{p\times 1} \qquad \mathrm{n}_2 = \begin{bmatrix} \mathbf{y}_{21} \\ \mathbf{y}_{22} \\ \vdots \\ \mathbf{y}_{2n_2} \end{bmatrix}_{p\times 1} \qquad \mu = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix}$$

ditransfer ke bentuk skalar untuk $n_1 + n_2$

$$\begin{bmatrix} z_{11} \\ z_{12} \\ \vdots \\ z_{1n_1} \end{bmatrix} \qquad \begin{bmatrix} z_{21} \\ z_{22} \\ \vdots \\ z_{2n_1} \end{bmatrix}$$

dengan $z_{11} = \mathbf{a}'\mathbf{y}_{11}$.

Nilai tengah untuk z sampel dari n variabel yaitu $\bar{z} = \frac{1}{n}\sum_{i=1}^{n} z_i = \mathbf{a}'\bar{\mathbf{y}}$ dengan vektor $\bar{\mathbf{y}}$ sebagai nilai tengah sampel $\mathbf{y}_1, \mathbf{y}_2, ... \mathbf{y}_n$. Sedangkan varian sampel (s^2) dari $z_i = \mathbf{a}'\mathbf{y}_i$ dapat ditemukan sebagai varian sampel $z_1, z_2, \ldots, z_n$ atau langsung diganti dengan **a** dan **S**, dengan **S** adalah kovarian sampel dari $\mathbf{y}_1, \mathbf{y}_2, ... \mathbf{y}_n$ maka $s_z^2 = \frac{\sum_{i=1}^{n}(z_i - \bar{z})^2}{n-1} = \mathbf{a}'\mathbf{S}\mathbf{a}$. Sebuah varian selain tidak negatif, karena $s_z^2 \geq 0$ dan $\mathbf{a}'\mathbf{S}\mathbf{a} \geq 0$, untuk setiap vektor **a**. Makna **S** ini adalah paling sedikit semi definit positif. Jika variabel kontinu dan tidak linier dengan $n - 1 > p$, maka **S** adalah definit positif.

Kita menentukan nilai tengah $\bar{z}_1 = \sum_{j=1}^{n_1} \frac{z_{1j}}{n_1} = \mathbf{a}_1'\bar{\mathbf{y}}_1$ dan $\bar{z}_2 = \mathbf{a}_2'\bar{\mathbf{y}}_2$ dan menentukan vektor **a** dengan perbedaan yang dibakukan secara maksimal $\frac{(\bar{z}_1 - \bar{z}_2)}{S_z}$. Sebab $\frac{(\bar{z}_1 - \bar{z}_2)}{S_z}$ akan menjadi negatif, kita menggunakan jarak yang dikuadratkan $\frac{(\bar{z}_1 - \bar{z}_2)^2}{s_z^2}$ karena $s_z^2 \geq 0$ dan definit positif, dapat ditulis sebagai berikut :

$$\frac{(\bar{z}_1 - \bar{z}_2)^2}{s_z^2} = \frac{[\mathbf{a}'(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)]^2}{\mathbf{a}'\mathbf{S}_{pl}\mathbf{a}} \qquad \text{.............................} \qquad (2.1)$$

dengan

s_z = simpangan baku dari z

$\mathbf{S}_{pl}$ = simpangan baku yang diratakan dari $\mathbf{y}_1$ dan $\mathbf{y}_2$

Untuk memperoleh $\mathbf{S}_{pl}$ kita menggunakan $\mathbf{W}_1$ dan $\mathbf{W}_2$ sebagai matrik jumlah kuadrat dan hasil kali (*sum of squares and cross products)* untuk dua sampel

$$\mathbf{W}_1 = \sum_{j=1}^{ni} (\mathbf{y}_{1j} - \overline{\mathbf{y}}_1)(\mathbf{y}_{1j} - \overline{\mathbf{y}}_1)' = (n_1 - 1)\mathbf{S}_1$$

$$\mathbf{W}_2 = \sum_{j=2}^{n2} (\mathbf{y}_{2j} - \overline{\mathbf{y}}_2)(\mathbf{y}_{21} - \overline{\mathbf{y}}_2)' = (n_2 - 1)\mathbf{S}_2$$

dengan $(n_1\text{-}1)\mathbf{S}_1$ adalah penduga tidak bias $(n_1\text{-}1)\Sigma$ dan $(n_2\text{-}1)\mathbf{S}_2$ adalah penduga tidak bias $(n_2\text{-}1)\Sigma$, kita dapat menggabung kedua penduga tersebut untuk memperoleh penduga tidak bias matrik peragam Σ populasi. Maksud tidak bias yaitu statistik yang secara rata-rata menduga paremeter sebenarnya.

$$\mathbf{S}_{\mathrm{pl}} = \frac{1}{n_1 + n_2 - 2}(\mathbf{W}_1 - \mathbf{W}_2) = \frac{1}{n_1 + n_2 - 2}[(n_1 - 1)\mathbf{S}_1 + (n_2 - 1)\mathbf{S}_2]$$

diasumsikan $\Sigma_1 = \Sigma_2$.

Persamaan (2.1) akan maksimum jika

$$\mathbf{a} = \mathbf{S}_{\mathrm{pl}}^{-1}(\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2), \qquad \text{...} \quad (2.2)$$

sehingga pemaksimalan vektor kolom **a** tidak tunggal. Tetapi arahnya, adalah tunggal; yaitu nilai relatif atau rasio dari $a_1, a_2, \ldots, a_p$ adalah tunggal. Karena untuk $\mathbf{S}_{\mathrm{pl}}^{-1}$ *exist* dan $n_1 + n_2 - 2 > p$.

Untuk menghasilkan hasil maksimal maka (2.2) disubtitusikan ke (2.1) sehingga diperoleh

$$\frac{(\overline{z}_1 - \overline{z}_2)^2}{s_z^2} = \frac{[\mathbf{a}'(\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2)]^2}{\mathbf{a}'\mathbf{S}_{\mathrm{pl}}\mathbf{a}}$$

$$= \frac{(\mathbf{a}')^2(\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2)^2}{\mathbf{a}'\mathbf{S}_{\mathrm{pl}}\dfrac{(\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2)}{\mathbf{S}_{\mathrm{pl}}}}$$

$$= \mathbf{a}'(\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2)$$

$$= \left[\frac{(\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2)}{\mathbf{S}_{pl}^{-1}}\right]^t (\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2)$$

$$\frac{(\overline{z}_1 - \overline{z}_2)^2}{s_z^2} = (\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2)'\ \mathbf{S}_{\mathrm{pl}}^{-1}(\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2) \qquad \text{......................................} \quad (2.3)$$

untuk $z = \mathbf{a}'\mathbf{y}$ dengan $\mathbf{a} = \mathbf{S}_{pl}^{-1}(\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2)$. Jadi $\frac{(\bar{z}_1 - \bar{z}_2)^2}{s_z^2} = \mathbf{a}'(\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2)$ dan petunjuk lain yang digambarkan oleh $\mathbf{a} = \mathbf{S}_{pl}^{-1}(\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2)$ akan mempunyai hasil jarak yang kecil antara $\mathbf{a}'\overline{\mathbf{y}}_1$ dan $\mathbf{a}'\overline{\mathbf{y}}_2$.

w sebuah variabel (yang diidentifikasikan sebagai kelompok 1 dan 2) dan didefinisikan $\mathbf{b} = (b_1, b_2, \ldots, b_p)'$ sebagai vektor koefisien regresi w terhadap y. Vektor $\mathbf{b}$ adalah proporsi dari fungsi diskriminan dengan vektor koefisien $\mathbf{a} = \mathbf{S}_{pl}^{-1}(\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2)$:

$$\mathbf{b} = \frac{n_1 n_2}{(n_1 + n_2)(n_1 + n_2 - 2) + n_1 n_2 D^2} \cdot \mathbf{a} \qquad \text{.............................. (2.4)}$$

dengan $D^2 = (\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2)' \; \mathbf{S}_{pl}^{-1}(\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2) = (\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2)\mathbf{a}$ yang merupakan jarak antara dua sampel. Koefisien korelasi berganda R (*multiple correlation coefficient* R),jika dihubungkan pada D^2, diperoleh

$$\begin{aligned} R^2 = (\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2)'\mathbf{b} &= (\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2)^t \frac{(n_1 n_2)}{(n_1 + n_2)(n_1 + n_2 - 2) + n_1 n_2 D^2}\mathbf{a} \\ &= \mathbf{a}(\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2)' \frac{(n_1 n_2)}{(n_1 + n_2)(n_1 + n_2 - 2) + n_1 n_2 D^2}\mathbf{a} \\ &= \frac{D^2 n_1 n_2}{(n_1 + n_2)(n_1 + n_2 - 2) + n_1 n_2 D^2} \qquad \text{.............................. (2.5)} \end{aligned}$$

Uji statistik untuk hipotesis q variabel dan $p + q$ variabel yang digunakan untuk pemisahan pengelompokan seperti melalui regresi

$$F = \frac{n_1 + n_2 - p - q - 1}{q} \; \frac{R_{p+q}^2 - R_p^2}{1 - R_{p+q}^2} \qquad \text{...................................... (2.6)}$$

dengan R^2_{p+q} dan R^2_p untuk masing-masing $p + q$ variabel dan p variabel seperti dalam regresi. Dari (2.3) kita melihat pemisahan nilai tengah ditransformasi, $\frac{(\bar{z}_1 - \bar{z}_2)^2}{s_z^2}$ yang dipakai oleh fungsi diskriminan untuk $z = \mathbf{a}'\mathbf{y}$ adalah ekuivalen untuk membakukan jarak antara vektor nilai tengah $\overline{\mathbf{y}}_1$ dan $\overline{\mathbf{y}}_2$. Jarak yang dibakukan adalah proporsional untuk T^2 dalam kasus 2 pengelompokan dimana T^2 adalah uji Hotteling. Koefisien vektor **a** pada fungsi diskriminan berbeda secara signifikan dari **0** jika T^2 signifikan.

b. Analisis Fungsi Diskriminan Untuk Beberapa Pengelompokan

Analisis fungsi diskriminan untuk beberapa pengelompokan dari pengamatan multivariat memperhatikan kombinasi linier variabel terbaik dalam pemisahan pengelompokan. Langkah-langkah analisis fungsi diskriminan untuk beberapa pengelompokan adalah (Alvin C Rencher;1996:303)

1. Menguji pemisahan pengelompokan dalam sebuah plot berdimensi 2 *(examine group separation in a two-dimensional plot)* . Ketika ada lebih dari 2 pengelompokan, diperlukan lebih dari satu fungsi diskriminan untuk mendeskripsikan pemisahan pengelompokan. Jika titik dalam ruang dimensi p diproyeksikan pada ruang dimensi 2 digambarkan oleh 2 fungsi diskriminan yang pertama. Kita menentukan kemungkinan yang terbaik bagaimana pengelompokan itu dipisahkan.
2. Menemukan bagian variabel asli yang memisahkan kelompok-kelompok sebaik *original set*.
3. Mengurutkan variabel sesuai kontribusi relatif untuk pemisahan pengelompokan. Kita bakukan koefisien fungsi diskriminan sehingga perbandingan yang lebih akurat bisa dibuat.
4. Menginterprestasikan dimensi baru yang digambarkan oleh fungsi diskriminan.
5. Mengikuti aturan MANOVA dengan pengaruh tetap.

Untuk k pengelompokan dengan n_i observasi dalam pengelompokan ke-i, kita transformasi masing-masing pengamatan/observasi vektor $\mathbf{y}_{ij}$ untuk menentukan $z_{ij} = \mathbf{a}'\mathbf{y}_{ij}$, $i = 1, 2, \ldots, k$; $j = 1, 2, \ldots, n_j$ dan nilai tengah $\bar{z}_i = \mathbf{a}'\bar{\mathbf{y}}_i$. Untuk mencari vektor **a** yang memisahkan secara maksimal $\bar{z}_1, \bar{z}_2, \ldots, \bar{z}_k$. Untuk menggambarkan pemisahan $\bar{z}_1, \bar{z}_2, \ldots, \bar{z}_k$, kita memperluas kriteria pemisahan (2.1) dibakukan adalah proporsional untuk T^2 dalam kasus 2 pengelompokan dimana T^2 adalah uji Hotteling. Koefisien vektor **a** pada fungsi diskriminan berbeda secara signifikan dari **0** jika T^2 signifikan.

b. Analisis Fungsi Diskriminan Untuk Beberapa Pengelompokan

Analisis fungsi diskriminan untuk beberapa pengelompokan dari pengamatan multivariat memperhatikan kombinasi linier variabel terbaik dalam pemisahan pengelompokan. Langkah-langkah analisis fungsi diskriminan untuk beberapa pengelompokan adalah (Alvin C Rencher;1996:303)

1. Menguji pemisahan pengelompokan dalam sebuah plot berdimensi 2 *(examine group separation in a two-dimensional plot)* . Ketika ada lebih dari 2 pengelompokan, diperlukan lebih dari satu fungsi diskriminan untuk mendeskripsikan pemisahan pengelompokan. Jika titik dalam ruang dimensi p diproyeksikan pada ruang dimensi 2 digambarkan oleh 2 fungsi diskriminan yang pertama. Kita menentukan kemungkinan yang terbaik bagaimana pengelompokan itu dipisahkan.
2. Menemukan bagian variabel asli yang memisahkan kelompok-kelompok sebaik *original set*.
3. Mengurutkan variabel sesuai kontribusi relatif untuk pemisahan pengelompokan. Kita bakukan koefisien fungsi diskriminan sehingga perbandingan yang lebih akurat bisa dibuat.
4. Menginterprestasikan dimensi baru yang digambarkan oleh fungsi diskriminan.
5. Mengikuti aturan MANOVA dengan pengaruh tetap.

Untuk k pengelompokan dengan n_i observasi dalam pengelompokan ke-i, kita transformasi masing-masing pengamatan/observasi vektor $\mathbf{y}_{ij}$ untuk menentukan $z_{ij} = \mathbf{a}'\mathbf{y}_{ij}$, $i = 1, 2, \ldots, k$; $j = 1, 2, \ldots, n_j$ dan nilai tengah $\bar{z}_i = \mathbf{a}'\bar{\mathbf{y}}_i$. Untuk mencari vektor **a** yang memisahkan secara maksimal $\bar{z}_1, \bar{z}_2, \ldots, \bar{z}_k$. Untuk menggambarkan pemisahan $\bar{z}_1, \bar{z}_2, \ldots, \bar{z}_k$, kita memperluas kriteria pemisahan (2.1) untuk k pengelompokan. Untuk $\mathbf{a}'(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2) = (\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)'\mathbf{a}$, kita dapat menggambarkan (2.1) dalam bentuk

$$\frac{(\bar{z}_1 - \bar{z}_2)^2}{s_z^2} = \frac{[\mathbf{a}'(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)]^2}{\mathbf{a}'\mathbf{S}_{pl}\mathbf{a}} = \frac{\mathbf{a}'(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)'\mathbf{a}}{\mathbf{a}'\mathbf{S}_{pl}\mathbf{a}} \quad \text{.................................. (2.7)}$$

Persamaan (2.7) untuk k pengelompokan, kita menggunakan matrik **H** dari MANOVA untuk $(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)'$ dan matrik **E** untuk $\mathbf{S}_{pl}$ sehingga diperoleh

$$\lambda = \mathbf{a}'\mathbf{H}\mathbf{a}/\mathbf{a}'\mathbf{E}\mathbf{a} \quad \text{.. (2.8)}$$

dengan

- **H** adalah matrik peragam antar kelompok dari jumlah kuadrat total diagonal untuk masing-masing p variabel.
- **E** adalah matrik peragam dalam kelompok dari jumlah kuadrat total untuk masing-masing variabel diagonal yang analog dengan jumlah hasil kali elemen di luar diagonal (*sum of products off-diagonal*).

Ini dapat ditulis dalam bentuk

$$\mathbf{a}'\mathbf{H}\mathbf{a} = \lambda\ \mathbf{a}'\mathbf{E}\mathbf{a} \text{ atau } \mathbf{a}'(\mathbf{H}\mathbf{a} - \lambda\ \mathbf{E}\mathbf{a}) = 0 \quad \text{.. (2.9).}$$

Untuk menguji nilai λ dan **a** solusi dari (2.9) agar menghasilkan nilai **a** adalah hasil maksimal λ. Solusi $\mathbf{a}^t = \mathbf{0}^t$ tidak diperbolehkan karena mengakibatkan $\lambda = \frac{0}{0}$ dalam (2.8), mungkin solusi yang lain ditemukan dari

$$\mathbf{H}\mathbf{a} - \lambda\mathbf{E}\mathbf{a} = \mathbf{0} \quad \text{.. (2.10)}$$

dapat ditulis dalam bentuk

$$(\mathbf{E}^{-1}\mathbf{H} - \lambda\mathbf{I})\mathbf{a} = \mathbf{0} \quad \text{.. (2.11).}$$

Solusi dari (2.11) untuk $\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_s$ sebagai vektor ciri (*eigenvektor*) yang berpadanan dengan $\lambda_1, \lambda_2, ..., \lambda_s$ yang merupakan akar ciri (*eigenvalue)* dari matrik $\mathbf{E}^{-1}\mathbf{H}$. Mengenai akar ciri dan diurutkan menjadi $\lambda_1 > \lambda_2 > ... > \lambda_s$. Nilai akar ciri ke-$s$ tidak nol diurutkan pada **H**, dapat ditemukan lebih kecil dari $k - 1$ atau p. Bila fungsi diskriminan $z_1 = \mathbf{a}_1'\mathbf{y}$ yang memaksimumkan nisbah antara matrik peragam antar kelompok dengan matrik peragam dalam kelompok maka yang ingin kita cari ialah $\mathbf{a}_1$ sehingga $\mathbf{a}_1'\mathbf{H}\mathbf{a}_1/\mathbf{a}_1'\mathbf{E}\mathbf{a}_1$ maksimum.

Dari matrik $\mathbf{E}^{-1}\mathbf{H}$ yaitu s vektor ciri $\mathbf{a}_1,\mathbf{a}_2,...,\mathbf{a}_s$ yang bersesuaian dengan akar ciri $\lambda_1,\lambda_2,...,\lambda_s$, diperoleh fungsi diskriminan $z_1 = \mathbf{a}_1'\mathbf{y}, z_2 = \mathbf{a}_2'\mathbf{y},...,z_s = \mathbf{a}_s^t\mathbf{y}$. Fungsi diskriminan ini tidak berkorelasi. Mereka menunjukkan dimensi yang berbeda untuk $\bar{\mathbf{y}}_1, \bar{\mathbf{y}}_2,..., \bar{\mathbf{y}}_k$, dengan $\bar{\mathbf{y}}_i = \sum_{j=1}^{n_i} \frac{\bar{\mathbf{y}}_{ij}}{n_i}$.

Masing–masing fungsi diskriminan mempertimbangkan akar ciri sebagai jumlah proporsi

$$\lambda_i \Big/ \sum_j \lambda_j \qquad \text{.. (2.12)}$$

jika akar ciri kecil dapat diabaikan dalam fungsi diskriminan.

Matrik $\mathbf{E}^{-1}\mathbf{H}$ tidak simetris. Untuk perhitungan akar ciri dan vektor ciri diperlukan matrik simetris. Jika sebuah matrik ditranspos maka hasilnya sama dengan matrik asalnya disebut matrik simetris.

Masalah di atas digambarkan dalam bentuk sampel pertidaksamaan $n_1,n_2,...,n_k$. Dalam aplikasi, situasi yang biasa dapat diganti dengan tidak mendapat kesulitan. Idealnya, n_i paling kecil akan dilebihkan dari angka p variabel. Ini tidak diperlukan secara matematis, tetapi akan lebih stabil dalam fungsi diskriminan.

Rasio antara jumlah kuadrat total untuk fungsi diskriminan yang pertama , $z_1 = \mathbf{a}_1'\mathbf{y}$:

$$\eta_\theta^2 = \theta = \frac{\lambda_1}{1+\lambda_1} = \frac{\mathrm{SSH}(z_1)}{\mathrm{SSE}(z_1)+\mathrm{SSH}(z_1)}$$

dengan

SSH = jumlah kuadrat matrik peragam antar kelompok

SSE = jumlah kuadrat matrik peragam dalam kelompok

η^2_θ adalah korelasi kuadrat maksimum antar fungsi diskriminan yang pertama dan kombinasi linier terbaik untuk dummy k–1 dari variabel anggota pengelompokan, korelasi maksimum disebut juga korelasi kanonik, dimana korelasi kanonik kuadrat dihitung untuk masing – masing fungsi diskriminan.

$$r_i^2 = \frac{\lambda_i}{1+\lambda_i}, \qquad i = 1,2,...,s \qquad \text{.. (2.13)}$$

Jika y tidak setaraf, kita membutuhkan koefisien a_r^* , yang diaplikasikan untuk variabel yang dibakukan. Untuk pengamatan ke-j vektor $\mathbf{y}_{ij}$ dalam pengelompokkan ke – i, kita dapat mengekspresikan fungsi diskriminan dengan variabel yang dibakukan sebagai berikut :

$z_{ij} = a_1^* \dfrac{y_{ij1} - \bar{y}_{i1}}{s_1} + a_2^* \dfrac{y_{ij2} - \bar{y}_{i2}}{s_2} + \ldots + a_p^* \dfrac{y_{ij2} - \bar{y}_{ip}}{s_p}$, $i = 1, 2$; $j = 1, 2, \ldots, n_i$ dengan $\bar{\mathbf{y}}_i' = (\bar{y}_{1j}, \bar{y}_{i2}, \ldots, \bar{y}_{ip})$ adalah vektor nilai tengah untuk pengelompokan ke–i, $i = 1, 2$ dan s_r adalah simpangan baku (*standard deviasi*) untuk variabel ke–r. Dijelaskan sebagai akar kuadrat ke–r dari elemen diagonal $\mathbf{S}_{\mathbf{pl}}$. Maka koefisien yang dibakukan harus dibentuk

$$a_r^* = s_r a_r \ , r = 1, 2, \ldots, p . \qquad \ldots\ldots (2.14)$$

Untuk beberapa kasus pengelompokan, kita dapat membakukan fungsi diskriminan dengan menunjukkan koefisien ke–r dalam fungsi diskriminan ke–m oleh a_{mr}, kemudian bentuk baku $a^*_{mr} = s_r . a_{mr}$, dengan s_r adalah simpangan baku yang dijelaskan dari diagonal $\mathbf{S}_{\mathbf{pl}} = \mathbf{E}/v_E$. Jika vektor ciri ke–m adalah tunggal untuk perkalian skalar, kita dapat membakukan dengan menggunakan $a_{mr}^* = \sqrt{e_{rr} \bullet a_{mr}}$; $r = 1, 2, \ldots, p$ dengan e_{rr} adalah elemen diagonal ke–r dari $\mathbf{E}$

Akar ciri sama seperti uji Λ Wilks digunakan untuk menguji vektor nilai tengah yang berbeda secara signifikan. Uji Λ Wilks dinyatakan dengan $\Lambda_1 = \prod_{i=1}^{s} \dfrac{1}{1+\lambda_i}$,yang berdistribusi Λ_p , $k - 1$, $N - k$ dengan $N = \Sigma_i n_i$ untuk model tidak seimbang atau $N = kn$ untuk kasus seimbang. Λ_1 adalah kecil jika salah satu atau lebih dari λ_i adalah besar. Uji Λ Wilks diberikan secara signifikan atas akar ciri untuk fungsi diskriminan. s akar ciri menggambarkan dimensi s dari pemisahan vektor nilai tengah $\bar{\mathbf{y}}_1, \bar{\mathbf{y}}_2, \ldots, \bar{\mathbf{y}}_k$. Jika terdapat banyak dimensi, maka signifikan. Dalam fungsi diskriminan, uji Λ Wilks lebih banyak digunakan daripada 3 uji statistik (MANOVA) yang lain, karena uji ΛWilks dapat digunakan pada bagian akar ciri. Kita dapat menggunakan χ^2 sebagai aproksimasi untuk Λ_1 dengan $v_E = N - k = \Sigma_i n_i - k$; $v_H = k - 1$ dan v adalah statistik V-Barlett:

$$v_1 = -\left[v_E - \tfrac{1}{2}(p - v_H + 1)\right] \ln \Lambda_1$$

$$= -\left[N - 1 - \tfrac{1}{2}(p + k)\right] \ln \prod_{i=1}^{s} \frac{1}{1+\lambda_i}$$

$$= \left[N - 1 - \tfrac{1}{2}(p + k)\right] \sum_{i=1}^{s} \ln(1+\lambda_i) \quad \text{.................................} (2.15)$$

dengan χ^2 mempunyai derajat kebebasan $p(k-1)$. Statistik uji Λ_1 dan pendekatan (2.15) menguji signifikan dari $\lambda_1, \lambda_2, \ldots, \lambda_s$. Jika uji menolak H_0, kita memutuskan paling sedikit satu λ yang signifikan dari 0. Misal, λ_1 paling besar, maka tingkat signifikannya sama dengan $z_1 = \mathbf{a}_1'\mathbf{y}$.

Untuk menguji tingkat signifikan dari $\lambda_2, \lambda_3, \ldots, \lambda_s$, kita menghapus λ_1 dari Λ Wilks dan menggunakan χ^2 sebagai aproksimasi untuk Λ_2 dengan

$$\Lambda_2 = \prod_{i=2}^{s} \frac{1}{1+\lambda_i}$$

$$v_2 = -\left[N - 1 - \tfrac{1}{2}(p + k)\right] \ln \Lambda_2$$

$$= \left[N - 1 - \tfrac{1}{2}(p + k)\right] \sum_{i=2}^{s} \ln(1+\lambda_i) \quad \text{.............................} (2.16)$$

dengan χ^2 mempunyai derajat kebebasan $(p-1)(k-2)$. Jika uji Λ Wilks menolak H_0, kita memutuskan signifikan yang terkecil, misal λ_2. Tingkat signifikansi tersebut dinyatakan dengan fungsi diskriminan $z_2 = \mathbf{a}_2'\mathbf{y}$. Selanjutnya, uji masing–masing λ_i sampai menolak H_0. Statistik uji pada langkah ke–m adalah $\Lambda_m = \prod_{i=m}^{s} \frac{1}{1+\lambda_i}$, dengan berdistribusi sebagai $\Lambda_{p-m+1,\, k-m,\, N-k-m+1}$, sehingga

$$v_m = -\left[N - 1 - \tfrac{1}{2}(p + k)\right] \ln \Lambda_m$$

$$= \left[N - 1 - \tfrac{1}{2}(p + k)\right] \sum_{i=m}^{s} \ln(1+\lambda_i) \quad \text{.............................} (2.17)$$

dengan approksimasi χ^2 berdistribusi $(p-m+1)(k-m)$. Jika $\frac{\lambda_i}{\sum_j \lambda_j}$ adalah kecil, mungkin pengelompokan fungsi diskriminan menjadi tidak menarik, meskipun signifikan.

Kita dapat juga menggunakan aproksimasi uji F untuk masing–masing Λ_i.

Untuk $\Lambda_i = \prod_{i=1}^{s} \frac{1}{1+\lambda_i}$ dengan $v_E = N - k$ dan $v_H = k - 1$, maka $F = \frac{1-\Lambda_1^{1/t}}{\Lambda_1^{1/t}} \frac{df_2}{df_1}$,

dengan $t = \sqrt{\frac{p^2(k-1)^2 - 4}{p^2 + (k-1)^2 - 5}}$, $w = N - 1 - ½ (p + k)$

$df_1 = p (k - 1)$ $\quad\quad$ $df_2 = wt - ½ [p (k - 1) - 2]$.

Untuk $\Lambda_m = \prod_{i=m}^{s} \frac{1}{1+\lambda_i}$, m =2, 3, ... , s maka $F = \frac{1-\Lambda_m^{1/t}}{\Lambda_m^{1/t}} \frac{df_2}{df_1}$,

dengan $p - m - 1$ dan $k - m$ di tempatkan pada p dan $k - 1$.

$$t = \sqrt{\frac{(p-m+1)^2(k-m)^2 - 4}{(p-m+1)^2 + (k-m)^2 - 5}}$$

$w = N - 1 - ½ (p + k)$

$df_1 = (p - m + 1)(k - m)$

$df_2 = wt - ½ [(p - m + 1)(k - m) - 2]$

(Alvin C Rencher ; 1996 : 303 - 313).

2.3 Interprestasi Fungsi Diskriminan

Secara khusus, kita ingin mempelajari penggunaan kontribusi setiap variabel untuk fungsi diskriminan, yang pada umumnya berkorelasi. Ada tiga metode yaitu.

1. Menguji koefisien fungsi diskriminan yang dibakukan.

Koefisien fungsi diskriminan dapat dibakukan dengan menggunakan (2.14). Persamaannya adalah

$$Z_{ij} = a_1^* \frac{y_{ij1} - \bar{y}_{i1}}{s_1} + a_2^* \frac{y_{ij2} - \bar{y}_{i2}}{s_2} + \ldots + a_p^* \frac{y_{ij2} - \bar{y}_{ip}}{s_p}$$

$j = 1, 2, \ldots , n_1$.

Variabel yang dibakukan $\frac{\bar{\mathbf{y}}_{ijr} - \bar{\mathbf{y}}_{ir}}{\mathbf{s}_r}$ adalah bebas skala dan koefisien yang dibakukan $a_1^*, a_2^*, \ldots, a_p^*$ bilamana gabungan kontribusi yang benar dari variabel fungsi diskriminan z pada pemisahan pengelompokan itu dengan maksimal. Vektor koefisien fungsi diskriminan adalah vektor ciri dari $\mathbf{E}^{-1}\mathbf{H}$ dan dihitung korelasi sampel variabel. Berdasarkan hal tersebut, koefisien menunjukkan pengaruh pada setiap variabel dalam munculnya suatu variabel yang lain.

Pendekatan fungsi diskriminan sama seperti dalam kombinasi linier pada regresi, untuk contoh (1) koefisien mungkin berubah secara khusus jika variabel lain ditambahkan atau dihapus, (2) koefisien mungkin tidak stabil dari sampel ke sampel kecuali ukuran sampel besar untuk angka–angka variabel.

2. Menghitung parsial uji F untuk masing–masing variabel.

Setelah menghitung parsial uji F untuk masing–masing variabel, maka variabel–variabel itu dapat diperingkat. Untuk kasus 2 pengelompokan, parsial F diberikan sebagai berikut :

$$F = (v - p + 1)\frac{T_p^2 - T_{p-1}^2}{v + T_{p-1}^2} \quad \text{...} (2.18)$$

dengan T_p^2 adalah 2 sampel T^2 Hotelling dengan semua p variabel, T^2_{p-1} adalah statistik T^2 dengan semua variabel kecuali y_k dan $v = n_1 + n_2 - 2$. Statistik uji dalam (2.18) berdistribusi $F_{1,\, v-p+1}$.

Untuk beberapa kasus pengelompokan, parsial Λ, untuk y_k disesuaikan untuk $p - 1$ variabel yang lain, yaitu :

$$\Lambda\left(y_k \middle| y_1, \cdots, y_{k-1}, y_{k+1}, \cdots, y_p\right) = \frac{\Lambda_p}{\Lambda_{p-1}} \quad \text{.....................................} (2.19)$$

dengan Λ_p adalah Λ Wilks untuk semua p variabel dan Λ_{p-1} meliputi semua variabel kecuali y_k. Sehingga parsial F diberikan sebagai berikut :

$$F = \frac{1-\Lambda}{\Lambda}\,\frac{v_E - p + 1}{v_H} \quad \text{...} (2.20)$$

dengan Λ didefinisikan (2.19), $v_E = N - k$, dan $v_H = k - 1$. Parsial statistik Λ dalam (2.19) berdistribusi Λ_1, v_H, $v_E - p + 1$, dan parsial F dalam (2.20) berdistribusi Fv_H , $Fv_E - p + 1$.

Sebuah indeks parsial pengelompokan untuk y_i yang sama maka ukuran yang menyeluruh untuk semua y dengan $\eta^2{}_\Lambda$ dapat didefinisikan oleh

$$R_i^2 = 1 - \Lambda_i \;;\quad i = 1, 2, \ldots, p \qquad \text{.................................(2.21)}$$

dengan Λ_i adalah parsial Λ dalam (2.19) untuk y_i . Parsial R^2 dikelompokkan ukurannya antara variabel pengelompokan dan y_i sesudah ada penyelesaian untuk $p - 1$ dari y.

3. Menghitung korelasi antara masing–masing variabel dan fungsi diskriminan

Metode yang ketiga direkomendasikan dengan luas dan bagian yang sering digunakan. Karena korelasi mempunyai informasi yang lebih daripada koefisien yang dibakukan bergabung dengan variabel kontribusi fungsi diskriminan. Korelasi biasanya dengan t atau F untuk masing–masing variabel dan mengabaikan variabel yang lain.

Dalam analisis multivariat, ketertarikan dipusatkan dalam penampilan kumpulan variabel pada salah satu pihak. Kontribusi masing – masing variabel independen untuk semua variabel yang lain adalah ekuivalen untuk menjaga indeks univariat mengabaikan variabel yang lain. Kita tidak akan heran ketika kita tidak mendapatkan korelasi yang tepat, sebab tidak ada daerah nilai tengah antara univariat dan multivariat secara nyata.

Nilai pengklasifikasian untuk sebuah kasus perlu mempertimbangkan beberapa hal, yaitu :

1. Jarak Mahalanobis *(Mahalanobis distance)*

Secara umum jarak Mahalanobis adalah jarak antara dua titik dalam suatu ruang didefinisikan oleh satu atau lebih variabel yang berkorelasi. Jika ada dua variabel yang tidak berkorelasi, kita dapat menggunakan skatter plot dua dimensi (*two dimentional scatterplot*). Jarak antara titik mahalanobis akan identik dengan jarak Euclidean. Jika ada tiga variabel yang tidak berkorelasi kita dapat menggunakan plot tiga dimensi untuk menentukan jarak antara titik. Jika ada lebih dari tiga variabel, kita tidak dapat lagi mempresentasikan jarak dalam plot. Sehingga variabel yang dikorelasikan dan diakses dalam plot, dapat difikirkan *non-orthogonal* atau tidak tegak lurus, sehingga mereka tidak dapat diposisiskan dalam sudut yang benar. Untuk kasus ini jarak euclidean tidak dapat mengukur secara tepat, sementara jarak mahalanobis cukup memadai untuk menghitung korelasi.

2. Jarak mahalanobis dan klasifikasi *(mahalanobis distance and classification)*

Untuk masing-masing pengelompokan, kita dapat menentukan posisi titik dengan menggambarkan nilai tengah semua variabel yang didefinisikan dalam model. Titik ini disebut centroid pengelompokan. Kemudian untuk masing-masing individu dapat dihitung jarak mahalanobis dari masing-masing centroid pengelompokan. Kita dapat mengklasifikasi individu yang termasuk dalam pengelompokan secara dekat dengan jarak mahalanobis paling kecil.

3. Probabilitas pengklasifikasian terakhir (*posterior classification probabilities*)

Jarak mahalanobis dipakai dalam pengklasifikasian, untuk memperoleh probabilitas. Probabilitas termasuk dalam bagian pengelompokan dengan dasar proposional untuk jarak mahalanobis dari centroid pengelompokan (hal ini tidak proposional dengan tepat karena kita mengasumsikan sebuah multivariat berdistribusi normal disekitar centroid). Untuk menghitung posisi masing-masing nilai prior suatu variabel dalam model disebut probabilitas posterior. Dengan kata lain probabilitas posterior adalah probabilitas yang didasarkan atas suatu nilai variabel yang lain. Beberapa softwere yang tersedia akan menghitung dengan otomatis probabilitas untuk semua individu (atau hanya untuk menyeleksi cara pembelajaran validasi silang/*cross validation*).

4. Probabilitas pengklasifikasian awal (*a prior classification probabilities*)

Hal ini adalah salah satu tugas tambahan untuk mempertimbangkan pengklasifikasian. Kadang-kadang, kita tahu terlebih dahulu bahwa ada pengamatan lebih dalam satu pengelompokan. Jadi probabilitas prior termasuk pengelompokan tertinggi. Secara spesifik, perbedaan probabilitas prior digunakan untuk mengatur kasus pengklasifikasian dan perhitungan probabilitas posterior dilakukan secara urut.

2.4 Menghitung rata-rata misklasifikasi

Untuk menghitung proses pengklasifikasian dalam memprediksikan anggota pengelompokan , biasanya dengan menggunakan probabilitas misklasifikasi yang disebut sebagai rata-rata kesalahan (*error rate*). Sedangkan komplemen dari rata-rata kesalahan adalah rata-rata pengklasifikasian yang benar (*correct classification rate*). Penduga rata-rata kesalahan dapat diperoleh dari hasil proses pengklasifikasian dengan data sama yang digunakan untuk menghitung fungsi klasifikasi. Biasanya ditunjukkan dengan mengganti kembali (*resubstitution*). Masing-masing vektor $\mathbf{y}_{ij}$ dipertimbangkan pada fungsi pengklasifikasian dan dimasukkan dalam suatu pengelompokan. Kemudian kita menghitung jumlah pengklasifikasian yang benar dan jumlah yang misklasifikasi. Hasil proporsi misklasifikasi dari *resubstitution* disebut rata-rata kesalahan yang nampak (*apparent error rate*). Sehingga

$$\text{Apparent error rate} = \frac{n_{12}+n_{21}}{n_1+n_2} = \frac{n_{12}+n_{21}}{n_{11}+n_{12}+n_{21}+n_{22}} \quad \ldots\ldots\ldots\ldots\ldots (2.22)$$

dengan :

n_1 = pengamatan yang pertama dalam G_1

n_{11} = hasil yang diklasifikasikan benar dalam G_1

n_{12} = hasil yang misklasifikasi dalam G_1

$n_1 = n_{11} + n_{12}$

n_2 = pengamatan yang kedua dalam G_2

n_{21} = hasil yang diklasifikasikan benar dalam G_2

n_{22} = hasil yang misklasifikasi dalam G_2

$n_2 = n_{21} + n_{22}$

Dengan cara yang sama, kita dapat mendefinisikan sebagai berikut:

$$\text{Apparent correct classification rate} = \frac{n_{11} + n_{22}}{n_1 + n_2} \qquad \text{..................... (2.23)}$$

maka

Apparent error rate = 1 – apparent correct classification rate

Kita dapat menggunakan (2.23) dan (2.24) untuk kasus beberapa pengelompokan.

2.5 Mengalokasikan individu baru ke dalam salah satu kelompok

Ada dua bentuk dalam mengalokasikan suatu individu baru kedalam kelompok baru untuk analisis diskriminan yaitu

1. Analisis diskriminan untuk dua pengelompokan

   Untuk mengalokasikan suatu individu baru ke dalam kelompok baru mempunyai kriteria penggolongan yaitu

   a. Fungsi linier diskriminan Fisher (F)

   $$z = \mathbf{a}'\mathbf{y} = (\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)'\mathbf{S}_{pl}^{-1}\mathbf{y} \qquad \text{.............................. (2.24)}$$

   dengan

   $\mathbf{a}'$ = vektor koefisien pembobot fungsi diskriminan

   $\bar{\mathbf{y}}_1$ = vektor rata-rata variabel acak dari kelompok 1

   $\bar{\mathbf{y}}_2$ = vektor rata-rata variabel acak dari kelompok 2

   $\mathbf{S}_{pl}^{-1}\mathbf{y}$ = invers peragam gabungan

   Titik tengah m diantara rata-rata kelompok 1 dan 2

   $$m = \frac{1}{2}(\bar{\mathbf{S}}_1 + \bar{\mathbf{S}}_2) = \frac{1}{2}(\bar{\mathbf{y}}_1 + \bar{\mathbf{y}}_2)'\mathbf{S}_{pl}^{-1}(\bar{\mathbf{y}}_1 - \bar{\mathbf{y}}_2)$$

   Maka :

   - mengalokasikan individu ke dalam kelompok 1 jika $z_0 > m$ atau $z_0 - m > 0$;
   - mengalokasikan individu ke dalam kelompok 2 jika $z_0 \leq m$ atau $z_0 - m \leq 0$.

   dengan z_0 adalah skor diskriminan dari individu tersebut.

   b. Statistik Wald Anderson (W)

$$W = \mathbf{y}'\mathbf{S}_{pl}^{-1}(\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2) - \frac{1}{2}(\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2)'\mathbf{S}_{pl}^{-1}(\overline{\mathbf{y}}_1 - \overline{\mathbf{y}}_2) \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (2.25)$$

Maka:

- mengalokasikan individu ke dalam kelompok 1 jika $w > 0$;
- mengalokasikan individu ke dalam kelompok 2 jika $w \leq 0$.

2. Analisis diskriminan untuk beberapa kelompok

   a. Fisher

   Mengalokasikan individu ke kelompok k, $k = 1,2,\ldots,g$ jika

$$\sum_{m=1}^{r}(z_m - \bar{z}_{km})^2 = \sum_{m=1}^{r}\left(\mathbf{a}_m^1(\mathbf{y} - \overline{\mathbf{y}}_k)\right)^2$$

$$\leq \sum_{m=1}^{r}\left(\mathbf{a}_m^1(\mathbf{y} - \overline{\mathbf{y}}_n)\right)^2$$

$$\forall\ n \neq k\ ;\ r \leq s\ ;\ s = \min(g-1, p) \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (2.26)$$

   dengan

   z_m = skor diskriminan ke-m dari individu tertentu

   $\bar{z}_{km}$ = rata-rata skor diskriminan ke-m dari kelompok ke-k,k = 1,2,…,g

   $\mathbf{a}_m^1$ = koefisien fungsi diskriminan ke-m

   $\mathbf{y}$ = vektor data pengamatan dari individu

   $\overline{\mathbf{y}}_k$ = vektor nilai rata-rata peubah dari kelompok ke-k

   r = banyaknya fungsi diskriminan yang dipergunakan dalam penggolongan

   s = banyaknya fungsi diskriminan yang mungkin dibentuk

   g = banyaknya kelompok

   p = banyaknya variabel

   b. Wald Anderson

$$W_{kh} = \mathbf{y}'\mathbf{S}_{pl}^{-1}(\overline{\mathbf{y}}_k - \overline{\mathbf{y}}_h) - \frac{1}{2}(\overline{\mathbf{y}}_k + \overline{\mathbf{y}}_h)'\mathbf{S}_{pl}^{-1}(\overline{\mathbf{y}}_k + \overline{\mathbf{y}}_h) \quad \ldots\ldots\ldots\ldots \quad (2.27)$$

   Mengalokasikan individu ke dalam kelompok ke-k, k = 1,2,…,g jika $W_{kh} > 0$; $\forall\ h \neq k$.

c. D^2 Mahalanobis

$$D_k^2 = (\mathbf{y} - \overline{\mathbf{y}}_\mathbf{k})' \mathbf{S}_{\mathbf{pl}}^{-1} (\mathbf{y} - \overline{\mathbf{y}}_\mathbf{k}); k = 1,2,\ldots,g \qquad \ldots\ldots\ldots\ldots\ldots\ldots \quad (2.28)$$

Mengalokasikan individu ke dalam kelompok ke-k

jika $D_k^2 = \min\left(D_1^2,\ldots,D_g^2\right)$.

BAB III
METODOLOGI PENELITIAN

3.1 Pengumpulan Data

Data yang digunakan dalam penelitian ini adalah data sekunder dari FMIPA Universitas Jember. Data sekunder pada penelitian ini berasal dari hasil NEM dan nilai rata-rata mata kuliah semua jurusan mahasiswa yang diambil sebagai sampel penelitian.

Peneliti mengambil kebijakan untuk mengambil sejumlah mahasiswa yaitu kurang lebih 100 orang pada angkatan 1998/1999. Adapun pelaksanaan pengambilan sampel dilakukan selama kurun waktu lebih kurang 2 minggu. Pada tanggal 28 Januari – 5 Februari 2002

3.2 Identifikasi Objek, Populasi dan Sampel Penelitian

Objek penelitian yang digunakan dalam penelitian ini adalah mahasiswa Fakultas MIPA Universitas Jember. Pada penelitian ini populasi penelitian di batasi pada mahasiswa Fakultas MIPA Universitas Jember angkatan 1998/1999. Sampel penelitian di ambil sebanyak mahasiswa yang ada pada angkatan tersebut.

3.3 Identifikasi Atribut –Atribut dan Variabel Penelitian

Dalam penelitian ini menggunakan atribut-atribut penelitian di mana identifikasi awal atribut-atribut tersebut diperoleh dari studi pustaka dan hasil penelitian.

Penelitian ini menggunakan 4 atribut yaitu:

a. matematika di beri kode 1;
b. fisika di beri kode 2;
c. kimia di beri kode 3;
d. biologi di beri kode 4.

Sedangkan untuk variabel independen dalam penelitian ini menggunakan:

y_1/nmath : nilai NEM matematika untuk semua jurusan;

y_2/nrmath : nilai rata-rata mata kuliah matematika untuk semua jurusan pada semester I dan II;

y_3/nfis : nilai NEM fisika untuk semua jurusan;

y_4/nrfis : nilai rata-rata mata kuliah fisika untuk semua jurusan pada semester I dan II;

y_5/nkim : nilai NEM kimia untuk semua jurusan;

y_6/nrkim : nilai rata-rata mata kuliah kimia untuk semua jurusan pada semester I dan II;

y_7/nbio : nilai NEM biologi untuk semua jurusan;

y_8/nrbio : nilai rata-rata mata kuliah biologi untuk semua jurusan pada semester I dan II;

y_9/nlain : nilai NEM lainnya untuk semua jurusan yaitu PPKn, Bahasa Indonesia dan Bahasa Inggris;

y_{10}/nrlain : nilai rata-rata mata kuliah lainnya yang tidak berhubungan dengan matematika, fisika, kimia dan biologi;

Pengolahan data dengan menggunakan bantuan SPSS versi 7.5. Alasan pemilihan SPSS, karena memiliki program-program statistik yang lebih lengkap daripada program lainnya dan pada umumnya lebih dikenal di bidang sosial maupun penelitian ilmiah.

3.4 Pengolahan Data

Dalam penelitian ini dilakukan pengolahan data yang terdapat VI bagian yang berurutan, untuk mencapai tujuan penelitian.

I. Mengelompokkan mahasiswa per jurusan dengan pemberian kode 1, 2, 3 & 4.

II. Memasukkan nilai NEM sesuai dengan jurusannya masing-masing dan menghitung nilai NEM lainnya.

III. Menghitung nilai rata-rata mata kuliah sesuai dengan jurusan yang berhubungan untuk Matematika, Fisika, Kimia, dan Biologi.

IV. Menghitung nilai rata-rata mata kuliah lainnya disesuaikan dengan jurusan masing-masing mahasiswa.

V. Mengolah data dari bagian I, II, III & IV untuk mengetahui misklasifikasi awal pengelompokan mahasiswa FMIPA Angkatan 1998/1999 Universitas Jember. Analisis diskriminan dibantu oleh paket program SPSS versi 7.5.

VI. Menganalisa misklasifikasi awal pengelompokan mahasiswa FMIPA angkatan 1998/1999 Universitas Jember dan menentukan langkah yang terbaik bagi pihak Fakultas MIPA Universitas Jember.

3.5 Keluaran SPSS

Keluaran SPSS untuk analisis diskriminan yang akan dipakai dalam penelitian adalah:

- *Canonical discriminant function standardized coefficients* yaitu nilai koefisien fungsi pembeda telah dibakukan, yang berarti nilai koefisien tersebut dikalikan dengan dugaan gabungan dari simpangan baku koefisien tersebut.
- *Eigenvalues* adalah akar ciri menunjukkan rasio jumlah kuadrat antar kelompok dan jumlah kuadrat dalam kelompok. Nilai ini dapat juga diinterprestasikan sebagai jumlah keragaman yang dapat dijelaskan oleh masing-masing fungsi pembeda diskriminan.
- *Wilks Lambda* yaitu nilai yang menjelaskan tentang proporsi keragaman antar kelompok yang tidak dapat dijelaskan oleh masing-masing fungsi pembeda diskriminan.
- *Test of equality of group means* digunakan untuk menguji nilai tengah masing-masing kelompok agar mengetahui ada tidaknya perbedaan antar nilai tengah kelompok mahasiswa FMIPA angkatan 1998/1999.
- *Structure matrix* berisi informasi tentang korelasi antar variabel asal dengan variabel kanonik (fungsi diskriminan). Besarnya nilai koefisien dalam korelasi akan ditunjukkan sebagai tingkat kepentingan suatu variabel terhadap fungsi diskriminannya. Tanda "*" menandakan bahwa variabel-variabel tersebut dikelompokkan menjadi satu kelompok. Sedangkan simbol "a" menandakan bahwa varibel tersebut tidak digunakan dalam model.

- *Canonical discriminant function* merupakan posisi masing-masing individu dan pusat pengelompokkannya didekati dengan 2 dimensi.
- *Classification results* merupakan informasi pengujian keanggotaan individu terhadap masing-masing kelompok, serta relokasi kembali individu yang diduga salah alokasi.
- *Casewise statistics* berisi informasi mengenai mahasiswa yang salah alokasi dan direlokasikan ke dalam kelompok baru dengan mengggunakan jarak Mahalanobis.s
- *Territorial Map* berisi informasi yang berkaitan dengan daerah pengelompokan masing-masing kelompok, dengan batas antar kelompok memiliki dua nilai yang merupakan pembatas masing-masing kelompok.
- *Variable Entered/Removed* berisi informasi tentang variabel-variabel yang berperan nyata dalam membedakan kelompok yang terbentuk, sementara itu variabel yang tidak signifikan (tidak membedakan), dapat diartikan semua kelompok memiliki nilai yang sama terhadap variabel-variabel tersebut.

BAB V
KESIMPULAN DAN SARAN

5.1 Kesimpulan

Dari pembahasan pada bab sebelumnya dapat diambil kesimpulan sebagai berikut.

1. Variabel yang membedakan misklasifikasi awal pengelompokan adalah nilai rata-rata matematika (y_2), nilai rata-rata fisika (y_4), nilai NEM kimia (y_5) , nilai rata-rata kimia (y_6)dan nilai rata-rata biologi (y_8).
2. Ada tiga fungsi diskriminan yang terbentuk, yakni:

$$Z_{score\_}1 = -0,863*y_2 + 0,471*y_4 + 0,522*y_5 + 0,464*y_6 - 0,426*y_8$$

$$Z_{score\_}2 = -0,080*y_2 - 0,656*y_4 + 0,330*y_5 + 0,391*y_6 + 0,704*y_8$$

$$Z_{score\_}3 = -0,307*y_2 + 0,284*y_4 + 0,679*y_5 - 1,045*y_6 + 0,479*y_8$$

3. Hasil penelitian menunjukkan terjadinya misklasifikasi mahasiswa FMIPA angkatan 1998/1999 Universitas Jember sebesar 46,4%. Angka yang relatif tinggi mungkin disebabkan tanpa memperhatikan profil kemampuan akademik dari mata kuliah yang terkait (baik NEM, mata kuliah dari kemipaan).
4. Ada perbedaan yang signifikan (kurang dari 1%) untuk ketiga fungsi diskriminan yang menerangkan perbedaan antara kelompok matematika, fisika, kimia dan biologi, sehingga analisis diskriminan layak untuk digunakan.

5.2 Saran

Dari pembahasan pada bab sebelumnya dapat diambil saran sebagai berikut.

1. Melihat adanya misklasifikasi yang relatif tinggi, sebaiknya pihak FMIPA Universitas Jember memberikan informasi dan wawasan yang luas mengenai prospek dan prasyarat dari masing-masing jurusan di masa yang akan datang untuk pelajar SMU yang akan mengikuti UMPTN serta memberikan motivasi cara belajar yang efektif untuk menghasilkan lulusan yang berkualitas.
2. Sistem penjurusan mahasiswa FMIPA Universitas Jember ada baiknya mempertimbangkan prestasi akademik awal setelah di Perguruan Tinggi selain minat dan profil NEM.
3. Penelitian dapat di kembangkan dengan menggunakan sampel yang lebih luas (berbagai angkatan) dan variabel lain (nilai STTB dan hasil skor UMPTN).

DAFTAR PUSTAKA

Coopeer. D.R. And William E. Morry. 1999. *Metode Penelitian Bisnis*. Jilid II. Erlangga Jakarta.

Huberty. J.C. 1994. *Applied Discriminant Analysis*. Inc. 605 Third Avene. New York.

Johnson. RA & Dear W. Wichern. 1992. *Applied Multivariate Statistical Analysis*. University of Winconsin Practice Hall. Inc. New York.

Rencher. A. C.1996. *Methods of Multivariate Analysis*. Inc. New York.

Santoso.S. 2002. *SPSS Analisis Multivariat*. Elex Media Komputindo Kelompok Gramedia. Jakarta.

Siswandi dan Budi Suharjo.1999. *Analisis Eksplorasi Data Peubah Ganda & SPSS versi 7.5* . Jurusan Matematika FMIPA-IPB

Lampiran 1

Group	y1	y2	y3	y4	y5	y6	y7	y8	y9	y10
1	6,95	2,83	5,65	1,5	7,84	3	6,1	2	6,3	3
1	8,55	3,83	5,05	3,5	6,07	3	6,3	4	51	2,75
1	8,65	2,83	4,95	2	7,84	4	6,4	4	6,8	3,5
1	7,7	2,17	5,6	1,5	7,81	2	6,4	3	7,8	2,75
1	6,65	2	4,8	2	6,41	3	6,2	1	6,8	3
1	5,55	2,33	5,7	2,5	5,71	3	6,5	3	7,3	3,25
1	5,2	3,33	5,35	2	7,59	4	8,1	3	7,1	3
1	5,9	3	3,2	2,5	5,08	2	4,6	4	6,8	3,5
1	4,9	2,67	3,04	2	4,22	2	4,2	3	6,4	3
1	8,05	3,17	5	2,5	6,54	4	5,5	3	6,7	2,75
1	5,25	3,17	4,15	2	7,38	4	6,9	4	7,6	3
1	7,3	2,67	4,6	2	6,78	3	6,5	3	7,8	3,25
1	5	1,83	3,3	2	5,82	2	4,5	1	6,1	3
1	6,05	3	4,5	2	5,69	3	6,1	3	6,6	3,25
1	5,55	2,5	4,75	1,5	6,49	2	4,9	3	6,4	3
1	4,5	2,83	4,37	2	5,18	3	5,9	4	6,5	3,5
1	6,6	3,17	5,8	2,5	7,35	4	6,2	4	6,9	3,25
1	8,45	3,67	4,7	3	6,18	4	5,2	3	7,6	3,5
1	5,2	2,83	5,2	2	6,72	2	6,1	3	6,6	3,25
1	7,45	3	5,05	2	7,38	3	5,7	3	7,1	3,25
1	3,05	2,17	3,95	1,5	4,06	1	4,9	2	6,5	3,5
1	7,8	3,33	3,9	2	5,95	3	5,8	3	6,9	3,25
1	5,45	2,5	3,25	1,5	6,38	3	6,1	3	6,7	2,75
1	8,4	3,67	4,95	1,5	7,98	4	7,2	3	7,5	3
1	5,95	2,5	3,95	2	5,25	3	4	2	6,1	3,25
1	5,35	2,5	6,3	1,5	6,29	3	5	3	7	3
2	5,8	1,5	5,75	2,5	5,33	2	7,2	2	5,7	2,86
2	5,35	1,75	5,85	2	4,96	3	7,4	2	7,6	3,57
2	5,6	2,5	4,45	2,5	5,57	3	6	2	6,6	2,57
2	4,7	2,5	3,9	2	4,75	2	5,3	3	7,8	3,14
2	4,45	2	4,2	2	4,58	3	5,8	1	6,2	2,71
2	6,95	3,25	6,9	3	7,91	4	7,4	3	7	2,71
2	5,95	3,25	4,75	3	7,45	4	4,9	3	6,8	2,28
2	6,4	3,5	5,75	2,5	6,02	4	5,7	3	7,5	3
2	4,6	2,5	3,9	2	4,26	1	4,9	1	6,1	2,43
2	4,05	2	3,95	2,5	4,58	2	5	2	6,1	2,28
2	5,05	2,5	4,55	2,5	6,14	2	5	1	6,5	2,14
2	5,2	2,25	5,85	2,5	6,85	4	5,9	2	7,3	3
2	5,88	2,25	6,12	2	7,45	3	6	2	6,9	3,29
2	4,1	2,5	3,4	2	5,26	2	5,2	1	6,3	2,14
2	6,05	2,5	4,25	2	4,76	4	5,1	2	7,4	2,14
2	5,35	2,75	5,65	2	7,03	2	5,8	2	6,8	2,71
2	6,95	3,25	6	2,5	6,06	3	6	3	6,3	2,71
2	5,2	3,25	5,4	2,5	4,71	4	5,4	2	6,4	3
2	7,7	2	5,4	2	7,03	3	6,1	2	7,6	3
2	5,3	2,75	5,84	2,5	5,84	4	6,6	2	6,6	2,71
2	6,45	3,5	2,85	2,5	7,25	3	6,3	3	7	3,29
2	4,85	2	4,62	2	6,59	2	5,7	2	6,4	3
2	5,6	1,75	4,6	2	6,42	3	5,6	3	7	2,71
2	4,2	2,25	3,55	2	5,82	3	5,9	2	6,6	2,86
3	4	2,33	3,47	2	5,44	1	5,3	2	7,1	2,71
3	8	2	5	2	6,36	2	5,3	3	7	3,57
3	6,3	2	5,4	3	7,56	3,5	6,9	4	6,7	3,29
3	7,35	2,33	5,3	3	7,77	3,5	5,8	2	7,2	3
3	7	2,33	7	2,5	6	2,5	7	3	7,3	3,71
3	5	1,67	4,25	2	6,84	2,5	5,4	3	6,2	3,43

Lampiran 1

3	8,15	3,33	5,4	3	7,87	3,5	7	3	7,4	3,29
3	4,8	2,33	3,95	1	5,65	2	5,8	2	7,3	3
3	6,7	2,33	5,45	3	7,16	4	6,6	3	7	2,86
3	6,15	2	5,75	2	6,9	3	5,3	2	7,5	3
3	4,9	3,33	4,25	3	6,98	2,5	5,8	2	6,8	2,43
3	4,9	2	6,15	2,5	6,18	1,5	6,2	3	7,6	3,14
3	8,05	2	5,9	3,5	8,09	3,5	6,1	3	7,2	3
3	4,5	2	4,8	2	5,26	2,5	6,5	2	7,5	3,57
3	6,8	2,33	3,75	2	8,56	3,5	6,5	2	7,4	2,43
3	5,2	2	5,05	3	7,1	2,5	6,9	2	6,9	3
3	5,8	2,33	5,25	2	6,6	3	5,8	4	6,9	3,14
3	6,3	2,67	4,2	1,5	7,2	3	6,5	1	6,6	3,14
3	5,65	2	4,3	1,5	6,22	2,5	5,7	3	6,6	3,29
3	6,2	3	4	1,5	5,95	2,5	5,2	3	7	2,71
3	5,05	2	5,45	2	7,2	2,5	6,3	3	7,6	3,29
3	7,55	3	5,25	2,5	8,76	4	6,9	3	7,9	3,14
3	3,85	1,67	4,4	2	5,67	3,5	5,1	3	6,6	3,14
3	6,45	2,33	4,2	1	7,49	2,5	6,1	1	6,9	2,43
3	8,25	3,33	5,5	4	8,8	3,5	6,1	3	7,1	2,57
4	6,05	2,67	3,6	2,5	7,91	2,5	6,8	3	6,3	2,83
4	5,3	2,33	3,6	1,5	7,74	2	4,5	2	7,1	2,83
4	4,15	2	4,05	2	3,87	1,5	5,1	1,5	5,6	2,83
4	5,55	2,67	3,3	3	4,38	1,5	6,8	2,5	6,8	3,17
4	4,45	2,33	3	3	4,11	3,5	5,9	3	7,3	3,5
4	4,95	2,33	4,7	1,5	6,38	2	6,2	1,5	6,9	2,83
4	6,25	2,33	5,5	2	7,61	1,5	6,6	3,5	7,5	2,67
4	6,95	3,67	5,15	2,5	6,48	3	6,7	3,5	7	3
4	4,45	3,67	3,95	2,5	5,33	2	5,9	3	6,7	3,33
4	4,45	3,67	4,13	2,5	7,49	3	7	2,5	6,8	3,17
4	4,7	1,33	4,25	2	4,22	1,5	4,8	1,5	7	3
4	4,55	2,33	3,1	3	3,71	1,5	4,9	2	6,4	2,67
4	5,35	3	4,6	3,5	5,95	3	7,3	3,5	7,4	2,67
4	3,4	2,67	2,85	1	3,42	2	6,3	2	6,8	3,33
4	4,25	2,67	4,44	1,5	7,03	2	6,9	2,5	6,4	3
4	3,3	2	3,6	1,5	4,47	1,5	6,3	1,5	6,4	2,83
4	4,7	3,33	4,25	2,5	5,07	2	6,1	2,5	6,4	2,67
4	6,5	3,33	5,25	3,5	6,48	1,5	7	2,5	8,1	3,33
4	6,05	2	3,85	1,5	6,78	2,5	6,1	2,5	6,6	2,67
4	4,1	2,33	4,6	1,5	4,23	1	6,8	2	7,2	3,17
4	4,55	2,33	4,1	2,5	5,08	2,5	5,4	2,5	6,5	2,83
4	5,4	3	4,48	2,5	5,51	1,5	6,9	2	7,2	3,33