UNIVERSITI SAINS MALAYSIA

Peperiksaan Semester Pertama
Sidang Akademik 1998/99

Ogos/September 1998

MSG 366/MSG 462 - Analisis Multivariat

Masa: [3 jam]

ARAHAN KEPADA CALON:

Sila pastikan bahawa kertas peperiksaan ini mengandungi EMPAT soalan di dalam SEMBILAN BELAS halaman dan TUJUH halaman lampiran yang bercetak sebelum anda memulakan peperiksaan ini.

Jawab **SEMUA** soalan. Semua soalan mesti dijawab dalam Bahasa Malaysia.

1. (a) Katakan $\underset{\sim}{A} = \begin{pmatrix} 5 & 1 \\ 4 & 2 \end{pmatrix}$.

   (i) Adakah $\underset{\sim}{A}$ tentu positif?

   (ii) Tentukan nilai-nilai eigen dan vektor-vektor eigen bagi $\underset{\sim}{A}$.

   (iii) Tuliskan penguraian spektrum bagi $\underset{\sim}{A}$.

   (b) Diberikan matriks data

$$\underset{\sim}{X} = \begin{pmatrix} -2 & 1 & 4 \\ 2 & 3 & 1 \\ -3 & 1 & 2 \end{pmatrix},$$

dan gabungan-gabungan linear

$$\underset{\sim}{b}'\underset{\sim}{x} = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

dan

$$\underset{\sim}{c}'\ \underset{\sim}{x} = \begin{pmatrix} 1 & 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix},$$

nilaikan min, varians dan kovarians sampel bagi $\underset{\sim}{b}'\ \underset{\sim}{x}$ dan $\underset{\sim}{c}'\ \underset{\sim}{x}$.

...2/-

(c) Cari anggaran kebolehjadian maksimum bagi vektor min $\underset{\sim}{\mu}$ dan matriks kovarians $\underset{\sim}{\Sigma}$ berdasarkan sampel rawak

$$\underset{\sim}{X} = \begin{pmatrix} 3 & 4 & 2 & 3 \\ 6 & 6 & 5 & 3 \end{pmatrix}$$

daripada suatu populasi normal bivariat.

(d) Katakan $\underset{\sim}{X}_1$, $\underset{\sim}{X}_2$, $\underset{\sim}{X}_3$, dan $\underset{\sim}{X}_4$ ialah vektor-vektor rawak tak bersandar yang tertabur $N_5\left(\underset{\sim}{\mu}, \underset{\sim}{\Sigma}\right)$.

(i) Cari taburan-taburan sut bagi vektor-vektor rawak

$$\underset{\sim}{V}_1 = \frac{1}{4}\underset{\sim}{X}_1 + \frac{1}{4}\underset{\sim}{X}_2 - \frac{1}{4}\underset{\sim}{X}_3 - \frac{1}{4}\underset{\sim}{X}_4 \quad \text{dan}$$

$$\underset{\sim}{V}_2 = \frac{1}{4}\underset{\sim}{X}_1 - \frac{1}{4}\underset{\sim}{X}_2 + \frac{1}{4}\underset{\sim}{X}_3 - \frac{1}{4}\underset{\sim}{X}_4 .$$

(ii) Cari fungsi ketumpatan tercantum bagi $\left(\underset{\sim}{V}_1, \underset{\sim}{V}_2\right)$ dengan $\underset{\sim}{V}_1$ dan $\underset{\sim}{V}_1$ tertakrif seperti di dalam (i).

(100/100)

2. (a) Cerapan-cerapan pada dua sambutan, X_1 dan X_2, dikutip bagi 3 rawatan. Vektor-vektor cerapan $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ adalah seperti berikut:

Rawatan 1 : $\begin{pmatrix} 5 \\ 6 \end{pmatrix}, \begin{pmatrix} 4 \\ 8 \end{pmatrix}, \begin{pmatrix} 7 \\ 5 \end{pmatrix}, \begin{pmatrix} 4 \\ 9 \end{pmatrix}$

Rawatan 2 : $\begin{pmatrix} 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 5 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

Rawatan 3 : $\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 4 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}$.

(i) Bina jadual MANOVA satu hala bagi data ini.

(ii) Nilaikan lambda Wilks, Λ^*, dan uji bagi kesan rawatan. Andaikan $\alpha = .05$.

(iii) Ulangkan ujian dengan menggunakan penghampiran khi-kuasa dua dengan pembetulan Bartlett.

Bandingkan kesimpulan-kesimpulan anda.

...3/-

(b) Prestasi (markah dari jumlah 100) yang dicapai oleh sekumpulan 10 pelajar di dalam tiga peperiksaan ijazah yang berlainan memberi suatu vektor min $\underset{\sim}{\bar{x}}' = (55, 48, 56)$ dan matriks kovarians

$$\underset{\sim}{S} = \begin{pmatrix} 35 & 3 & 15 \\ 3 & 20 & 12 \\ 15 & 12 & 24 \end{pmatrix}$$

Dengan menganggapkan bahawa cerapan-cerapan adalah tertabur *normal trivariat*, ujikan hipotesis bahawa min populasi $\underset{\sim}{\mu}' = (50, 50, 50)$ melawan hipotesis alternatif bahawa $\underset{\sim}{\mu}' \neq (50, 50, 50)$. Gunakan $\alpha = .05$.

(100/100)

3. (a) Tulis nota pendek tentang tajuk-tajuk di bawah:

(i) Komponen prinsipal
(ii) Analisis pembezalayan
(iii) Kaedah kelompok berhierarki (*hierarchical*)

(b) Diberi lima objek mempunyai jarak berikut:

$$\underset{\sim}{D} = \{d_{ik}\} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{pmatrix} 0 & 4 & 12 & 7 & 10 \\ 4 & 0 & 3 & 10 & 8 \\ 12 & 3 & 0 & 8 & 11 \\ 7 & 10 & 8 & 0 & 6 \\ 10 & 8 & 11 & 6 & 0 \end{pmatrix} \end{array}$$

gunakan (i) kaedah pautan tunggal,
(ii) kaedah pautan lengkap

daripada prosedur "*agglomerative hierarchical*" untuk memperoleh gambarajah-gambarajah dendogram untuk lima objek itu.

4. Bagi setiap bahagian yang berikut, tulis suatu perenggan yang menghuraikan kesimpulan-kesimpulan anda. *Output-output* bagi setiap bahagian dilampirkan pada akhir soalan ini.

(a) Andaikan jabatan periklanan telah menyediakan tiga iklan bagi mengemukakan suatu produk baru dan jabatan tersebut berminat mengecamkan iklan yang terbaik. Iklan pertama menggunakan "*humorous appeal*", iklan kedua menggunakan "*emotional appeal*", dan iklan ketiga menggunakan "*comparative approach*".
Dipercayai bahawa jantina responden memang ada kesan pada "*preference*"nya bagi jenis iklan.

...4/-

Suatu ujikaji dijalankan dengan 12 orang lelaki dan 12 orang perempuan didedahkan kepada satu daripada iklan-iklan tersebut. Kumpulan lelaki dibahagi secara rawak ke dalam tiga subkumpulan dengan empat orang setiap satu. Setiap subkumpulan didedahkan kepada suatu iklan yang berbeza dan responden-responden ditanya menilaikan iklan itu terhadap bagaimana "*informative*" (Y_1) dan "*believable*" (Y_2) iklan itu pada suatu skala sebelas titik dengan 1 mewakili "*low informativeness*" dan "*low believability*", dan 11 mewakili: "*high informativeness*" dan "*high believebility*".
Prosedur di atas diulangi bagi kumpulan perempuan. Data dikutip dianalisis melalui prosedur MANOVA daripada SPSS, dan *output* ditunjuk dalam OUTPUT(a).

(b) Penggunaan protin di negeri-negeri Eropah dikelaskan ke dalam sembilan kumpulan makanan (daging merah, daging putih, ..., buah-buahan dan sayur-sayuran). Data dikutip dianalisis melalui prosedur FACTOR ANALYSIS daripada SPSS. *Output* yang diperolehi ditunjuk dalam OUTPUT (b).

(c) Pertimbangkan data dalam (b). Prosedur HIERARCHICAL CLUSTER ANALYSIS dari SPSS dijalankan. *Output* yang diperolehi ditunjuk dalam OUTPUT (c).

(d) Data kewangan daripada 12 syarikat "*most-admired*" dan 12 syarikat "*least-admired*" dikutip. Pembolehubah-pembolehubah dikutip adalah: EBITASS (*Earnings before interest and taxes to total asset*") dan ROTC (*Return on total capital*).
Prosedur DISCRIMINANT ANALYSIS daripada SPSS dijalankan, dan *output* yang didapati ditunjuk dalam OUTPUT (d).

(100/100)

OUTPUT (a)

```
* * * * * * A n a l y s i s   o f   V a r i a n c e * * * * * *

        24 cases accepted.
         0 cases rejected because of out-of-range factor values.
         0 cases rejected because of missing data.
         6 non-empty cells.

         1 design will be processed.

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

* * * * * * A n a l y s i s   o f   V a r i a n c e -- design   1 * * * * * *

EFFECT .. IKLAN BY JANTINA
Multivariate Tests of Significance (S = 2, M = -1/2, N = 7 1/2)

Test Name         Value  Approx. F Hypoth. DF   Error DF  Sig. of F

Pillais          .92596    7.75922       4.00      36.00       .000
Hotellings      6.65546   26.62185       4.00      32.00       .000
Wilks            .12409   15.62986       4.00      34.00       .000
Roys             .86832
Note.. F statistic for WILKS' Lambda is exact.

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
EFFECT .. IKLAN BY JANTINA (Cont.)
Univariate F-tests with (2,18) D. F.

Variable   Hypoth. SS   Error SS Hypoth. MS   Error MS          F  Sig. of F

Y1          156.00000   24.00000   78.00000    1.33333   58.50000       .000
Y2          149.33333   48.00000   74.66667    2.66667   28.00000       .000
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -


EFFECT .. JANTINA
Multivariate Tests of Significance (S = 1, M = 0, N = 7 1/2)

Test Name         Value    Exact F Hypoth. DF   Error DF  Sig. of F

Pillais          .23226    2.57143       2.00      17.00       .106
Hotellings       .30252    2.57143       2.00      17.00       .106
Wilks            .76774    2.57143       2.00      17.00       .106
Roys             .23226
Note.. F statistics are exact.

- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
EFFECT .. JANTINA (Cont.)
Univariate F-tests with (1,18) D. F.

Variable   Hypoth. SS   Error SS Hypoth. MS   Error MS          F  Sig. of F

Y1            6.00000   24.00000    6.00000    1.33333    4.50000       .048
Y2            2.66667   48.00000    2.66667    2.66667    1.00000       .331
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
```

13 Jul 98 SPSS for MS WINDOWS Release 6.1 Page 43

* * * * * * A n a l y s i s   o f   V a r i a n c e -- design 1 * * * * * *

EFFECT .. IKLAN
Multivariate Tests of Significance (S = 2, M = -1/2, N = 7 1/2)

Test Name	Value	Approx. F	Hypoth. DF	Error DF	Sig. of F
Pillais	.37696	2.09032	4.00	36.00	.102
Hotellings	.60504	2.42017	4.00	32.00	.069
Wilks	.62304	2.26867	4.00	34.00	.082
Roys	.37696				

Note.. F statistic for WILKS' Lambda is exact.

- -

EFFECT .. IKLAN (Cont.)
Univariate F-tests with (2,18) D. F.

Variable	Hypoth. SS	Error SS	Hypoth. MS	Error MS	F	Sig. of F
Y1	12.00000	24.00000	6.00000	1.33333	4.50000	.026
Y2	5.33333	48.00000	2.66667	2.66667	1.00000	.387

- -

OUTPUT (b)

4 Jul 98 SPSS for MS WINDOWS Release 6.0 Page 1

- - - - - - - - - - - - F A C T O R   A N A L Y S I S - - - - - - - - - - - -

nalysis number 1   Listwise deletion of cases with missing values

	Mean	Std Dev	Label
EREAL	32.24800	10.97479	
AGING_M	9.82800	3.34708	daging merah
AGING_P	7.89600	3.69408	daging putih
VEG	4.13600	1.80390	fruits, vegetables
KAN	4.28400	3.40253	
UTS	3.07200	1.98568	pulses, nuts, oil-seeds
TARCHY	4.27600	1.63408	makanan 'starchy'
USU	17.11200	7.10542	
ELUR	2.93600	1.11762	

umber of Cases = 25

orrelation Matrix:

	CEREAL	DAGING_M	DAGING_P	F_VEG	IKAN	NUTS	STARCHY
EREAL	1.00000						
AGING_M	-.49988	1.00000					
AGING_P	-.41380	.15300	1.00000				
VEG	.04655	-.07422	-.06132	1.00000			
KAN	-.52423	.06096	-.23401	.26614	1.00000		
UTS	.65100	-.34945	-.63496	.37497	-.14715	1.00000	
TARCHY	-.53326	.13543	.31377	.08441	.40385	-.47431	1.00000
USU	-.59274	.50293	.28148	-.40836	.13788	-.62109	.22241
ELUR	-.71244	.58561	.62041	-.04552	.06557	-.55978	.45223

	SUSU	TELUR
USU	1.00000	
ELUR	.57553	1.00000

- - - - - - - - - - - - F A C T O R A N A L Y S I S - - - - - - - - - - - -

Extraction 1 for analysis 1, Principal Components Analysis (PC)

Initial Statistics:

Variable	Communality	*	Factor	Eigenvalue	Pct of Var	Cum Pct
		*				
CEREAL	1.00000	*	1	4.00644	44.5	44.5
DAGING_M	1.00000	*	2	1.63500	18.2	62.7
DAGING_P	1.00000	*	3	1.12792	12.5	75.2
F_VEG	1.00000	*	4	.95466	10.6	85.8
IKAN	1.00000	*	5	.46384	5.2	91.0
NUTS	1.00000	*	6	.32513	3.6	94.6
STARCHY	1.00000	*	7	.27161	3.0	97.6
SUSU	1.00000	*	8	.11629	1.3	98.9
TELUR	1.00000	*	9	.09911	1.1	100.0

PC extracted 3 factors.

Factor Matrix:

	Factor 1	Factor 2	Factor 3
CEREAL	-.87619	-.29855	.10187
DAGING_M	.60571	-.07193	-.31604
DAGING_P	.62161	-.30286	.66260
F_VEG	-.22102	.68561	.43284
IKAN	.27152	.82707	-.34120
NUTS	-.84135	.18325	-.05776
STARCHY	.59497	.45115	.25805
SUSU	.75606	-.23603	-.40958
TELUR	.85404	-.04518	.19279

Final Statistics:

Variable	Communality	*	Factor	Eigenvalue	Pct of Var	Cum Pct
		*				
CEREAL	.86722	*	1	4.00644	44.5	44.5
DAGING_M	.47193	*	2	1.63500	18.2	62.7
DAGING_P	.91716	*	3	1.12792	12.5	75.2
F_VEG	.70626	*				

14 Jul 98 SPSS for MS WINDOWS Release 6.0 Page 3

- - - - - - - - - - - - FACTOR ANALYSIS - - - - - - - - - - - - -

Variable	Communality	*	Factor	Eigenvalue	Pct of Var	Cum Pct
IKAN	.87419	*				
NUTS	.74478	*				
STARCHY	.62412	*				
SUSU	.79510	*				
TELUR	.76860	*				

VARIMAX rotation 1 for extraction 1 in analysis 1 - Kaiser Normalization.

VARIMAX converged in 8 iterations.

Rotated Factor Matrix:

	Factor 1	Factor 2	Factor 3
CEREAL	-.49201	-.51398	-.60081
DAGING_M	.19630	.62549	.20533
DAGING_P	.93782	.06741	-.18198
F_VEG	.07908	-.69116	.47150
IKAN	-.15648	.08812	.91757
NUTS	-.63954	-.56755	-.11687
STARCHY	.54010	.03549	.57545
SUSU	.25137	.84635	.12487
TELUR	.72907	.43278	.22308

Factor Transformation Matrix:

	Factor 1	Factor 2	Factor 3
Factor 1	.68518	.63795	.35149
Factor 2	-.11759	-.37936	.91775
Factor 3	.71882	-.67015	-.18491

Hi-Res Chart # 1:Factor plot of factors 1, 2, 3

- - - - - - - - - - - F A C T O R   A N A L Y S I S - - - - - - - - - - -

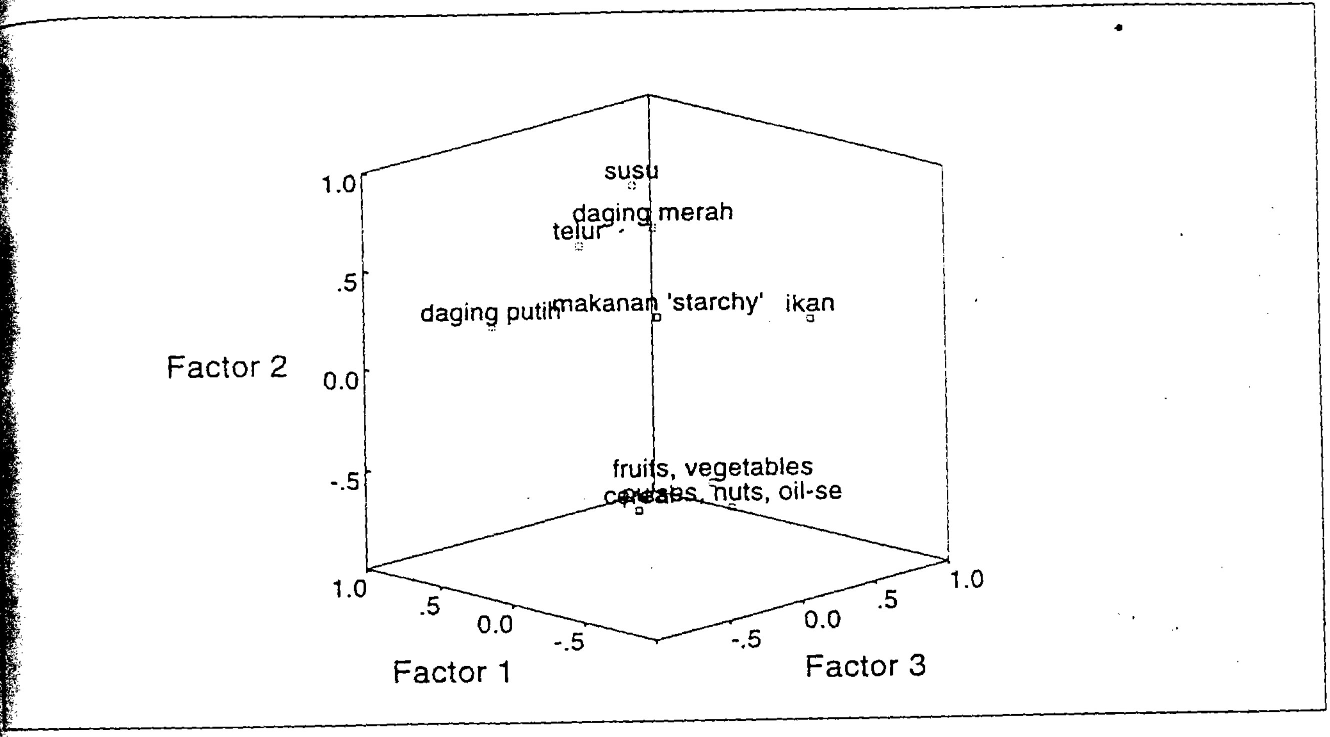
Factor Score Coefficient Matrix:

	Factor 1	Factor 2	Factor 3
CEREAL	-.06345	-.13077	-.26115
DAGING_M	-.09265	.30091	.06458
DAGING_P	.55036	-.22443	-.22409
F_VEG	.18874	-.45144	.29449
IKAN	-.23050	.05406	.54400
NUTS	-.19388	-.14217	.03852
STARCHY	.23376	-.16326	.26313
SUSU	-.11475	.41851	.00099
TELUR	.27217	.03193	.01796

Covariance Matrix for Estimated Regression Factor Scores:

	Factor 1	Factor 2	Factor 3
Factor 1	1.00000		
Factor 2	.00000	1.00000	
Factor 3	.00000	.00000	1.00000

Factor Plot in Rotated Factor Space



OUTPUT (c)

* * * * * * * * * * * * * P R O X I M I T I E S * * * * * * * * * * * * * *

Data Information

25 unweighted cases accepted.
0 cases rejected because of missing value.

Squared Euclidean measure used.

\- -

* * * * * * H I E R A R C H I C A L C L U S T E R A N A L Y S I S * * * * * *

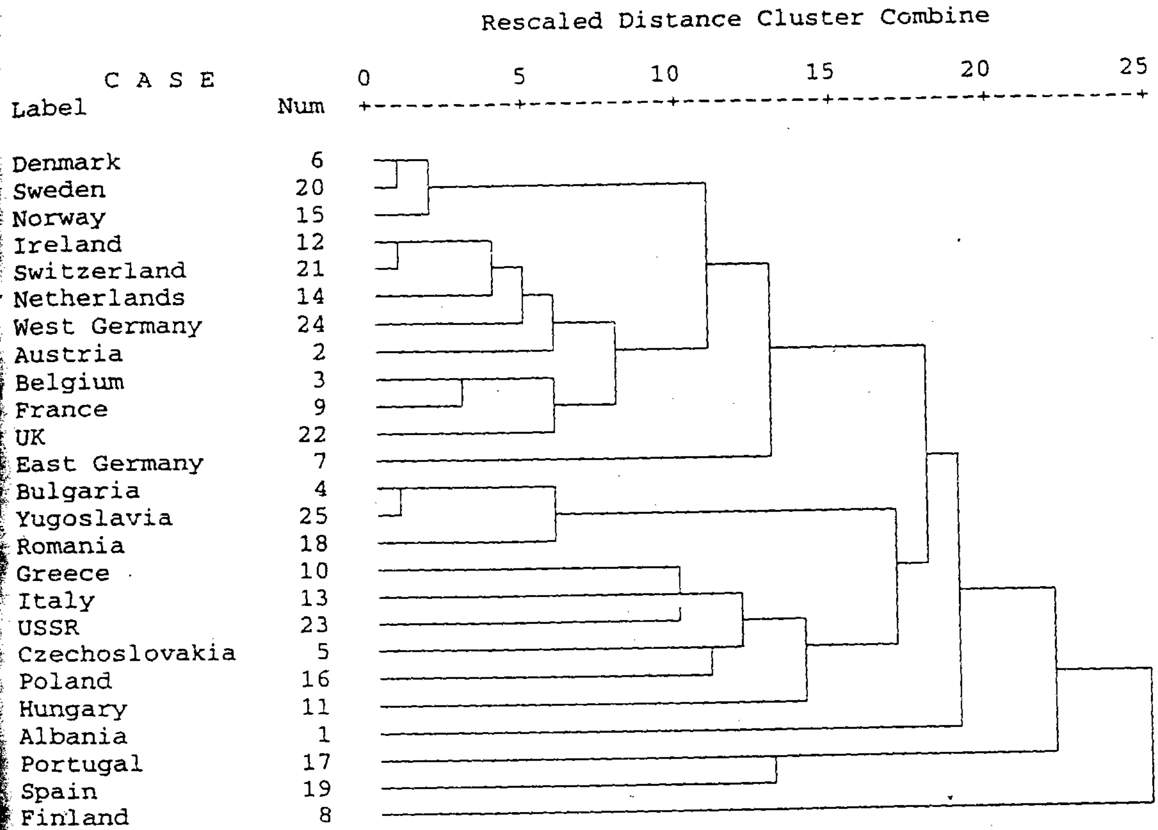
Agglomeration Schedule using Single Linkage

Stage	Clusters Combined Cluster 1	Cluster 2	Coefficient	Stage Cluster 1st Appears Cluster 1	Cluster 2	Next Stage
1	6	20	23.000000	0	0	4
2	4	25	23.770000	0	0	10
3	12	21	25.980000	0	0	6
4	6	15	30.889999	1	0	15
5	3	9	36.090000	0	0	9
6	12	14	39.660000	3	0	7
7	12	24	42.580002	6	0	8
8	2	12	45.750000	0	7	11
9	3	22	46.650002	5	0	11
10	4	18	47.759998	2	0	20
11	2	3	56.639999	8	9	15
12	10	13	63.630001	0	0	13
13	10	23	66.839996	12	0	16
14	5	16	68.150002	0	0	16
15	2	6	69.889999	11	4	18
16	5	10	75.779999	14	13	19
17	17	19	77.239998	0	0	23
18	2	7	79.720001	15	0	21
19	5	11	84.269997	16	0	20
20	4	5	97.620003	10	19	21
21	2	4	98.849998	18	20	22
22	1	2	106.489998	0	21	23
23	1	17	119.040001	22	17	24
24	1	8	133.850006	23	0	0

13 Jul 98 SPSS for MS WINDOWS Release 6.1 Page 8

* * * * * * H I E R A R C H I C A L   C L U S T E R   A N A L Y S I S * * * * * *

Dendrogram using Single Linkage



OUTPUT (d)

--------- D I S C R I M I N A N T   A N A L Y S I S ---------

On groups defined by GROUP   Firm

24 (Unweighted) cases were processed.
0 of these were excluded from the analysis.
24 (Unweighted) cases will be used in the analysis.

Number of cases by group

GROUP	Number of cases Unweighted	Weighted	Label
1	12	12.0	Most Admired
2	12	12.0	Least Admired
Total	24	24.0	

Group means

GROUP	EBITASS	ROTC
1	.19133	.18350
2	.00333	.00125
Total	.09733	.09238

Group standard deviations

GROUP	EBITASS	ROTC
1	.05324	.03022
2	.04492	.06852
Total	.10743	.10652

Wilks' Lambda (U-statistic) and univariate F-ratio
with 1 and 22 degrees of freedom

Variable	Wilks' Lambda	F	Significance
EBITASS	.20108	87.4076	.0000
ROTC	.23638	71.0699	.0000

3 Jul 98 SPSS for MS WINDOWS Release 6.1 Page 23

- - - - - - - - D I S C R I M I N A N T   A N A L Y S I S - - - - - - - - -

On groups defined by GROUP Firm

Analysis number 1

Direct method: all variables passing the tolerance test are entered.

Minimum tolerance level.................. .00100

Canonical Discriminant Functions

Maximum number of functions............. 1
Minimum cumulative percent of variance... 100.00
Maximum significance of Wilks' Lambda.... 1.0000

Prior probability for each group is .50000

Classification function coefficients
(Fisher's linear discriminant functions)

GROUP =	1 Most Admired	2 Least Admire d
EBITASS	61.2374430	2.5511703
ROTC	21.0268971	-1.4044441
(Constant)	-8.4807470	-.6965214

Canonical Discriminant Functions

Fcn	Eigenvalue	Pct of Variance	Cum Pct	Canonical Corr		After Fcn	Wilks' Lambda	Chi-square	df	Sig
					:	0	.195162	34.312	2	.0000
1*	4.1239	100.00	100.00	.8971	:					

* Marks the 1 canonical discriminant functions remaining in the analysis.

Standardized canonical discriminant function coefficients

	Func 1
EBITASS	.74337
ROTC	.30547

Structure matrix:

Pooled within-groups correlations between discriminating variables
and canonical discriminant functions
(Variables ordered by size of correlation within function)

	Func 1
EBITASS	.98154
ROTC	.88506

Canonical discriminant functions evaluated at group means (group centroids)

Group	Func 1
1	1.94429
2	-1.94429

13 Jul 98 SPSS for MS WINDOWS Release 6.1 Page 25

Case Number	Mis Val	Sel	Actual Group		Highest Probability Group	P(D/G)	P(G/D)	2nd Highest Group	P(G/D)	Discrim Scores
1			1		1	.6088	.9962	2	.0038	1.4326
2			1		1	.6807	.9999	2	.0001	2.3558
3			1		1	.7930	.9998	2	.0002	2.2067
4			1		1	.1008	1.0000	2	.0000	3.5853
5			1		1	.8884	.9997	2	.0003	2.0846
6			1		1	.6329	.9999	2	.0001	2.4220
7			1		1	.5606	.9950	2	.0050	1.3624
8			1		1	.2669	1.0000	2	.0000	3.0545
9			1		1	.0567	.5373	2	.4627	.0384
10			1		1	.3375	.9788	2	.0212	.9853
11			1		1	.9502	.9993	2	.0007	1.8818
12			1		1	.9823	.9994	2	.0006	1.9222
13			2		2	.6764	.9999	1	.0001	-2.3617
14			2		2	.4286	.9888	1	.0112	-1.1528
15			2		2	.4693	.9914	1	.0086	-1.2207
16			2		2	.1512	1.0000	1	.0000	-3.3795
17			2		2	.1190	1.0000	1	.0000	-3.5032
18			2		2	.9312	.9996	1	.0004	-2.0307
19			2		2	.8081	.9987	1	.0013	-1.7014
20			2	**	1	.0616	.5727	2	.4273	.0753
21			2		2	.3096	1.0000	1	.0000	-2.9605
22			2		2	.3218	.9761	1	.0239	-.9535
23			2		2	.5563	.9999	1	.0001	-2.5326
24			2		2	.7384	.9981	1	.0019	-1.6104

Symbols used in plots

Symbol	Group	Label
1	1	Most Admired
2	2	Least Admired

```
                Histogram for group 1  Most Admired

                     Canonical Discriminant Function 1
      4 +                                                                 +
        |                                                                 |
  F     |                                                                 |
  r   3 +                                                                 +
  e     |                                                                 |
  q     |                                                                 |
  u     |                                                                 |
  e   2 +                                  1  1 1                         +
  n     |                                  1  1 1                         |
  c     |                                  1  1 1                         |
  y     |                                  1  1 1                         |
      1 +                              1 1 1111  1  1                     +
        |                        1     1 1 1111  1  1                     |
        |                        1     1 1 1111  1  1                     |
        |                        1     1 1 1111  1  1                     |
        |                        1                                        |
        X---------+---------+---------+---------+---------+---------------X
      out       -4.0      -2.0        .0       2.0       4.0            out
  Class    222222222222222222222222222222211111111111111111111111111111111
Centroids                                       1

                Histogram for group 2  Least Admired

                     Canonical Discriminant Function 1
      4 +                                                                 +
        |                                                                 |
  F     |                                                                 |
  r   3 +                                                                 +
  e     |                                                                 |
  q     |                                                                 |
  u     |                                                                 |
  e   2 +                        2                                        +
  n     |                        2                                        |
  c     |                        2                                        |
  y     |                        2                                        |
      1 +           22 2 22 222 22     2                                  +
        |           22 2 22 222 22     2                                  |
        |           22 2 22 222 22     2                                  |
        |           22 2 22 222 22     2                                  |
        X---------+---------+---------+---------+---------+---------------X
      out       -4.0      -2.0        .0       2.0       4.0            out
  Class    222222222222222222222222222222211111111111111111111111111111111
Centroids                        2
```


...19

```
                         All-groups Stacked Histogram

                      Canonical Discriminant Function 1
   4 +                                                                 +
     |                                                                 |
     |                                                                 |
     |                                                                 |
   3 +                                                                 +
     |                                                                 |
     |                                                                 |
     |                                                                 |
   2 +                        2     2       1  1 1                     +
     |                        2     2       1  1 1                     |
     |                        2     2       1  1 1                     |
     |                        2     2       1  1 1                     |
   1 +             22 2 22 222 22   1     1 1 1111   1  1              +
     |             22 2 22 222 22   1     1 1 1111   1  1              |
     |             22 2 22 222 22   1     1 1 1111   1  1              |
     |             22 2 22 222 22   1     1 1 1111   1  1              |
     X---------+---------+---------+---------+---------+---------------X
    out      -4.0      -2.0        .0       2.0       4.0            out
 Class  222222222222222222222222222222221111111111111111111111111111111
entroids                  2                 1
```

lassification results -

Actual Group	No. of Cases	Predicted Group Membership 1	2
roup 1 ost Admired	12	12 100.0%	0 .0%
roup 2 east Admired	12	1 8.3%	11 91.7%

ercent of "grouped" cases correctly classified: 95.83%

assification processing summary

24 (Unweighted) cases were processed.
0 cases were excluded for missing or out-of-range group codes.
0 cases had at least one missing discriminating variable.
24 (Unweighted) cases were used for printed output.

- oooOooo -

MSG 466 - ANALISIS MULTIVARIAT

LAMPIRAN

Tatatanda adalah seperti di dalam kuliah.

1. Penguraian spektrum bagi suatu matriks simetrik k x k, $\underset{\sim}{A}$ diberikan oleh

$$\underset{\sim}{A} = \lambda_1 \underset{\sim}{e}_1 \underset{\sim}{e}_1' + \lambda_2 \underset{\sim}{e}_2 \underset{\sim}{e}_2' + \cdots + \lambda_k \underset{\sim}{e}_k \underset{\sim}{e}_k'$$

di mana $\lambda_1, \lambda_2, \ldots, \lambda_k$ adalah nilai-nilai eigen $\underset{\sim}{A}$ dan $\underset{\sim}{e}_1, \underset{\sim}{e}_2, \ldots, \underset{\sim}{e}_k$ adalah vektor-vektor eigen terpiawai yang berkaitan.

2. Katakan $\underset{\sim}{X}$ mempunyai $E(\underset{\sim}{X}) = \underset{\sim}{\mu}$ dan Kov $(\underset{\sim}{X}) = \underset{\sim}{\Sigma}$. Maka $\underset{\sim}{c}'\underset{\sim}{X}$ mempunyai min, $\underset{\sim}{c}'\underset{\sim}{\mu}$ dan varians, $\underset{\sim}{c}'\underset{\sim}{\Sigma}\underset{\sim}{c}$.

3. f.k.k. normal bivariat:

$$f(x_1, x_2) = \frac{1}{2\pi\sqrt{\sigma_{11}\sigma_{22}(1-\rho_{12}^2)}} \times \exp\left\{ -\frac{1}{2(1-\rho_{12}^2)} \left[\left(\frac{x_1 - \mu_1}{\sqrt{\sigma_{11}}}\right)^2 + \left(\frac{x_2 - \mu_2}{\sqrt{\sigma_{22}}}\right)^2 - 2\rho_{12}\left(\frac{x_1 - \mu_1}{\sqrt{\sigma_{11}}}\right)\left(\frac{x_2 - \mu_2}{\sqrt{\sigma_{22}}}\right)\right] \right\}$$

4. f.k.k. normal multivariat:

$$f(\underset{\sim}{x}) = \frac{1}{(2\pi)^{p/2} |\Sigma|^{1/2}} e^{-(1/2)(\underset{\sim}{x} - \underset{\sim}{\mu})'\underset{\sim}{\Sigma}^{-1}(\underset{\sim}{x} - \underset{\sim}{\mu})}$$

5. Jika $\underset{\sim}{X} \sim N_p(\underset{\sim}{\mu}, \underset{\sim}{\Sigma})$, maka $\underset{\sim}{A}\underset{\sim}{X} \sim N_q(\underset{\sim}{A}\underset{\sim}{\mu}, \underset{\sim}{A}\underset{\sim}{\Sigma}\underset{\sim}{A}')$.

6. Satu sampel:

(a) $T^2 = n(\bar{\underset{\sim}{X}} - \underset{\sim}{\mu})' \underset{\sim}{S}^{-1}(\bar{\underset{\sim}{X}} - \underset{\sim}{\mu})$

$$\bar{\underset{\sim}{X}} = \frac{1}{n}\sum_{j=1}^{n} \underset{\sim}{X}_j, \quad \underset{\sim}{S} = \frac{1}{n-1}\sum_{j=1}^{n}(\underset{\sim}{X}_j - \bar{\underset{\sim}{X}})(\underset{\sim}{X}_j - \bar{\underset{\sim}{X}})'$$

$$T^2 \sim \frac{(n-1)p}{n-p} F_{p,n-p}$$

(b) Lambda Wilks $\Lambda^{2/n} = |\hat{\underset{\sim}{\Sigma}}| \Big/ |\hat{\underset{\sim}{\Sigma}}_0| = \left[1 + \frac{T^2}{(n-1)}\right]^{-1}$

(c) Selang keyakinan serentak 100(1-α)% bagi $\underset{\sim}{\ell}'\underset{\sim}{\mu}$:

$$\underset{\sim}{\ell}'\bar{\underset{\sim}{X}} \pm \sqrt{\frac{p(n-1)}{n(n-p)} F_{p,n-p}(\alpha)\ \underset{\sim}{\ell}'\underset{\sim}{S}\,\underset{\sim}{\ell}}$$

(d) Selang keyakinan serentak Bonferroni 100(1-α)% bagi

μ_i, $i = 1, \ldots, p$:

$$\bar{X}_i \pm t_{n-1}\left(\frac{\alpha}{2p}\right)\sqrt{\frac{S_{ii}}{n}}$$

7. Dua sampel tak bersandar:

(a) $$T^2 = \left[\bar{\underset{\sim}{X}}_1 - \bar{\underset{\sim}{X}}_2 - (\underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2)\right]' \left[\left[\frac{1}{n_1} + \frac{1}{n_2}\right]\underset{\sim}{S}_p\right]^{-1} \left[\bar{\underset{\sim}{X}}_1 - \bar{\underset{\sim}{X}}_2 - \left(\underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2\right)\right]$$

$$T^2 \sim \frac{\left(n_1 + n_2 - 2\right)p}{\left(n_1 + n_2 - p - 1\right)} F_{p,\, n_1 + n_2 - p - 1}$$

(b) Selang keyakinan serentak 100(1-α)% bagi $\underset{\sim}{\ell}'(\underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2)$:

$$\underset{\sim}{\ell}'\left[\underset{\sim}{\bar{X}}_1 - \underset{\sim}{\bar{X}}_2\right] \pm c \sqrt{\underset{\sim}{\ell}'\left[\frac{1}{n_1} + \frac{1}{n_2}\right] \underset{\sim}{S}_p \underset{\sim}{\ell}}$$

$$\text{di mana } c^2 = \frac{\left[n_1 + n_2 - 2\right]p}{n_1 + n_2 - p - 1} F_{p,\, n_1 + n_2 - p - 1}$$

8. <u>MANOVA satu-hala:</u>

(a) $$\underset{\sim}{B} = \sum_{\ell=1}^{g} n_\ell (\underset{\sim}{\bar{x}}_\ell - \underset{\sim}{\bar{x}})(\underset{\sim}{\bar{x}}_\ell - \underset{\sim}{\bar{x}})'$$

$$\underset{\sim}{W} = \sum_{\ell=1}^{g} \sum_{j=1}^{n_\ell} \left[\underset{\sim}{x}_{\ell j} - \underset{\sim}{\bar{x}}_\ell\right]\left[\underset{\sim}{x}_{\ell j} - \underset{\sim}{\bar{x}}_\ell\right]' = (n_1 - 1)S_1^2 + \ldots + (n_g -$$

$$\Lambda^* = \frac{|\underset{\sim}{W}|}{|\underset{\sim}{B} + \underset{\sim}{W}|}$$

(b) Selang keyakinan serentak 100(1-α)% bagi $\tau_{ki} - \tau_{\ell i}$:

$$\bar{X}_{ki} - \bar{X}_{\ell i} \pm t_{n-g}\left[\frac{\alpha}{pg(g-1)}\right] \sqrt{\frac{W_{ii}}{n-g}\left[\frac{1}{n_k} + \frac{1}{n_\ell}\right]}\ ,$$

$i = 1, 2, \ldots, p \quad , \quad \ell < k = 1, 2, \ldots, g$

9. Andaikan $\underset{\sim}{E}$ mempunyai d.k. m_E dan $\underset{\sim}{H}$ mempunyai d.k. m_H.

$$\text{Katakan } \Lambda = \frac{|\underset{\sim}{E}|}{|\underset{\sim}{E} + \underset{\sim}{H}|}$$

Maka (1) Untuk p = 1,

$$\left[\frac{1 - \Lambda}{\Lambda}\right] \frac{m_E}{m_H} \sim F_{m_H, m_E} \quad \text{bagi sebarang } m_H .$$

(2) Untuk $m_H = 1$,

$$\left[\frac{1 - \Lambda}{\Lambda}\right] \frac{m_E+1-p}{P} \sim F_{P, m_E+1-p} \quad \text{bagi sebarang P.}$$

(3) Untuk p = 2,

$$\left[\frac{1 - \Lambda^{1/2}}{\Lambda^{1/2}}\right] \left[\frac{m_E - 1}{m_H}\right] \sim F_{2m_H, 2(m_E-1)}$$

untuk $m_H \geq 2$.

(4) Untuk $m_H = 2$,

$$\left[\frac{1 - \Lambda^{1/2}}{\Lambda^{1/2}}\right] \left[\frac{m_E + 1 - p}{p}\right] \sim F_{2p, 2(m_E+1-p)}$$

untuk $p \geq 2$.

Pembetulan Bartlett: Katakan $n_o = m_E + m_H$.

Bagi m_E besar,

$$-f \log \Lambda \sim X^2_{pm_H}$$

di mana
$$f = m_E - \frac{1}{2}\left[p - m_H + 1\right]$$
$$= n_o - \frac{1}{2}\left[p + m_H + 1\right]$$

10. MANOVA dua-hala:

$$SSP_{faktor\ 1} = \sum_{\ell=1}^{g} bn\left[\bar{\underset{\sim}{x}}_{\ell.} - \bar{\underset{\sim}{x}}\right]\left[\bar{\underset{\sim}{x}}_{\ell.} - \bar{\underset{\sim}{x}}\right]'$$

$$SSP_{faktor\ 2} = \sum_{k=1}^{b} gn\left(\bar{\underset{\sim}{x}}_{.k} - \bar{\underset{\sim}{x}}\right)\left(\bar{\underset{\sim}{x}}_{.k} - \bar{\underset{\sim}{x}}\right)'$$

$$SSP_{\substack{tindakan\\bersaling}} = \sum_{\ell=1}^{g}\sum_{k=1}^{b} n\left(\bar{\underset{\sim}{x}}_{\ell k} - \bar{\underset{\sim}{x}}_{\ell.} - \bar{\underset{\sim}{x}}_{.k} + \bar{\underset{\sim}{x}}\right)$$

$$\left(\bar{\underset{\sim}{x}}_{\ell k} - \bar{\underset{\sim}{x}}_{\ell.} - \bar{\underset{\sim}{x}}_{.k} + \bar{\underset{\sim}{x}}\right)'$$

$$SSP_{residual} = \sum_{\ell=1}^{g}\sum_{k=1}^{b}\sum_{r=1}^{n}\left(\underset{\sim}{x}_{\ell k r} - \bar{\underset{\sim}{x}}_{\ell k}\right)\left(\underset{\sim}{x}_{\ell k r} - \bar{\underset{\sim}{x}}_{\ell k}\right)'$$

11. Komponen Prinsipal

(a) $\underset{\sim}{Y}_i = \underset{\sim}{e}'_i \underset{\sim}{X}$, $i = 1, 2, \ldots, p.$

$$\rho_{Y_i, X_k} = \frac{e_{ki}\sqrt{\lambda_i}}{\sqrt{\sigma_{kk}}} , \qquad i, k = 1, 2, \ldots, p.$$

(b) $\underset{\sim}{Y}_i = \underset{\sim}{e}'_i \underset{\sim}{Z}$

$$\rho_{Y_i, Z_k} = e_{ki}\sqrt{\lambda_i} , \qquad i, k = 1, 2, \ldots, p.$$

12. Analisis Faktor

(a) $\underset{\sim}{X} - \underset{\sim}{\mu} = \underset{\sim}{L}\,\underset{\sim}{F} + \underset{\sim}{\epsilon}$

(b) $Kov(\underset{\sim}{X}) = \underset{\sim}{L}\,\underset{\sim}{L}' + \underset{\sim}{\Psi}$

$Kov(\underset{\sim}{X}, \underset{\sim}{F}) = \underset{\sim}{L}$

(c) $h_i^2 = \ell_{i1}^2 + \ell_{i2}^2 + \ldots + \ell_{im}^2$, $i = 1, 2, \ldots, p$.

$\sigma_{ii} = h_i^2 + \psi_i$, $i = 1, 2, \ldots, p$.

(d) Kriterium varimax: Pilih transformasi ortogon T yang menjadikan

$$V = \frac{1}{p} \sum_{j=1}^{m} \left[\sum_{i=1}^{p} \tilde{\ell}_{ij}^{*4} - \left(\sum_{i=1}^{p} \tilde{\ell}_{ij}^{*2} \right)^2 \Big/ p \right]$$

sebesar yang mungkin.

13. Analisis Pembezalayan

(a) $Y = \underset{\sim}{\ell}'\underset{\sim}{X} = (\underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2)' \underset{\sim}{\Sigma}^{-1} \underset{\sim}{X}$

$$m = \frac{1}{2}\left(\underset{\sim}{\mu}_1 - \underset{\sim}{\mu}_2\right)' \underset{\sim}{\Sigma}^{-1} \left(\underset{\sim}{\mu}_1 + \underset{\sim}{\mu}_2\right)$$

(b) $y = \hat{\underset{\sim}{\ell}}'\underset{\sim}{x} = \left(\bar{\underset{\sim}{x}}_1 - \bar{\underset{\sim}{x}}_2\right)' \underset{\sim}{S}_p^{-1} \underset{\sim}{x}$

$$\hat{m} = \frac{1}{2}\left(\bar{\underset{\sim}{x}}_1 - \bar{\underset{\sim}{x}}_2\right)' S_p^{-1} \left(\bar{\underset{\sim}{x}}_1 + \bar{\underset{\sim}{x}}_2\right)$$

(c) Petua peruntukan:

$$\text{Untukkan } \underset{\sim}{x}_o \text{ kepada } \begin{cases} \pi_1 & \text{jika } y_o \geq \hat{m} \\ \pi_2 & \text{jika } y_o < \hat{m} \end{cases}$$

(d) $\underset{\sim}{B}_o = \sum_{i=1}^{g} \left[\underset{\sim}{\mu}_i - \bar{\underset{\sim}{\mu}}\right] \left[\underset{\sim}{\mu}_i - \bar{\underset{\sim}{\mu}}\right]'$

$\lambda_1, \ldots, \lambda_s$ nilai eigen dan

$\underset{\sim}{e}_1, \ldots, \underset{\sim}{e}_s$ vektor eigen $\underset{\sim}{\Sigma}^{-1}\underset{\sim}{B}_o$.

$\underset{\sim}{\ell}_i \underset{\sim}{X} = \underset{\sim}{e}_i \underset{\sim}{X}$ pembezalayan ke-i, $i = 1, 2, \ldots, s$.

(e) $\hat{\underset{\sim}{B}}_o = \sum_{i=1}^{g} \left[\bar{\underset{\sim}{x}}_i - \bar{\underset{\sim}{x}}\right]\left[\bar{\underset{\sim}{x}}_i - \bar{\underset{\sim}{x}}\right]'$

$\underset{\sim}{W} = \sum_{i=1}^{g} \sum_{j=1}^{n_i} \left[\underset{\sim}{x}_{ij} - \bar{\underset{\sim}{x}}_i\right]\left[\underset{\sim}{x}_{ij} - \bar{\underset{\sim}{x}}_i\right]'$

$\hat{\underset{\sim}{\ell}}_i \underset{\sim}{x} = \hat{\underset{\sim}{e}}_i \underset{\sim}{x}$ pembezalayan sampel ke-i, $i = 1, \ldots, s$.

(f) Petua peruntukan:

Untukkan $\underset{\sim}{x}$ kepada π_k jika

$$\sum_{j=1}^{r} \left(\hat{y}_j - \bar{y}_{kj}\right)^2 = \sum_{j=1}^{r} \left[\hat{\underset{\sim}{\ell}}'_j \left(\underset{\sim}{x} - \bar{\underset{\sim}{x}}_k\right)\right]^2$$

$$\leq \sum_{j=1}^{r} \left[\hat{\underset{\sim}{\ell}}'_j \left(\underset{\sim}{x} - \bar{\underset{\sim}{x}}_i\right)\right]^2 .$$

bagi semua $i \neq k$, $r \leq s$.

\- oooOooo -