UNIVERSITI SAINS MALAYSIA

Peperiksaan Semester Pertama
Sidang Akademik 2003/2004

September / Oktober 2003

MSG 366 – ANALISIS MULTIVARIAT
Masa : 3 jam

Sila pastikan bahawa kertas peperiksaan ini mengandungi **LIMA [5]** muka surat dan **SEMBILAN BELAS [19]** lampiran yang bercetak sebelum anda memulakan peperiksaan ini.

Jawab **semua empat** soalan.

1. (a) Katakan $\mathbf{X}' = (X_1, X_2, X_3, X_4)$ tertabur $N_4(\mu, \Sigma)$ di mana

$$\Sigma = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 4 & 3 & 2 \\ 1 & 3 & 2 & 4 \\ 1 & 2 & 4 & 5 \end{pmatrix}.$$

Cari taburan bagi $4X_1 - 3X_2 + 2X_3 - X_4$.

[20 markah]

(b) Diberikan matriks data

$$\mathbf{X} = \begin{pmatrix} 1 & 4 & 4 \\ 2 & 1 & 0 \\ 5 & 6 & 4 \end{pmatrix}$$

dan gabungan-gabungan linear

$$\mathbf{b}'\mathbf{X} = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} \text{ dan } \mathbf{c}'\mathbf{X} = \begin{pmatrix} 1 & 2 & -3 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix}.$$

Dapatkan min, varians dan kovarians sampel bagi $\mathbf{b}'\mathbf{X}$ dan $\mathbf{c}'\mathbf{X}$.

[20 markah]

(c) Cari anggaran kebolehjadian maksimum bagi vector min μ dan matriks kovarians Σ berdasarkan pada sampel rawak

$$\mathbf{X} = \begin{pmatrix} 5 & 3 & 7 & 9 \\ 3 & 7 & 4 & 6 \end{pmatrix}$$

dari suatu populasi normal bivariat.

[20 markah]

(d) Biarkan matriks

$$A = \begin{pmatrix} 1 & 2 \\ 2 & -2 \end{pmatrix}.$$

(i) Adakah **A** bersimetri?

(ii) Adakah **A** tentu positif?

(iii) Dapatkan pasangan nilai eigen – vector eigen **A**.

(iv) Cari $\mathbf{A}^{-1}$.

[40 markah]

2. (a) (i) Dengan menggunakan data

$$\mathbf{X} = \begin{pmatrix} 14 & 10 & 10 & 11 \\ 2 & 8 & 6 & 8 \end{pmatrix}$$

nilaikan T^2 Hotelling untuk menguji $H_0: \mu' = (7, 12)$.

(ii) Nyatakan taburan T^2 bagi keadaan dalam (i) di atas. Berikan anggapan-anggapan yang telah anda menggunakan.

(iii) Dengan menggunakan (i) dan (ii) di atas, ujikan H_0 pada paras $\alpha = 0.01$. Apakah kesimpulan anda?

[40 markah]

(b) Cerapan-cerapan pada dua balasan, X_1 dan X_2, dikutip bagi tiga rawatan. Vektor-vektor cerapan $\mathbf{x}' = (x_1 \quad x_2)$.

Rawatan 1 : $(2 \quad 6)$ $(3 \quad 6)$ $(4 \quad 7)$
Rawatan 2 : $(3 \quad 4)$ $(3 \quad 5)$
Rawatan 3 : $(4 \quad 4)$ $(5 \quad 4)$ $(5 \quad 5)$

(i) Binakan suatu jadual MANOVA satu hala bagi data ini.

(ii) Nilaikan Lambda Wilks dan ujikan bagi kesan rawatan. Gunakan $\alpha = 0.05$.

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

[60 markah]

...4/-

3. (a) Andaikan **X** tertabur $N_3(\mu, \Sigma)$ dengan $\mu' = \begin{pmatrix} -3 & 1 & 6 \end{pmatrix}$ dan

$$\Sigma = \begin{pmatrix} 1 & -2 & 0 \\ -2 & 5 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Di antara pembolehubah-pembolehubah rawak berikut, yang manakah tak bersandar? Terangkan.

a. X_1 dan X_2
b. X_2 dan X_3
c. (X_1, X_2) dan X_3
d. $\dfrac{X_1 + X_2}{2}$ dan X_3

[20 markah]

(b) Tuliskan nota pendek tentang tajuk-tajuk di bawah:

(i) Analisis komponen prinsipal
(ii) Analisis faktor
(iii) Analisis pembezalayan
(iv) Analisis kelompok

[80 markah]

4. Bagi setiap bahagian yang berikut, huraikan kesimpulan-kesimpulan anda.

(a) Data makanan yang berkhasiat dikutip. Lima jenis khasiat: 'calories', 'protein', 'fat', dan 'iron' dipilih untuk 27 butir makanan seperti berikut:

Braised beef, Hamburger, Roast beef, Beef steak, Canned beef, Broiled chicken, Canned chicken, Beef heart, Roast lamb leg, Roast lamb shoulder, Smoked ham, Roast pork, Beef tongue, Veal cutlet, Baked bluefish, Raw clams, Canned clams, Canned crabmeat, Fried haddock, Broiled mackerel, Canned mackerel, Fried perch, Canned salmon, Canned sardines, Canned tuna dan *Canned shrimp.*

Hasil yang diperoleh dari SAS melalui prosedur FASTCLUS adalah seperti yang diberikan di dalam Lampiran 1.

[30 markah]

(b) Skor-skor ujian pelajar untuk kursus-kursus: *Matematik (M), Fizik (P), Kimia (C), Inggeris (E), Sejarah (H) dan Perancis (F)* tersedia ada. Analisis faktor dijalankan untuk menerangkan prestasi pelajar dalam kursus yang tersebut. Pakej SAS digunakan dan hasil-hasilnya adalah di dalam Lampiran 2.

[30 markah]

...5/-

(c) Analisis pembezalayan dijalankan pada data kewangan untuk 12 buah firma '*most admired*' dan 12 buah firma '*least admired*'. Nisbah-nisbah kewangan adalah EBITASS, *earnings before interest and taxes to total assets*, dan ROTC, *return on total capital*. Pakej SPSS digunakan dan hasil-hasilnya adalah seperti diberikan di dalam Lampiran 3.

[40 markah]

-oooOooo-

FASTCLUS Procedure

Initial Seeds

Cluster	CALORIES	PROTEIN	FAT	CALCIUM	IRON
1	331.111	19.000	27.556	8.778	2.467
2	161.667	20.500	7.500	14.250	1.925
3	100.000	14.800	3.400	114.000	3.000

Minimum Distance Between Seeds = 117.4876

Iteration	Change in Cluster Seeds 1	2	3
1	10.84747	6.464462	44.25093
2	0	0	0

Cluster Summary

Cluster	Frequency	RMS Std Deviation	Maximum Distance from Seed to Observation	Nearest Cluster	Centroid Distance
1	8	20.8936	78.8882	2	175.1
2	13	19.1728	82.2338	3	152.6
3	6	54.0996	221.3	2	152.6

Statistics for Variables

Variable	Total STD	Within STD	R-Squared	RSQ/(1-RSQ)
CALORIES	101.2078	44.933808	0.818048	4.495961
PROTEIN	4.251696	4.003804	0.181424	0.221634
FAT	11.257033	4.736894	0.836553	5.118185
CALCIUM	78.034254	50.445526	0.614244	1.592313
IRON	1.461305	1.439748	0.103956	0.116017
OVER-ALL	57.409577	30.345806	0.742091	2.877336

Pseudo F Statistic = 34.53
Approximate Expected Over-All R-Squared = 0.71972
Cubic Clustering Criterion = 0.640
WARNING: The two above values are invalid for correlated variables.

Cluster Means

Cluster	CALORIES	PROTEIN	FAT	CALCIUM	IRON
1	341.875	18.750	28.875	8.750	2.437
2	168.077	20.538	8.231	13.846	1.985
3	113.333	16.000	4.333	156.167	3.167

Cluster Standard Deviations

Cluster	CALORIES	PROTEIN	FAT	CALCIUM	IRON
1	46.363	1.669	5.463	0.707	0.207
2	41.460	3.799	4.585	9.017	1.520
3	50.563	6.197	3.933	109.631	2.085

CLUSTER=1

OBS	NAME	CLUSTER	DISTANCE	CALORIES	PROTEIN	FAT	CALCIUM	IRON
1	Braised beef	1	2.4357	340	20	28	9	2.6
2	Roast beef	1	78.8882	420	15	39	7	2.0
3	Beef steak	1	33.2744	375	19	32	9	2.6
4	Roast lamb leg	1	77.3963	265	20	20	9	2.6
5	Roast lamb shoulder	1	42.0616	300	18	25	9	2.3
6	Smoked ham	1	2.4311	340	20	28	9	2.5
7	Roast pork	1	1.9132	340	19	29	9	2.5
8	Simmered pork	1	13.1779	355	19	30	9	2.4

CLUSTER=2

OBS	NAME	CLUSTER	DISTANCE	CALORIES	PROTEIN	FAT	CALCIUM	IRON
9	Hamburger	2	77.5775	245	21	17	9	2.7
10	Canned beef	2	12.6616	180	22	10	17	3.7
11	Broiled chicken	2	53.6594	115	20	3	8	1.4
12	Canned chicken	2	5.3630	170	25	7	12	1.5
13	Beef heart	2	10.9935	160	26	5	14	5.9
14	Beef tongue	2	38.0812	205	18	14	7	2.5
15	Veal cutlet	2	17.8056	185	23	9	9	2.7
16	Baked bluefish	2	35.2199	135	22	4	25	0.6
17	Canned crabmeat	2	82.2338	90	14	2	38	0.8
18	Fried haddock	2	33.5954	135	16	5	15	0.5
19	Broiled mackerel	2	33.5174	200	19	13	5	1.0
20	Fried perch	2	27.4520	195	16	11	14	1.3
21	Canned tuna	2	8.5208	170	25	7	7	1.2

CLUSTER=3

OBS	NAME	CLUSTER	DISTANCE	CALORIES	PROTEIN	FAT	CALCIUM	IRON
22	Raw clams	3	86.155	70	11	1	82	6.0
23	Canned clams	3	107.322	45	7	1	74	5.4
24	Canned mackerel	3	41.958	155	16	9	157	1.8
25	Canned salmon	3	7.746	120	17	5	159	0.7
26	Canned sardines	3	221.254	180	22	9	367	2.5
27	Canned shrimp	3	58.778	110	23	1	98	2.6

Correlations

	M	P	C	E	H	F
M	1.00000	0.62000	0.54000	0.32000	0.28400	0.37000
P	0.62000	1.00000	0.51000	0.38000	0.35100	0.43000
C	0.54000	0.51000	1.00000	0.36000	0.33600	0.40500
E	0.32000	0.38000	0.36000	1.00000	0.68600	0.73000
H	0.28400	0.35100	0.33600	0.68600	1.00000	0.73500
F	0.37000	0.43000	0.40500	0.73000	0.73500	1.00000

Initial Factor Method: Iterated Principal Factor Analysis

Partial Correlations Controlling all other Variables

	M	P	C	E	H	F
M	1.00000	0.44624	0.30877	0.01370	-0.03204	0.06098
P	0.44624	1.00000	0.20254	0.05115	0.02581	0.09909
C	0.30877	0.20254	1.00000	0.04790	0.03147	0.08633
E	0.01370	0.05115	0.04790	1.00000	0.31717	0.41595
H	-0.03204	0.02581	0.03147	0.31717	1.00000	0.45155
F	0.06098	0.09909	0.08633	0.41595	0.45155	1.00000

Kaiser's Measure of Sampling Adequacy: Over-all MSA = 0.81296561

M	P	C	E	H	F
0.768868	0.812442	0.866921	0.831964	0.812097	0.796664

Prior Communality Estimates: ONE

Preliminary Eigenvalues: Total = 6 Average = 1

	1	2	3	4	5	6
Eigenvalue	3.367086	1.194179	0.507006	0.371847	0.313156	0.246726
Difference	2.172907	0.687173	0.135159	0.058691	0.066430	
Proportion	0.5612	0.1990	0.0845	0.0620	0.0522	0.0411
Cumulative	0.5612	0.7602	0.8447	0.9067	0.9589	1.0000

2 factors will be retained by the MINEIGEN criterion.

Scree Plot of Eigenvalues

```
E    |
i  4 +
g    |
e    |          1
n    |
v    |
a  2 +
l    |
u    |                    2
e    |                              3         4         5         6
s    |
   0 +
     -------+---------+---------+---------+---------+---------+---------+------
            0         1         2         3         4         5         6
                                        Number
```

Iter	Change	Communalities					
1	0.359383	0.76582	0.71564	0.64062	0.79670	0.81162	0.83086
2	0.127702	0.69838	0.62622	0.51292	0.72426	0.74467	0.78387
3	0.042180	0.67946	0.59762	0.47074	0.69784	0.71918	0.77400
4	0.013512	0.67487	0.58806	0.45722	0.68774	0.70846	0.77438
5	0.005134	0.67443	0.58455	0.45287	0.68358	0.70333	0.77691
6	0.002800	0.67509	0.58304	0.45140	0.68171	0.70053	0.77935
7	0.001878	0.67593	0.58224	0.45084	0.68078	0.69884	0.78122
8	0.001343	0.67670	0.58173	0.45059	0.68028	0.69775	0.78257
9	0.000931	0.67734	0.58136	0.45044	0.67999	0.69702	0.78350

Convergence criterion satisfied.

Eigenvalues of the Reduced Correlation Matrix:
Total = 3.86960282 Average = 0.6449338

	1	2	3	4	5	6
Eigenvalue	3.028416	0.841238	0.001570	0.001117	-0.001250	-0.001488
Difference	2.187178	0.839668	0.000453	0.002367	0.000239	
Proportion	0.7826	0.2174	0.0004	0.0003	-0.0003	-0.0004
Cumulative	0.7826	1.0000	1.0004	1.0007	1.0004	1.0000

Factor Pattern

	FACTOR1	FACTOR2
M	0.63577	0.52263
P	0.65778	0.38559
C	0.59807	0.30457
E	0.76220	-0.31471
H	0.74932	-0.36815
F	0.83151	-0.30345

Lampiran 2

Variance explained by each factor

FACTOR1	FACTOR2
3.028416	0.841238

Final Communality Estimates: Total = 3.869654

M	P	C	E	H	F
0.677344	0.581359	0.450445	0.679993	0.697015	0.783498

Residual Correlations With Uniqueness on the Diagonal

	M	P	C	E	H	F
M	0.32266	0.00028	0.00059	-0.00011	0.00001	-0.00006
P	0.00028	0.41864	-0.00084	-0.00001	0.00006	0.00005
C	0.00059	-0.00084	0.54956	0.00001	-0.00002	0.00012
E	-0.00011	-0.00001	0.00001	0.32001	-0.00099	0.00072
H	0.00001	0.00006	-0.00002	-0.00099	0.30298	0.00021
F	-0.00006	0.00005	0.00012	0.00072	0.00021	0.21650

Root Mean Square Off-diagonal Residuals: Over-all = 0.00042558

M	P	C	E	H	F
0.000298	0.000396	0.000461	0.000551	0.000456	0.000342

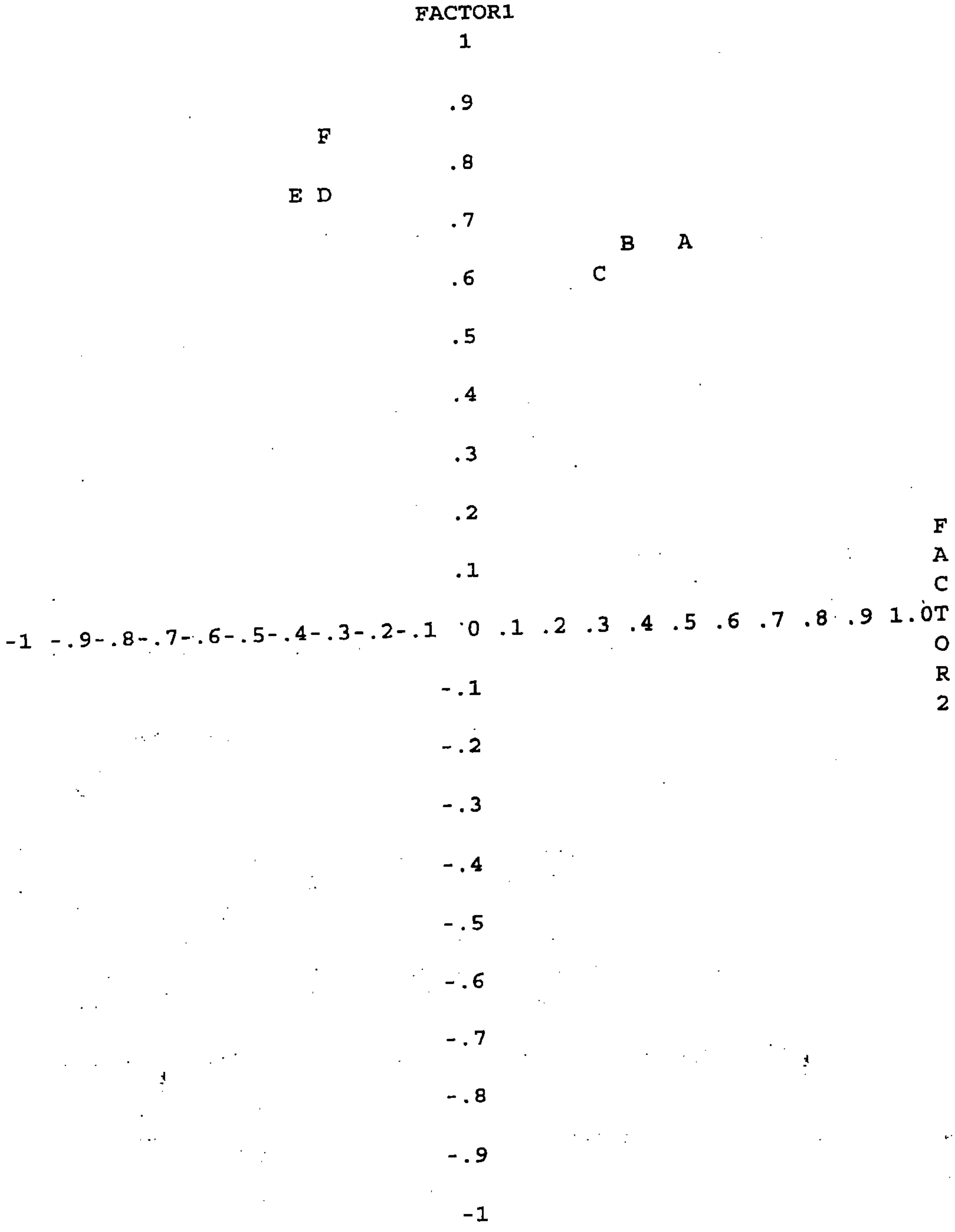
Partial Correlations Controlling Factors

	M	P	C	E	H	F
M	1.00000	0.00076	0.00140	-0.00033	0.00002	-0.00023
P	0.00076	1.00000	-0.00174	-0.00003	0.00017	0.00017
C	0.00140	-0.00174	1.00000	0.00001	-0.00005	0.00034
E	-0.00033	-0.00003	0.00001	1.00000	-0.00319	0.00274
H	0.00002	0.00017	-0.00005	-0.00319	1.00000	0.00084
F	-0.00023	0.00017	0.00034	0.00274	0.00084	1.00000

Root Mean Square Off-diagonal Partials: Over-all = 0.00127326

M	P	C	E	H	F
0.000737	0.000858	0.001013	0.001886	0.001478	0.001296

Plot of Factor Pattern for FACTOR1 and FACTOR2



M =A P =B C =C E =D H =E
F =F

Rotation Method: Varimax

Orthogonal Transformation Matrix

	1	2
1	0.76675	0.64194
2	-0.64194	0.76675

Rotated Factor Pattern

	FACTOR1	FACTOR2
M	0.15198	0.80885
P	0.25683	0.71791
C	0.26305	0.61745
E	0.78645	0.24798
H	0.81087	0.19875
F	0.83236	0.30112

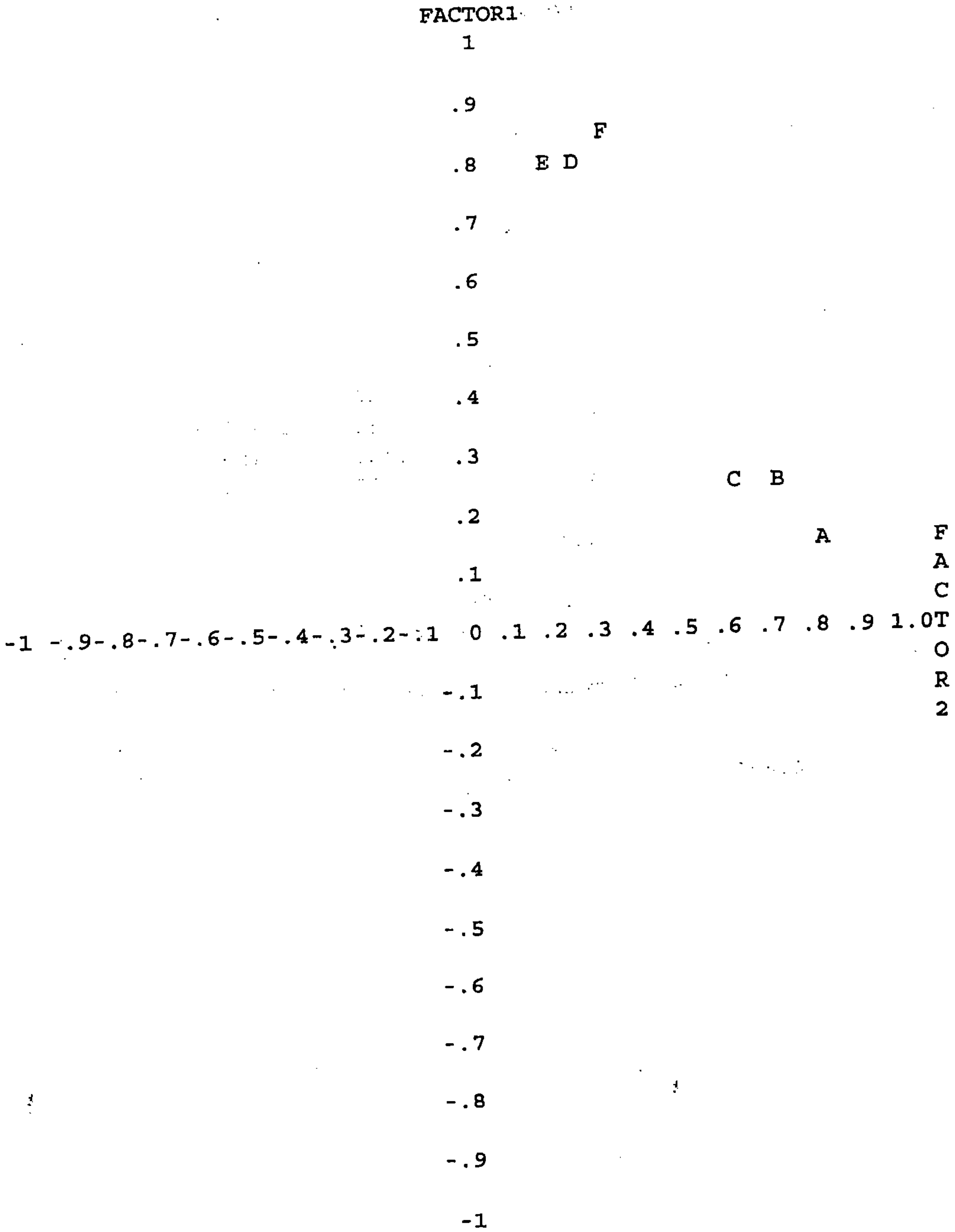
Variance explained by each factor

FACTOR1	FACTOR2
2.127097	1.742556

Final Communality Estimates: Total = 3.869654

M	P	C	E	H	F
0.677344	0.581359	0.450445	0.679993	0.697015	0.783498

Plot of Factor Pattern for FACTOR1 and FACTOR2



M =A P =B C =C E =D H =E
F =F

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

On groups defined by EXCELL   Excell

Number of cases by group

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

Group means

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

Group standard deviations

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

Pooled within-groups covariance matrix with 22 degrees of freedom

	EBITASS	ROTC
EBITASS	2.4261515E-03	
ROTC	2.0336818E-03	2.8041477E-03

Pooled within-groups correlation matrix

	EBITASS	ROTC
EBITASS	1.00000	
ROTC	.77969	1.00000

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

Covariance matrix for group 1, Most-Admired

	EBITASS	ROTC
EBITASS	.0028	
ROTC	.0011	.0009

Covariance matrix for group 2, Least-Admired

	EBITASS	ROTC
EBITASS	.0020	
ROTC	.0030	.0047

Total covariance matrix with 23 degrees of freedom

	EBITASS	ROTC
EBITASS	.0115	
ROTC	.0109	.0113

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)

EXCELL =	1 Most-Admired	2 Least-Admired
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

Unstandardized canonical discriminant function coefficients

	Func 1
EBITASS	15.0919163
ROTC	5.7685027
(Constant)	-2.0018120

Canonical discriminant functions evaluated at group means (group centroids)

Group	Func 1
1	1.94429
2	-1.94429

Test of Equality of Group Covariance Matrices Using Box's M

The ranks and natural logarithms of determinants printed are those of the group covariance matrices.

Group Label	Rank	Log Determinant
1 Most-Admired	2	-13.516047
2 Least-Admired	2	-14.107651
Pooled within-groups covariance matrix	2	-12.834397

Box's M	Approximate F	Degrees of freedom		Significance
21.50395	6.46365	3,	87120.0	.0002

Symbols used in plots

Symbol	Group	Label
1	1	Most-Admired
2	2	Least-Admired

All-groups Stacked Histogram

Canonical Discriminant Function 1

```
       4
  F
  r    3
  e
  q
  u                                  2      2       1  1 1
  e    2                             2      2       1  1 1
  n                                  2      2       1  1 1
  c                                  2      2       1  1 1
  y                  22 2 22 222 22  1     1 1 1111  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

          XX
         out     -4.0      -2.0      .0      2.0      4.0       out
    Class
  Centroids                  2                 1
```

Classification results -

Actual Group	No. of Cases	Predicted Group Membership 1	2
Group 1 Most-Admired	12	12 100.0%	0 .0%
Group 2 Least-Admired	12	1 8.3%	11 91.7%

Percent of "grouped" cases correctly classified: 95.83%

MSG 366 - ANALISIS MULTIVARIAT

LAMPIRAN

Tatatanda adalah seperti di dalam kuliah.

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

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

di mana $\lambda_1, \lambda_2, \dots, \lambda_k$ adalah nilai-nilai eigen $\underset{\sim}{A}$ dan $\underset{\sim}{e}_1, \underset{\sim}{e}_2, \dots, \underset{\sim}{e}_k$ 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) $T^2 = n\,(\underset{\sim}{\bar{X}} - \underset{\sim}{\mu})'\ \underset{\sim}{S}^{-1}(\underset{\sim}{\bar{X}} - \underset{\sim}{\mu})$

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

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

(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 $\underset{\sim}{\ell}'\underset{\sim}{\mu}$:

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

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

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

(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. MANOVA satu-hala:

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

$$\underset{\sim}{W} = \sum_{\ell=1}^{g} \sum_{j=1}^{n_\ell} \left[\underset{\sim}{x}_{\ell j} - \underset{\sim}{\bar{x}}_\ell\right]\left[\underset{\sim}{x}_{\ell j} - \underset{\sim}{\bar{x}}_\ell\right]'$$

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

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

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

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

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

Katakan $\Lambda = \dfrac{|\underset{\sim}{E}|}{|\underset{\sim}{E} + \underset{\sim}{H}|}$.

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:

$$SSP_{faktor\ 1} = \sum_{\ell=1}^{g} bn\left(\bar{x}_{\ell.} - \bar{x}\right)\left(\bar{x}_{\ell.} - \bar{x}\right)'$$

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

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

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

11. Komponen Prinsipal

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

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

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

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

12. Analisis Faktor

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

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

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

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

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

(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:

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