



KENYATTA UNIVERSITY
UNIVERSITY EXAMINATIONS 2010/2011
INSTITUTIONAL BASED PROGRAMME (DECEMBER SESSION)
EXAMINATION FOR THE DEGREE OF MASTER OF SCIENCE

SMA/SS/565 : MULTIVARIATE DATA ANALYSIS

DATE: THURSDAY 28TH APRIL 2011

TIME: 9.00 A.M. - 12.00 NOON

INSTRUCTIONS: Answer ANY THREE questions. Each Questions is 23 marks

Question one

- a) Let $\mathbf{X}$ be a random vector of size p , $p \geq 2$. Suppose $\mathbf{X}$ has a p.d.f given by $f(\mathbf{x})$. Take a random sample of size n from this population.
- Define the following terms:
 - Data Matrix
 - Mean Vector
 - Sample Variance Matrix
- b) Show that the mean vector is an unbiased estimator of the population mean and that the sample variance is a biased estimator of the population variance.
- c) Given the following data calculate the sample mean and ample variance matrix

X_1	8	8	4	7	4
X_2	1	9	3	7	1
X_3	6	6	1	3	7
X_4	1	4	6	6	8

Question Two

- Define the Mahalanobis distance between two vectors.
- Using the given data, determine which canine is closest to the modern dog using Mahalanobis distance.

Question Three

- Discuss the problems encountered in displaying many variables in two dimension.
- Briefly discuss the draftsman's plot and the star plot.
- Using the given data, draw a draftsman's plot.

Question Four

- Briefly discuss the univariate and multivariate tests bringing out their advantages and disadvantages.
- Using the data given and an appropriate test, conclude whether the Pgi frequencies are similar for colonies that are close.

Question Five

- Clearly explain the problem of principal component analysis.
- Given the following data obtain the first principal component and show how much it explains of the total variation.

Q2

Chapter one: The material of multivariate analysis

9

Table 1.4 Mean Mandible Measurements for Seven Canine Groups

Group	X_1 (mm)	X_2 (mm)	X_3 (mm)	X_4 (mm)	X_5 (mm)
Modern dog	9.7	21.0	19.4	7.7	32.0
Golden jackal	8.1	16.7	18.3	7.0	30.3
Chinese wolf	13.5	27.3	26.8	10.6	41.9
Indian wolf	11.5	24.3	24.5	9.3	40.0
Cuon	10.7	23.5	21.4	8.5	28.8
Dingo	9.6	22.6	21.1	8.3	34.4
Prehistoric dog	10.3	22.1	19.1	8.1	32.2

MBU/

Q4

Table 1.3 Environmental Variables and Phosphoglucose Isomerase (Pgi) Gene Frequencies for Colonies of the Butterfly *Euphydryas editha* in California and Oregon^a

Colony	Altitude (ft)	Annual Precipitation (in.)	Temperature (°F)		Frequencies of Pgi Mobility Genes (%) ^b						
			Maximum	Minimum	0.4	0.6	0.8	1	1.16	1.3	
SS	500	43	98	17	0	3	22	57	17	1	
SB	808	20	92	32	0	16	20	38	13	13	
WSB	570	28	98	26	0	6	28	46	17	3	
JRC	550	28	98	26	0	4	19	47	27	3	
JRH	550	28	98	26	0	1	8	50	35	6	
SJ	380	15	98	28	0	2	19	44	32	3	
CR	930	21	99	28	0	0	15	50	27	8	
UO	650	10	101	27	10	21	40	25	4	0	
LO	600	10	101	27	14	26	32	28	0	0	
DP	1,500	19	99	23	0	1	6	80	12	1	
PZ	1,750	22	101	27	1	4	34	33	22	6	
MC	2,000	58	100	18	0	7	14	66	13	0	
IF	2,500	34	102	16	0	9	15	47	21	8	
AF	2,000	21	105	20	3	7	17	32	27	14	
GH	7,850	42	84	5	0	5	7	84	4	0	
CL	10,500	50	81	-12	0	3	1	92	4	0	

^a The data source was McKethnie et al. (1975), with the environmental variables rounded to integers for simplicity. The original data were for 21 colonies, but for the present example, five colonies with small samples for the estimation of gene frequencies have been excluded to make all estimates about equally reliable.

^b The numbers 0.40, 0.60, etc. represent different genetic types of Pgi so that the frequencies for a colony (adding to 100%) show the frequencies of the different types for the *E. editha* at that location.

Source: Adapted from McKethnie, S.W. et al. (1975), *Genetics*, 81: 571-594.

Table 1.1 Body Measurements of Female Sparrows

Bird	X_1 (mm)	X_2 (mm)	X_3 (mm)	X_4 (mm)	X_5 (mm)
1	156	245	316	185	205
2	154	240	304	179	196
3	153	240	310	184	206
4	153	236	309	177	202
5	155	243	315	186	203
6	163	247	320	190	209
7	157	238	309	184	202
8	155	239	328	186	212
9	164	248	327	191	211
10	158	238	310	188	220
11	158	240	313	186	220
12	160	244	311	186	205
13	161	246	323	193	218
14	157	245	320	191	200
15	157	235	315	181	198
16	156	237	309	180	203
17	158	244	314	185	216
18	153	238	305	182	209
19	155	236	303	185	201
20	163	246	325	186	219
21	159	236	315	180	215
22	155	240	314	180	207
23	160	242	326	188	217
24	156	240	315	182	206
25	152	232	303	172	198
26	160	250	317	188	225
27	155	237	310	185	200
28	157	245	322	195	214
29	165	245	331	198	227
30	153	231	301	173	198
31	162	239	303	180	231
32	162	243	316	188	213
33	159	245	318	185	217
34	159	247	309	181	190
35	155	243	309	185	213
36	162	252	319	191	222
37	152	230	304	173	186
38	159	242	308	182	205
39	155	238	312	179	193
40	163	249	334	195	228
41	163	242	310	181	207
42	156	237	317	182	203
43	159	238	313	184	203
44	161	245	321	191	208
45	153	235	307	177	196

(continued)

Table 1.1 (continued) Body Measurements of Female Sparrows

Bird	X_1 (mm)	X_2 (mm)	X_3 (mm)	X_4 (mm)	X_5 (mm)
46	162	247	319	191	204
47	153	237	306	186	204
48	162	245	325	185	211
49	164	248	323	188	209