UNIVERSITY OF CAMBRIDGE

M. PHIL. IN STATISTICAL SCIENCE

Wednesday 2 June, 2004 9:00 to 11:00

Applied Multivariate Analysis

Attempt **THREE** *questions.*

There are **five** *questions in total.*

The questions carry equal weight.

You may not start to read the questions printed on the subsequent pages until instructed to do so by the Invigilator.

1 Suppose that the p-dimensional vector X is distributed as $N_p(\mu, V)$.

(i) Show that if we partition X into components X_1, X_2, so that $X^T = (X_1^T, X_2^T)$, then the covariance matrix of X_1 conditional on $X_2 = x_2$ is $V_{11} - V_{12}V_{22}^{-1}V_{21}$, where

$$V = \begin{pmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{pmatrix}.$$

(ii) Hence or otherwise find an expression for the variance of $(X_1|X_2 = x_2)$ in terms of V^{-1}, when X_1, X_2 are of dimensions $1, p-1$ respectively.

(iii) If now $p = 3$, and $X^T = (X_1, X_2, X_3)$, derive an expression for the correlation of X_1, X_2 conditional on $X_3 = x_3$, in terms of (ρ_{ij}), where $\rho_{ij} = \operatorname{corr}(X_i, X_j)$ for $1 \le i < j \le 3$.

2 (i) Let $x_1, \ldots, x_n$ be a random sample from the distribution $N_p(\mu, V)$. Find an expression for the loglikelihood function $\ell(\mu, V)$ in terms of the standard statistics $\bar{x}, S$, and state without proof the form of the maximum likelihood estimators $\hat{\mu}, \hat{V}$.

(ii) Now suppose that we have independent observations from g distinct groups, with

$$x_1^{(\nu)}, \ldots, x_{n_\nu}^{(\nu)}$$

being the sample from the νth group, which is assumed to be a random sample from $N(\mu^{(\nu)}, V)$, for $1 \le \nu \le g$. Using the results of (i) above, show that the generalized likelihood ratio test of

$$H_0 : \mu^{(1)} = \ldots = \mu^{(g)} = \mu \text{ say}$$

with μ, V both unknown, is of the form:

reject H_0 if

$$log \frac{|W + B|}{|W|} > \text{ constant},$$

where W, B are matrices that you should define.

(iii) Describe briefly the use of the matrices W, B in discriminant analysis.

3 Interpret the commands and the corresponding output, giving appropriate sketch graphs. (Formal proofs are not required.)

```
> a _ read.table("students", header=T)
>a
          meat coffee beer UKres Cantab Fem sports driver Left.h
    Taeko    1      1    0     1      1   1      0      0      0
 Luitgard    0      1    0     0      1   1      1      1      0
     Alet    1      1    1     0      0   1      0      1      0
      Tom    1      1    1     1      1   0      1      1      0
   LinYee    1      1    0     0      0   0      1      1      0
      Pio    1      1    0     0      0   0      1      0      0
 LingChen    1      0    0     0      0   1      1      0      0
  HuiChin    1      1    0     0      0   1      1      1      0
   Martin    1      1    1     1      0   0      1      1      0
  Nicolas    1      1    1     0      0   0      1      1      1
 Mohammad    1      1    0     0      0   0      0      1      0
      Meg    1      1    0     0      0   1      1      0      0

> d _ dist(a, metric="binary") ; round(dist2full(d),2)
      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10] [,11] [,12]
 [1,] 0.00 0.57 0.57 0.50 0.71 0.67 0.67 0.57 0.62  0.78  0.67  0.50
 [2,] 0.57 0.00 0.57 0.50 0.50 0.67 0.67 0.33 0.62  0.62  0.67  0.50
 [3,] 0.57 0.57 0.00 0.50 0.50 0.67 0.67 0.33 0.43  0.43  0.40  0.50
 [4,] 0.50 0.50 0.50 0.00 0.43 0.57 0.75 0.50 0.14  0.38  0.57  0.62
 [5,] 0.71 0.50 0.50 0.43 0.00 0.25 0.60 0.20 0.33  0.33  0.25  0.40
 [6,] 0.67 0.67 0.67 0.57 0.25 0.00 0.50 0.40 0.50  0.50  0.50  0.25
 [7,] 0.67 0.67 0.67 0.75 0.60 0.50 0.00 0.40 0.71  0.71  0.80  0.25
 [8,] 0.57 0.33 0.33 0.50 0.20 0.40 0.40 0.00 0.43  0.43  0.40  0.20
 [9,] 0.62 0.62 0.43 0.14 0.33 0.50 0.71 0.43 0.00  0.29  0.50  0.57
[10,] 0.78 0.62 0.43 0.38 0.33 0.50 0.71 0.43 0.29  0.00  0.50  0.57
[11,] 0.67 0.67 0.40 0.57 0.25 0.50 0.80 0.40 0.50  0.50  0.00  0.60
[12,] 0.50 0.50 0.50 0.62 0.40 0.25 0.25 0.20 0.57  0.57  0.60  0.00

> h _  hclust(d, method="compact")
```

4 Write brief essays, which should include appropriate sketch graphs, on **two** of the following three S-Plus functions,

`princomp()`

`tree()`

`cmdscale()`

The second function may be replaced by

`rpart()`

if you prefer.

5 Suppose we have two known classes, C_1 and C_2, and our observation x is known to have arisen from one of C_1 or C_2, with prior probabilities π_1, π_2 respectively. The corresponding known probability densities are those of $N(\mu_1, V)$, $N(\mu_2, V)$ respectively.

(i) Show that the Bayes rule for assigning x to C_1 or C_2 is of the form:

assign x to C_1 if $a^T x > b$

where you should determine a, b.

(ii) In the case $\pi_1 = \pi_2 = 1/2$, show that the above rule has error probabilities

$$P(\text{assign } x \text{ to } C_1 | x \text{ is from } C_2) = P(\text{assign } x \text{ to } C_2 | x \text{ is from } C_1) = p$$

say, where $p = p(\delta)$, $\delta > 0$ and

$$\delta^2 = (\mu_1 - \mu_2)^T V^{-1} (\mu_1 - \mu_2).$$