STAT 501   **EXAM I**   NAME _____
Spring 1999

Instructions: You may use only your calculator and the attached tables and formula sheet. You can detach the tables and formula sheet from the rest of this exam. Show your formulas and answers in the space provided on this exam. Credit will be given for exhibiting a correct method or correct formulas, even when the final numerical answer is missing or incorrect. If you need more room use the back of the page or attach extra sheets of paper, but clearly indicate where this is done. Put your name on any pages you attach to this exam.

1.   Suppose $\mathbf{X} = (X_1, X_2, X_3, X_4)'$ has a multivariate normal distribution with mean vector and covariance matrix

$$\mathbf{m} = \begin{bmatrix} 1 \\ -3 \\ -2 \\ 4 \end{bmatrix} \qquad \mathbf{S} = \begin{bmatrix} 28 & 4 & 24 & -14 \\ 4 & 22 & 12 & -7 \\ 24 & 12 & 78 & -35 \\ -14 & -7 & -35 & 97 \end{bmatrix}$$

The eigenvalues for Σ are $\lambda_1 = 132$, $\lambda_2 = 55$, $\lambda_3 = 20$ and λ_4 .

(a)   What is the distribution of $\begin{bmatrix} X_1 - 2X_4 \\ X_1 + X_2 + 3X_3 \end{bmatrix}$?

(b)   What is the value of the smallest eigenvalue λ_4 ?

(c)   Is $\mathbf{a} = \left(\dfrac{-1}{\sqrt{5}}, \dfrac{2}{\sqrt{5}}, 0, 0 \right)'$ an eigenvector for Σ ? Explain.

(d)   What is the conditional distribution of X_1 and X_3 given that

$$\begin{bmatrix} \mathbf{X}_2 \\ \mathbf{X}_4 \end{bmatrix} = \begin{bmatrix} -2 \\ 4 \end{bmatrix} ?$$

(e)   What is the partial correlation between X_1 and X_3 given $X_2 = -1$ and $X_4 = 3$?

(f)   List properties of $|\Sigma|$ that make it a reasonable measure of dispersion for

$X = (X_1, X_2, X_3, X_4)'$.

2.   A manufacturer of electrical components buys electrodes from two suppliers named supplier 1 and supplier 2. In a quality control study of these suppliers, $n_1 = 20$ electrodes were randomly selected from those purchased from supplier 1 and $n_2 = 80$ electrodes were randomly selected from those selected from supplier 2. Three measurements were made on each electrode:
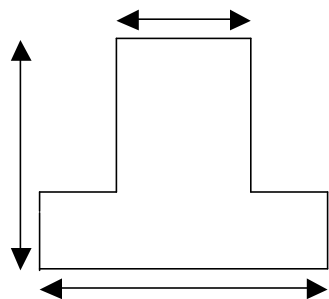
These measurements, shown in the diagram, are all made in units of 10^{-3} millimeters. The sample covariance matrices for the samples from the two suppliers are:

$$S_1 = \begin{bmatrix} 2.48 & 0.72 & 3.27 \\ 0.72 & 2.42 & 3.70 \\ 3.27 & 3.70 & 20.00 \end{bmatrix} \qquad S_2 = \begin{bmatrix} 0.75 & 0.29 & 0.58 \\ 0.29 & 1.22 & 0.82 \\ 0.58 & 0.82 & 15.37 \end{bmatrix}$$

$|S_1| = 67.2585$ $\qquad\qquad\qquad\qquad$ $|S_2| = 12.1321$

trace $(S_1) = 24.90$ $\qquad\qquad\qquad$ trace $(S_2) = 17.34$

You may assume that checks of normality showed that the joint distribution of (X_1, X_2, X_3) is essentially tri-variate normal for each supplier.

(a)   Test the null hypothesis that the covariance matrices for (X_1, X_2, X_3) are the same for the two suppliers. Report values for the test statistic and its degrees of freedom. Clearly state your conclusion.

(b)   Let $(\mu_{11}, \mu_{21}, \mu_{31})$ and $(\mu_{12}, \mu_{22}, \mu_{32})$ denote the vectors of true means for supplier 1 and supplier 2 respectively. Show how you would test H_0: $(\mu_{11}, \mu_{21}, \mu_{31}) = (\mu_{12}, \mu_{22}, \mu_{32})$ against H_A: $(\mu_{11}, \mu_{21}, \mu_{31}) \neq (\mu_{12}, \mu_{22}, \mu_{32})$. Give a formula for your test statistic, the associated degrees of freedom, and the .05 critical value for your test statistic.

3.   The following study was performed to compare how two different raters judged the English proficiency of students enrolled at ISU who are non-native English speakers. Each student in a sample of 30 non-native English speakers read the same passage to the two raters. Each rater gave scores on syllable structure (X_1), segmental errors (X_2), and super segmental errors (X_3) for each of the 30 students (these are measurements of the quality of different aspects of speech). Consequently, 30 sets of values for (X_1, X_2, X_3) were provided by rater 1 and 30 sets of values for (X_1, X_2, X_3) were provided by rater 2.

(a)   Give a formula for the T^2 test you would use to test the null hypothesis that the factor of mean ratings $\blacksquare$ = (μ_{11}, μ_{21}, μ_{31})$'$ for rater 1 is the same as the vector of mean ratings μ_2 = (μ_{12}, μ_{22}, μ_{32})$'$ for rater 2. Give the degrees of freedom and $\alpha = .05$ critical value for the test.

(b)   Which of the following assumptions were used in the development of your test in part (a)? (circle each correct answer).

(i)   Any X_1 score given by rater 1 is independent of any X_1 score given by rater 2.

(ii)   Scores given by rater 1 have a 3-dimensional normal distribution.

(iii)   The "true" covariance matrix for the 3 scores given by rater 1 is the same at the corresponding "true" covariance matrix for the 3 scores given by rater 2.

(iv)   The correlation between the X_1 and X_2 scores given by rater 1 is the same as the correlation between the X_1 and X_3 scores given by rater 1.

(c)   Suppose the researcher wanted to construct simultaneous 95% confidence intervals for the differences in rater means for each of the three aspects of speech, i.e., $\mu_{11}-\mu_{12}$, $\mu_{21}-\mu_{22}$, and $\mu_{31}-\mu_{32}$. Give a formula for using the Bonferroni method to do this.

4. Thirty-two adult males were randomly divided into four groups with eight men in each group. Each group was given a different diet. These diets are labeled as diet 1, diet 2, diet 3, and diet 4. Each diet was low in riboflavin, but otherwise adequate. The riboflavin concentration in each subject's blood was measured at three time points during the study. Measurements were made at 10 days, 20 days, and 30 days after the subject began using the prescribed diet. Measurements are recorded as

$$\text{(mg. of riboflavin)}/\text{(100 ml. of blood)} - 10 \ .$$

Vectors of sample means are as follows:

	Diet 1	Diet 2	Diet 3	Diet 4
10 days	6.2	6.4	10.3	13.4
20 days	5.0	4.9	9.6	12.3
30 days	3.8	4.0	8.2	11.4

(a) Test the null hypothesis that vectors containing the mean responses at 10, 20, or 30 days are the same for all four diets. For this null hypothesis, the H and E matrices produced by the MANOVA option of PROC GLM in SAS are

$$H = \begin{bmatrix} 285.0 & 298.8 & 302.8 \\ 298.8 & 317.2 & 317.7 \\ 301.8 & 317.7 & 319.6 \end{bmatrix} \quad \text{and} \ E = \begin{bmatrix} 122.1 & 71.2 & 56.1 \\ 71.2 & 135.1 & 61.0 \\ 56.1 & 61.0 & 98.7 \end{bmatrix}$$

and the value of $|E| / |H+E|$ is 0.197. Can the null hypothesis of no diet effects, be rejected at the $\alpha = .05$ level? Show how you reached your conclusion.

(b) Specify conditions under which the test in part (a) is appropriate. (Assume these conditions are true for the rest of this problem.)

(c)   The value of Wilk's criterion for testing that the riboflavin blood concentration profiles are parallel for all 4 diets is 0.961. Convert this into the value for an F-test and report the degrees of freedom.

(d)   Test the null hypothesis that the mean response at 30 days is the same for all 4 diets. State your conclusion.

(e)   Consider testing the null hypothesis that the mean response has a straight line relationship with time for each of the four diets. What are the degrees of freedom for the F-test derived from the Wilks criterion?

(f)   Consider a likelihood ratio test of the null hypothesis that the correlation matrices are the same for all four diets. What are the degrees of freedom for the large sample chi-square approximation for the distribution of $-2 \log$(likelihood ratio)?

(g)   An analysis of variance table of the form

Source of Variation	d.f.
Diets	3
Men within Diets	28
Time	2
Diet $\times$ Time Interaction	6
Residuals	56
Corrected total	95

was suggested by one researcher. When, if ever, is

$F = MS_{\text{Diet} \times \text{Time Interaction}} / MS_{\text{Residuals}}$, on $(6, 56)$ degrees of freedom,

an appropriate test for parallel profiles?

(h)   What are conservative degrees of freedom for the F-test for interaction in part (g)?

(i)   What differences, if any, are there between the test for no diet effects in part (a) and the F-test for no diet effects provided by the ANOVA table in part (g)?

EXAM SCORE _____

X_2

X_3

$X_1 = $ diameter at the base

$X_2 = $ diameter at the top

$X_3 = $ length from the base to the top

X_1