Stat 501 **Solutions and Comments on Exam 1**
Spring 2005

1. (a) (5 points) $\underset{\sim}{Y} \sim N\left(\begin{bmatrix} -4 \\ -1 \end{bmatrix}, \begin{bmatrix} 20 & -4 \\ -4 & 34 \end{bmatrix}\right)$

   (b) (5 points) $X_3 \,|\, (X_1, X_2) = (5,8) \sim N(11.25, 20.9375)$

   (c) (10 points, 2 for each part) (i) , (ii), and (v) are true. (iii) is false because the correlations are non-zero. (iv) is false, the volume is proportional to $|\Sigma|^{1/2}$

2. (4 points) The Anderson-Darling test is a goodness-of-fit test. We used it to test the hypothesis that a sample was drawn from a population with a univariate normal distribution. It is a weighted average of the squared deviations between the sample cdf that assigns probability 1/n to each observation in the observed sample and the hypothesized cdf evaluated at estimates of the unknown parameters (we used unbiased estimators for the unknown mean and variance.) The weights give more emphasis to deviations in the tail of the distribution. The sample cdf is a consistent estimator of the true cdf, but the hypothesized cdf evaluate at estimates of unknown parameters will not be a consistent estimator unless the hypothesized cdf coincides with the truth.

3. (a) (8 points) There is more than one way to correctly express the null hypothesis in matrix form. One way to state the null hypothesis is

$$H_0: \; C\underset{\sim}{\mu} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix} \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \\ \mu_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Reject the null hypothesis that the mean nitrogen concentration is the same at all three depths difference if

$$F = \frac{(n\text{-}4)\,T^2}{4(n-1)} = \frac{n(n\text{-}4)}{4(n\text{-}1)}(C\bar{\underset{\sim}{X}}\text{-}\underset{\sim}{0})'(CSC')^{-1}(C\bar{\underset{\sim}{X}}\text{-}\underset{\sim}{0}) > F_{(4,n\text{-}4),\alpha}$$

(d) (8 points) The null hypothesis is

$$H_0: \mu_2 = .5\mu_1 \text{ and } \mu_3 = .5\mu_2 \text{ and } \mu_4 = .5\mu_3 \text{ and } \mu_5 = .5\mu_4$$

There are various ways to express this in matrix form. One way is

$$H_0: \; C\underset{\sim}{\mu} = \begin{bmatrix} 0.5 & -1 & 0 & 0 & 0 \\ 0 & 0.5 & -1 & 0 & 0 \\ 0 & 0 & 0.5 & -1 & 0 \\ 0 & 0 & 0 & 0.5 & -1 \end{bmatrix} \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \\ \mu_4 \\ \mu_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Reject the null hypothesis that the mean nitrogen concentration is the same at all three depths difference if

$$F = \frac{(n\text{-}4)\,T^2}{4(n-1)} = \frac{n(n\text{-}4)}{4(n\text{-}1)} (C\bar{\underset{\sim}{X}}\text{-}\underset{\sim}{0})'(CSC')^{-1}(C\bar{\underset{\sim}{X}}\text{-}\underset{\sim}{0}) > F_{(4,n\text{-}4),\alpha}$$

(c) (4 points) Under the null hypothesis the covariance matrix has the form

$$\begin{bmatrix} \sigma_{11} & \alpha\sqrt{\sigma_{11}\sigma_{22}} & \beta\sqrt{\sigma_{11}\sigma_{33}} & \gamma\sqrt{\sigma_{11}\sigma_{44}} & \delta\sqrt{\sigma_{11}\sigma_{44}} \\ \alpha\sqrt{\sigma_{11}\sigma_{22}} & \sigma_{22} & \alpha\sqrt{\sigma_{22}\sigma_{33}} & \beta\sqrt{\sigma_{22}\sigma_{44}} & \gamma\sqrt{\sigma_{22}\sigma_{55}} \\ \beta\sqrt{\sigma_{33}\sigma_{11}} & \alpha\sqrt{\sigma_{33}\sigma_{22}} & \sigma_{33} & \alpha\sqrt{\sigma_{33}\sigma_{44}} & \beta\sqrt{\sigma_{33}\sigma_{55}} \\ \gamma\sqrt{\sigma_{44}\sigma_{11}} & \beta\sqrt{\sigma_{44}\sigma_{22}} & \alpha\sqrt{\sigma_{44}\sigma_{33}} & \sigma_{44} & \alpha\sqrt{\sigma_{44}\sigma_{55}} \\ \delta\sqrt{\sigma_{55}\sigma_{11}} & \gamma\sqrt{\sigma_{55}\sigma_{22}} & \beta\sqrt{\sigma_{55}\sigma_{33}} & \alpha\sqrt{\sigma_{55}\sigma_{44}} & \sigma_{55} \end{bmatrix}$$

Under the null hypothesis the covariance matrix has the form

$$\begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} & \sigma_{14} & \sigma_{15} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} & \sigma_{24} & \sigma_{25} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} & \sigma_{34} & \sigma_{35} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_{44} & \sigma_{45} \\ \sigma_{51} & \sigma_{52} & \sigma_{53} & \sigma_{54} & \sigma_{55} \end{bmatrix}$$

The difference in the number of parameters in these covariance matrices is 15-9=6. With respect to the α level of significance, reject the null hypothesis if

$$-2\ell og(\Lambda) > \chi^2_{6,.05}.$$

This is uses the large sample chi-square approximation to the null distribution of the $-2\ell og(\Lambda)$.

4. (a) (2 points) The experimental units are the 60 pigs.

(b) (2 points) The repeated measures factor is time with levels 2, 4, 6, 8, weeks

(c) (i) (6 points) You can set this up using an effect model or a cell means model. Here we will display results for an effects model. For the k-th pig in the j-th treatment (riboflavin) group, assume that the distribution of muscle tissue increases at 2, 4, 6, and 8 weeks is

$$\underset{\sim}{X}_{jk} = \begin{bmatrix} X_{1jk} \\ X_{2jk} \\ X_{3jk} \\ X_{4jk} \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_1 + \alpha_{1j} \\ \mu_2 + \alpha_{2j} \\ \mu_3 + \alpha_{3j} \\ \mu_4 + \alpha_{4j} \end{bmatrix}, \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} & \sigma_{24} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} & \sigma_{34} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_{44} \end{bmatrix} \right)$$

for j=1,2,3 and k=1,2,…,20, and assume that responses from different pigs are independent. Include the parameter restriction that

$$\begin{bmatrix} \alpha_{14} \\ \alpha_{24} \\ \alpha_{34} \\ \alpha_{44} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Then, the null hypothesis of no interaction between level of riboflavin and time can be expressed as

$$H_0: \underset{\sim}{0} = C\beta M = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} \mu_1 & \mu_2 & \mu_3 & \mu_4 \\ \alpha_{11} & \alpha_{21} & \alpha_{31} & \alpha_{41} \\ \alpha_{12} & \alpha_{22} & \alpha_{32} & \alpha_{42} \end{pmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & -1 & -1 \end{bmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

(c) (ii) (2points) Compute k=2, r=3, p=4, u=3, n=60, a=56, b=2, c =2,

$$F = \frac{1-\sqrt{0.835}}{\sqrt{0.835}} \frac{(56)(2)-2}{(2)(3)} = 1.73 \text{ with } (6,\ 110) \text{ df}$$

Since $1.73 < F_{(6,110),.05} = 2.18$, the null hypothesis is not rejected.

(d) (6 points) The conditions are given in part (c) (i). Responses from different pigs are independent, covariance matrices are homogeneous, and the biweekly gains in muscle tissue have a four dimensional normal distribution.

(e) (2 points) Since b=2, the Wilks criterion provides an exact F-test.

(f) (i) (6 points) Using the model set up in part (c), this null hypothesis is expressed as

$$H_0:\ \underset{\sim}{0} = C\beta M = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} \mu_1 & \mu_2 & \mu_3 & \mu_4 \\ \alpha_{11} & \alpha_{21} & \alpha_{31} & \alpha_{41} \\ \alpha_{12} & \alpha_{22} & \alpha_{32} & \alpha_{42} \end{pmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

(f) (ii) (2points) Compute k=2, r=3, p=4, u=4, n=60, a=55.5, b=2, c =3,

$$F = \frac{1-\sqrt{0.632}}{\sqrt{0.632}} \frac{(55.5)(2)-3}{(4)(2)} = 3.48 \text{ with } (8,\ 108) \text{ df}$$

Since $3.48 > F_{(8,108),.05} = 2.03$, the null hypothesis is rejected. The sets of mean biweekly gains in muscle tissue are not the same for all levels of riboflavin.

(g) (i) (6 points) Using the model set up in part (c), this null hypothesis is expressed as

$$H_0:\ \underset{\sim}{0} = C\beta M = \begin{bmatrix} 1 & 1 & 0 \end{bmatrix} \begin{pmatrix} \mu_1 & \mu_2 & \mu_3 & \mu_4 \\ \alpha_{11} & \alpha_{21} & \alpha_{31} & \alpha_{41} \\ \alpha_{12} & \alpha_{22} & \alpha_{32} & \alpha_{42} \end{pmatrix} \begin{bmatrix} 1 & 0 \\ -2 & 1 \\ 1 & -2 \\ 0 & 1 \end{bmatrix} = \begin{pmatrix} 0 & 0 \end{pmatrix}.$$

(g) (ii) (2points) Compute k=1, r=3, p=4, u=2, n=60, a=56, b=1, c =0,

$$F = \left(\frac{1-0.891}{0.891}\right)\left(\frac{56-0}{(1)(2)}\right) = 3.43 \text{ with } (2,\ 56) \text{ df}$$

Since $3.43 > F_{(2,56),.05} = 3.16$, the null hypothesis is rejected. When no riboflavin is added to the diets of the pigs, the mean biweekly increases in muscle tissue do not follow a straight line trend across time.

(h) (i) (6 points) The i–th biweekly muscle tissue increase measured on the k-th pig given the j-th level of riboflavin may be modeled as

$$X_{ijk} = \mu + \alpha_j + P_{jk} + \tau_i + (\tau\alpha)_{ij} + \varepsilon_{ijk}$$

where $P_{jk} \sim NID(0, \sigma_p^2)$ are random pig effects, $\varepsilon_{ijk} \sim NID(0, \sigma_\varepsilon^2)$ are random errors, and any P_{jk} is independent of any ε_{ijk}. With regard to fixed effects, α_j corresponds to a riboflavin level, τ_i corresponds to the j-th level of riboflavin and $(\tau\alpha)_{ij}$ corresponds to interaction between riboflavin level and time. The precise interpretation of the fixed effects parameters depends on the restrictions you place on those parameters.

(h) (ii) (4 points) The F-test statistic is computed as

$$F = \frac{MS_{\text{riboflavin levels}}}{MS_{\text{pigs(riboflavin levels)}}} \quad \text{on } (2, 57) \text{ degrees of freedom.}$$

This is a test of the hypothesis that the mean increases in muscle tissue, averaging across inspection times, are the same for all levels of riboflavin. The test in part (f) tests the null hypothesis that the vectors of mean muscle tissue increases for the four inspection times are the same for all levels of riboflavin.

(h) (iii) (4 points) One set of conditions is given by the model described in part (h) (i). A more general condition on the covariance structure for repeated measurements on a pig is given by Mauchly's condition. If the model in part (h)(i) is correct, the test in part (h)(ii) would have more power for detecting interaction than the test in part (c).

(5) (6 points) Since the correlation estimates are not independent, you will not be able to apply the Fisher z-transformation to each correlation and make a confidence interval for the difference in the transformed correlations using the assumption that the correlations are independent. You could describe an appropriate bootstrap procedure. Alternatively, you could describe a likelihood ratio test using a large sample chi-square approximation with one degree of freedom for the null distribution of $-2\ell og(\Lambda)$.

There are 100 possible points for your exam. There should be a score marked on your paper for each part of the exam. Also check if your total score was correctly recorded on page 7 of your exam paper. Please return the exam to me if you think any errors were made in recording your score.

Exam Scores (100 possible points**):**

```
10 | 0
 9 | 7 9
 9 | 0 0 0 1 2 2 2 4
 8 | 5 7 9
 8 | 1 2 3 4 4
 7 | 5 9 9
 7 | 0 1 2 3
 6 | 5
 6 | 0 3
 5 | 6 8 9
 5 | 0 1
 4 | 9
 4 | 0
 3 |
 3 |
 2 | 9
 2 |
 1 | 7
```