The exam will be offered on November 19th in our regular class meeting, or November 15th, 8-9:15 am, in this room, a class meeting that is offered in place of the November 21st meeting. Additional review today at 5:00 pm in my office, or as agreed today in class.

STA 5707-STA 4702 - Sample Exam 2

November 14, 2007

Student's name:

Notes: yes no

Book: yes no

This is a show your work, 70 minute exam. There is a deduction of 15% from your score if you use your personal class notes or your book, and of 25% if you use them both. A one page formula sheet with no hints of solutions of any type of exercise is ok. You should fully justify your answers. Data files, some printouts and SAS codes can be found on http://stat.fsu.edu/ vic/5707-4702/. You should save all your SAS or MINITAB work in a folder that will be copied on my flash drive. Good luck!

EXERCISE 1. *Assume the sampling matrix* $\mathbf{x}$ *for a random sample of size* $n = 4$ *from a bivariate normal population of mean vector* μ_0 *with* $\mu_0^T = (9, 5)$ *is given by* $\mathbf{x} = \begin{pmatrix} 6 & 10 & 8 & 8 \\ 9 & 6 & 3 & 2 \end{pmatrix}$.

a. What is the sampling distribution of T^2 *?*

b. Calculate the value of T^2 *for the given data.*

c. Is the test $\mu = \mu_0$ *statistically significant?*

EXERCISE 2. *The Alaska Fish and Game Department monitors grizzly bears. Bears are shot with a dart to induce sleep, the measured for length (in centimeters) and weighted (in kilograms) Table T1_4.dat gives the weights then lengths of seven female bears at 2, 3, 4 and 5 years of age.*

a. Obtain the equation of a 95% T^2 *confidence ellipse for the mean increases in weight from 2 to 3 years and from 4 to 5 years.*

b. Obtain 95% T based simultaneous Bonferroni confidence intervals for the mean increases in weight from 2 to 3 years and from 4 to 5 years.

EXERCISE 3. *Consider the data in Table $T6\_18.dat$, where the spectral reflectance is measured at wavelength 560 nm (first column), respectively at 720 nm (second column), for three species of seedlings: SS (sitka spruce), JL (Japanese larch) and LP (lodgepole pine) at different times during the growing season : 1 - Julian day 150, 2 - Julian day 235 and 3 - Julian day 320.*
Assuming the MANOVA assumptions hold true, perform a two-factor MANOVA using this data. Test for species effect, time effect and interaction effect.

EXERCISE 4. *Find the proportion of the total population variance explained by the first principal component, when the covariance matrix is given by* $\Sigma = \sigma^2 \begin{pmatrix} 1 & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho \\ \rho & \rho & 1 & \rho \\ \rho & \rho & \rho & 1 \end{pmatrix}$

EXERCISE 5. *The national track records for men in $T8\_6.dat$ give the records in seconds in the races of 100 m, 200 m, 400 m and then the records in minutes in the 800 m, 1500 m, 3000 m and Marathon races.*
a. Perform a principal components analysis for this standardized data. How many principal components do we need to explain at least 90% of the total variance?
b. Rank the nations based on their first principal component. Does this ranking correspond to any extent to your intuitive notion of athletic excellence for the various countries?

EXERCISE 6. *The weekly rates of return for five stocks (Allied Chemical, Du Pont, Union Carbide, Exxon, Texaco) are given it $T8\_4.dat$. Determine the proportion of the total sample variance explained by the first three principal components. Based on this data do you think at least 90% of the variability in these stock rates of return can be explained by these three principal components?*

EXERCISE 7. *Textbook: 5.1-5.3, 5.6, 5.18, 5.20-5.21, 5.23.*

EXERCISE 8. *Textbook: 6.1, 6.3, 6.4.a, 6.6 a,b, 6.24, 6.25, 6.33a*

EXERCISE 9. *Textbook: 8.5, 8.14, 8.17, 8.18, 8.20, 8.23.*