ASSIGNMENT #5
DUE MONDAY APRIL 2, 2001

1. Perform a classification / discriminant analysis using the SAS procedures DISCRIM and STEPDISC on the BANKRUPTCY DATA set in Table 11.4 of Johnson & Wichern. (See Problem 11.24 (page 711) for a description of the data.) The data are in

/afs/asu.edu/users/a/t/d/atdly/pub/stp533/jw4/T11-4.DAT

The SAS infile and input statements are:

infile '/afs/asu.edu/users/a/t/d/atdly/pub/stp533/jw4/T11-4.DAT';
input X1 X2 X3 X4 BANKRUPT;

For BANKRUPT, a 0 indicates YES (bankrupt) and a 1 indicates NO (not bankrupt).

In PROC DISCRIM make a run with each of the following options:

POOL = YES  (to obtain linear class. functions)
POOL = TEST (to test equality of cov. matrices and do quad. class. analysis if appropriate.)

Your analysis should also include plots of the data.

You should write a report of the results of your analysis including such points as: which variables are “most” important in the analysis; how well can one classify / discriminate; are there problems with the assumptions of the classification model; conclusions you can draw from the analysis as to differences between populations.

2. Perform the same analysis as in Problem #1 and write a report of your analysis of the CRUDE OIL DATA set in Table 11.7 of Johnson and Wichern. (See Problem 11.30 (page 717) for a description of the data.) The data are in

/afs/asu.edu/users/a/t/d/atdly/pub/stp533/jw4/T11-7.DAT

The SAS infile and input statements are:

infile '/afs/asu.edu/users/a/t/d/atdly/pub/stp533/jw4/T11-7.DAT';
input VANADIUM IRON BERYLUM STAHCAR AROMHCAR ZONE $;

The values of ZONE are coded: Wilhelm, Submuli, and Upper.