Spring 2003 (Newton) 15.075 Applied Statistics

Homework # 7 assigned 7 April 2003, due 16 April 2003

Read Chapter 11 in Tamhane and Dunlop and Chapter 10 (pages 247-260, 270-275) in the S-Plus On-line Doc Guide to Statistics, Volume 1. Then do the following problems.

Use the function **lm** in S-Plus to fit a linear model. One option for doing these problems is to extract the x and y variables from the dataframe by saying **y<-dataf[,1]** (if dataf is the name of the dataframe and y is in column 1) and **x1<-dataf[,2]** (if x1 is in column 2), etc. Then you can something like **my.lm<-lm(*y~x1+x2)*** or **my.lm<-lm(log(y)~x1+x2)**, if you want to take the log of the response variable. To fit an interaction term say **my.lm<-lm(y~x1+x2+x1:x2).**

Then **summary(my.lm)** gives regression coefficients, **summary.aov(my.lm)** gives an analysis of variance table, **plot(my.lm)** gives several diagnostic plots, **fitted(my.lm)** gives the fitted values, **resid(my.lm)** gives the residuals.

S-Plus will give you most of the numbers you need to answer the questions involving modeling, but please show where these numbers come from.

Type 3 sums of squares are sums of squares for each term in the model adjusted for all the other terms in the model. If my.lm is an object returned by the lm function, the type 3 sums of squares can be obtained by saying **summary.aov(my.lm, ssType=3)**. The t-statistics in **summary(my.lm)** are equal to the square root of the f-statistics in **summary.aov(my.lm, ssType=3).**

pairs(dataframename) is a good first look at a dataframe.
solve(x) gives the inverse of a square matrix x.
x[,-i] is a matrix x without the ith column.
x[-i,] is a matrix x without the ith row.

1. 11.4
2. 11.10
3. 11.14
4. 11.20
5. 11.22
6. 11.26
7. 11.34 cor(Ex11.3) will give the correlations. The diagonal elements of solve(cor(Ex11.3[,-1])) are the VIFs.
8. 11.38
9. 11.42