STA 302/1001 Summer 2001 Assignment 5

Quiz on June 20th. Do this assignment in preparation for the quiz. Bring a calculator. Bring printouts of the log and list files to the quiz. **Quiz 5 will be open book** (including photocopies), so make sure you can locate any formulas you needed to do this assignment.

1. Use SAS to do 6.6 *a* and *b* but not *c*, and 6.8. For part *a* of 6.8, also give the $\mathbf{C}$, $\boldsymbol{\beta}$ and $\mathbf{h}$ matrices of $H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{h}$. As usual, the data are on the Web as well as well as on your data disk.

2. Do exercise 6.26. I seem to recall that there is a very quick proof, but I cannot recover it. If you can re-discover the quick way, great. Otherwise, these hints may help. Start with the formula for r, and square it. In the denominator, use the formulas for $SSTO$ and SSR, recalling $\sum_{i=1}^n Y_i = \sum_{i=1}^n \hat{Y}_i$. Expand the numerator and work on it, using $\mathbf{e}'\hat{\mathbf{Y}} = \mathbf{0}$.

3. This item starts with a reading guide to Chapter 7. First read Sect. 7.1, but don't get too caught up in the book's strange "extra sum of squares" notation. It's easy to decode. For example, $SSR(X_3, X_4, X_5|X_1, X_2)$ refers to $SSR(F) - SSR(R)$, where the full model has $X_1, \ldots, X_5$, and the reduced model has just X_1 and X_2.

   Read 7.2 and 7.3. You might want to take a look at Section 2.8 starting on P. 78 for some background.

   Section 7.4 you can skip. A coefficient of partial determination is the proportion of the remaining variation explained by the additional variables in a full model. The concept is important, but the notation and formulas are not.

   Skim or skip Sect. 7.5, read Sect 7.6 on mulitcolinearity, read Sect 7.7 on polynomial regression, read Sect 7.8 on interactions (we did more detail in lecture), and skip Sect 7.9.

   Do problems 7.1, 7.2, 7.20, 7.46, 7.47. For 7.46 and 7.47, also give the $\mathbf{C}$, $\boldsymbol{\beta}$ and $\mathbf{h}$ matrices of $H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{h}$.

4. In Chapter 11, read 11.1-11.4, skip 11.5, read 11.6 and skip 11.7. Then do 11.1, 11.3, 11.10, 11.12. In 11.12b, let's assume they are asking us to state H_0. Please do this in symbols, writing more than one equation in terms of β values. Then state H_0 in matrix terms, giving the $\mathbf{C}$, $\boldsymbol{\beta}$ and $\mathbf{h}$ matrices of $H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{h}$.

5. This should be quick. Going back to the SMSA data set you used in Assignment 4, create indicator dummy variables to represent geographic region. Then, treating number of serious crimes as the dependent variable,

   (a) Test physicians and hospital beds controlling for all the other independent variables.

   (b) Test geographic region controlling for all the other independent variables.

   Please compute F values with the SAS `test` statement, not a calculator.