Homework 5
BIOST 515
Due: February 17, 2004 in class
This assignment is worth 20 points.

1. Assuming a sample size of 100, simulate 20 predictors from normal distributions. You may use whatever means and variances you choose. Also, simulate an outcome from a normal distribution (indepedent of the predictors and again choose your own means and variances). Use stepwise regression with AIC to choose a model. At each step of the stepwise procedure, explain the selection that is made (whether to add a variable, drop a variable or do nothing). Comment on the final model. Repeat (with comments) using forward and backward selection (these are options in stepAIC()).

2. Using the SMSA data, model log(NOx) as a predictor of Mortality using cubic B splines with equally spaced knots. Plot the fitted values of Mortality over the scatterplot of log(NOx) vs. Mortality. Comment. Refit with knots based on quantiles and compare. For both models, do the appropriate F test for log(NOx) as a predictor for Mortality.