1 Review problem set 3 solutions

Here are solutions to review problem set 3:

12.12 a) see the printout   b) No, r $= -.24$ c) $2(k+1)/n = 2(3)/22 = .27$ and the plot of leverage values shows two points above it d) $\widehat{y_i} = 144.39 + .17*$Distance $+1.09*$Pop . The plots show problems due to these two leverage points e) Both slopes look to be affected by the leverage points, but they may have the correct sign.

12.15) a) Yes, $F = 14.65$ on 2 and 19 degrees of freedom, $p < .0001$. b) The test statistic $t = 1.11$ on 19 df, $p = .28$, fail to reject $H_0 : \beta_1 = 0$. c) The confidence interval is given by: $\widehat{\beta} \pm t_{\alpha/2, n-k-1} s.e.(\widehat{\beta})$, and $t_{.025,19} = 2.093$, so for the Distance slope we have .169 $\pm(2.093)(.152)$, giving $(-.15, .49)$. d) Note that this has a one-sided alternative hypothesis. The test statistic is $t = (\widehat{\beta} - \beta_0)/s.e.(\widehat{\beta}) = (1.09 - .5)/.202 = 2.92.$Since $t_{.05,19} = 1.729$, our t value is greater than the tabled value, so we can reject H_0 and conclude that the slope for population size is greater than .5.

12.17) a) see the printout, not much relationship is evident for either variable b) $\widehat{y_i} = 184.55 + .14*$Distance $+.62*$Pop. The fit is better but there is still one leverage point. Neither variable is significant in this model. c) Had we not inspected plots, we would have inferred an incorrect relationship due to the effect of one or two data points.

12.53 a) $\widehat{y_i} = 102.71 - .83 * PROTEIN - 4.00 * ANTIBIO - 1.38 * SUPPLEM$ b) $\widehat{\sigma}_\varepsilon = s_\varepsilon = 1.71$ c) $R^2 = .9007$ d) There is no collinearity problem, as seen in the correlation matrix which shows that all x variables are uncorrelated. This is also reflected in the VIF values of 1 in the printout.

12.55 a) $\widehat{y_i} = 89.83 - .83*PROTEIN$ b) $R^2 = .506$ c) $F = \frac{(371.08 - 208.33)/2}{2.92} = 81.38/2.92 = 27.87$. Since $F_{2,14,.05} = 3.74$, we reject the null and conclude at least one slope is not equal to 0.

13.3 Check the computer program for details. a) There is moderate collinearity for both A and C, as seen by VIF values, but both are below 10.   b) The best model according to multiple criteria (Cp, AIC, BIC) is the model with the A, I, and G terms c) Any other variables that may affect a person's view on crime.