STAT 3611 Lab10-Procedure S2015 Maghsoodloo

1. Enter the design matrix of Example 13.16 on page 566 of Devore from Canvas onto a Minitab worksheet starting in C1. Go to Stat → Regression → General Regression, response: y; Model (or regressors): x_1 x_2 x_3 x_4; Prediction interval for new observations: 7 4 180 5, Results: display CIs; Storage: Fits, Residuals, Studentized Residuals, and C = $(X'X)^{-1}$, then ok. Determine if this model is adequate based on the $\mathbf{R}^2_{\mathbf{Model}}$ and why? To obtain the matrix C go to Data → Display Data, and double click on XPXI. Compare your regression coefficients with those of Devore near the bottom of p. 566 of Devore.

2. Go back to your MTW and use Minitab's Calc to Compute the quadratic components $x_{11} = \mathbf{x}_1^2$ storing it in C9, $x_{22} = \mathbf{x}_2^2$ in C10, $x_{33} = \mathbf{x}_3^2$ in C11 , … , $x_{12} = x_1x_2$ in C13, ..., and $x_{34} = x_3x_4$ in C18. Although these columns are unnecessary because Minitab will accept x_1*x_1, but they are needed for Minitab's Best-subset. Now regress y on the full quadratic model but W/O the CI, PI, residuals, and fits, etc. Compare your answers with those of Devore at the bottom of p. 566. Do you see a typo near the bottom of p. 566 for full quadratic model? Examine the *P*-values of all the t-statistics of coefficients to determine if you observe model over-fitting? Test to see if the net contributions of all the quadratic variables is significant at the 5% level.

3. Use Minitab's Best subsets Regression to obtain the best k-variable fits to the data, limiting the number of regressors from k = 5 to 10. Select the best model. Hint: All models whose C_p index exceed k must first be eliminated from considerations. Usually, the model with the smallest C_p index relative to k (i.e., the smaller C_p index is relative to k the better is the model), the smallest S, and with the largest R-Sq is generally the best model.

4. Next obtain the selected model of part (3), whose coefficients must be significant at the 10% level, storing $\hat{Y}$, e, r vectors, and also obtain the 95% CI for μ_0 and PI for y_0 at x_1 = 3, x_2 = 9, x_3 = 120 and x_4 = 4.

5. Compute the correlation coefficient between vectors $\mathbf{Y}$ and $\hat{\mathbf{Y}}$ of parts 3 & 4 above and describe the relationship between r($\mathbf{Y}$, $\hat{\mathbf{Y}}$) and $\mathbf{R}^2_{\mathbf{Model}}$ of part 4.