

Name: _____

Instructor: D. Kushary

1. A researcher was investigating the performance of a chemical process as a server controllable process variable where Y is the CO_2 product while X_1 and X_2 are the total solvent and hydrogen consumption(in months)(Call it Model-1). Then a multiple linear regression was fitted using twenty seven data points and the following results were found as:

$$\hat{Y} = 36.77 - 0.8117 \times X_1 + 2.0996 \times X_2 \quad \& \quad R^2 = 55.44\% \quad S_{b_1} = 0.17 \quad S_{b_2} = 0.19$$

- a) Is the variable X_2 significant at 5% level?
- b) If you drop X_2 from the model, what will be the R^2 for the new reduced model ? (10 points)
- c) If the 7th point in the data set has $Y = 19.07$, $X_1 = 18.9$, and $X_2 = 1.71$, find the press residual for the 7th point (Given $h_{77} = 0.0386$).
2. A model was established to predict the TOTSALLES(y) using PARTSERV(X_1), RETUSED(X_2), RETNEW(X_3) and YRSINBUS(X_4) (These are all variable names, Model-1). The R-output is given below.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
PARTSERV	1	2.92	2.92	301.03	0.0000
RETUSED	1	1.22	1.22	125.77	0.0000
RETNEW	1	0.15	0.15	15.46	0.0004
YRSINBUS	1	0.05	0.05	5.15	0.0281
Residuals	29	0.28	0.0097		

- (a) Let's say, Model-2 only kept 'PARTSERV' and 'RETUSED,' as independent variables. Find $adj - R^2$ for the Model-2.
- (b) Is Model-2 significant at 5% level?
- (c) Is 'RETUSED' significant in Model-2 at 5%?

(WRITE R-PROGRAM AND SUBMIT IN DROPBOX -
FILE NAME SHOULD BE TEST2.LASTNAME.FIRSTNAME.R)

3. The attached data set has 9 independent variables (x_1 - x_9) and one dependent variable (y). Use forward ($\alpha = 0.05$), backward($\alpha = 0.05$) and stepwise (alpha.enter=0.05, and alpha.remove=0.05) selection procedure to find the best models.
- (a) Let say X_m is the last variable dropped from the model in the backward selection procedure above. Now, if α value is changed from the 0.10 to 0.05 in backward selection, X_m will still be dropped from the model (no need to run with the new alpha)?
- (b) Calculate the $R^2_{prediction}$ and BIC for the final model.
- (c) Calculate $adj - R^2$, C_p , SSE and BIC for all possible models using only the variables in your final models.