KING FAHD UNIVERSITY OF PETROLEUM & MINERALS
DEPARTMENT OF MATHEMATICS & STATISTICS
DHAHRAN, SAUDI ARABIA

STAT 310: Linear Regression
Semester 161
Third Major Exam (Mathematical)
Sunday January 01, 2017
7:00 pm

Name: ID #:

Question No	Full Marks	Marks Obtained
1	**07**	
2	**15**	
3	**10**	
4	**10**	
Total	**42**	

Q.No.1:- (7 points) The Cook's D statistic is given as $D_i = \frac{(\hat{\boldsymbol{\beta}}_{(i)} - \hat{\boldsymbol{\beta}})' \boldsymbol{M} (\hat{\boldsymbol{\beta}}_{(i)} - \hat{\boldsymbol{\beta}})}{p(MSE)}$ by definition. If we know that $\left(\hat{\boldsymbol{\beta}}_{(\boldsymbol{i})} - \hat{\boldsymbol{\beta}}\right) = \frac{(\boldsymbol{X}'\boldsymbol{X})^{-\mathbf{1}} \boldsymbol{x}_i e_i}{1 - h_{ii}}$, then mathematically show that

$$D_i = \frac{r_i}{p} \frac{h_{ii}}{(1 - h_{ii})}$$

where $r_i = \frac{e_i}{\sqrt{MSE(1 - h_{ii})}}$ are studentized residuals.

Q.No.2:- (2+6+4+3 = 15 points) The excel file provides the data on life expectancy (Y), people per TV (X_1) and people per doctor (X_2) for 38 countries. Transform the data using square roots of both the response and the regressors.

(a) Write down the fitted model and predict the life expectancy of Saudi Arabia if people per TV is 35 and people per doctor is 623.

(b) Find the Cook's D statistic for Ethiopia and Sudan. Comment on the calculated values.

(c) Calculate the covariance ratio ($COVRATIO_i$) for Sudan and comment on the value. Should Sudan be removed from the data? Why?

(d) Calculate the joint Cook's D statistic for Spain, Turkey and Germany. Comment on the value.

Q.No.3:- (1+3+6 = 10 points) The carbonation level of a soft drink beverage is affected by the temperature of the product and the filter operating pressure. Twelve observations were obtained and the resulting data are shown below.

Carbonation, y	*Temperature, x_1*	*Pressure, x_2*
2.6	31	21
2.4	31	21
17.32	31.5	24
15.6	31.5	24
16.12	31.5	24
5.36	30.5	22
6.19	31.5	22
10.17	30.5	23
2.62	31	21.5
2.98	30.5	21.5
6.92	31	22.5
7.06	30.5	22.5

(a) Fit a second-order polynomial and write down the model.

(b) Test the overall significance of fitted model.

H_0:

H_1:

F_0= with v_1= and v_2=

P-value=

Decision and conclusion:

(c) Does the second-order terms contribute significantly to the model?

H_0:

H_1:

F_0= with v_1= and v_2=

P-value=

Decision and conclusion:

Q.No.4:- (2+3+5 = 10) Smith et al. (1992) discuss a study of the ozone layer over the Antarctic. These scientists develop a measure of the degree to which oceanic phytoplankton production is inhibited by exposure to ultraviolet radiation (UVB). The response is INHIBIT. The regressors are UVB and SURFACE, which is depth below which the ocean's surface from which the sample was taken.

(a) Fit a multiple regression of INHIBIT on UVB and SURFACE and write down the model.

(Note: If you define any new variables, write their description)

(b) Make a plot of residuals against fitted values. What is wrong with the residual plot? Do you suggest any transformation on y? Mention the transformation.

(c) After making the transformation, fit the model again and report all the 3 estimated coefficients. Also, interpret the 3 estimated coefficients.

Some useful formulas

$$SST = \boldsymbol{y}'\boldsymbol{y} - \frac{(\sum y_i)^2}{n}, \quad SSR = SS_{Regression} = \widehat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{y} - \frac{(\sum y_i)^2}{n}, \quad SSE = SS_{Residuals} = \boldsymbol{y}'\boldsymbol{y} - \widehat{\boldsymbol{\beta}}'\boldsymbol{X}'\boldsymbol{y}$$

$$R^2_{adj} = 1 - \frac{SSE/(n-p)}{SST/(n-1)}, \quad \hat{\beta}_j \pm t_{\frac{\alpha}{2},n-k-1} se(\hat{\beta}_j), \quad PRESS = \sum\left(\frac{e_i}{(1-h_{ii})}\right)^2, \quad R^2_{prediction} = 1 - \frac{PRESS}{SST}$$

$$d_i = \frac{e_i}{\sqrt{MSE}}, \qquad r_i = \frac{e_i}{\sqrt{MSE(1-h_{ii})}}, \qquad t_i = \frac{e_i}{\sqrt{S^2_{(i)}(1-h_{ii})}} \text{ where } S^2_{(i)} = \frac{(n-k-1)MSE - \frac{e_i^2}{(1-h_{ii})}}{n-k-2}$$

$$COVRATIO_i = \frac{|(\mathbf{X}'_{(i)}\mathbf{X}_{(i)})^{-1} S^2_{(i)}|}{|(\mathbf{X}'\mathbf{X})^{-1} MS_{\mathrm{Res}}|} = \frac{(S^2_{(i)})^p}{MS^p_{\mathrm{Res}}}\left(\frac{1}{1-h_{ii}}\right)$$

With the Best Wishes