KING FAHD UNIVERSITY OF PETROLEUM & MINERALS
DEPARTMENT OF MATHEMATICS AND STATISTICS
Term 142

STAT 310 Introduction to Linear Regression
Final Exam

May 18, 2015

Name: _____ID #:_____

Important Note:

- Show all your work including formulas, intermediate steps and final answer
- In questions requiring a test of a hypothesis, it is important to specify the null and alternative hypotheses, the testing procedure, the decision and the conclusion.

Question No	Full Marks	Marks Obtained
1	18	
2	17	
3	4	
4	6	
Total	45	

1) Answer the following:

i) Write down a second order hierarchical model in three variables. *(2 pts.)*

ii) Name two drawback of using a polynomial model relative to using a linear model with transformed variables. Explain. *(4 pts.)*

iii) In the quadratic model $y_i = \beta_0 + \beta_1 x_i + \beta_{11} x_i^2 + \varepsilon_i$, does testing for the quadratic effect necessarily establish that a linear function is appropriate? Explain. *(2 pts.)*

iv) Explain what the DFBETAS, DFFITS, and Cook's Distance are and what are they used for? *(6 pts.)*

v) Explain the forward selection procedure. *(2 pts.)*

vi) Explain the backward elimination procedure. *(2 pts.)*

2) Given the following data

X	0	0	1	1	2	2	3	3	4	4	5	5	6	6
Y	508.1	498.4	568.2	577.3	651.7	657.0	713.4	697.5	755.3	758.9	787.6	792.1	841.4	831.8

The analyst fit a second order model $y_i = \beta_0 + \beta_1 x_i + \beta_{11} x_i^2 + \varepsilon_i$ and got the following results

Coefficient	Estimated Value	Standard Error
β_0	705.474	3.208
β_1	54.893	1.050
β_{11}	-4.249	0.606

i) Write down the estimated model. *(1 pt.)*

ii) Complete the ANOVA Table *(5 pts.)*

ANOVA Table

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square
Regression	171773		
x	168741		
$x^2 \mid x$			
Error			
Total	172453		

iii) Test for lack of fit. *(5 pts.)*

iv) Test for quadratic effect in two ways. *(4 pts.)*

v) What is the value of the coefficient of determination? Explain its meaning. *(2 pts.)*

3) A researcher fit the following model

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_{11} x_{i1}^2 + \beta_{22} x_{i2}^2 + \beta_{12} x_{i1} x_{i2} + \varepsilon_i \quad i = 1, \dots, 10$$

ANOVA TABLE

Source of Variation	Sum of Squares
Regression	55366
x_1	18704
$x_2 \mid x_1$	34201
$x_1^2 \mid x_1, x_2$	1646
$x_2^2 \mid x_1, x_2, x_1^2$	285
$x_1 x_2 \mid x_1, x_2, x_1^2, x_2^2$	529
Error	5240
Total	60606

Help the researcher test if a first order model is adequate. *(4 pts.)*

4) In a production process, the engineer wanted to study the relationship between the amount of product as a function of line speed and production line. He considered the model

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_{12} x_{i1} x_{i2} + \varepsilon_i \quad i = 1, \dots, 10$$

where X_1 denotes the line speed, and

X_2 is a qualitative variable with two levels, denoting the line.

And got the following fitted model $\hat{Y} = 7.57 + 1.322X_1 + 90.39X_2 - 0.176X_1X_2$

Source of Variation	Sum of Squares
Regression	169165
x_1	149661
$x_2 \mid x_1$	18694
$x_1x_2 \mid x_1, x_2$	810
Error	9904
Total	179069

i) Test if the regression lines are identical. *(3 pts.)*

ii) Test if the regression lines are parallel. *(3 pts.)*