TEST 2
STAT 572
Spring 1995

Name____________________

1. An experiment was conducted to investigate the effect of injecting CO_2 into oil pockets to recover underground crude oil. The experiment was conducted in the laboratory using three levels of pressure and three different dipping angles. The data appears below.

Pressure (x_1, pounds per square inch)	**Dipping Angle** (x_2, degrees)	**Oil Recovery** (y, percentage)
1,000	0	60.58
1,000	15	72.72
1,000	30	79.99
1,500	0	66.83
1,500	15	80.78
1.500	30	89.78
2,000	0	69.18
2,000	15	80.31
2,000	30	91.99

The results of fitting a quadratic model were:

Response: Recovery

Summary of Fit

RSquare	0.991618
RSquare Adj	0.983236
Root Mean Square Error	1.353834
Mean of Response	76.90667
Observations (or Sum Wgts)	9

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	23.643333	8.185442	2.89	0.0446
Pressure	0.0494167	0.011541	4.28	0.0128
Pressure*Pressure	-0.000013	0.000004	-3.48	0.0253
Angle	0.9301111	0.132853	7.00	0.0022
Angle*Angle	-0.006867	0.004255	-1.61	0.1818

Effect Test

Source	Nparm	DF	Sum of Squares	F Ratio	Prob>F
Pressure	1	1	33.605600	18.3350	0.0128
Pressure*Pressure	1	1	22.244450	12.1364	0.0253
Angle	1	1	89.838001	49.0150	0.0022
Angle*Angle	1	1	4.774050	2.6047	0.1818

a) Comment on the fit of the model. (5 p)

b) The model with interaction was fitted and resulted in an R^2=0.995. Test the significance of this term. (10 p)

c) Comment on the collinearity structure of these data. (5 p)

d) A second replicate was taken at each point. Outline the complete ANOVA table for this model. Breakdown the SSE into two components and explain how a test for lack of fit would be conducted. (10 p)

e) Estimate the optimal settings for the two factors in order to maximize the response. How reliable are these estimates? (5 p)

2. A data set that consists of 28 properties is analyzed. The response is the price of the house and several regressors are used to explain it. The results follow.

Response: Price

Summary of Fit

RSquare	0.937441
RSquare Adj	0.919568
Root Mean Square Error	40.14948
Mean of Response	381.5714
Observations (or Sum Wgts)	28

Lack of Fit

Parameter Estimates

Effect Test

Source	Nparm	DF	Sum of Squares
x1	1	1	9282.788
x2	1	1	4970.581
x3	1	1	697.544
x4	1	1	12376.003
x5	1	1	1869.166
x6	1	1	1678.305

Sequential (Type 1) Tests

Source	Nparm	DF	Seq SS
x1	1	1	454291.71
x2	1	1	34441.78
x3	1	1	4606.07
x4	1	1	11559.39
x5	1	1	690.01
x6	1	1	1678.30

a) Explain why, except for the last one, the sums of squares in the "Effect Test" frame are different to the ones for the "Sequential Tests." (5 p)

b) If the regressors were orthogonal (uncorrelated) to each other, how would the two sets of sums of squares compare? (5 p)

c) Test the contribution of x5 and x6 (jointly) to the model that contains all other regressors [Hint: S_{yy} = 541,118.86]. (10 p)

d) Test the contribution of x3 to a model that contains x1 and x2 only. (10 p)

e) Could you test the joint contribution of x3 and x5 to the model that contains all other regressors? Explain. (5 p)

3. Determine the truthfulness or falseness of the following propositions. (3 p each)

T F If there is strong collinearity, the model is useless for prediction

T F If an observation is deleted, then the SSE will necessarily decrease

T F If the value of DFBETAS for certain coefficient and certain observation is negative, it means that if the observation is deleted the coefficient will decrease

T F When the F-test is significant, then at least one of the individual t-tests must be significant

T F If model 1 has an R^2 larger than model 2, then the MSE of model 1 is smaller than that for model 2.