Department of Mathematics and Statistics
Semester 121

STAT310 Midterm Exam Tuesday November 13, 2012

Name: _____ ID #: _____

1) A company that sells home insurance would like to know whether advertising increases sales. Let Y denote the number of policies sold in a month, and X the amount spent on advertising, in thousands of dollars. The company has data on advertising costs and policy sales for 42 months, and they fit a simple linear regression with the following results:

	Mean	Standard Deviation	Minimum	Maximum
Policies Sold	373.5	124.04	146	609

Parameter Estimates

Predictor	Coef	Stdev	t-ratio
Constant	48.40	17.61	_____
Ad Spending	10.25	0.52	_____

S = _____ R^2 = _____

Analysis of Variance

SOURCE	DF	SS	MS	F ratio
Regression	__	_____	_____	____
Error	__	59413	_____	
Total	__	_____	_____	

i) What is the linear regression relation between sales and advertising expenditures?

ii) Complete the tables above.

iii) What percentage of policy sales is explained by advertising expenditures? Does the model do a good job in this respect? Explain.

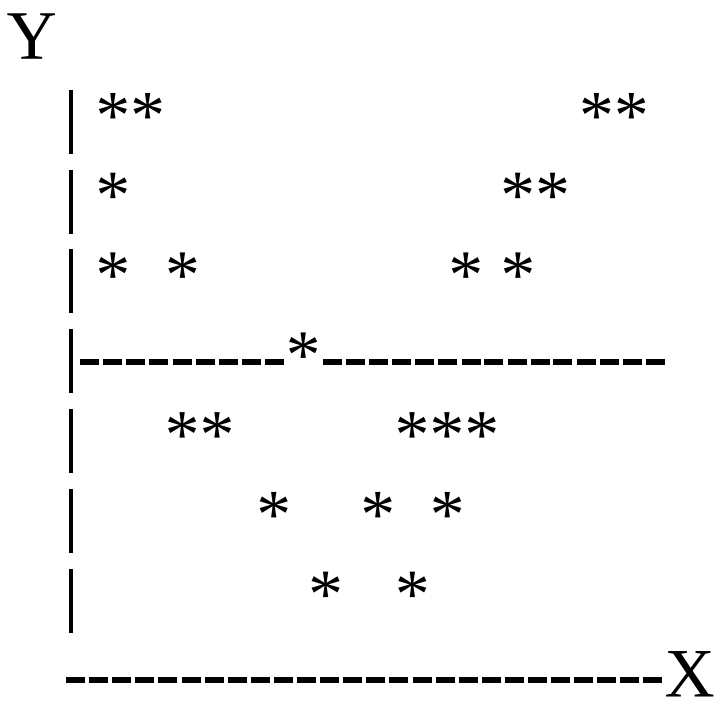
iv) Do advertising expenditures give good predictions for the number of policies sold? Briefly justify your answer using appropriate numbers from the tables above

v) Is there a significant linear relationship between sales, and advertising? Justify your answer by performing two statistical tests. Clearly state your hypotheses, your test statistics, your conclusion and decision. Is there any difference between the two tests? And Why?

vi) Find a 99% confidence interval for the slope of the regression line, and explain what it tells you about the relationship between advertising and television sales.

vii) Suppose your company makes a $100 profit per television sold before taking advertising costs into account. According to your best estimate, do the ads appear to be paying for themselves? Explain.

2) Consider the residual plot of a simple linear regression given here.



For each statement below, indicate if it is true or false and why.

a) The mean 0 assumption is correct because there are approximately as many points above the line as below it.

b) The constant variance assumption is violated because there is a curved pattern to the points.

c) The errors for this data set are approximately normally distributed.

d) A linear model is not appropriate for this data set because of the curved pattern in the points.

3) Consider the data set problem3 in Blackboard. Use Minitab to solve the parts below. Save all Minitab output used in a word file called "yourlastname", and email it to me in blackboard or to walid@kfupm.edu.sa before the end of the exam session.

i) Fit a multiple linear regression relating y to the two regressors.

ii) Test for the significance of the regression. Write down the hypotheses, test statistic, decision and conclusion.

iii) Asses the contribution of each regressor. What conclusions can you draw?

iv) Use a partial F test to determine the contribution of x_2.

v) Use a partial regression plot to assess whether x_2 is necessary.

vi) Assess the normality assumption of the model.

vii) Construct and interpret a plot of the residuals versus the predicted response.

viii) What do we need to perform a lack of fit test? Explain.

ix) Can you perform a lack of fit test for this data set? Explain if not, otherwise perform the test.