STAT 2509B
Test#1

(February 6th, 2012)

All solutions must be shown clearly.

1. A manufacturer of laundry detergent was interested in testing a new product prior to market release. One area of concern was the relationship between the height of the detergent suds in a washing machine as a function of the amount of detergent added in the wash cycle. For a standard size washing machine tub filled to the full level, random assignments of amounts of detergent were made and tested on the washing machine.

Height (Y)	**Amount** (X)
28.1	6
32.3	7
34.8	8
38.2	9
43.5	10
60.3	6
63.7	7
65.4	8
69.2	9
72.9	10
88.2	6
89.3	7
94.1	8
95.7	9
100.6	10

$\sum y_i$ = 358.9 $\quad \sum x_i$ = 80

$\sum y_i^2$ = 13 227.03 $\quad \sum x_i^2$ = 660 $\quad \sum x_i y_i$ = 2 952.5

[1] (a) The response variable, y, is: _____

[1] (b) The explanatory variable, x, is: _____

[6] (c) State a SLR model making sure you give all assumptions necessary for statistical inference.

[5] (d) Find the least squares estimates of β_0 and β_1. Find the least squares fitted regression line.

Assuming no violations of the assumptions, answer the following questions:

[6] (e) Find s^2, an estimate of σ^2.

[6] (f) Use the t-test to test whether there is a significant linear relationship between the height of the detergent suds in a washing machine and the amount of detergent added. Use α = 0.05.

[4] (g) Find a 90% confidence interval for the true population slope, β_1.

[23] (h) Complete the following ANOVA table and hence test whether there is a significant linear relationship between the height of the detergent suds in a washing machine and the amount of detergent added. Use α = 0.05.

Source	*d.f.*	*SS*	*MS*	*F*
Regression				
Error				
Total				

[5] (i) Find the values of the coefficient of correlation, r, and coefficient of determination, r^2, and interpret their meanings in this problem. What is your conclusion about the model?

[5] (j) Find a 95% Confidence Interval of the average height of the detergent suds if the amount of the added detergent is 5.

[5] (k) Find a 95% Prediction Interval of the height of the detergent suds if the amount of the added detergent is 5.

2. Refers to question 1.

Amount x_i	**Height** y_{ij}	n_i	$\bar{y}_i$	$\sum_j \left(y_{ij} - \bar{y}_i\right)^2$
6	28.1, 27.6	2	27.85	0.125
7	32.3, 33.2	2	32.75	0.405
8	34.8, 35.0	2	34.9	0.02
9	38.2, 39.4	2	38.8	0.72
10	43.5, 46.8	2	45.15	5.445

[5] (a) Decompose SSE into the sum of squares due to the pure error, SSPE, and sum of squares due to the lack of fit, SSLF.

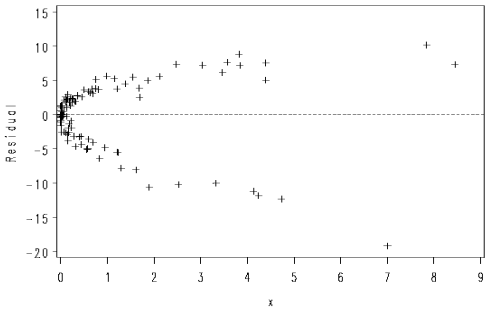
<u>Hint</u>: $SSPE = \sum_i \sum_j \left(y_{ij} - \bar{y}_i\right)^2$ = 6.715

$\sum x_i$ = 80 $\sum x_i^2$ = 660 $\sum y_i$ =358.9 $\sum y_i^2$ = 13 227.03 $\sum x_i y_i$ = 2 952.5

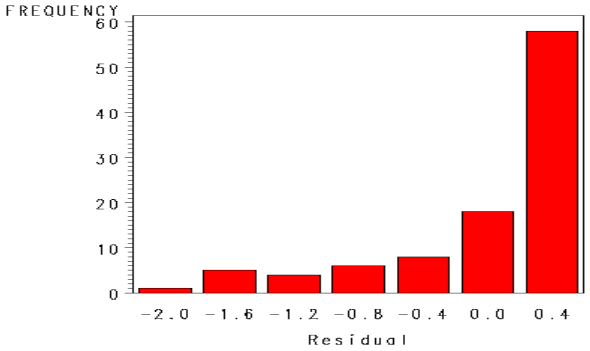
[6] (b) Test whether the linear model $y = \beta_0 + \beta_1 x + \varepsilon$ is appropriate.
Use α = 0.05.

3. State which violations of the SLR model (if any) are indicated by each of the following residual plots. Give reasons for your answer.

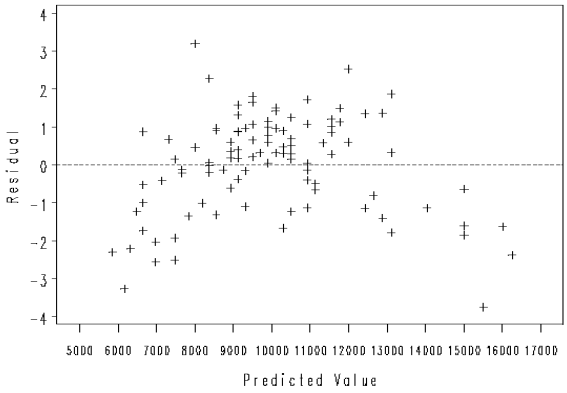
[3] (a)



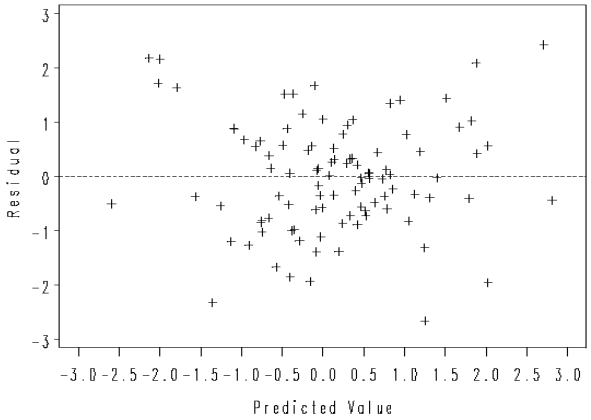
[3] (b)


FREQUENCY
60
50
40
30
20
10
0
-2.0 -1.6 -1.2 -0.8 -0.4 0.0 0.4
Residual

[3] (c)


Residual
4
3
2
1
0
-1
-2
-3
-4
5000 6000 7000 8000 9000 10000 11000 12000 13000 14000 15000 16000 17000
Predicted Value

[3] (d)


Residual
3
2
1
0
-1
-2
-3
-3.0 -2.5 -2.0 -1.5 -1.0 -0.5 0.0 0.5 1.0 1.5 2.0 2.5 3.0
Predicted Value