

MTH 4230 Lab 2 Answer Sheet

Due Wed., Feb. 12

1 Part A: Lack of Fit Test

1.1 Chemistry Experiment Data Set

1. NA
2. *Don't* print the plot. Just **describe the pattern** in the plot.
3. NA
4. *Don't* print the plot. Just **indicate how well the line fits the data** (based on the plot).
5. *Don't* print the plot. Is there **any indication that the linear model doesn't fit the data** (based on the plot)? Explain.
6. NA
7. NA
8. (Refer to **R Notes 2** in R Instructions for Regression MTH 4230 on the course website for help on interpreting the output from `anova()`).

Give the values of the following statistics:

$$\text{SSPE (or equivalently SSE(F))} = \text{-----}$$

$$\text{SSLF (or equivalently SSE(R) - SSE(F))} = \text{-----}$$

$$F = \frac{\text{SSLF}/(df_R - df_F)}{\text{SSPE}/df_F} = \text{-----}$$

$$\text{P-value} = \text{-----}$$

Based on the results of the lack of fit test, is a straight line regression model appropriate, or should a more sophisticated model be used?

2 Part B: Transformations

2.1 Chemistry Experiment Data Set (Cont'd)

1. NA
2. *Don't* print the plot. Just **describe the pattern** in the plot.
3. Give the equation of the fitted regression line.
4. NA (*don't* print the scatterplot).
5. *Don't* print the histogram or normal probability plot. Just **indicate whether the normality assumption appears to be met** (based on the plots).
6. *Don't* print the plot. Just **indicate whether the constant standard deviation assumption appears to be met** (based on the plots).
7. *Don't* print the plot. Just **describe the pattern of the points and the line**.

3 Part C: General Linear F Test

3.1 Chemistry Experiment Data Set (Cont'd)

1. NA
2. NA
3. Refer to **R Notes 2** in R Instructions for Regression MTH 4230 on the course website for help on interpreting the output from `anova()`.

Give the following values:

$$\text{SSE(F)} = \text{-----}$$

$$\text{SSE(R)} = \text{-----}$$

$$F = \frac{(\text{SSE(R)} - \text{SSE(F)}) / (df_R - df_F)}{\text{SSE(F)} / df_F} = \text{-----}$$

$$\text{P-value} = \text{-----}$$

Based on the results of the general linear F test, is a model with just an intercept sufficient for describing these data, or should the model include a slope?