**228.371 Assignment 3**

**Assessed lab: June 18, 2014**

**Question 1**

There are 9 parts to this question, worth a total of 25 marks. The parts vary in difficulty; please do not panic if you find some parts particularly challenging.

Chowdhury, Gijo, and Raghavan conducted an experiment to decrease the number of defects on a printed circuit assembly encoder. The factors involve the washing and drying of the circuit after the copper has been etched, and the coating of solder to prevent the copper from oxidizing. The factors are bath temperature, wave height, overhead preheater, the temperature difference between two side preheaters, the number of air knife passes, and vibration of the solder wave. Response (labeled “Response” in the data set) was total number of defects on a unit. Load the data in the comma delimited file “circuit.csv.”

Please insert your answer between <begin answer here> and < end answer here> for each numbered question below, using as much space as needed. Include relevant R code and output only (eg not all the lines where you didn’t spell ‘circuit’ right and kept getting an error). Please use a fixed width font , such as courier for your R output to preserve the formatting, and select a font size so that the output lines do not wrap onto the next line. Please do not use a screen shot for code.

1.1 What data features can Poisson Regression accommodate that regular linear regression cannot (list at least two). Explain why poisson regression is more suitable for this data.

[3 marks]

<begin answer here>

**Poisson Regression accommodated features that arise in count data: the data are non-negative, and only intergers can be observed. It also accommodates changes in the variance as the mean changes. This data is clr=early count data, although an increase of the variance with the mean is not obvious.**

<end answer here>

1.2 Fit an initial Poisson Regression, using just the main effects and treating all the variables other than temperature difference as factors. Show your R code and output.

[2 marks]

<begin answer here>

<end answer here>

1.3 Do you think an interaction between the vibration of the solder wave and the wave height would be a useful addition to this model? Give both a relevant plot and hypothesis test, and comment.

[3 marks]

<begin answer here>

<end answer here>

1.4 For Wave Height, OverheadPreHeat, AirKnife and VibSolWave decide whether they should be treated as factors or continuous predictors, using R output to justify you choices. Use each variable in the mode you have selected for the remainder of this question.

[3 marks]

<begin answer here>

<end answer here>

1.5 The model treating BathTemp as a continuous predictor is equivalent to the one treating it as a factor. Explain when this happens; then explain why it happens, using geometric considerations.

[3 marks]

<begin answer here>

<end answer here>

1.6 Use “step” for an initial model selection. Give your code for the step function, the summary for the mode selected by that function. Are there any other variables you think should be included or excluded? Explain why. Give your final model.

[4 marks]

<begin answer here>

<end answer here>

1.7 Make the 2x2 array of diagnostic plots for the model selected in (4). Include them here, and comment on whether anything needs further investigation.

[2 marks]

<begin answer here>

<end answer here>

1.8 Explain the difference between the linear regression residual and the standardized Pearson’s residual, and give this computation for point 13 in the dataset.

[3 marks]

<begin answer here>

<end answer here>

1.9 Suggest the optimal settings for each relevant variable, based on your final model. Restrict yourself to the range of each variable used in the experiment.

[2 marks]

<begin answer here>

<end answer here>