Paige Carr

Correlation, Regression, Data Analysis

Due: November 20, 2009

A.

1. According to the correlation coefficient for the relation between the rank and population,

-0.702597025, suggest a pretty negative linear relation. However, the residual plot shows a clear curve, a pattern, which indicates that a linear model is not an appropriate representation.

2. The correlation coefficient r= -0.954391826 indicates a strong, negative linear relation.

However, like in the relation between x and y, the residual plots for the relation between ln(y)

and x has a clear pattern so a exponential model is not appropriate.

3. For the relation between ln(y) and ln(x) r=-0.984565661, which is an indication of an even stronger linear relation, however the residuals show don’t show a pattern, therefore a power function is an appropriate model.

4. When you repeat one through three, removing the first eight terms, a linear regression model is appropriate and the residuals for x and y no longer have a pattern. There is no “lack of fit”.

5. With the first eight populations, rank and population follow a power law trend, but when you remove the first eight populations, the model fits a linear regression.

B.

1. The statistic to measure uniformity would be standard deviation.
2. It makes sense to exclude the site 13 from each wafer because it changes the SD of each dramatically. If you compare the SD of the wafers with all site and the SD of the wafers without site 13 there is a significant difference. The SDs without site 13 are significantly lower than the SD with site 13. For example, for wafer A17 has an SD of approximately 85. However, excluding site 13, wafer A17 has an SD of approximately 48, nearly half of the first standard deviate with site 13.

4 . The residuals of the data from (3) is a good for the data because a pattern would indicate a “lack of fit”, but these residuals have no apparent pattern. Therefore, it is a good fit.

5. The signs coefficients indicate whether the x variables are increasing or decreasing. A negative sign means it is decreasing and a positive sign means it is increasing. The oxide thickness is decreasing by approximately -0.072x and the deposition time is increasing by approximately 1.21x.