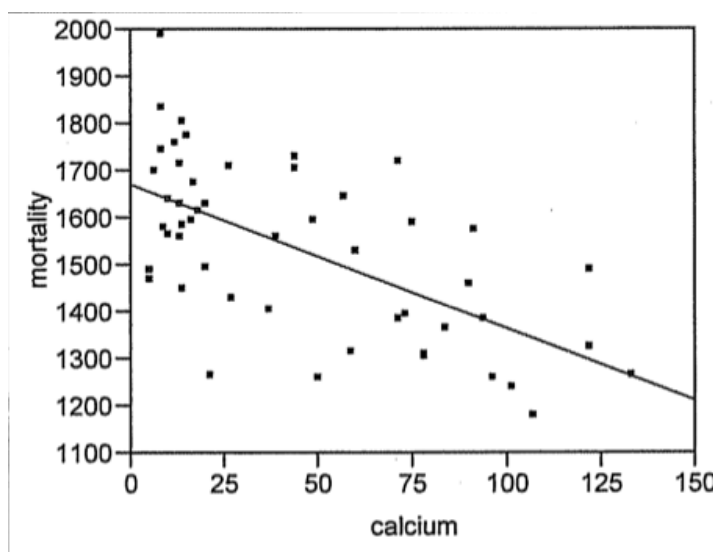


Math 107 Final Exam Practice Problems

Prof. Adam Loy

June 1, 2016

1. In an investigation of environmental causes of disease, data were collected on the annual mortality rate (deaths per 100,000) for males in large towns in England and Wales. One environmental factor that was measured was the water hardness. Hard water is water that has a high mineral concentration, usually consisting of calcium, magnesium ions, and possibly other dissolved compounds. The water hardness for the towns was recorded as the calcium concentration (parts per million, ppm) in the drinking water. R was used to fit the least squares regression line for predicting annual mortality rate (deaths per 100,000) from calcium concentration (ppm). Use the below R output to answer the following questions.



```
Call: lm(formula = mortality ~ calcium, data = mortality)
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	1668.584	32.349	51.58	<.0001
calcium	-3.071	0.553	-5.56	<.0001

Residual standard error: 144.682 on 48 degrees of freedom

Multiple R-squared: 0.392, Adjusted R-squared: 0.379

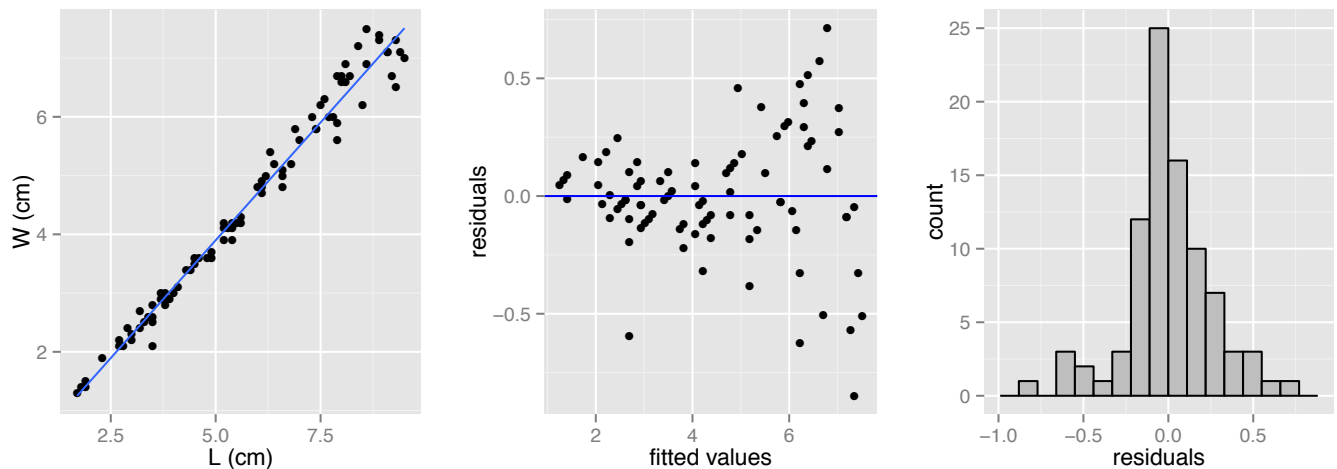
F-statistic: 30.884 on 1 and 48 DF, p-value: <.0001

- (a) (4 pts) Describe the association between calcium concentration (ppm) and annual mortality rate.

- (b) (3 pts) Report the least squares regression equation for predicting annual mortality rate from calcium concentration.
- (c) (5 pts) Interpret the slope of the least squares regression line in the context of the problem.
- (d) (4 pts) Is the slope significantly different from 0? To answer this question you may assume that all assumptions necessary for inference are valid.
- (e) (4 pts) What is the value of R^2 for this model? Give the interpretation of R^2 in the context of the problem.
- (f) (2 pts) Use the regression equation to predict the annual mortality rate for a town that has water hardness (calcium concentration) of 60 ppm.
- (g) (2 pts) Would you use the least squares regression line based on these data to predict the annual mortality rate for a town that had water hardness of 160 ppm? Briefly justify your answer.

- (h) (2 pts) Based on the results of this regression analysis, could we make the claim that hard water causes a reduction in the annual mortality rate? Explain briefly.

2. (4 pts) The website for the Quantitative Environmental Learning Project (funded by the National Science Foundation) describes data they collected on the lengths and widths of Puget Sound Butter Clams. A scatterplot of the data (with the regression line), a plot of the residuals against fitted (predicted) values, and a histogram of the residuals are provided.



Use the above plots to check each of the conditions for using a linear model with this data. Is using a linear model appropriate for these data? Explain why or why not.