

Stat 401 F/XW Homework #8:, Due Oct 26

Conceptual problems: **Chapter 8:** 2, 3, 6, 7; **Chapter 9:** 2, 4, 8

Computational problems:

1) Chapter 8: 16/16, ([Meat data thru 24 hr.](#)) Answer the book's questions.

Note: remember to use $\log(\text{time})$ as the X variable in the regression.

2) The data in [diversity.txt](#) are described in Chapter 8, problem 22/22. Ecological theory suggests that the appropriate regression model is $\text{diversity} = B_0 + B_1 * \log(\text{area})$.

These data come from an experimental study where the area of the patch was randomly assigned to plots of land.

a) Estimate the regression coefficients, then test whether the slope is significantly different from 0. Report the estimates and the test statistic and p-value for the slope test.

b) Write a one sentence interpretation of the effect of area on diversity that includes a number quantifying that effect.

c) Use the ANOVA lack of fit test to evaluate lack of fit of the regression model.

d) The investigators are interested in using this model to predict the number of species on new patches of forest. What is the area of a patch that has the smallest standard error for predictions of the mean number of species? What is the area of a patch that has the smallest standard error for predictions of number of species in an individual patch?

Hint: look at the formulae for se's of predictions.

e) The investigators would really like to make predictions of diversity in an individual patch that have standard errors less than 20 species. Is that possible with this model and data?

f) The study that produced these data has attracted a lot of international attention. Imagine that this study can be repeated with 10 times as many plots (160 instead of the 16 here). All else about the system remains the same: the regression intercept, slope, and error variance are the same as here. Will this study be able to predict diversity in an individual patch with a standard error less than 20 species? Briefly explain why or why not.

3) Chapter 9:14/14 ([Pace of life data](#)) Answer the book's questions (a-d).

New part e) Write a one sentence interpretation/description of the association between walk and heart based on your results from the multiple regression including bank walk and talk as predictor variables.

Hint: The following SAS code will draw a scatterplot matrix for question 3a. This assumes you named the variables bank walk talk and heart as in the book. Because heart is the last variable, the bottom row of plots has heart on the Y axis, as requested.

```
proc sgscatter;  
  matrix bank walk talk heart;  
run;
```

