ES/STT 7140: Homework 3

Brandon M. Greenwell

February 18, 2018

Question 1 (autocorrelation) The data set `co2.csv` contains monthly mean atmospheric CO_2 concentrations measured at Mauna Loa, Hawaii, from January 1959 to December 2003. Atmospheric CO_2 concentrations show a distinct seasonal pattern, reflecting the annual cycle of plant activities. The data set has four columns: `CO2` (monthly CO_2 concentrations in ppm), `Month` (calendar month), `Year`, and `Months` (months since January 1959). The data set can be loaded using the following lines of code:

```
url <- paste0("https://raw.githubusercontent.com/bgreenwell/",
              "stt7140-env/master/data/co2.csv")
co2 <- read.csv(url, header = TRUE)
```

a. Using the following code to construct a scatterplot of `CO2` versus `Months` and describe the relationship using one–two complete sentences. **Note:** specifying `type = "l"` in the call to `plot()` requests lines instead of points.

```
plot(CO2 ~ Months, data = co2, type = "l")
```

b. Fit a simple linear regression using `CO2` as the response variable and `Months` as the predictor variable. Quantify the temporal trend (i.e., monthly or annual rate of increase in CO_2 concentration) and discuss any potential problems with the model (**Hint:** look at the residuals!). It will be helpful to construct a scatterplot of the residuals and specify `type = "l"` in the call to `plot()`—what assumptions of the simple linear regression model appear to be violated?

Question 2 (transformations) Mercury contamination of edible freshwater fish poses a direct threat to our health. Large mouth bass were studied in 53 Florida lakes to examine the factors that influence the level of mercury contamination. Water samples were collected from the surface of the middle of each lake in August 1990 and then again in March 1991. The pH level, the amount of chlorophyll, calcium, and alkalinity were measured in each sample. The average of the August and March values were used in the analysis Next, a sample of fish was taken from each lake with sample sizes ranging from 4–44 fish, each used to measure a mercury concentration and the average concentration is reported. The authors of the study [Lange et. al., 1993] indicated that alkalinity is the best predictor of the average mercury concentration and a linear model was suggested. The data can be loaded using the following lines of code:

```
url <- paste0("https://raw.githubusercontent.com/bgreenwell/",
              "eesR/master/R/Data/HgBass.txt")
hg_bass <- read.table(url, header = TRUE)
```

a. Discuss an appropriate sampling plan that could have been used to sample the fish from each lake (be sure to explain why you think this sampling plan is appropriate).

b. Fit a simple linear regression model using the average mercury concentration (`Avg.Mercury`) as the response and alkalinity (`Alkalinity`) as the predictor. Discuss the model fit and interpret the estimated slope parameter and R^2 value in the context of the problem. Be sure to discuss any potential problems with the model (**Hint:** look at the residual plots!). The following lines of code may be useful:

```
# PLOT YOUR DATA FIRST!
plot(Avg.Mercury ~ Alkalinity, data = hg_bass)

# Fit a simple linear regression model
mod1 <- lm(Avg.Mercury ~ Alkalinity, data = hg_bass)
summary(mod1)  # print model summary
```

```
# Residual plots
par(mfrow = c(1, 2))
plot(mod1, which = 1L:2L)
```

c. Use a log-transformation on one or both of the variables to see if the model in (b) can be improved. (You should try all three and select the model you think fits best.) Discuss the "best" fitting model (including why you think it's the best). Interpret the slope of your chosen model and provide a 95% confidence interval for it (**Hint:** use the `confint()` function). Be sure to interpret the confidence interval. The following lines of code may be useful:

```
mod2 <- lm(log(Avg.Mercury) ~ Alkalinity, data = hg_bass)
mod3 <- lm(Avg.Mercury ~ log(Alkalinity), data = hg_bass)
mod4 <- lm(log(Avg.Mercury) ~ log(Alkalinity), data = hg_bass)
```