Homework 8

Statistics 411: Fall 2017

Due: In class Monday, March 26

As usual, provide all R-code in an Appendix. You do not need to format this HW as a *Statistical Report*. You must work in a group of 2-3 students and submit a single solution set. You must form a group that is different than the group with whom you worked on the previous HWs.

Consider exercises 26 and 27 on pages 233-234 re: lifespans of mammals. Don't worry about the questions that your book poses, instead answer the questions below. You may want to refer to Lab 9 (solutions are on the course web site). To get the data:

```
library(Sleuth3)
d = ex0826
```

1. Plot lifespans vs metabolic rate using a scatterplot then fit a SLR. Give the equation of the line using proper notation: $\hat{\mu}\{Y|X\} = ...$ Now use the `abline()` function to add the fit line to the scatterplot.
2. Diagnose the fit of the line using `diagANOVA()` from the course website, `source("http://www.math.montana.edu/parker/courses/STAT411/diagANOVA.r")`. You will see that there is a problem with model fit. Which assumption is violated?
3. As in Lab 4, try to log-transform metabolic rate (x) and/or log-transform lifespan (y) then fit an SLR (i.e., consider 3 models: y vs $\log_{10}(x)$; $\log_{10}(y)$ vs x and $\log_{10}(y)$ vs $\log_{10}(x)$). Which model best satisfies the regression assumptions? Justify your conclusion using the appropriate plots.
4. Write out the final fit SLR equation using proper notation that incorporates whatever transform that you decided upon, which will be one of: $\hat{\mu}\{\log_{10}(Y)|X\} = ...$ or $\hat{\mu}\{Y|\log_{10}(X)\} = ...$ or $\hat{\mu}\{\log_{10}(Y)|\log_{10}(X)\} = ...$ depending on your answer to #3.
5. Report the correlation and R^2 for your final model.
6. Plot the transformed data in a scatterplot, and use the `abline()` function to add the fit line to the scatterplot.
7. Back-transform the response to re-write the regression equation using proper notation in terms of the <u>median</u> response: $\widehat{Median}\{Y|...\} = ...$ *Hint:* See #3 in Lab 9.
8. Test whether there is an association between metabolic rate and lifespan. Report on all 6 steps of the hypothesis test including the parameter of interest (either β_0 or β_1), a conclusion that includes the test statistic and p-value and a 95% CI for the parameter <u>in terms of the problem</u>.
9. What is the *Scope of Inference*?
10. Use the "shift trick" to construct an individual 95% CI for the <u>median</u> lifespan of a mammal with a metabolic rate of 100 (i.e., $\log_{10} x = 2$). Interpret the CI in terms of the problem.
11. Construct a band of Workman-Hotelling CIs that maintain a family-wise confidence level of 95%. Add these Workman-Hotelling CIs and to the scatterplot in #6.
12. Why is the Workman-Hotelling 95% CI for the median lifespan of mammals with a metabolic rate of 100 ($\log_{10} x = 2$) LARGER than the 95% CI you found in #10?
13. Construct a band of individual 95% PIs for lifespans. Add both these individual 95% PIs to the scatterplot in #6.