Statistics 512: Homework 2

Divisions 1, 2, 3, and 4: Due Monday, February 9, 2015

Division 5: Due Tuesday, February 10, 2015

A reminder – Please do not hand in any unlabeled or unedited SAS output. Include in your write-up only those results that are necessary to present a complete solution. In particular, questions must be answered in order (including graphs), and all graphs must be fully labeled (main title should include question number, and all axes should be labeled). Don't forget to put all necessary information (see course policies) on the first page. Include the SAS input for all questions at the very end of your homework. You will often be asked to continue problems on successive homework assignments. So save all your SAS code.

Important Note *– Every graph or plot you create should have your name printed as a subtitle. Consequently, any graph with no name will result in a* **20% points off** *on that question. Also, please attach your code at the end; any homework with no code provided will result in a* **50% points off** *on the entire assignment, NO EXCEPTIONS.*

1. KNNL Problem 2.17, page 92 (white text) or page 98 (blue text).
2. KNNL Problem 2.22, page 92 (white text) or page 98 (blue text).
3. Given that $R^2 = SSM/SST$, it can be shown that $R^2/(1-R^2) = SSM/SSE$. If you have $n = 22$ cases and $R^2 = 0.4$, what is the F-statistic for the test that the slope is equal to zero?
4. Calculate power for the slope using the results of text Problem 1.22 as follows. Assume $n = 16$, $\sigma^2 = MSE$, and $SS_X = 1280$. (Note: this last value could be obtained with SAS using

   ```
   proc univariate data = (dataset name);
   var time;
   ```

   and looking at the output titled "`Corrected SS`" in the `Moments` section.)

   (a) Find the power for rejecting the null hypothesis that the regression slope is zero using an $\alpha = 0.02$ significance test when the alternative is $\beta_1 = 0.6$.

   (b) Plot the power as a function of β_1 for values of β_1 between -2.8 and $+2.8$ in increments of 0.2.

The next 5 problems continue the analysis of the plastic hardness data begun in the first homework.

5. Plot the data using `proc gplot`. Include a smoothed function on the plot by using the `i = smnn` option on the `symbol1` statement, where `nn` is a number between 1 and 99. Please use the number 70. Is the relationship approximately linear?
6. Plot the 94% bounds (confidence band) for the mean (use `i=rlclm` on the `symbol1` statement).
7. Plot the 94% bounds for individual observations (using `i=rlcli`).

8. Give an estimate of the *mean* hardness that you would expect after 36 and 43 hours; and a 94% confidence interval for each estimate. Which confidence interval is wider and why is it wider?

9. Give a prediction for the hardness that you would expect for an *individual* piece of plastic after 43 hours; give a 94% prediction interval for this quantity.