Spring 2016 Test

You may use SAS help at `sas.support.com` as well as **R** online help; either package is appropriate for any question. You may also review your PROC IMPORT statements for tab-delimited text files with headers. Both data sets used in this exam (Fall2008.txt and Tuna.txt) are available in Blackboard and on the course website as tab-delimited text files.

1. Download the 2008 GPA data in file Fall2008.txt. Create a subset of the data for which Registration Status (regstat) is either 'Continuing', 'Readmit' or 'New Transfer' (the other registration status categories tend not to list the response variable for this exercise: GPA (cltotgpa)).

   (a) Plot GPA against Math SAT (SATM) and overlay a loess fit. Does the relationship appear linear?

   (b) Use PROC TRANSREG to test whether GPA should be transformed and use any appropriate transformation for the analysis that follows.

   (c) Plot (transformed) GPA on Math SAT using Registration Status as a grouping variable and overly regression lines (this is easiest in PROC SGSCATTER). Does a parallel slopes (additive) model appear as though it might be appropriate? Comment.

   (d) Fit an ANCOVA model with nonparallel slopes for the regression of (transformed) GPA on Math SAT and Class and test whether the interaction term can be dropped. Include the /s option in your model statement in order to generate a table of Parameter Estimates.

   (e) In the plot from part (c), it appears as though New Transfer and Readmit have similar slopes. Conduct a formal test whether categories New Transfer and Readmit have the same slope. Can you find confirmation of the results of your test by referring to the table of Parameter Estimates from part (d)?

2. The data set in Tuna.txt includes pounds of yellowfin tuna caught while trolling (a form of fishing in which a boat moving at a near-constant speed pulls lures or baits behind it) off the South Carolina coast. Factors include Bait (Artificial, Live Bait) and Trolling Speed in RPM (1000, 1500, 2000).

   (a) Is this a balanced two-factor study?

   (b) Construct an interaction plot and comment.

   (c) Fit a model with an interaction term. If interaction is present, construct appropriate Bonferroni comparisons ($\alpha = 0.05$) for each of the main effects while holding the level of the other factor constant. Include graphics.

   (d) Write a formula for $E(\bar{Y}_{1..})$ in terms of μ_{ij} and n_{ij}, assuming i indexes RPM (Level 1=1000 RPM) and j indexes Bait (Level 1=Artificial). Be sure to plug in the actual values of the n_{ij} to help with interpretation.

   (e) Is $E(\bar{Y}_{1..})$ estimated by the MEANS statement or the LSMEANS statement? Report the output for RPM for both MEANS and LSMEANS and discuss any differences (Hint: look at MEANS for RPM*Bait).