Statistics 601, Spring 2018

Assignment 1

On the course web page in the folder Handouts and Extras is a copy of a 1935 paper by C.J. Bliss that is one of the classic papers on the analysis of quantal response data. This paper contains the so-called 'Bliss Beetle Data' – see Table IV in the paper.

Using data from both acute toxicity experiments (called 'Series' in the paper), conduct an analysis using a generalized linear model having a binomial random component and using Pregibon's one-parameter link as given in expression (1.20) in the course notes. Use log dose as the covariate – but note that the column labeled "log of dosage" in Table IV in the paper gives log base 10. So, use the usual logarithm applied to the column labeled "CS_2 mg per litre". Also note that in the course notes Table 1.1 presents these data *summed across* the two individual experiments. You should use the individual values and conduct an analysis using the separate Series as two data sets. As part of your analysis:

1. Test the model with estimated link against the simpler model with a fixed logit link. Do this separately for each of the 'Series' for which data are given in the paper. Provide details about the procedure you used to accomplish this. Indicate how you determined starting values for glms that you fit (with either logit or estimated links).

2. Determine whether there are differences between the two Series. That is, would you prefer two different models, one for each Series, or just one model that applies to both? Use whatever statistical procedure you believe is appropriate to make this determination. Provide details about the procedure you select.

3. Present your final model along with estimated response curve(s). Also plot

the estimated tolerance distribution(s). Provide relevant details about how this was accomplished (if not already provided).

4. Assess your model fits for whatever final model you have arrived at (some combination of one versus two groups and logit versus estimated link). For example, if you determine there is no difference between the two Series and a model having a logit link is your preferred model, then assess a basic glm with binomial random component and logit link fit to the combined data set. Provide details of the procedure you choose to accomplish this assessment.

5. Provide a sentence or two that indicates how you accomplished computation for this assignment, such as writing your own code, writing code to expand on functions provided in Stat 520, searching the web for an R package that someone wrote to accomplish this, etc.

IMPORTANT NOTE: In Series II, both of the two highest doses produced complete mortality. Drop the highest dose from these data (Series II only) so that there are only 7 dose levels (and 7 responses).