# Statistical Approach

The survey data is analyzed in the R language using the R survey package written by Thomas Lumley and described in his book *Complex Surveys: A Guide to Analysis in R* (Wiley, 2010.) Professor Lumley was kind enough to make some small enhancements to his package for our benefit. They have since been published for all in the CRAN repository of R packages.

A specific focus of effort was on the best possible computation of confidence intervals. Several methods have solid theoretical underpinnings, and Professor Lumley’s package implements five. Of these, the one he named “xlogit”, which closely mirrors the method used by STATA and SUDAAN, was chosen. Degrees of freedom were adjusted to correspond to the study as a whole rather than to specific questions, again mirroring STATA. As a result, the confidence intervals align precisely with those produced by existing software on the YRBSS site, insuring that there will be no unpleasant surprises for researchers using the two platforms together.