

Rank Tests with Data from a Complex Survey

Alastair Scott* and Thomas Lumley

Department of Statistics, University of Auckland

e-mail: a.scott@auckland.ac.nz, t.lumley@auckland.ac.nz

ABSTRACT.

Data from complex multistage surveys such as NHANES are increasingly important in the health and social sciences, and researchers would like analyze such data using the same familiar statistical techniques that they would use with independent observations. This is now possible for many standard techniques but one significant gap is the lack of rank tests such as the Wilcoxon test. Except in a few special cases, design-based analogues of such tests have not been developed, and many researchers are using the standard versions designed for independent data instead. We show how to construct general rank tests under complex sampling, both for comparing groups within a survey and for using a national survey as a reference distribution, and look at their performance in some examples from the literature where ordinary rank tests have been applied naively.

Key Words: Wilcoxon tests; multistage sampling, survey weights.