Hypotheses Testing from Complex Survey Data Using Bootstrap Weights: A Unified Approach

Zhonglei Wang

Xiamen University
E-mail: wangzl@xmu.edu.cn

Abstract: Standard statistical methods that do not take proper account of the complexity of survey design can lead to erroneous inferences when applied to survey data due to unequal selection probabilities, clustering, and other design features. In particular, the actual type I error rates of tests of hypotheses using standard methods can be much bigger than the nominal significance level. Methods that take account of survey design features in testing hypotheses have been proposed, including Wald tests and quasi-score tests that involve the estimated covariance matrices of parameter estimates. In this paper, we present a unified approach to hypothesis testing that does not require computing the covariance matrices by constructing bootstrap approximations to weighted likelihood ratio statistics and weighted quasi-score statistics and establish the asymptotic validity of the proposed bootstrap tests. In addition, we also consider hypothesis testing from categorical data and present a bootstrap procedure for testing simple goodness of fit and independence in a two-way table. In the simulation studies, the type I error rates of the proposed approach are much closer to their nominal level compared with the naive likelihood-ratio test and quasi-score test. An application to data from an educational survey under a logistic regression model is also presented.