ADJUSTMENTS TO THE PEARSON AND MANTEL-HAENSZEL CHI-SQUARED TESTS FOR DATA FROM COMPLEX SURVEYS

BY

D. R. Weerasekera

Department of Statistics & Computer Science

Thesis submitted for the degree of

Doctor of Philosophy of the University of Colombo, Sri Lanka

FEBRUARY 1990

ABSTRACT

The usual chi-squared testing procedure to test independence between variables in two-way contingency tables assumes that the observations are independent or come from simple random sampling. This assumption is invalid for data from complex survey designs, and hence some adjustments to this test have been developed in the recent past.

In the first part of the thesis, after a review of the theory developed to obtain the modification factors for the Pearson chi-squared test, further investigations on the behaviour of the modification factor obtained by Bedrick (1983) are carried out in various situations which arise in practice, such as the collapsing categories or the analysis of data within subclasses. The computation of this modification factor needs the values of the design effects of the cell proportions and marginal proportions, but the design effects of the cell proportions are not easy to obtain. Our results show that an approximate adjustment based on the square root of the smaller of the two marginal average design effects gives satisfactory results in practice.

In the second part of the thesis, two different adjustments to the Mantel-Haenszel chi-squared test are developed for complex survey data.

i) A modification of the effective sample size for each row of each table, using the design effect. This is an extension of the method proposed by Donald & Donner (1987) for a special type of single-stage sample design.

ii) As the denominator of the Mantel-Haenszel statistic is the variance of the square root of the numerator, the adjusted statistic is obtained by calculating this variance for a complex survey design by a Taylor-series expansion using the method of Woodruff (1971).

The above two adjusted statistics are evaluated and compared using both simulated and real data from actual complex surveys. The results show that the unadjusted Mantel-Hanenszel test can give very misleading results, but the two adjusted tests perform equally well.

We also discuss the suitability of using a possible approximate adjustment based only on the design effects of the prevalences of each row of each table.

Finally, Woodruff's method is proposed to adjust the odds ratio variance estimator for complex survey data.