

ERROR ESTIMATION OF SURVEY DATA

W.M. Liyanage

Dept. of Statistics & Computer Science, University of Colombo.

Assumption of simple random sampling is made in the analysis of survey data, specially because of the unavailability of software which take into account the complexity of the sampling scheme. However if the sample is not self-weighting the error estimation under the above assumption is incorrect, although the estimation of simple descriptive statistics, viz., means, proportions and aggregates, is valid.

The effect of the complexity of a particular design of sampling on the precision of the estimates derived from its data, is defined as design effect or 'deff' (Kish 1963). This can be expressed in terms of rate of homogeneity (roh), which measures the degree of similarity of variables within clusters formed by the design in a complex survey.

The rohs and deffs for descriptive statistics for 31 variables of three surveys carried out under the Expanded Programme of Immunisation (EPI) were studied.

The roh and deff for health variables are higher than those of immunisation variables as immunisation is more homogeneous than health status over the population. In order to carry out error estimation of EPI surveys the roh values of .3 and .2 can be used for health and immunisation variables respectively.