SUBSAMPLE AND HALF-SAMPLE METHODS

GUTTI JOGESH BABU

Department of Statistics, 319 Classroom Building, The Pennsylvania State University, University Park, PA 16802, U.S.A.

(Received October 31, 1990; revised November 18, 1991)

Abstract. Hartigan's subsample and half-sample methods are both shown to be inefficient methods of estimating the sampling distributions. In the sample mean case the bootstrap is known to correct for skewness. But irrespective of the population, the estimates based on the subsample method, have skewness factor zero. This problem persists even if we take only samples of size less than or equal to half of the original sample. For linear statistics it is possible to correct this by considering estimates based on subsamples of size λn, when the sample size is n. In the sample mean case λ can be taken as $0.5(1-1/\sqrt{5})$. In spite of these negative results, the half-sample method is useful in estimating the variance of sample quantiles. It is shown that this method gives as good an estimate as that given by the bootstrap method. A major advantage of the half-sample method is that it is shown to be robust in estimating the mean square error of estimators of parameters of a linear regression model when the errors are heterogeneous. Bootstrap is known to give inconsistent results in this case; although, it is more efficient in the case of homogeneous errors.

Key words and phrases: Half-sample method, bootstrap, variance estimation, linear models, asymptotic relative efficiency, Bahadur's representation, quantiles.