A composite Estimator for Cut-Off sampling

Hee-Jin Hwang*
The Bank of Korea, Seoul, Republic of Korea, hjhwang@bok.or.kr

Key-Il Shin
Hankuk University of Foreign Studies, Seoul, Republic of Korea, keyshin@hufs.ac.kr

The Cut-off sampling method has been widely used for a highly skewed population like a business survey by discarding a part of population, so called take-nothing stratum and taking all samples from take-all stratum. Then the estimated population total is obtained by the summation of the estimated totals of the take-all stratum, sample stratum and take-nothing stratum. In some business survey the precision of the result might be improved by excluding take-nothing stratum because of difficulty in survey and the cost.
However if the proportion of the total of take-nothing stratum in comparison with the population total is not negligible, then the population total have to be estimated by including that of take-nothing stratum, not yet surveyed. In that case, because the precision of the estimation for the take-nothing stratum could greatly affect that of the population total, it is necessary to increase the precision for the take-nothing stratum.
Several methods in order to increase the precision of the estimation for the take-nothing stratum have been suggested with auxiliary information or administrative data. Recently Hwang and Shin (2012) suggested a composite estimator which uses information of take-nothing stratum and sample stratum. In this method, only few samples in the take-nothing stratum are used for the estimate of that stratum total. Therefore even though it is not easy to survey samples in take-nothing stratum, the composite estimator should be considered to improve the precision of the estimated total.
In this paper, we suggest an estimator, a generalized version of the composite estimator, which is obtained by use of the results of take-nothing stratum and a part of sample stratum. For this, we divide sample stratum into k sub-strata. There are some stratifying methods to divide a population into k sub-population and in this paper, the well-known L-H(Lavallee-Hidiroglou) algorithm is used. Then we choose one sample sub-stratum, for example i-th sub-stratum, which is most correlated with the take-nothing stratum. Then we obtain a composite estimator combining information of the only i-th sub-stratum and the take-nothing stratum.
Small simulation studies are conducted to compare performances of the composite estimator suggested by Hwang and Shin (2012) and the composite estimator suggested in this study. To establish population data of the simulation, we use the same steps as in Lee et. al. (1995). Also for real data analysis, we use the Korea briquette consumption survey.

Key words : Take-nothing stratum, L-H algorithm, best linear unbiased predictor, ratio estimator