Inferential Approaches in Finite Populations

Chaudhuri A*

Applied Statistics Unit, Indian Statistical Institute, India

Submission: February 05, 2018; **Published:** March 12, 2018

***Corresponding author:** Arijit Chaudhuri, Applied Statistics Unit, Indian Statistical Institute, Kolkata, India; Email: arijitchaudhuri1@rediffmail.com

Keywords: Finite population; Pre-assigned probabilities; Sampling design; Linear unbiased estimator; Unsurveyed values

Opinion

Neyman developed the classical Design-based inferential theory in finite populations. A finite population containing a known number of objects, finite in number, and identifiable and tagged with labels is supposed to have a real variable defined on it with unknown values having a total which is required to be estimated. For this one may carry out a survey to ascertain these values. Instead of a complete enumeration a sample of the individuals are better chosen and surveyed especially in case the population is big enough and observation is difficult and costly. Once a sample is chosen with pre-assigned probabilities according to a sampling design a plan is to collect the survey-data and choose a suitable function thereof to be called a statistic to be employed as an estimator. Then one does not care to examine how close to or far from the population total or mean is the realized value of the statistic used as an estimator but to evaluate the utility of the statistic in terms of some of its theoretical performance characteristics, defined in terms of the sampling design called expectation, bias (the expectation minus the parameter), mean square error, variance, standard deviation, called standard error, coefficient of variation etc.

The combination of the design and the estimator, together called a strategy, is judged to be good if the bias is numerically small enough and the mean square error is also under control in magnitude. In particular an unbiased estimator is good if its variance is the minimum when based on a particular sampling design. But Basu showed that for no non-census design an unbiased estimator exists with a uniformly minimum variance. Godambe introduced for a population total a homogeneous linear unbiased estimator and showed that among all such estimators no one is available with the least variance for a general class of designs. Basu prescribed how to get the minimal sufficient statistics, given the raw survey data.. Also it is known how to construct a complete class within the class of unbiased estimators for a finite population total such that given one in the wider class there exists one in the narrower class one that has a smaller variance. Using this concept it is now established

that a very restricted class of sampling designs admits a unique estimator with the uniformly smallest variance among Godambe's homogeneous linear unbiased estimators. But this is such a restrictive result that this Neyman based classical theory is quite deficient in giving us acceptably strong and useful results. So, a revolutionary step for effective rectification is really needed.

One such effort was due to Godambe and Mary Thompson who considered the vector of variate values of the population as a random vector which has a probability distribution which is called the super-population. Since the design-based variance of an unbiased estimator for a finite population total cannot be desirably controlled, now with the super-population modeling approach the following is attempted. The probability-distribution of the population vector of variate-values is not required to be pinpointed; rather, only a class of such distributions is hit upon admitting only low order moments like expectation, variance, covariance etc. This class of distributions is called a model and such a loose specification is called modeling and one feels satisfied with controlling the model-based expectation of the design-based variance of an unbiased estimator of the population total. Based on a given sampling design if a design-based estimator for a population total has the least value for the model-based expectation of its design-based variance, then it is taken as an optimal estimator for the finite population total. Here the insistence on design-based unbiasedness of the estimator is considered an artificial restriction.

A third inferential approach as follows is considered a natural approach. A population total is expressed as the sum of the sampled variate-values and the aggregate of the unsampled and unsurveyed values. Any proposed estimator for the total is also the aggregate of the observed sample-values plus the difference between the estimator proposed minus the sampled sum. If the estimator is to come close to the population total, the estimand parameter, then the unsampled sum cannot be claimed to be close, for a surveyed sample, to the estimated value minus

the sampled summed values unless the variate-values in the population vector of values are suitably inter-related among themselves. So, this third inferential approach starts with the requirement that the population vector of variate-values as a random vector. This necessitates the population total also to be a random variable. Hence the finite population total is not a constant and hence cannot be estimated. But one may proceed to predict it by estimating the model-based expectation of the finite population total. This is known as the prediction approach and originated with the research by Brewer and Royall. In order to discriminate among the proposed predictors one chooses the criterion of minimizing the model -based expectation of the squared difference between the proposed predictor and the finite population total restricting the choice of the predictor as a linear function of the variate-values in the sample and subject to the condition that the model expectation of the linear estimator is exactly equal to the model expectation of the finite population total. Deriving such an optimal linear model-unbiased predictor is easy as in applying Gauss-Markov theory of linear estimation.

Unfortunately both the prediction approach and the super-population modeling approach are handicapped by infeasibility of solution because both involve unknowable model-based parameters. Brewer offers a solution with his asymptotic formulation in the context of finite populations. The model parameters hard to handle are replaced by some assignable constants so that the revised predictors satisfy the condition of asymptotic design unbiasedness and the limiting value of the asymptotic design expectation of the model expectation of the squared difference between the predictor and the finite population total is minimized. This provides handy solutions in both the prediction approach and the super population modeling approach. This is known as Model assisted inference approach in finite population sampling. Bayesian and empirical Bayesian inferential approaches are also now-a-days quite popular in survey sampling. To cut the story short readers are invited to peruse the author's recent monograph entitled "Modern Survey Sampling" published in 2014 by CRC Press, Taylor & Francis where one may find all the references cited in the text above.

This work is licensed under Creative Commons Attribution 4.0 License
DOI: [10.19080/BBOAJ.2018.05.555669](https://doi.org/10.19080/BBOAJ.2018.05.555669)

Your next submission with Juniper Publishers

will reach you the below assets

- Quality Editorial service
- Swift Peer Review
- Reprints availability
- E-prints Service
- Manuscript Podcast for convenient understanding
- Global attainment for your research
- Manuscript accessibility in different formats (Pdf, E-pub, Full Text, Audio)
- Unceasing customer service

Track the below URL for one-step submission

<https://juniperpublishers.com/online-submission.php>