Estimation of current population mean using two-occasion successive sampling with one auxiliary variable

R. Zoramthanga

Department of Statistics, Pachhunga University College, Aizawl 796001, India

In this study, two-occasion successive sampling for ratio-to-regression estimator was used to determine the current estimate of the population mean using only the matched part and one auxiliary variable, which is available on both the occasions. The data used were based on the total number of female workers in villages in Mizoram with the total number of literate female in villages in Mizoram as an auxiliary variables. The data were gotten from Census of India 2001 and 2011. The optimum mean square error of the combined ratio-to-regression and ratio estimator has been compared with (i) the optimum mean square error of the chain-type ratio estimator (ii) mean per unit estimator and (iii) combined estimator when no auxiliary information is used at any occasion. This result showed that the combined ratio-to-regression and ratio estimator is more efficient than the other three existing estimators.

Key words: Ratio-to-regression estimator, two-occasion successive sampling, mean square error (MSE), relative efficiency.

Introduction

In a sample survey, it is often seen that sample surveys are not limited to one time enquiries. If the value of study character of a finite population is subject to change over time, a survey carried out on a single occasion will provide information about the characteristics of the surveyed population on the given occasion only and cannot give any information on the nature or the rate of change of the characteristics over all the occasions or more recent occasion. Data regarding changing properties of the population of cities or countries, such as unemployment statistics, are collected regularly on a sample basis, to estimate the change from one occasion to the next or to estimate the average over certain period. An important aspect of continuous survey is the structure of the sample on each occasion. To meet these requirements, successive sampling provides a strong tool for generating reliable estimates on different occasions. When the same population is sampled repeatedly the sampler is in an ideal position to make realistic estimates, both of costs and of variances and to apply the technique that leads to optimum efficiency of sampling. The successive sampling is a known technique that can be used in longitudinal survey to estimate the population parameters and measurements of difference or change of a study variable.

Successive sampling is used extensively in
Materials and Method

Data used

The data used for this study is from the records of the total number of female workers and the total number of literate female in villages in the state of Mizoram, Census of India 2001 and 2011.

Methodology

The variables $x(y)$ were defined as the total number of female workers in villages in the state of Mizoram, India in 2001 (2011) and $z$ is defined as an auxiliary variable which is the total number of literate female in villages in the
state of Mizoram, India.

Consider a population consisting of N units. Let a character under study on the first (second) occasion be denoted by \( x(y) \), respectively. It is assumed that the information on an auxiliary variable \( z \) is available on the first as well as on the second occasion. It is also assumed that the population to be large enough, and the sample size is constant on each occasion. Using simple random sampling without replacement (SRSWO) select a sample of size \( n \) on the first occasion. Of these \( n \) units, a sub-sample of size \( m = n \lambda \) is retained on the second occasion. This sub-sample is supplemented by selecting an SRSWO of \( u = (n - m) = n \mu \) units afresh from the units that were not selected on the first occasion.

**Results**

From the data used, using two occasion successive sampling a random sample of 70 villages was selected from a population of 669 villages on each occasion, this comprises of 35 matched villages and 35 unmatched villages in the state of Mizoram.

**Table 1** | Relative Efficiency (%) of \( T_{p_1} \) with respect to estimators \( T, \bar{y}_n \) and \( \bar{y}_2 \).

<table>
<thead>
<tr>
<th>Estimators</th>
<th>Estimates</th>
<th>MSE</th>
<th>Efficiency %</th>
</tr>
</thead>
<tbody>
<tr>
<td>( T_{p_1} )</td>
<td>146</td>
<td>107.49</td>
<td>100</td>
</tr>
<tr>
<td>( T )</td>
<td>139</td>
<td>344.65</td>
<td>320.62</td>
</tr>
<tr>
<td>( \bar{y}_n )</td>
<td>130</td>
<td>373.54</td>
<td>347.50</td>
</tr>
<tr>
<td>( \bar{y}_2 )</td>
<td>130</td>
<td>304.81</td>
<td>283.56</td>
</tr>
</tbody>
</table>

where

- \( T_{p_1} \) is a combined estimator proposed by Ralte and Das,\(^{17} \)
- \( T \) is a combined estimator proposed by Singh,\(^{18} \)
- \( \bar{y}_n \) is mean per unit estimator, and
- \( \bar{y}_2 \) is a combined estimator suggested by Cochran\(^{1} \) when no auxiliary information is used at any occasion.

**Conclusion**

From the table in the above section, the combined ratio-to-regression and ratio estimator i.e. \( T_{p_1} \) is more efficient than the other existing three estimators viz. \( T, \bar{y}_n \) and \( \bar{y}_2 \) with maximum gain in efficiency occurring while comparing with mean per unit estimator, which is very obvious. Hence, the estimator \( T_{p_1} \) is recommended for further practical use.

**References**


