CHAPTER 6
DETERMINATION OF SAMPLE SIZE FOR DUAL TO RATIO ESTIMATOR

6.1 Introduction

In sample surveys, sample size is generally taken as smaller than half of the population size. In such situations Srivenkataramana (1980) and Bandyopadhyay (1980) proposed a dual to ratio estimator for population mean $\bar{Y}$

$$\bar{y}_s = \bar{y} \left( \frac{N\bar{X} - n\bar{x}}{(N - n)\bar{X}} \right)$$

(6.1)

for positively correlatively variables. Under SRSWOR procedure this estimator has bias

$$B(\bar{y}_s) = -\frac{\bar{Y}}{N} C_{11}$$

(6.2)

and mean squar error

$$M(\bar{y}_s) = 0_1 \bar{Y} (C_{20} + g^2 C_{02} - 2g C_{11})$$

(6.3)

$$g = \frac{n}{(N - n)}$$

The estimator $\bar{y}_s$ is more efficient than sample mean estimator $\bar{y}$ and $\bar{y}_r$, ratio estimate if

$$K > \frac{g}{2}$$

(6.4)
\[ K < \frac{1+g}{2} \quad (6.5) \]

According to Srivenkataramana (1980) the estimator \( \bar{y}_s \) attains minimum MSE

\[ M_0(\bar{y}_s) = \theta_1 \bar{Y}^2 C_{20} (1 - \rho^2) \quad (6.6) \]

if

\[ g = K \quad (6.7) \]

Expression (6.6) is equal to MSE of regression estimator \( \bar{y}_{1r} \).

Simplifying (6.7) we have

\[ n_0 = \frac{NK}{1+K} \]

However in section 6.2, we proposed another sample size for which the estimator \( \bar{y}_s \) has smaller MSE than \( M_0(\bar{y}_s) \).

6.2 Optimum Choice Of Sample Size

Writing \( \theta_1 = \frac{1}{Ng} \), it can be seen that \( M(\bar{y}_s) \) may be expressed as

\[ M(\bar{y}_s) = \frac{1}{N} \left( \frac{C_{20}}{g} + g C_{02} - C_{11} \right) \quad (6.8) \]

Differentiating (6.8) with respect to \( n \), we find the optimum value of \( n \) as
\[ n_t = \frac{NK}{K + \rho} \quad (6.9) \]

In this case the minimum value of \( M(\bar{y}_s) \) is obtained as

\[ M_{01} (\bar{y}_s) = \frac{2}{N} \frac{\bar{Y}^2 (1 - \rho) \sqrt{C_{20} C_{02}}}{\sqrt{C_{20} C_{02}}} \quad (6.10) \]

### 6.3 Efficiency Comparision

Substituting the value of \( n = n_0 \) in (6.6), it can be seen that \( M_0 (\bar{y}_s) \) may be expressed as

\[ M_0 (\bar{y}_s) = \frac{\bar{Y}^2 (1 - \rho) \sqrt{C_{20} C_{02}}}{N \rho} \quad (6.11) \]

Comparing (6.10) and (6.11) we find that

\[ M_{01} (\bar{y}_s) \leq M_0 (\bar{y}_s) \quad (6.12) \]

while

\[ n_1 > n_0 \quad (6.13) \]

In section 6.4 efficiency of \( \bar{y}_s \) is studied for live data.
\[ n_1 = \frac{NK}{K + \rho} \]  \hspace{1cm} (6.9)

In this case the minimum value of \( M(\bar{y}_s) \) is obtained as

\[ M_{01}(\bar{y}_s) = \frac{2}{N} \bar{Y}^2 (1 - \rho) \sqrt{C_{20} C_{02}} \]  \hspace{1cm} (6.10)

### 6.3 Efficiency Comparison

Substituting the value of \( n = n_0 \) in (6.6), it can be seen that \( M_0(\bar{y}_s) \)
may be expressed as

\[ M_0(\bar{y}_s) = \frac{\bar{Y}^2 (1 - \rho) \sqrt{C_{20} C_{02}}}{N \rho} \]  \hspace{1cm} (6.11)

Comparing (6.10) and (6.11) we find that

\[ M_{01}(\bar{y}_s) \leq M_0(\bar{y}_s) \]  \hspace{1cm} (6.12)

while

\[ n_1 > n_0 \]  \hspace{1cm} (6.13)

In section 6.4 efficiency of \( \bar{y}_s \) is studied for live data.
6.4 Numerical Illustration

Let us consider the data from Tripathi and Ahmed (1995), where the population consists of 278 villages under Gajole Police Station, West Bengal.

Here

\( y = \) number agricultural labourers in 1971,

\( x = \) population size of village in 1961.

\( \bar{Y} = 30.07, \bar{X} = 339.95, C_{20} = 3.5249, \)

\( C_{02} = 1.0680, \rho = 0.7705, K = 1.3997. \)

For this data \( n_0 = 162, n_1 = 179. \) Percentage values of relative MSE \((M(.) / \bar{Y}^2)\) of various estimators for different sample sizes is given in the following table:

<table>
<thead>
<tr>
<th>Estimator</th>
<th>50</th>
<th>100</th>
<th>125</th>
<th>150</th>
<th>162</th>
<th>179</th>
<th>200</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \bar{y} )</td>
<td>5.7819</td>
<td>2.2569</td>
<td>1.5519</td>
<td>1.0819</td>
<td>0.9079</td>
<td>0.7013</td>
<td>0.4945</td>
</tr>
<tr>
<td>( \bar{y}_s )</td>
<td>4.7906</td>
<td>1.3973</td>
<td>0.7903</td>
<td>0.4567</td>
<td>0.3689</td>
<td>0.3204</td>
<td>0.4040</td>
</tr>
<tr>
<td>( \bar{y}_r )</td>
<td>2.6293</td>
<td>1.0264</td>
<td>0.7057</td>
<td>0.4920</td>
<td>0.4129</td>
<td>0.3189</td>
<td>0.2249</td>
</tr>
<tr>
<td>( \bar{y}_{lr} )</td>
<td>2.3493</td>
<td>0.9171</td>
<td>0.6036</td>
<td>0.4396</td>
<td>0.3689</td>
<td>0.2849</td>
<td>0.2009</td>
</tr>
</tbody>
</table>
From the above table it is clear that though for $n_0 = 162$, MSE of dual to ratio estimator is same as MSE of regression estimator but it is not its minimum MSE. It attains minimum at $n_1 = 179$, in which case it is less efficient than ratio and regression estimators. Beyond this as sample size increases, MSE of dual to ratio estimator also increases while it decreases for other estimators.