Sample Size of Cut-Off Method with Systematic Sampling

Lee, Sang Eun
Kyonggi University, Department of Applied and Information Statistics
Jin, Young
National Statistical Office, Statistical Information Division
Shin, Ming Woong
Hankuk University of Foreign Study, Department of Statistics

1. Introduction

Cut-off sampling method was given by Dalenius(1952) and Glasser(1962) and in Hidiroglou's paper, 1986, it showed the exact and approximation cut-off values are presented with desired level of precision rather than the sample size. That has been widely used in sampling technique , especially Establishment Survey. For practical use of cut-off design, we usually do the systematic sampling in take some stratum. However, method of getting sample size of cut-off which Hidiroglou suggested was assumed to be under simple random sampling. Here, in this study, we suggest the new formular for sample size of cut-off design with systematic sampling.

2. Sample Size of Cut-Off Sampling

Consider the population of N units;

$$y_{(1)}, y_{(2)}, ..., y_{(N)}$$

with $y_{(i)} \leq y_{(i+1)}$ for $i = 1, 2, ..., N-1$. There are t units in the take-all stratum and N-t "small" units in the take-some stratum, with

$$Y = \sum_{i=1}^{N-t} y_{(i)} + \sum_{i=N-t+1}^{N} y_{(i)}$$

Let total sample of size, $n(t)$, is to be selected from the whole population and $N(t)$ is composed of t units from take all stratum and $N(t) - t$ units from take some stratum.
Now estimator of the total Y would be

$$\hat{Y} = \frac{N-t}{N(t)-t} \sum_{i=1}^{n(t)-t} z_i + \sum_{i=N-t+1}^{N} y_{(i)}$$

where $y_{(1)} \leq z_i \leq y_{(N-t)}$ for $i = 1, 2, ..., n(t) - t$

2.1 With Simple Random Sampling(with S.R.S)

In 1986, Hidiroglou suggested the total sample size, $n(t)_s$,is as following ;

$$n(t)_s = N - \frac{(n(t)-t)c^2Y^2}{c^2Y^2 + (N-t)S^2_{[N-t]}}$$

where,
$S^2_{[N-t]} = \frac{1}{(n-t-1)} \sum_{i=1}^{N-t}(y_{(i)} - \mu_{[N-t]})$, $\mu_{[n-t]} = \frac{1}{N-t} \sum_{i=1}^{N-t} y_{(i)}$ and $Var(\hat{Y}) = c^2Y^2$

2.2 With Systematic Sampling(with Syst.S)

In this section,
using the variance of systematic sampling which is $S^2_{syst} = (N-t)^2[\frac{N-t-1}{N-t}S^2_{[N-t]} - \frac{k(n(t)-t)}{N-t}S^2_w]$, the total sample size , $n(t)_{syst}$, is as following;

$$n(t)_{syst} = \frac{z^2N^2S^2_w}{E^2Y^2 - N(N-1)z^2S^2_{[n-t]} + z^2N^2S^2_w}$$

where,
$S^2_w = \frac{1}{k(n(t)-t)} \sum_{i=1}^{n(t)-t} \sum_{j=1}^{k}(y_{ij} - \bar{y}_{i.}), k = \frac{N-t}{n(t)-t}$, z is confidential coefficient and E is allowed error$(= \frac{z^2\sigma^2_y}{\hat{y}})$.

3. Data Description and Summary

In this study, we use Mining and Manufacturing Survey, 1998 data and size of population is 312.

	with S.R.S	with Syst.S
t	97	71
n(t)-t	34	37
n(t)	131	108

From the above table, we can easily see that the total sample was reduced form 131 to 108. Interestingly, in this case, mostly reduced sample size was happened from take all stratum which from 97 to 71. We are suspecting that it will be related to value of skewness and kurtosis of data which are 5.86 and 45.02 respectively.

REFERENCES

Hidiroglou, M.A(1986). The Construction of a Self-Representing Stratum of Large Units in Survey Design. *The American Statistician* , **40** , 27-31.
National Statistical Office , Korea . Technical Report of Cut-Off Sampling Method, 1991