International Journal of Scientific Research in Mathematical and Statistical Sciences
Volume-7, Issue-3, pp.89-92, June (2020)

Research Paper

E-ISSN: 2348-4519

Auxiliary Information and Its Implication in Indirect Method of Domain Estimation

Shreya Agarwal[1*], Sarla Pareek[2], Piyush Kant Rai[3]

[1,2]Department of mathematics and statistics, Banasthali university
[3]Department of mathematics and Statistics, Banaras Hindu University

**Corresponding Author: shreya.agarwal3011@gmail.com*

Available online at: www.isroset.org

Received: 20/Apr/2020, Accepted: 05/June/2020, Online: 30/June/2020

***Abstract*—**The governmental studies regularly provide precise measurements to large domain. The issue begins when we talk about the exact insights for the sub population called "domains ". The quantity of perceptions that happen to fall in the domain is customarily arbitrary and in some cases small. These are the highlights that give domain estimation its specific flavor. To expand the effectiveness of assessments for small domains in which the example estimate is small, assistant factors from authoritative records are domain estimation in factual writing are for the most part plan and model based and assistant variable have extraordinary effect in the advancement of aberrant strategy for domain estimation under above said condition. The present paper portrays the utilization of two auxiliary variables in different proposed estimators for domains which are plan and model based. Likewise, a recreated work is done to think about the proficiency of various proposed estimators under two assistant factors as far as absolute relative bias ((ARB) and Simulated Relative Standard Error (SRSE) for various domains.

***Keywords*—**Auxiliary variables, Domain estimation, Indirect method of domain estimation

I. INTRODUCTION

A small domain is one that represents just a minor part of the entire population. In such a domain, the analyst probably do not to have many perceptions. This confusion makes "the small domain estimation issue". Small domain estimation estimation is another term frequently in this association. Based on different techniques used, domains are of three types: Planned Domain, Unplanned Domain, and Cross Classes. Based on size of population or size of tests, domains are of four types:

- Major Domain
- Minor Domain.
- Mini Domain
- Rare Domain

Because of the inadequate precision of customary assessments at lower level which are commonly deficient to give solid information. you need was felt to create elective estimators to give small territory measurements used auxiliary data.

Gonzalez and Waksberg (1973) examined the estimators which are called synthetic estimator, if a dependable direct estimator for a bigger zone, covering a few have the same characteristics as the small domain and it is utilized to infer a circuitous estimator for the small zone under the presumption that they have the same characteristics as the large domain.

Schaible, Brock, Casady and Schnack (1977) resulted that if small domain sample sizes are relatively small the synthetic estimator performs better than the simple direct estimators, whereas the sample sizes are large the direct estimators performs better. Auxiliary variables are used to increase the precision of the estimators.

Tikkiwal and Ghiya (2000) characterized a summed up class of synthetic estimators; utilizing helper data, under basic irregular testing and stratified random sampling design . Further they thought about the overall execution of process, proportion and item engineered estimators with the relative comparing direct estimators observationally through a simulation study.
Tikkiwal and Ghiya (2007) found that ratio estimator is a standout amongst the most regularly utilized estimators among others for the populacion mean or populace aggregate with the help of an auxiliary variable.
Rai and Pandey (2013) talked about the diverse part of the summed up class of synthetic estimators for small territory estimation issue when more than one auxiliary variables are is accessible.

II. UTILITY OF AUXILIARY VARIABLE IN DOMAIN ESTIMATION

In review research, there are times when data is accessible on each unit in the population. On the off chance that a variable that is known for each unit of the population is anything but a variable of intrigue however is rather utilized to improve the inspecting plan or to upgrade estimation of the factors of intrigue, it is called an auxiliary variable. Estimators which depend on auxiliary data. The term synthetic was utilized in light of the fact that these assessments were not taken directly from survey results. this term is presently utilized, all the more explicitly, to allude to this specific technique for getting data from comparative small domains so as to build the exactness of the resulting estimates.

III. PLANNED AND UNPLANNED DOMAIN STRUCTURE

Different domain structures can show up in practical uses of domain estimation. Sampling design might be founded on learning of domain choice of the population unit that full with in domain. That the testing configuration is stratified, domains being the strata, the domains are called planned. For planned domain structures, the population domains can be viewed as independent subpopulations. Along these lines, standard populace estimators are relevant in that capacity. The domain estimate in each domain is regularly accepted known and the example measured in domain test is fixed advance of time. Stratified examining in association with an appropriate allotment plan, for example, ideal or power assignment is regularly utilized in down to earth applications, so as to acquire command over domain test sizes. Depict allotment procedures to accomplish sensible exactness for small domains, as yet holding great precision for extensive domains. Propose test adjusting and coordination strategies for cases with a substantial number of various stratification structures to be tended to in domain estimation.

That the domain membership isn't consolidated into the sampling plan, the sizes d_s n of domain tests will be irregular. The domains are then called unplanned. Unplanned domain structures commonly cut crosswise over plan strata. The property of arbitrary domain test sizes presents an expansion in the fluctuation of domain estimators. Likewise, incredibly small number (even zero) of test components in a domain can be acknowledged, if the small number in the population is small.

IV. LITERATURE SURVEY

Lessler et al. (2006) in this paper, conducted experiments intended to incite nonresponse, yet the nonresponse predispositions for the appraisals in these trials were not emotional by and large. We presume that huge nonresponse inclinations are bound to happen for a set number of insights from an overview, and the measurements with generous predispositions being those that are exceedingly associated to an immediate reason for nonresponse.

Small (1991) in this paper, arrive base at a similar end. Since the genuine sources of nonresponse inclination are regularly not well-known in surveys in reviews, the examination additionally proposes that it is fitting to incorporate however many factors as could be expected under the circumstances in the adjustment loads. This equivalent kind of guidance is given for developing affinity models for making causal inferences in observational studies.

. Madow et al. (2006) in this paper, give a case of an survey estimate with large nonresponse bias that is because of the powerlessness to contact a specific gathering of tested people. This work stimulates a few thoughts that can be utilized practically speaking. The reproductions propose that present stratifying on the fullest degree conceivable is an imperative technique that may lessen nonresponse inclination over a wide array of statistics.

Montaquila et al. (2010) in this paper, Nonresponse inclination can be significant and is a major issue in review look into, including official insights. Adjustment weighting techniques can decrease nonresponse predisposition, however including at least one helper variable that is identified with differential reaction affinities does not ensure the inclinations for all measurements will be eliminated.

Scheuren (2011) et al. in this paper, Biases in evaluations for the full populace might be influenced uniquely in contrast to those for domains when the reaction affinities change by the domain. Assessed sums are most defenseless to this sort of predisposition when contrasted with methods and medians. The logical and recreation results for proportions propose that measurements that are elements of two factors may be increasingly touchy to nonresponse inclination. For proportions, the covariance between the variable and the reaction affinities must be zero for both the numerator and denominator factors to acquire unbiased estimates.

Rizzo et al.(2009) in this paper, Choices in which auxiliary factors are incorporated and how they are utilized in the alignment influences the nonresponse inclination in the appraisals. The recreations affirmed that the specialized decision of the alignment strategy (e.g., raking or straight adjustment) isn't vital, in any event for the measurements analyzed. In the event that the factors that impact reaction penchants are between related, at that point incorporating these connections in the adjustment plot is fundamental to lessening nonresponse inclination when all is said in done.

V. NOTATION AND TERMINOLOGY

Assume that a finite population U = (1, 2... I... N) is partitioned into 'A' non overlapping small domains Ua of given size N_a (a = 1, 2, ..., A) for which estimates are required. We signify the characteristic under study by 'y'. We further accept that the auxiliary information is accessible and denote this by 'x'. An random sample s of size n is chosen through straightforward arbitrary inspecting plan from population U with the end goal that

N_a units in the example 's' originates from small domain U_a (a = 1, 2, ..., A). Therefore,

$$\sum_{a=1}^{A} N_a = N \text{ and } \sum_{a=1}^{A} n_a = n$$

we consider the instance of summed up synthetic estimator for estimating the populace mean Y a for domain 'a' under two auxiliary variables x_1 and x_2;

$$y_{syn,a}=W_1 y\left[\left(\frac{X_1}{X_{1,a}}\right)\beta_1 + W_2 y\left(\frac{X_2}{X_{2a}}\right)\beta_2\right] \quad \ldots\ldots\ldots\ldots\ldots(1)$$

Here W_1 and W_2 are the weights with the end goal that $W_1 + W_2 = 1$ and β_1, β_2 are reasonably picked constants. The structure predisposition and MSE of $Y_{syn,}$ a is given by

$$B\,(y_{syn,a})W_1 Y\left(\frac{X_1}{X_{1a}}\right)d1(1+\frac{f}{n}\left(\frac{\beta_1(\beta_1-1)}{2!}\,C_2^1+\beta_1 C_{01}\right)$$

$$W_2 Y\left(\frac{X_2}{X_{2a}}\right)d2(1+\frac{f}{n}\left(\frac{\beta_2(\beta_2-1)}{2!}\,C_2^2+\beta_2 C_{02}\right)+Y_{a\ldots(2)}$$

$$\text{MSE}=(y_{syn,a})=y^2\left[W_1\left(\frac{X_1}{X_{1,a}}\right)d1+W_2\left(\frac{X_2}{X_{2a}}\right)d2\right]^2$$

$$+y^2 W_1^2\left(\frac{X_1}{X1a}\right)2d1[\frac{f}{n}\{C_0^2+\beta_1(2\beta_1 C_1^2-C_1^2+4Co_1)\}]$$
$$+y^2 W_1^2\left(\frac{X_1}{X1a}\right)2d2[\frac{f}{n}\{C_0^2+\beta_2(2\beta_2 C_2^2-C_2^2+4Co_2)\}]$$
$$+2W_1W_2y^2\left(\frac{X_a}{X_{1a}}\right)d1\left(\frac{X_2}{X_{2a}}\right)(\frac{f}{n}(C_0^2+\beta_1(2C_{01}+\frac{(\beta_1-1)}{2!}C_1^2$$
$$+\beta_2\left(2C_{o2}+\frac{(\beta_2-1)}{2!}C_2^2\right)+\beta_1\beta_2C_{12}))$$
$$-2y_a[W_1Y\left(\frac{X_1}{X_2}\right)d1(1+\frac{f}{n}(\frac{\beta_1(\beta_1-1)}{2!}\,C_1^2+\beta_1C_{01})$$

$$W_2 Y\left(\frac{X_2}{X_{2a}}\right)d2(1+\frac{f}{n}\left(\frac{\beta_2(\beta_2-1)}{2!}\,C_2^2+\beta_2 C_{02}\right)+Y^2_{a\ldots\ldots(3)}$$

VI. ESTIMATORS UNDER STUDY

The following estimators are considered in our study by putting the distinctive estimations of β_1 and β_2 in (1) that we put $\beta_1 = \beta_2 = 0$, we get simple synthetic estimator

$$t_{1,a} = y_{syn,a} = y$$

- if we put $\beta_1 = \beta_2 = -1$, we get ratio synthetic estimator

$$t_{2,a} = y_{syn,r,a} = W_1\left(\frac{y}{x_1}\right)x_{1a} + W_2(\frac{y}{x_2})X_{2a}$$

- if we put $\beta1 = \beta2 = +1$, we get product synthetic estimator

$$t_{3,a} = y_{syn.p.a} = W_{1y}\left(\frac{X_1}{X_{1a}}\right) + W_{2y}(\frac{x_2}{X_{2a}})$$

We can get the loss of predisposition and MSE for the above examined estimators with the assistance of condition (2), (3). Since, the predisposition and MSE are not in diagnostic structure for the proposed estimators in this way we attempt to contrast the result of these estimators and the help of simulation.

VII. SIMULATION STUDY

Simulation study is done utilizing simple random sampling without replacement (SRSWOR) method. Here we have three domains of sizes 300, 400, 300 separately with there various methods and the population is considered as the blend of these three domains. Likewise we have drawn 500 free basic irregular examples for every size of 40, 70 and 100 from the created population of size 1000 utilizing the product R and to survey the general execution of the estimators under thought, their absolute relative bias (ARB) and simulated relative standard error (SRSE) are determined as follows:

$$\text{ARB}=\frac{\frac{1}{500}\sum_{s=1}^{500} t_{k,a}^s - Y_a}{Y_a}\times 100$$

$$\text{SRSE}=\sqrt{\frac{ASE(t_{k,a})}{E(t_{k,a})}}\times 100$$

Where

$$\text{ASE}\,(t_{k,a})=\frac{1}{500}\sum_{s=1}^{500}\,(t_{k,a}^s - Y_a)^2$$

For k=1,..,3 and a= 1,..,3

VIII. RESULTS AND DISCUSSION

Table 1: ARB under different estimators

Sample Size	Domain	t_1	t_2	t_3
	I	29.24807	0.101162	67.52363
40	II	72.04588	0.123856	196.7687
	III	32.59712	0.037357	54.37675
	I	29.19275	0.071982	67.15932
70	II	71.97224	0.09467	196.1233
	III	32.62597	0.066496	54.47596
	I	29.22957	0.070035	67.09577
100	II	72.02125	0.092722	196.0107
	III	32.60677	0.06844	54.49327

Table 2: SRSE under different estimators

Sample Size	Domain	t_1	t_2	t_3
	I	23.46174	0.337435	42.16354
40	II	42.33155	0.344895	89.76231
	III	48.75663	0.825927	62.57644
	I	23.08615	0.258203	41.27128
70	II	42.11759	0.265407	89.03217
	III	48.65565	0.654289	62.68556
	I	22.89577	0.210402	40.77306
100	II	42.01818	0.218981	88.62593
	III	48.51317	0.534912	62.64411

the notable points for the three proposed estimators t_1,t_2 and t_3 which are simple, ratio and product synthetic estimator separately are the following: he ARB estimate for t_1 changes from 29.22 to 72.04, for t_2 it fluctuates from 0.03 to 0.12 while for t_3 estimator it is 54.73 to 196.12 for the three distinct domains I, II, III and test sizes 40, 70 and 100. ARB value for t_2 (ratio synthetic estimator r) is very lower than other two estimators, additionally the SRSE value for t_1,t_2 and t_3 shifts from (29.89, 48.75), (0.21, 0.82) and (40.77, 89.76) individually for the domain I, II, and III and test sizes 40,70 and 100. In this manner, according to SRSE is concerned it is additionally small for t_2 ratio synthetic estimator in contrast with other two estimators under study.

IX. CONCLUSION

In this paper, a small domain is one that represents just a minor part of the entire population. In such a domain, the analyst is probably going to have not many perceptions. This difficulty makes "the small domain estimation issue". Small domain estimation is another term regularly heard in this association. It is found that ratio synthetic estimates outperform the simple and product synthetic estimates.

REFERENCES

[1] Rai, P.K. and Pandey, K.K.(2013). Synthetic estimators using Auxiliary information in small domains. Statistics in Transition-new series, spring 2013. Vol. 14, No. 1, pp. 31 - 44.

[2] Pandey, K.K. and Tikkiwal, G.c.(2010). Generalized class of synthetic estimators for small domains under synthetiv sampling scheme.Statistics in Transition-new series, july 2010. Vol. 11, No. 1, pp. 75 – 89

[3] Tikkiwal,G.C.and pandey, K.K.(2007). On synthetic and composite stimators for small domain estimation under lahiri-midzuno sampling scheme.Statistics in Transition-new series, april 2007. Vol. 8, No. 1, pp. 111 – 123

[4] Tikkiwal,G.C.and Ghiya, A.(2000). A generalized class of synthetic estimators with application to crop acreage estimation for small domains.Biometrical Journal. Vol. 7, pp. 865 - 876.

[5] Noel J. Purcell and Leslie Kish(1979)Estimation for small domains.Biometrics,Jun., 1979, Vol. 35, No. 2 , pp. 365-384.

[6] Gonzalez, M. E. and Waksberg, J. (1973).Estimation of the error of synthetic estimates.paper presented at first meeting of the International association of survey statisticians,august 1973, Vienna, Austria, pp.18-25.

[7] Cochran, W.G. (1977).Sampling Techniques,Wiley & Sons. Singh, D. and Chaudhary, F.S.(1986). Theory and Analysis of Sample Survey Design.Wiley Eastern Limited..

[8] Lessler, J. T., and Kalsbeek, W. D. (1992) Nonsampling Error in Surveys, Wiley, New York.

[9] Small, R. J. (1986) Survey Nonresponse Adjustments for Estimates of Means, International Statistical Review, 54, 139-157.

[10] Madow, W. G., Nisselson, H., and Olkin, I. (Editors) (1983) Incomplete data in sample surveys, Volume 2 Theory and Bibliographies, Academic Press, New York.

[11] Montaquila, J. M., Brick, J. M., Hagedorn, M. C., Kennedy, C., and Keeter, S. (2008) Aspects of Nonresponse Bias in RDD Telephone Surveys, Chapter 25, In: J. M. Lepkowski.

[12] H. L., and Scheuren, F. J. (1983) Weighting Adjustment for Unit Nonresponse, Chapter 13, In: W. G. Madow, H. Nisselson, and I. Olkin (eds.): “Incomplete data in sample surveys, Volume 2 Theory and Bibliographies”, Academic Press, New York, 143-184.

[13] Rizzo, L., Kalton, G., and Brick, J. M. (1996) A Comparison of Some Weighting Adjustment Methods for Panel Nonresponse, Survey Methodology, 22 (1), 43-53.

