ORIGINAL PAPER

DETERMINATION OF OPTIMUM SAMPLE SIZE AND VARIANCE IN MULTIVARIATE STRATIFIED SAMPLING WITH NON-LINEAR TIME FUNCTION

ABRAR AHMAD[1], QUAZZAFI RABBANI[1], A.H. ANSARI[2]

Manuscript received: 20.08.2022; Accepted paper: 04.04.2023;
Published online: 30.06.2023.

Abstract. *Stratified sample size which determined in Stratified Random Sampling (SRS) with the help of different allocation techniques. The multivariate stratified sampling with linear or nonlinear time function can be optimized using generalized model of time function. After formulation as generalized model with time function (linear or nonlinear), in which the information given that there is only one varying time (time of transportation) and different time involved are constants or insignificant. The formulated generalized model was solved using different techniques but in present paper we use Lagrange multiplier technique to get the optimal solution (i.e. by minimizing the variance) with the involved time constraints. The generalized model with time of transportation is derived under optimum allocation and other involved time is fixed, when the sample size and its variances for a given time is unknown, then different allocation techniques are required for proportionality constant and a given time.*

Keywords: *optimum allocation; compromise allocation; time functions; sample size; proportionality constant.*

1. INTRODUCTION

Sampling technique is one of the best techniques to obtain information from a larger population (universe). Utility of this technique is determined by its appropriate application in a given context of real life examples. Technical aspects of its application increase the chances of accuracy of information obtained. It is believed that this technique may yield almost the same result of a survey covered the entire universe. So increasing sample size is not the only solution to get accuracy rather sampling technique is the best alternative available.

For a relatively bigger and diversified universe, it is important to consider the inner homogeneity and number while dividing or forming the strata within universe. Each stratum should be based on most possible homogenous group that may be called sub-population and non-overlapping with each other. From each stratum a fixed number of representatives are selected. These selections must be based on some logical and convincing grounds that may ensure the representation of all possible diversities within the given strata. In some cases each stratum may be divided again in sub-strata depending on the requirement of the study and size of the population. The most important thing to be remembered while executing all these steps

[1]Integral University, Department of Mathematics & Statistics, 226026 Lucknow, India.
E-mail:abrar.stats@gmail.com, quazzafi@gmail.com.
[2] Kejriwal Institute of Management and Development Studies, 834010 Ranchi, India.
E-mail:athar_hussain1@rediffmail.com.

is the logic and objectives. Each step must have a clear and logical base. The use of this type of sampling technique in sample literature requires the solution of basic three types of problems as below:

i) To find the number of strata
ii) To find the stratum boundaries, and
iii) To find the sample sizes to be selected from various strata.

In the literature related with sampling or stratified sampling the main problem is to find the size of samples to be picked from each stratum out of different allocation techniques like equal, proportional and optimum etc.

The method that is considered the best is optimum allocation. But in real situations, the implementation of optimum allocation is not possible. In present case it seems suitable to find near optimal allocation or compromise allocation. In case of multivariate sampling problem the optimal allocation method does not give the optimal solution for each variable and then researcher need to adapt in solution up to some extent by which the solution gives the optimal allocation in some context.

In the sampling literature the researchers are interested to get high precision of given character under study with the help of stratification and reduce the heterogeneity of the population as per stratification technique required.

In [1], the author explained very nicely for the univariate population optimum allocation in stratified random sampling literature. Later [2], explained how to determine strata boundaries in multivariate situation and also discuss about compromise allocation which is not optimal for all characters individually. The time of a character with in different stratum may differ in different situations.

The problem of determining the optimum allocation was formulated as a generalized programming problem and solved using Lagrange multiplier technique to minimize the variance subject to a given cost [3]. Immense literature is available on the cost of sampling surveys, but the time factor involved in it has not been given due importance in the process of data collection. It is true that budget is an important aspect in all researches but in some cases time acquires the primary importance. Situation may arise where the responses are determined by situation and may change if not captured immediately or under the same situation. Therefore, in such situations *fixation of time becomes more important than budgetary allocation*.

The simple form of linear time function is most appropriate to use when a main part of the time is about measurements of all units involved. Suppose t_0 represents the overhead time and t_h is the time of enumerating a character per unit. The total time incurred can be expressed as:

$$T = t_0 + \sum_{h=1}^{L} (t_h n_h).$$

In above equation the time function is Linear and simple in the form of sample size and gets the minimum variance estimator by this way. This type of time function is suitable if the selection time for each unit from h^{th} stratum is not significantly different.

In this paper, we considered different allocation techniques using a generalized model to get optimum size of allocation and variance for time function as nonlinear by using langrage multiplier technique.

In research methods, the money has acquired the dominant place in deciding samples and various techniques to be used. It is of course true, that the cost is the most important factor involved in gathering information. Sometimes cost becomes the dominant factor and determines all other factors such as time and energy. It happens when a researcher needs to gather information about a social behavior in a given situation which is not static for example the behaviour of people during a pandemic or disaster.

It is accepted truth that behaviors is shaped and get impacted by pandemic or any other similar disaster towards social issues. Those impacts don't last beyond a particular period or in any case could not maintain its intensity. Therefore, it is important rather essential for any researcher who is willing to capture the real impact of such pandemic or behaviour to complete the gathering of information within the given limited time or any emergent situations such as coronavirus disease (COVID-19) is one of the best examples of such situation. In such situations time acquires more important factor than money where a researcher has fixed or limited time to gather information.

The idea is to formulate a generalized model having time function as nonlinear time function to get the optimum solution as size of stratum sample for given only one varying time (transportation time for enumerating the samples) and other time incurred are constant or insignificant.

Table 1. Time required for different activities during the sample survey

Fixed time (t_0)	time for survey planning
	time for colleting basic information of selected filed study (town/village)
	time for listing and selecting households in the villages
	time for development of the survey design
	time for survey management
	time for introducing the purpose for the survey to the selected samples
	time required for the administrative, professional and supervisory personnel
	time for data processing, analysis and presentation of results
Varying time (t_{rh})	transportation time for enumerating the samples (field travel/time to travel between the samples etc.)

Nonlinear time function: If we are considering a practical situation then we conclude that the number of sampling unit selected and transportation time is not directly proportional. For this reason, practically the time function considered is nonlinear in nature.

When the **transportation time** is considered as nonlinear then, the **time function** is as follows:

$$T = t_{0v} + \sum_{h=1}^{L}(t_h n_h) + \sum_{h=1}^{L}(t_{rh}\sqrt{n_h}), \tag{1}$$

If the travel time between units are substantial the time of enumerating a character per unit t_h is nearly constant (time for enumerating the individual samples are supposed to be equal), hence the time required for enumerating the sample $t_e = \sum_{h=1}^{L}(t_h n_h)$ which is not varying time and can be represented as t_e . Hence the total overhead time can be represented as ($t_0 = t_{0v} + t_e$) then the time function of equation (1) can be written as approximately:

$$T = t_0 + \sum_{h=1}^{L}(t_{rh}\sqrt{n_h}). \tag{2}$$

Where,

T total time

t_0 overhead time,

t_h time of enumerating a character per unit

n_h sample size in various strata
t_{rh} time of transportation or travel time per unit

The equation (2) can be expressed as:

$$T = t_0 + \sum_{h=1}^{L} (t_{rh} n_h^{\beta}) \,, \tag{3}$$

Where β be the constant of proportionality defined on a set of real numbers, such that β =1/2 for a case in which it is only varying travel time t_{rh} that is incurred.

(i) The constant of proportionality indicates the effect on selection of one sampling unit in time function from h^{th} stratum.
(ii) If constant of proportionality is less than 1(β <1), the affected time function will be less than one unit in the selection of one sampling unit from the strata.
(iii) If constant of proportionality is larger than 1(β >1), the affected time function will be greater than one unit the selection of one sampling unit from the strata.
(iv) Constant of proportionality is a constant positive value determined by the researcher.
(v) If the selection time of one unit differs among strata, then it is suggested that the time function in equation (3) be used.

2. GENERALIZED MODEL OF OPTIMUM STRATIFICATION

The optimum solution can be obtained by minimizing either total time or variance of the estimate when the other is subject to constraint. The time and variance of a survey are major factors of sample allocation to various strata. Most of the surveys are based on a multi stage stratified sampling design.

We either minimize the time for a given variance or minimize the variance of the sample mean with respect to time.

A sample is drawn (by the method of simple random sampling without replacement) independently in different strata the variance of the estimate $\bar{y}_{str}$ is

$$V(\bar{y}_{str}) = \sum_{h=1}^{L} \left(\frac{W_h^{\,2} S_{hj}^{\,2}}{n_h} \right) - \sum_{h=1}^{L} \left(\frac{W_h^{\,2} S_{hj}^{\,2}}{N_h} \right), \tag{4}$$

Hence, let $V(\bar{y}_{str})$ be the optimization function, given a proportionality constant β, let the total time incurred T be the constraint function. Then we minimize

$$V(\bar{y}_{str})_{\min} = \sum_{h=1}^{L} \left(\frac{W_h^{\,2} S_{hj}^{\,2}}{n_h} \right) - \sum_{h=1}^{L} \left(\frac{W_h^{\,2} S_{hj}^{\,2}}{N_h} \right), \tag{5}$$

Subject to, $T = t_0 + \sum_{h=1}^{L} (t_{rh} n_h^{\beta}) \,,\; n_h \geq 1$.

Where,

N Population size

$W_h = \frac{N_h}{N}$ Stratum weight

S_h^2 Stratum mean squire

n_h Sample size of h^{th} stratum

N_h Population size of h^{th} stratum

β proportionality constant

n sample size

Using Lagrange multiplier technique, let

$$G(n_h, \lambda) = V(\bar{y}_{str}) + \lambda[t_0 + \sum_{h=1}^{L}(t_{rh} n_h^{\beta}) - T].$$

Differentiating G with respect to n_h ,

$$\frac{\delta G}{\delta n_h} = -\sum_{h=1}^{L}\left(\frac{W_h^{\ 2} S_{hj}^{\ 2}}{n_h^{\ 2}}\right) + \lambda\beta\sum_{h=1}^{L}(t_{rh} n_h^{\beta-1}),$$

Setting the above equation to zero $\sum_{h=1}^{L}[\lambda\beta(t_{rh} n_h^{\beta-1}) - \left(\frac{W_h^{\ 2} S_{hj}^{\ 2}}{n_h^{\ 2}}\right)] = 0$. This implies,

$$\lambda\beta(t_{rh} n_h^{\beta-1}) - \left(\frac{W_h^{\ 2} S_{hj}^{\ 2}}{n_h^{\ 2}}\right)] = 0\ ;\ \lambda\beta(t_{rh} n_h^{\beta-1}) = \left(\frac{W_h^{\ 2} S_{hj}^{\ 2}}{n_h^{\ 2}}\right);\ n_h^{\beta+1} = \frac{1}{\lambda}\left(\frac{W_h^{\ 2} S_{hj}^{\ 2}}{\beta t_{rh}}\right);$$

Making n_h the subject of formula,

$$n_h = \left(\frac{1}{\lambda}\right)^{1/(\beta+1)}\left(\frac{W_h^{\ 2} S_{hj}^{\ 2}}{\beta t_{rh}}\right)^{1/(\beta+1)} \tag{6}$$

We know that $\sum_{h=1}^{L} n_h = n$; and then summing up the equation (6) both side, then the equation becomes:

$$n = \left(\frac{1}{\lambda}\right)^{1/(\beta+1)}\sum_{h=1}^{L}\left(\frac{W_h^{\ 2} S_{hj}^{\ 2}}{\beta t_{rh}}\right)^{1/(\beta+1)}\ . \tag{7}$$

From equation (6) & (7) we can deduce that,

$$\frac{n_h}{n} = \frac{\left(\frac{W_h^{\ 2} S_{hj}^{\ 2}}{\beta t_{rh}}\right)^{1/(\beta+1)}}{\sum_{h=1}^{L}\left(\frac{W_h^{\ 2} S_{hj}^{\ 2}}{\beta t_{rh}}\right)^{1/(\beta+1)}}\ ;\ n_h = \frac{n\left(\frac{W_h^{\ 2} S_{hj}^{\ 2}}{\beta t_{rh}}\right)^{1/(\beta+1)}}{\sum_{h=1}^{L}\left(\frac{W_h^{\ 2} S_{hj}^{\ 2}}{\beta t_{rh}}\right)^{1/(\beta+1)}}\ ; \tag{8}$$

$$n_h = \frac{n \dfrac{{W_h}^{2/(1+\beta)} {S_{hj}}^{2/(1+\beta)}}{{t_{rh}}^{1/(1+\beta)}}}{\sum_{h=1}^{L} \dfrac{{W_h}^{2/(1+\beta)} {S_{hj}}^{2/(1+\beta)}}{{t_{rh}}^{1/(1+\beta)}}} \quad ;$$

In the above equation n_h be the size of sample in the h^{th} stratum with given the constant of proportionality β. On replacing the value of n_h from above equation (8) in equation (4) we get generalized variance $(\bar{y}_{str})$

$$V(\bar{y}_{str})_{gen} = \frac{1}{n} \sum_{h=1}^{L} \left(\frac{{W_h}^{2/(1+\beta)} {S_{hj}}^{2/(1+\beta)}}{{t_{rh}}^{1/(1+\beta)}} \right) \sum_{h=1}^{L} \left({W_h}^{2\beta/(1+\beta)} {S_{hj}}^{2\beta/(1+\beta)} {t_{rh}}^{1/(1+\beta)} \right) - - \frac{1}{N} \sum_{h=1}^{L} (W_h {S_{hj}}^2). \tag{9}$$

In the above generalized variance shown that there is only one varying time (transportation time) and other time incurred are insignificant or fixed given a constant of proportionality β.

Remarks:

(a) Generalized stratum size for β=1, substituting (β=1) in the equation (8) reduces to:

$$n_h = \frac{n \dfrac{W_h S_{hj}}{{t_{rh}}^{1/2}}}{\sum_{h=1}^{L} \dfrac{W_h S_{hj}}{{t_{rh}}^{1/2}}} \quad ;$$

(b) Generalized variance for β=1, substituting (β=1) in the equation (9) reduces to:

$$V(\bar{y}_{str}, \beta=1)_{gen} = \frac{1}{n} \sum_{h=1}^{L} \left(\frac{W_h S_{hj}}{{t_{rh}}^{1/2}} \right) \sum_{h=1}^{L} \left(W_h S_{hj} {t_{rh}}^{1/2} \right) - \frac{1}{N} \sum_{h=1}^{L} (W_h {S_{hj}}^2)$$

3. OPTIMUM ALLOCATION OF SAMPLE SIZE AND VARIANCE

We discussed the allocation technique of total sample size *n* in different stratum, i.e. the number of samples that should be taken from the different stratum. The optimum size of allocation from different strata can be considered either by reducing the survey time, keeping the variance (efficiency) of the estimator to a certain level or by minimizing variance (efficiency of the survey) keeping the time of the survey to a certain level.

The total size of sample *n* required for the optimum size of sample within strata. The solution for the value *n* depends on whether the sample is chosen to meet a specified total time T or to give a specified variance V for $\bar{y}_{str}$.

3.1. CASEI: OPTIMUM VARIANCE WHEN THE TIME IS FIXED

If the travel time t_{rh} of collecting information from h^{th} stratum is constant for a given constant of proportionality β; $t_{rh}= t,$ then the equation (3) reduces to:

$$T = t_0 + t\sum_{h=1}^{L} n_h^{\beta}, \tag{10}$$

and when $t_{rh}= t$ then the equation (8) can be written as

$$n_h = \frac{\frac{n}{t^{1/(1+\beta)}} W_h^{\,2/(1+\beta)} S_{hj}^{\,2/(1+\beta)}}{\frac{1}{t^{1/(1+\beta)}} \sum_{h=1}^{L} W_h^{\,2/(1+\beta)} S_{hj}^{\,2/(1+\beta)}} \;\; ; \; n_h = \frac{n \;\; W_h^{\,2/(1+\beta)} S_{hj}^{\,2/(1+\beta)}}{\sum_{h=1}^{L} W_h^{\,2/(1+\beta)} S_{hj}^{\,2/(1+\beta)}} \;\; ; \tag{11}$$

Substituting the value of n_h of equation (11) in the equation (9), then the equation of generalized variance becomes from $V(\bar{y}_{str})_{\text{gen}}$ to $V(\bar{y}_{str})_{\text{opt}}$ under optimum allocation for fixed time.

$$V(\bar{y}_{str})_{opt} = \frac{1}{n}\sum_{h=1}^{L}\left(W_h^{\,2/(\beta+1)} S_{hj}^{\,2/(\beta+1)}\right)\sum_{h=1}^{L}\left(W_h^{\,2\beta/(\beta+1)} S_{hj}^{\,2\beta/(\beta+1)}\right) - \frac{1}{N}\sum_{h=1}^{L}(W_h S_{hj}^{\,2}) \tag{12}$$

Where $V(\bar{y}_{str})_{\text{opt}}$ is the variance of optimum allocation for given the proportionality constant β and n_h is the size of sample of the h^{th} stratum and t is the constant time.

For β=1, substituting (β=1) in the equation (12) reduces to:

$$V(\bar{y}_{str}, \beta{=}1)_{opt} = \frac{1}{n}\sum_{h=1}^{L}\left(W_h S_{hj}\right)\sum_{h=1}^{L}\left(W_h S_{hj}\right) - \frac{1}{N}\sum_{h=1}^{L}(W_h S_{hj}^{\,2})$$

3.2. CASE II: UNKNOWN SAMPLE SIZE FOR A GIVEN TIME

For a given constant of proportionality β, the equation (6) can be written as

$$n_h^{\beta} = \left(\frac{1}{\lambda}\right)^{\beta/(\beta+1)} \left(\frac{W_h^{\,2} S_{hj}^{\,2}}{\beta t_{rh}}\right)^{\beta/(\beta+1)} \tag{13}$$

We know that the equation (3) of time function that is $T - t_0 = \sum_{h=1}^{L}(t_{rh} n_h^{\beta})$. Substituting the equation (13) into the above value T - t_0

$$T - t_0 = \left(\frac{1}{\lambda}\right)^{\beta/(\beta+1)} \left(\frac{1}{\beta}\right)^{\beta/(\beta+1)} \sum_{h=1}^{L}\left(t_{rh}^{1/(\beta+1)} (W_h^{\,2} S_{hj}^{\,2})^{\beta/(\beta+1)}\right) \tag{14}$$

The equation (13) can be re-written as:

$$\left(\frac{1}{\lambda}\right)^{\beta/(\beta+1)} = n_h^{\beta}(\beta)^{\beta/(\beta+1)}\left(\frac{t_{rh}}{W_h{}^2 S_{hj}{}^2}\right)^{\beta/(\beta+1)}; \tag{15}$$

On substituting equation (15) in equation (14), we get

$$T-t_0 = n_h^{\beta}(\beta)^{\beta/(\beta+1)}\left(\frac{t_{rh}}{W_h{}^2 S_{hj}{}^2}\right)^{\beta/(\beta+1)}\left(\frac{1}{\beta}\right)^{\beta/(\beta+1)}\sum_{h=1}^{L}\left(t_{rh}^{1/(\beta+1)}(W_h{}^2 S_{hj}{}^2)^{\beta/(\beta+1)}\right);$$

$$= n_h^{\beta}\left(\frac{t_{rh}}{W_h{}^2 S_{hj}{}^2}\right)^{\beta/(\beta+1)}\sum_{h=1}^{L}\left(t_{rh}^{1/(\beta+1)}(W_h{}^2 S_{hj}{}^2)^{\beta/(\beta+1)}\right);$$

Making $_h{}^{\beta}$ subject of formula

$$n_h^{\beta} = \left(\frac{W_h{}^2 S_{hj}{}^2}{t_{rh}}\right)^{\beta/(\beta+1)}\frac{T-t_0}{\sum_{h=1}^{L}\left(t_{rh}^{1/(\beta+1)}(W_h{}^2 S_{hj}{}^2)^{\beta/(\beta+1)}\right)};$$

This implies:

$$n_h = \left(\frac{W_h{}^2 S_{hj}{}^2}{t_{rh}}\right)^{1/(\beta+1)}\left(\frac{T-t_0}{\sum_{h=1}^{L}\left(t_{rh}^{1/(\beta+1)}(W_h{}^2 S_{hj}{}^2)^{\beta/(\beta+1)}\right)}\right)^{1/\beta}. \tag{16}$$

We know that $n=\sum_{h=1}^{L} n_h$ and substituting the value of optimum n_h of equation (16), then we can obtained the value of n (sample size for a given time)

$$n = \sum_{h=1}^{L} n_h = (T-t_0)^{1/\beta}\frac{\sum_{h=1}^{L}\left(W_h{}^{2/(\beta+1)} S_{hj}{}^{2/(\beta+1)} t_{rh}^{-1/(\beta+1)}\right)}{[\sum_{h=1}^{L}\left((W_h{}^{2\beta/(\beta+1)} S_{hj}{}^{2\beta/(\beta+1)} t_{rh}^{1/(\beta+1)}\right)]^{1/\beta}}. \tag{17}$$

For β=1, substituting (β=1) in the equation (17) reduces to:

$$n = (T-t_0)\frac{\sum_{h=1}^{L}\left(W_h S_{hj} t_{rh}^{-1/2}\right)}{[\sum_{h=1}^{L}\left((W_h S_{hj} t_{rh}^{1/2}\right)]^{1/2}}, \quad for\,\beta=1.$$

3.3. CASEIII: UNKNOWN SAMPLE SIZE FOR A SPECIFIED VARIANCE V_0

The equation (9) can be written as $(\bar{y}_{str})_{\text{gen}} = V_0$ (say) for a given proportionality constant β,

$$V_0 = \frac{1}{n}\sum_{h=1}^{L}\left(W_h^{2/(\beta+1)} S_{hj}^{2/(\beta+1)} t_{rh}^{-1/(\beta+1)}\right)\sum_{h=1}^{L}\left(W_h^{2\beta/(\beta+1)} S_{hj}^{2\beta/(\beta+1)} t_{rh}^{1/(\beta+1)}\right) - \frac{1}{N}\sum_{h=1}^{L} W_h S_{hj}^2.$$

This implies,

$$V_0 + \frac{1}{N}\sum_{h=1}^{L} W_h S_{hj}^2 = \frac{1}{n}\sum_{h=1}^{L}\left(W_h^{2/(\beta+1)} S_{hj}^{2/(\beta+1)} t_{rh}^{-1/(\beta+1)}\right)\sum_{h=1}^{L}\left(W_h^{2\beta/(\beta+1)} S_{hj}^{2\beta/(\beta+1)} t_{rh}^{1/(\beta+1)}\right). \qquad (18)$$

Hence,

$$n = \frac{\sum_{h=1}^{L}\left(W_h^{2/(\beta+1)} S_{hj}^{2/(\beta+1)} t_{rh}^{-1/(\beta+1)}\right)\sum_{h=1}^{L}\left(W_h^{2\beta/(\beta+1)} S_{hj}^{2\beta/(\beta+1)} t_{rh}^{1/(\beta+1)}\right)}{V_0 + \frac{1}{N}\sum_{h=1}^{L} W_h S_{hj}^2}. \qquad (19)$$

If travel time is constant for a specified variance V_0, then the equation (19) reduces as and get the value of n (sample size for a specified variance)

$$n = \frac{\sum_{h=1}^{L}\left(W_h^{2/(\beta+1)} S_{hj}^{2/(\beta+1)}\right)\sum_{h=1}^{L}\left(W_h^{2\beta/(\beta+1)} S_{hj}^{2\beta/(\beta+1)}\right)}{V_0 + \frac{1}{N}\sum_{h=1}^{L} W_h S_{hj}^2}. \qquad (20)$$

For β=1, substituting (β=1) in the equation (20) reduces to:

$$n = \frac{\sum_{h=1}^{L}\left(W_h S_{hj}\right)\sum_{h=1}^{L}\left(W_h S_{hj}\right)}{V_0 + \frac{1}{N}\sum_{h=1}^{L} W_h S_{hj}^2}; \quad for\,\beta=1.$$

4. EQUAL ALLOCATION: SAMPLE SIZE FOR A GIVEN TIME *T*

If there is only information is available regarding the size of strata the strata are approximately equal sizes, and no information related to the variability or response distribution within the strata, then allocation is considered as equal allocation which is best choice with researcher. If the researcher decides to make partition into L disjoint group/ stratum given a constant of proportionality β.

Then $n_h = \frac{n}{L}$, from equation (3) of time function that is $T - t_0 = \sum_{h=1}^{L}(t_{rh} n_h^{\beta})$.

Substituting n_h as $\frac{n}{L}$ then $T - t_0 = \sum_{h=1}^{L} t_{rh}\left(\frac{n}{L}\right)^{\beta} = \left(\frac{n}{L}\right)^{\beta}\sum_{h=1}^{L} t_{rh} = \frac{n^{\beta}}{L^{\beta}}\sum_{h=1}^{L} t_{rh}.$

Making n^{β} subject of formula $n^{\beta} = \frac{L^{\beta}(T - t_0)}{\sum_{h=1}^{L} t_{rh}}$.

Then,

$$n = L[\frac{(T - t_0)}{\sum_{h=1}^{L} t_{rh}}]^{1/\beta} \tag{21}$$

The above equation is used to calculate the size of sample.

For β=1, then substituting (β=1) in the equation (21) reduces to: $n = L[\frac{(T - t_0)}{\sum_{h=1}^{L} t_{rh}}]$

5. PROPORTIONAL ALLOCATION: SAMPLE SIZE FOR A GIVEN TIME *T*

The proportional allocation is meant the ratio of stratum sizes equal to sample sizes. If we interested to allocate the size of sample based on set proportion for each stratum given a constant of proportionality β.

Then $n_h = n(N_h/N) = nW_h$, from equation (3) of time function that is $T - t_0 = \sum_{h=1}^{L}(t_{rh} n_h{}^{\beta})$, Substituting n_h as nW_h into $T - t_0$; then $T - t_0 = \sum_{h=1}^{L} t_{rh}(nW_h{}^{\beta})$, $T - t_0 = n^{\beta} \sum_{h=1}^{L} t_{rh} W_h{}^{\beta}$,

The above expression may be rewritten as:

$$n^{\beta} = \frac{T - t_0}{\sum_{h=1}^{L} t_{rh} W_h{}^{\beta}},$$

Then,

$$n = (\frac{T - t_0}{\sum_{h=1}^{L} t_{rh} W_h{}^{\beta}})^{1/\beta} \tag{22}$$

This is applicable when the sample size for a given time is not known. For β=1, then substituting (β=1) in the equation (22) reduces to:

$$n = (\frac{T - t_0}{\sum_{h=1}^{L} t_{rh} W_h})$$

6. COMPROMISE ALLOCATION: SAMPLE SIZE FOR A GIVEN TIME *T*

We assume that L homogenous strata are constructed in the desired population. The stratum size N_h, stratum weight $W_h = \left(N_h / N\right)$, and sample size n_h of stratum h where h =1,2,..., L. The authors [4-5], mentioned a compromise allocation method utilizing convex combination of equal allocation ($n_h = \frac{n}{L}$) and proportional methods ($n_h = n\left(N_h / N\right) = nW_h$). The allocation is given as Compromised form

$$n_h^{comp} = k(nW_h) + (1-k)\frac{n}{L} = k(\frac{N_h}{N}n) + (1-k)\frac{n}{L} \tag{23}$$

Where, k is a constant in ($0 \le k \le 1$) range. This method is converted into proportional allocation when k=1 and equal allocation if k=0. From equation (3) of time function that is $T - t_0 = \sum_{h=1}^{L}(t_{rh} n_h{}^{\beta})$, substituting n_h as n_h^{comp} into $T - t_0$; then

$$T - t_0 = \sum_{h=1}^{L} t_{rh}(n_h^{comp})^{\beta} \text{ , } T - t_0 = \sum_{h=1}^{L} t_{rh}\left[k(nW_h) + (1-k)\frac{n}{L}\right]^{\beta} \text{ ,} \tag{24}$$

After simplification it can be written as:

$$\begin{aligned}
&(T - t_0) = \left(\frac{n}{L}\right)^{\beta} \sum_{h=1}^{L} t_{rh}\left[\frac{kLW_h + (1-k)}{L}\right]^{\beta} \text{ ,}\\
&\left(\frac{n}{L}\right)^{\beta} = \frac{T - t_0}{\sum_{h=1}^{L} t_{rh}\left[k(LW_h - 1) + 1)\right]^{\beta}} \text{ ,}\\
&n^{\beta} = \frac{L^{\beta}(T - t_0)}{\sum_{h=1}^{L} t_{rh}\left[k(LW_h - 1) + 1)\right]^{\beta}} \text{ ,}\\
&n = L\left[\frac{(T - t_0)}{\sum_{h=1}^{L} t_{rh}\left[k(LW_h - 1) + 1)\right]^{\beta}}\right]^{1/\beta} \text{ ,}
\end{aligned} \tag{25}$$

For β=1, then substituting (β=1) in the equation (25) reduces to:

$$n = L\left[\frac{(T - t_0)}{\sum_{h=1}^{L} t_{rh}\left[k(LW_h - 1) + 1)\right]}\right]$$

7. CONCLUSIONS

The method of Stratification will reduce variance as well as reduce sampling error to increase precision. In stratification we reduce the heterogeneity of universe by making different strata with high level of homogeneity. Several authors discussed their different problem of allocation and no one has taken care of non-linear time function in multivariate stratified sampling whether it has an important role in time optimization.

The purpose of study is to maximize the precision. So, we have formulated a general model for time as a variable in nonlinear form and a given constant of proportionality β for stratum size of sample and variance if we have only one time as a variable (transportation time) and some time included in model are constant or insignificant, if time is constant in each stratum and when size of sample for the survey time for a given variance is unknown.

REFERENCES

[1] Cochran, W.G., *Sampling Techniques*, 3rd Edition, John Wiley and Sons, New York, 1977.
[2] Ahsan, M.J., Khan, S.U., *Journal of Indian Society of Statistics and Operations Research*, **4**, 1, 1983.
[3] Fajana, O.B., Fadiji, F.A., *International Journal of Science and Research,* **8**(11), 1406, 2019.
[4] Costa, A., Satorra, A., Venture, E., *Statistics and Operations Research Transactions,* **28**, 69, 2004.
[5] Sahin Tekin, S.T., Ozdemir, Y.A., Metin, C.B., *Journal of Science*, **30**(3), 181, 2017.
[6] Ghufran S., Khowaja S., Ahsan, M.J., *Communications in Statistics - Simulation and Computation*, **41** (5), 598, 2012.
[7] Gupta, N., Bari A., *Journal of Statistical Computation and Simulation*, **87**(12), 2372, 2017.
[8] Khowaja, S., Ghufran, S., Ahsan, M.J., *Communications in Statistics - Simulation and Computation*, **40** (5), 710, 2011.
[9] Ali, I., Raghav, Y.S., Bari, A., *Journal of Statistical Computation and Simulation*, **83**(5), 962, 2013.
[10] Díaz-García, J.A., Garay-Tápia, M.M., *Computational Statistics & Data Analysis*, **51**(6), 3016, 2007.
[11] Díaz-García, J.A., Ramos-Quiroga, R., *Journal of Mathematical Modelling and Algorithms in Operations Research*, **13**(3), 315, 2014.
[12] Díaz-García, J.A., Ramos-Quiroga, R., *Optimization*, **63**(11), 1665, 2014.
[13] Alshqaq, S.S.A., Ahmadini, A.A.H., Ali, I., *Mathematical Problems in Engineering*, **2022**, 2502346, 2022.
[14] Haseen, S., Ali, I., Bari, A., *Communications in Statistics-Simulation and Computation,* **45**(8), 2810, 2014.
[15] Khan, M.F., Ali, I., Raghav, Y.S., Bari, A. *American Journal of Operations Research*, **2**(1), 122, 2012.