ON APPROXIMATELY OPTIMUM STRATA BOUNDARIES USING TWO AUXILIARY VARIABLES

Faizan Danish*[1], S.E.H. Rizvi** and Carlos Bouza***
*Research Consultation Servives, Doha, Qatar
**Division of Statistics and Computer Science,Faculty of Basic Sciences, SKUAST-JAMMU, (J&K), India
***Universidad de La Habana, Cuba

ABSTRACT
In the present investigation, a methodology has been developed for obtaining Approximately Optimum Strata Boundaries (AOSB), appropriate for surveys involving single study variable (Y), on the basis of two auxiliary variables (X and Z) closely related to the study variable. For theoretical development, regression model has been considered as Y = C (X, Z) + e , where C (X, Z) is a function of X & Z and 'e' is error term. Minimal equations have been obtained, under certain assumptions, by minimizing the variance of the estimation variable. Due to implicit nature of these equations, a Cum $\sqrt[3]{D_1(x,z)}$ rule has been proposed for finding out AOSB. Comparisons have been made empirically, using certain density functions, with cube root method due to Singh and Sukhatme (1969) for single auxiliary variable. It showed remarkable gain in efficiency in case two auxiliary variables are used as the basis of stratification.

KEYWORDS: Optimum Stratification, Minimal Equations, Strata Boundaries.

MSC:62D05

RESUMEN
En la presente investigación, una metodología ha sido desarrollada para obtener Fronteras Aproximadamente Óptimas de los Estratos (Approximately Optimum Strata Boundaries, AOSB), que sean apropiadas para encuestas en el estudio de una variable (Y), sobre la base de tener dos variables auxiliares (X y Z) relacionadas fuertemente con la bajo estudio. Para el desarrollo teórico, el modelo de regresión considerado fue Y = C (X, Z) + e , donde C (X, Z) es una función de X & Z y 'e' es el término del error. Ecuaciones minimales han sido obtenidas bajo ciertas asunciones, minimizando la variancia de la variable de estimación. Debidos a la naturaleza de estas ecuaciones, una regla Cum $\sqrt[3]{D_1(x,z)}$ ha sido propuesta para hallar las AOSB. Comparaciones han sido levada a cabo empíricamente, usando ciertas funciones de densidad, mediante el método de la raíz cúbica debida a Singh and Sukhatme (1969) para una sola variable auxiliar. Se muestra que una ganancia en eficiencia notable en el caso de las dos variables auxiliares son usadas como base para la estratificación.

PALABRAS CLAVE: Estratificación Òptima , Ecuaciones Minimales, Fronteras de los Estratos

1. INTRODUCTION

In stratified random sampling, a proper choice of strata boundaries is one of the important factors as regards to the efficiency of estimator of the population characteristic under study. Dalenius (1950) first considered the univariate problem treating estimation variable itself as stratification variable. Singh-Sukhatme (1969) provided some approximate solutions for the strata boundaries, by using single auxiliary variable as the stratification variable, under optimum and proportional allocations, Allende-Bouza (1987) proposed a method for optimum stratification in the multivariate case using mathematical programming algorithms. Lavallee-Hidiroglou (1988) proposed an algorithm to construct stratum boundaries for a power allocated stratified sample of non-certainty sample units. Niemiro (1999) proposed a random search method in the stratification problem but the algorithm did not guarantee that it leads to global optimum. Furthermore, it would go wrong in a case of a large population, as it requires too many iteration steps (see Kozak 2004). Rizvi *et al*. (2002) developed a method for obtaining approximately optimum strata boundaries (AOSB) on an auxiliary, closely related with the study variables, appropriate for surveys involving more than one study variable. Danish *et al.* (2017) proposed a technique for one study variable having an auxiliary variable with varying cost for each unit. Presuming the fact that the efficiency of the estimator may be improved by using more auxiliary information, in the present study a method has been developed for obtaining AOSB for an estimation variable using two auxiliary variables as the basis of stratification.

[1] corresponding Author: danishstat@gmail.com

Let there be a finite population consisting of N units, for which it is required to estimate the total or mean for the characteristic Y under study, using simple random sampling technique. In order to have this, we divide the whole population into L × M strata on the basis of two auxiliary variables, say, X and Z, such that the number of units in the $(h, k)^{th}$ stratum is N_{hk} so that

$$\sum_{h=1}^{L}\sum_{k=1}^{M} N_{hk} = N$$

A random sample of size 'n' is to be drawn from the whole population allocating n_{hk} units to the $(h, k)^{th}$ stratum such that

$$\sum_{h=1}^{L}\sum_{k=1}^{M} n_{hk} = n$$

Let the value of population units in the $(h, k)^{th}$ stratum be denoted by y_{hki} (i = 1, 2, …, N_{hk}) so that the population total is

$$Y = \sum_{h=1}^{L}\sum_{k=1}^{M}\sum_{i=1}^{N_{hk}} y_{hki}$$

The unbiased estimator of population mean $\bar{Y}$ is given by

$$\bar{y}_{st} = \sum_{h=1}^{L}\sum_{k=1}^{M} W_{hk}\bar{y}_{hk}$$

Where, $W_{hk} = \frac{N_{hk}}{N}$ denotes the stratum weight and $\bar{y}_{hk} = \frac{1}{n_{hk}}\sum_{i=1}^{n_{hk}} y_{hki}$ is the sample mean for the $(h, k)^{th}$ stratum. The sampling variance of the estimator $\bar{y}_{st}$ is obtained as:

$$V\left(\bar{y}_{st}\right) = \sum_{h}\sum_{k}\left(1 - f_{hk}\right)\frac{W_{hk}^2\sigma_{hky}^2}{n_{hk}}$$

Where, $f_{hk} = \frac{n_{hk}}{N_{hk}}$ denotes the sampling fraction and σ_{hky}^2 represents the population variance for the character Y in the $\left(h, k\right)^{th}$ stratum defined as

$$\sigma_{hky}^2 = \frac{1}{N_{hk}}\sum_{i=1}^{N_{hk}}\left(y_{hki} - \bar{y}_{hk}\right)^2$$

$\bar{y}_{hk}$ being the population mean of all the N_{hk} units in the $\left(h, k\right)^{th}$ stratum.

If the finite population correction (f.p.c) is ignored in each stratum, the variance can be expressed as

$$V\left(\bar{y}_{st}\right) = \sum_{h}\sum_{k}\frac{W_{hk}^2\sigma_{hky}^2}{n_{hk}}$$

2. METHOD OF OPTIMUM ALLOCATION

In this method, the sample sizes n_{hk} are determined in such a way that for the given total sample size (which amounts to fixed total cost when the cost of observing a unit in each stratum is same) the variance of the estimator $\bar{y}_{st}$ is minimized. Thus, we have to minimize

$$V\left(\bar{y}_{st}\right) = \sum_{h}\sum_{k}\frac{W_{hk}^2\sigma_{hky}^2}{n_{hk}}$$

Subject to the constraint

$$\sum_{h=1}^{L}\sum_{k=1}^{M} n_{hk} = n \tag{2.1}$$

We select n_{hk} and the Lagrangian multiplier 'λ' so as to minimize the function

$$\psi(n_{hk},\lambda) = \sum_h \sum_k \frac{W_{hk}^2 \sigma_{hky}^2}{n_{hk}} + \lambda\left(\sum_h \sum_k n_{hk} - n\right)$$

Differentiating it partially with respect to n_{hk} and equating to zero, we get

$$\frac{\partial}{\partial n_{hk}}\psi(n_{hk},\lambda) = -\frac{W_{hk}^2 \sigma_{hky}^2}{n_{hk}^2} + \lambda = 0$$

This gives, that the sample size is given by

$$n_{hk} = \frac{W_{hk}\sigma_{hky}}{\sqrt{\lambda}}$$

Using it in equation (2.1), we have

$$n = \sum_h \sum_k \frac{W_{hk}\sigma_{hky}}{\sqrt{\lambda}}$$

Therefore, the optimal strata`s sample sizes are:

$$n_{hk} = n\frac{W_{hk}\sigma_{hky}}{\sum_h \sum_k W_{hk}\sigma_{hky}}$$

Thus, for optimum method of allocation the variance of the estimator becomes

$$V\left(\bar{y}_{st}\right)_{opt} = \frac{\left(\sum_h \sum_k W_{hk}\sigma_{hky}\right)^2}{n} \tag{2.2}$$

3. VARIANCE EXPRESSION

Let the regression model of Y on X and Z be given as Y = C (X, Z) + e, where C (X, Z) is a function of X and Z and 'e' is error term such that

$E\left(e/_{x,z}\right) = 0$. and $V\left(e/_{x,z}\right) = \eta(x,z) = 0$, $\forall$ $x\varepsilon(a,b)$ $z\varepsilon(c,d)$, b-a<∞ ,c-d<∞

If the joint density function of (Y, X, Z) in the super population is f(y, x, z), joint marginal pdf of X and Z is f(x, z) and the marginal density function of X and Z are f(x) and f(z), respectively, then under above regression model, we have the weight of the $(h,k)^{th}$ stratum as

$$W_{hk} = \int_{x_{h-1}}^{x_h}\int_{z_{k-1}}^{z_k} f(x,z)\,\partial x \partial z$$

$\mu_{hky} = \mu_{hkc} = \frac{1}{W_{hk}} \int_{x_{h-1}}^{x_h}\int_{z_{k-1}}^{z_k} c(x,z)\, f(x,z)\,\partial x \partial z$ denotes the mean of the $(h,k)^{th}$ stratum and $\sigma^2_{hky} = \sigma^2_{hkc} + \mu_{hk\eta}$, where $(x_{h-1}, x_h, z_{k-1}, z_k)$ are the boundary points of $(h,k)^{th}$ stratum and $\mu_{hk\eta}$ is the expected value of the function $\eta(x,z)$ in the $(h,k)^{th}$ stratum and σ^2_{hkc} is given as

$$\sigma^2_{hkc} = \frac{1}{W_{hk}} \int_{x_{h-1}}^{x_h} \int_{z_{k-1}}^{z_k} c^2(x,z) f(x,z) \partial x \partial z - (\mu_{hkc})^2$$

Using these relations, the variance for the estimator $\bar{y}_{st}$ as given in (2.1) and (2.2) can be expressed in terms of the population parameters of the function $C(x,z)$ and $\eta(x,z)$. The variance expression for the case of optimum allocation is, therefore, given by

$$V(\bar{y}_{st})_{opt} = \frac{\left(\sum_h \sum_k W_{hk} \sqrt{\sigma^2_{hkc} + \mu_{hk\eta}} \right)^2}{n} \tag{3.1}$$

4. MINIMAL EQUATIONS

Let $[x_h, z_k]$ denotes the set of optimum points of stratification on the range (a, b) and (c, d) of X and Z , respectively, then corresponding to these strata boundaries the variance of the estimator $\bar{y}_{st}$ should be minimum. These points $[x_h, z_k]$ are the solutions of the minimal equations which are obtained by equating to zero the partial derivatives of $V(\bar{y}_{st})$ with respect to x_h *and* z_k .Before deriving the minimal equations ,let us first find out the expression for some partial derivatives which will be helpful in obtaining the equations.

The minimization of the variance expression as given in (3.1) is equivalent to minimization of the expression

$$\sum_h \sum_k W_{hk} \sqrt{\sigma^2_{hkc} + \mu_{hk\eta}} \tag{4.1}$$

Equating to zero, the partial derivative of this expression with respect to x_h ,we get

$$\sum_k \left[W_{hk} \frac{\partial}{\partial x_h}\left(\sqrt{h}\right) + \left(\sqrt{h}\right)\frac{\partial}{\partial x_h} W_{hk} + W_{ik} \frac{\partial}{\partial x_h}\left(\sqrt{i}\right) + \left(\sqrt{i}\right)\frac{\partial}{\partial x_h} W_{ik} \right] = 0 \tag{4.2}$$

where i=h+1, $(h) = \sigma^2_{hkc} + \mu_{hk\eta}$ and $(i) = \sigma^2_{ikc} + \mu_{ik\eta}$

Note that

$$\frac{\partial}{\partial x_h}(h) = \frac{1}{W_{hk}} \int_{z_{k-1}}^{z_k} f(x_h, z) \left\{ \left[c(x_h, z) - \mu_{hk\eta} \right]^2 - \sigma^2_{hkc} + \eta(x_h, z) - \mu_{hk\eta} \right\} \partial z$$

Similarly, we have

$$\frac{\partial}{\partial x_h}(i) = -\frac{1}{W_{hk}} \int_{z_{k-1}}^{z_k} f(x_h, z) \left\{ \left[c(x_h, z) - \mu_{ik\eta} \right]^2 - \sigma^2_{ikc} + \eta(x_h, z) - \mu_{ik\eta} \right\} \partial z$$

On simplifications of (4.2) we get the minimal equations as

$$\sum_k \left[\frac{\int_{z_{k-1}}^{z_k} \left\{ f(x_h, z)\left[c(x_h, z) - \mu_{hkc} \right]^2 + \sigma_{hkc}^2 + \eta(x_h, z) - \mu_{hk\eta} \right\} \partial z}{\sqrt{\sigma_{hkc}^2 + \mu_{hk\eta}}} \right]$$

$$= \sum_k \left[\frac{\int_{z_{k-1}}^{z_k} \left\{ f(x_h, z)\left[c(x_h, z) - \mu_{ikc} \right]^2 + \sigma_{ikc}^2 + \eta(x_h, z) - \mu_{ik\eta} \right\} \partial z}{\sqrt{\sigma_{ikc}^2 + \mu_{ik\eta}}} \right] \tag{4.3}$$

Similarly, once again equating to zero, the partial derivative of the (4.1) w.r.t z_k ,we shall get

$$\sum_h \left[\frac{\int_{x_{h-1}}^{x_h} \left\{ f(x, z_k)\left[c(x, z_k) - \mu_{hkc} \right]^2 + \sigma_{hkc}^2 + \eta(x, z_k) - \mu_{hk\eta} \right\} \partial x}{\sqrt{\sigma_{hkc}^2 + \mu_{hk\eta}}} \right] = \sum_h \left[\frac{\int_{x_{h-1}}^{x_h} \left\{ f(x, z_k)\left[c(x, z_k) - \mu_{ikc} \right]^2 + \sigma_{ikc}^2 + \eta(x, z_k) - \mu_{ik\eta} \right\} \partial x}{\sqrt{\sigma_{ikc}^2 + \mu_{ik\eta}}} \right] \tag{4.4}$$

Now, differentiating partially equation (4.4) w.r.t x_h , we have

$$\frac{f(x_h, z_k)\left[c(x_h, z_k) - \mu_{hkc} \right]^2 + \sigma_{hkc}^2 + \eta(x_h, z_k) - \mu_{hk\eta}}{\sqrt{\sigma_{hkc}^2 + \mu_{hk\eta}}} = \frac{f(x_h, z_k)\left[c(x_h, z_k) - \mu_{ijc} \right]^2 + \sigma_{ijc}^2 + \eta(x_h, z_k) - \mu_{ij\eta}}{\sqrt{\sigma_{ijc}^2 + \mu_{ij\eta}}} \tag{4.5}$$

where i=h+1,j=k+1,h=1,2,…,L and k=1,2,…,M

Here, it should be noted that the system of equations that we can obtain from (4.5) gives the stratification points $\left[x_h, z_k\right]$ which correspond to the minimum of the variance $V\left(\bar{y}_{st}\right)_{opt}$ if the function

$$\lambda(x, z) = f(x, z) \frac{\left[4\eta(x, z) c'^2(x, z) + \eta'^2(x, z) \right]}{\left[\eta(x, z) \right]^{\frac{3}{2}}}$$

belongs to the class of Ω of functions for x in [a, b] and z in [c, d] and if

$$\lambda(x, z) \in \Omega \quad \forall x \in [a, b] \quad \& \quad z \in [c, d]$$

then the system of equations (4.5) gives $\left[x_h, z_k\right]$ that minimizes the variance $V\left(\bar{y}_{st}\right)_{opt}$.

These equations on solving give the optimum points of stratification $\left[x_h, z_k\right]$.As the parameter involved (4.5) are themselves function of x_h *and* z_k ,the exact solutions are, therefore, very difficult to obtain. Due to this difficulty, it becomes extremely desirable to find some approximate solutions of this system of equations and then use some iteration procedure to obtain better approximations to $\left[x_h, z_k\right]$,if so desired.

Lemma 1: If the function $I_{ij}(x, z)$ is defined as

$$I_{ij}(x, z) = \int_{z_1}^{z_2} \int_{x_1}^{x_2} (t_1 - x_1)^i (t_2 - z_1)^j f(t_1, t_2) \partial t_1 \partial t_2 \quad , \qquad x_1 < x_2 \ \& \ z_1 < z_2$$

where $f(t_1, t_2)$ is a function of two variables, then

$$I_{ij}(x, z) = \left[\begin{array}{l} \dfrac{k_1^{i+1} k_2^{j+1}}{(i+1)(j+1)} f + \dfrac{k_1^{i+1} k_2^{j+1}}{(i+1)(j+1)} f_x + \dfrac{k_1^{i+1} k_2^{j+2}}{(i+1)(j+2)} f_z \\ + \dfrac{1}{2!} \left[\dfrac{k_1^{i+3} k_2^{j+1}}{(i+3)(j+1)} f_{xx} + 2 \dfrac{k_1^{i+2} k_2^{j+2}}{(i+2)(j+2)} f_{xz} + \dfrac{k_1^{i+1} k_2^{j+3}}{(i+1)(j+3)} f_{zz} \right] + O\left(k^{i+j+5}\right) \end{array} \right] \tag{4.6}$$

where $f\left(t_1,t_2\right)=f,\frac{\partial f}{\partial t_1}=f_x,\frac{\partial f}{\partial t_2}=f_z,\frac{\partial^2 f}{\partial t_1^2}=f_{xx},\frac{\partial^2 f}{\partial t_2^2}=f_{zz},\frac{\partial^2 f}{\partial t_1\partial t_2}=f_{xz}$

and k_1 , k_2 denote the ranges of 1st and 2nd strata.

Lemma 2: Let $\mu_\eta\left(x,z\right)$ denotes the conditional expectation of the function $\eta\left(t_1,t_2\right)$,so that

$$\mu_\eta\left(x,z\right)=\frac{\int_{z_1}^{z_2}\int_{x_1}^{x_2}\eta\left(t_1,t_2\right)f\left(t_1,t_2\right)\partial t_1\partial t_2}{\int_{z_1}^{z_2}\int_{x_1}^{x_2}f\left(t_1,t_2\right)\partial t_1\partial t_2}$$

Then, the series expansion of $\mu_\eta\left(x,z\right)$at point (t_1 , t_2)is given by

$$\mu_\eta\left(x,z\right)=\eta\begin{bmatrix}1+\frac{\eta'}{2\eta}\left(k_1+k_2\right)+\left(\frac{\eta'\left(f_x+f_z\right)+2f\eta''}{12f\eta}\right)\left(k_1+k_2\right)^2\\+\left(\frac{f\left(f_{xx}+f_{zz}+f_{xz}\right)\eta'+f\left(f_x+f_z\right)\eta''+f^2\eta'''-\eta'\left(f_x+f_z\right)^2}{12f^2\eta}\right)\left(k_1+k_2\right)^3\\+O\left(\left(k_1+k_2\right)^4\right)\end{bmatrix}\tag{4.7}$$

Lemma 3: If $\sigma_\eta^2\left(x,z\right)$ denotes the conditional variance of the function $\eta\left(t_1,t_2\right)$ in the interval (x, z) ,so that $\sigma_\eta^2\left(x,z\right)=\mu_\eta^2\left(x,z\right)-\left(\mu_\eta\left(x,z\right)\right)^2$. Then,

$$s_h^2\left(x,z\right)=\frac{\left(k_1+k_2\right)^2}{12}\left(h'\right)^2\left[1+\frac{h''}{h'}\left(k_1+k_2\right)+O\left(k_1+k_2\right)^2\right]\quad=\quad\frac{\left(k\right)^2}{12}\left(h'\right)^2\left[1+\frac{h''}{h'}\left(k\right)^1+O\left(k\right)^2\right]\tag{4.8}$$

where $\left(k\right)^1$ *and* $\left(k\right)^2$ denote all k_i's with power '1' and '2' respectively.

Note: If we take the function of $\eta\left(t_1,t_2\right)=\left(t_1,t_2\right)$,we obtain

$$\sigma_{t_1,t_2}^2\left(x,z\right)=\frac{\left(k\right)^2}{12}\left[1+O\left(k\right)^2\right]\tag{4.9}$$

$$\Rightarrow k=\sigma_\eta\left(x,z\right)\sqrt{12}\left(1+O\left(k\right)^2\right)\tag{4.10}$$

Lemma 4:

$$I_{00}\left(x,z\right)\sqrt{\sigma_c^2\left(x,z\right)+\mu_\eta\left(x,z\right)}=\int_{z_1}^{z_2}\int_{x_1}^{x_2}\sqrt{\eta\left(t_1,t_2\right)}f\left(t_1,t_2\right)\partial t_1\partial t_1+n_1\left[1+O\left(k_1+k_2\right)^2\right]\tag{4.11}$$

Where $n_1=\frac{\left(k_1+k_2\right)^2}{96}\int_{z_1}^{z_2}\int_{x_1}^{x_2}\left(\frac{4\eta c'^2+\eta'^2}{\sqrt[3]{\eta}}\right)_{\left(t_1,t_2\right)}f\left(t_1,t_2\right)\partial t_1\partial t_1$

5. MINIMAL EQUATIONS AND THEIR APPROXIMATE SOLUTIONS

In this section, we shall find the series expansion of the system of equations given in (4.5) about the point $\left(x_h,z_k\right)$, the common boundary points of $\left(h,k\right)^{th}$ and $\left(h+1,k+1\right)^{th}$ strata in order to obtain their approximate solutions. To find the expansion of (4.5) we shall use the relations obtained in different

Lemmas by replacing (x,z) by $((x_{h-1},x_h),(z_{k-1},z_k))$. Let us consider the development of right hand side. The corresponding expansion for the L.H.S of the equation can be obtained from the expansion of the right hand side by merely changing the signs of the coefficients of odd powers of k_i *and* k_j ,where $k_i = x_{h+1} - x_h$ and $k_j = z_{k+1} - z_k$, although the same result will be obtained if we develop the expansion of this side independently.

We have (4.6.4) after replacing x and z by x_h *and* x_{h+1} , respectively, we get

$$\mu_{ikc} = c\left[1+\frac{c'}{2c}k_i+\left(\frac{c'f_x+2fc''}{12fc}\right)k_i^2+\left(\frac{ff_{xx}c'+f^2c'''-c'f_x^2}{24f^2c}\right)k_i^3+O(k_i)^4\right]$$

Similarly, replacing x and z by z_k *and* z_{k+1} respectively, we get

$$\mu_{hjc} = c\left[1+\frac{c'}{2c}k_j+\left(\frac{c'f_z+2fc''}{12fc}\right)k_j^2+\left(\frac{ff_{zz}c'+f^2c'''-c'f_z^2}{24f^2c}\right)k_j^3+O(k_j)^4\right]$$

where the functions c, f and their derivatives like f_x, f_z, f_{xx}, f_{zz} etc are evaluated at x_h *and* z_k .Therefore, we get

$$\mu_{ikc} - c(x_h,z) = c\left[\frac{c'}{2c}k_i+\left(\frac{c'f_z+2fc''}{12fc}\right)k_i^2+\left(\frac{ff_{zz}c''+f^2c'''-c'f_z^2}{24f^2c}\right)k_i^3+O(k_i)^4\right]$$

$$\left[\mu_{ikc} - c(x_h,z)\right]^2 = \left(\frac{c'k_i}{2}\right)^2\left[\begin{array}{l}1+2\left\{\left(\frac{c'f_x+2fc''}{6fc'}\right)k_i+\left(\frac{ff_{xx}+ff_xc''+f^2c'''-c'f_x^2}{12f^2c'}\right)k_i^2\right\}\\ +\left(\frac{\left(c'f_x+2fc''\right)^2}{36f^2c'^2}\right)k_i^2+O\left(k_i^3\right)\end{array}\right] \quad (5.1)$$

where $f_5 = \dfrac{6f_{xx}c'^2+10ff_xc'c''+6f^2c'c'''-5f_x^2c'^2+4f^2c''^2}{36f^2}$

Also (4.8) can be written as

$$\sigma_{ikc}^2 = \frac{k_i^2}{4}\left[\frac{c'}{3}+\frac{c'c''}{3}k_i+O\left(k_i^2\right)\right] \quad (5.2)$$

Adding (5.1) and (5.2) ,we get

$$\left[\mu_{hjc} - c(x_h,z)\right]^2+\sigma_{ikc}^2 = \frac{k_i^2}{12}\left[4c'^2+\left(\frac{c'f_x+3fc'c''}{f}\right)k_i+O\left(k_i^2\right)\right]$$

Also from (4.7) ,we can write as

$$\eta(x_h,z)+\mu_{ik\eta} = \eta\left[2+\frac{\eta'}{2\eta}k_i+\left(\frac{\eta'f_x+2f\eta''}{12f\eta}\right)k_i^2+\left(\frac{ff_{xx}\eta'+ff_x\eta''+f^2\eta'''-f_x^2\eta'}{24f^2\eta}\right)k_i^3\right]+O\left(k_i^4\right)$$

where, on the right hand side of this equation the functions f,η and their derivatives are evaluated at point x_h .However, if they are evaluated at z_k it would take the form as

$$\eta(x,z_k)+\mu_{hj\eta}$$
$$=\eta\left[2+\frac{\eta'}{2\eta}k_j+\left(\frac{\eta' f_z+2f\eta''}{12f\eta}\right)k_j^2+\left(\frac{ff_{zz}\eta'+ff_z\eta''+f^2\eta'''-f_z^2\eta'}{24f^2\eta}\right)k_j^3\right]+O\left(k_j^4\right)$$

Thus, we get

$$\left[\mu_{hjc}-c(x_h,z)\right]^2+\sigma_{ikc}^2+\left(\eta(x_h,z)+\mu_{ik\eta}\right)$$
$$=2\eta\left[\begin{matrix}1+\dfrac{\eta'}{4\eta}k_i+\left(\dfrac{4c'^2 f+\eta' f_x+2f\eta''}{24f\eta}\right)k_i^2\\ +\left(\dfrac{2fc'^2 f_x+6f^2c'c''\eta+ff_{xx}\eta'+ff_x\eta''+f^2\eta'''-f_x^2\eta'}{48f^2\eta^2}\right)k_i^3\end{matrix}\right]+O\left(k_i^4\right) \tag{5.3}$$

Also, we have

$$\sigma_{ikc}^2+\mu_{ik\eta}=\eta\left[\begin{matrix}1+\dfrac{\eta'}{2\eta}(k_i)+\left(\dfrac{f_x\eta'+2f\eta''+fc'^2}{12f\eta}\right)(k_i)^2\\ +\left(\dfrac{ff_{xx}\eta'+ff_x\eta''+f^2\eta'''-\eta' f_x^2+2f^2c'c''}{24f^2\eta}\right)(k_i)^3+O(k_i)^4\end{matrix}\right]$$

so that

$$\left(\sigma_{ikc}^2+\mu_{ik\eta}\right)^{-\frac{1}{2}}=\frac{1}{\sqrt{\eta}}\left[1-\frac{\eta'}{4\eta}k_i+\frac{\left(9f\eta'^2-4f_x\eta\eta'-8\eta\eta''-4f\eta c'^2\right)}{96f\eta^2}k_i^2-d_1k_i^3+O(k_i)^4\right] \tag{5.4}$$

where

$$d_1=\frac{\left(\begin{matrix}8f\eta^2 f_{xx}\eta'+8f\eta^2 f_x\eta''+8f^2\eta^2\eta'''-8\eta^2 f_x^2\eta'+16f^2\eta^2c'c''\\ -12f\eta f_x\eta'^2-24f^2\eta\eta'\eta''-12f^2\eta\eta' c'^2+15f^2\eta'^3\end{matrix}\right)}{384f^2\eta^3}$$

Now, using the expression in (5.3) and (5.4) ,and on multiplying them, we get

$$\frac{\left[\mu_{ikc}-c(x_h,z)\right]^2+\sigma_{ikc}^2+\left(\eta(x_h,z)+\mu_{ik\eta}\right)}{\left(\sigma_{ikc}^2+\mu_{ik\eta}\right)^{\frac{1}{2}}}=2\sqrt{\eta}\left[1+(M_1)k_i^2+(M_2)k_i^3+O\left(k_i^4\right)\right]$$

where

$$M_1=\frac{4\eta c'^2+\eta'^2}{32\eta^2}$$

$$M_2=\frac{8f_x\eta^2c'^2+16f\eta^2c'c''+2\eta f_x\eta'^2+4f\eta\eta'\eta''-4f\eta\eta' c'^2-3f\eta'^2}{192f\eta^3}$$

that can be further expressed as

$$M_2=\frac{1}{96f\sqrt{\eta}}\frac{\partial}{\partial x_h}\left(\frac{4f\eta c'^2+f\eta'^2}{\eta^{\frac{3}{2}}}\right)$$

However, the R.H.S of the equation obtained in (4.3) can be expressed as

$$\frac{\left[\mu_{ikc}-c\left(x_h,z\right)\right]^2+\sigma_{ikc}^2+\left(\eta\left(x_h,z\right)+\mu_{ik\eta}\right)}{\left(\sigma_{ikc}^2+\mu_{ik\eta}\right)^{\frac{1}{2}}}$$

$$=2\sqrt{\eta}\left[1+\left(\frac{4\eta c'^2+\eta'^2}{32\eta^2}\right)k_i^2+\frac{k_i^3}{96f\sqrt{\eta}}\frac{\partial}{\partial x_h}\left(\frac{4f\eta c'^2+f\eta'^2}{\eta^{\frac{3}{2}}}\right)+O\left(k_i^4\right)\right] \tag{5.5}$$

In the similar way the expansion of L.H.S of the same equation can be obtained. The expansion is given by

$$\frac{\left[\mu_{hkc}-c\left(x_h,z\right)\right]^2+\sigma_{hkc}^2+\left(\eta\left(x_h,z\right)+\mu_{hk\eta}\right)}{\left(\sigma_{hkc}^2+\mu_{hk\eta}\right)^{\frac{1}{2}}}$$

$$=2\sqrt{\eta}\left[1+\left(\frac{4\eta c'^2+\eta'^2}{32\eta^2}\right)k_h^2-\frac{k_h^3}{96f\sqrt{\eta}}\frac{\partial}{\partial x_h}\left(\frac{4f\eta c'^2+f\eta'^2}{\eta^{\frac{3}{2}}}\right)+O\left(k_h^4\right)\right] \tag{5.6}$$

where the function f,η,c f and their derivatives are evaluated at x_h.

Similarly, we would get the results when the functions f,η,c and their derivatives are evaluated at z_k.
Thus equations (5.5) and (5.6) can take the form as

$$\frac{\left[\mu_{hjc}-c\left(x,z_k\right)\right]^2+\sigma_{hjc}^2+\left(\eta\left(x,z_k\right)+\mu_{hj\eta}\right)}{\left(\sigma_{hjc}^2+\mu_{hj\eta}\right)^{\frac{1}{2}}}=2\sqrt{\eta}\left[1+\left(\frac{4\eta c'^2+\eta'^2}{32\eta^2}\right)k_j^2+\frac{k_j^3}{96f\sqrt{\eta}}\frac{\partial}{\partial z_k}\left(\frac{4f\eta c'^2+f\eta'^2}{\eta^{\frac{3}{2}}}\right)+O\left(k_j^4\right)\right] \tag{5.7}$$

and

$$\frac{\left[\mu_{hkc}-c\left(x,z_k\right)\right]^2+\sigma_{hkc}^2+\left(\eta\left(x,z_k\right)+\mu_{hk\eta}\right)}{\left(\sigma_{hkc}^2+\mu_{hk\eta}\right)^{\frac{1}{2}}}=2\sqrt{\eta}\left[1+\left(\frac{4\eta c'^2+\eta'^2}{32\eta^2}\right)k_k^2-\frac{k_k^3}{96f\sqrt{\eta}}\frac{\partial}{\partial z_k}\left(\frac{4f\eta c'^2+f\eta'^2}{\eta^{\frac{3}{2}}}\right)+O\left(k_k^4\right)\right] \tag{5.8}$$

The equation (4.3) ,after cancelling the common terms on both the sides and multiplying both sides by $f\left(x_h,z\right)$,can be put as

$$k_h^2\left[\left(\frac{4\eta c'+f\eta'^2}{\eta^2}\right)-\frac{k_h}{3}\frac{\partial}{\partial x_h}\left(\frac{4f\eta c'^2+f\eta'^2}{\eta^{\frac{3}{2}}}\right)+O\left(k_h\right)^2\right]=k_i^2\left[\left(\frac{4\eta c'+f\eta'^2}{\eta^2}\right)+\frac{k_i}{3}\frac{\partial}{\partial x_h}\left(\frac{4f\eta c'^2+f\eta'^2}{\eta^{\frac{3}{2}}}\right)+O\left(k_i\right)^2\right] \tag{5.9}$$

where i=h+1 and the functions f,η,c and their derivatives are to be taken at x_h.

Similarly when the functions f,η,c and their derivatives are evaluated at z_k then we have (4.3) by substituted values obtained in (5.7) and (5.8) as

$$k_k^2\left[\left(\frac{4\eta c'+f\eta'^2}{\eta^2}\right)-\frac{k_k}{3}\frac{\partial}{\partial z_k}\left(\frac{4f\eta c'^2+f\eta'^2}{\eta^{\frac{3}{2}}}\right)+O\left(k_k\right)^2\right]=k_j^2\left[\left(\frac{4\eta c'+f\eta'^2}{\eta^2}\right)+\frac{k_j}{3}\frac{\partial}{\partial z_k}\left(\frac{4f\eta c'^2+f\eta'^2}{\eta^{\frac{3}{2}}}\right)+O\left(k_j\right)^2\right] \tag{5.10}$$

Combining (5.9) and (5.10) and simplifying or in other words the functions f,η,c and their derivatives are evaluated at x_h and z_k, we get the results as

$$\left(\frac{f\eta'^2+4f\eta c'^2}{\eta^{\frac{3}{2}}}\right)\left[\left(k_h k_k\right)^2\int_{x_{h-1}}^{x_h}\int_{z_{k-1}}^{z_k} m_1(t_1,t_2)\, f(t_1,t_2)\,\partial t_1\partial t_2\left[1+O\left(k_h k_k\right)^2\right]\right]^{\frac{2}{3}} \tag{5.11}$$

$$=\left(\frac{f\eta'^2+4f\eta c'^2}{\eta^{\frac{3}{2}}}\right)\left[\left(k_i k_j\right)^2\int_{x_h}^{x_{h+1}}\int_{z_k}^{z_{k+1}} m_1(t_1,t_2)\, f(t_1,t_2)\,\partial t_1\partial t_2\left[1+O\left(k_i k_j\right)^2\right]\right]^{\frac{2}{3}}$$

where

$$m_1(t_1,t_2)=\frac{\eta'^2(t_1,t_2)+4\eta(t_1,t_2)c'^2(t_1,t_2)}{\left(\eta(t_1,t_2)\right)^{\frac{3}{2}}} \tag{5.12}$$

$i=h+1, h=1,2,...,L$

$j=k+1, k=1,2,...,M$

and we have assumed that the function $f(x,z), m_1(x,z)\in\Omega\,\forall x\in[a,b], z\in[c,d]$

If we have large number of strata so that the strata widths k_h *and* k_k are small then the higher powers of the widths can be neglected and the system of minimal equations given in (4.3) can be written as

$$\left[\left(k_h k_k\right)^2\int_{x_{h-1}}^{x_h}\int_{z_{k-1}}^{z_k} m_1(t_1,t_2)\, f(t_1,t_2)\,\partial t_1\partial t_2\right]^{\frac{2}{3}}=\left[\left(k_i k_j\right)^2\int_{x_h}^{x_{h+1}}\int_{z_k}^{z_{k+1}} m_1(t_1,t_2)\, f(t_1,t_2)\,\partial t_1\partial t_2\right]^{\frac{2}{3}} \tag{5.13}$$

In other words it can be written that

$$\left(k_h k_k\right)^2\int_{x_{h-1}}^{x_h}\int_{z_{k-1}}^{z_k} m_1(t_1,t_2)\, f(t_1,t_2)\,\partial t_1\partial t_2 = \text{constant} \tag{5.14}$$

where terms of $O\left(\underset{(a,b)}{Sup}\left(k_h\right)\right)^4$ *and* $O\left(\underset{(c,d)}{Sup}\left(k_k\right)\right)^2$ have been neglected on both the sides of the equation. Since $\int_{x_{h-1}}^{x_h}\int_{z_{k-1}}^{z_k} m_1(t_1,t_2)\, f(t_1,t_2)\,\partial t_1\partial t_2 = O\left(\underset{((a,b),(c,d))}{Sup}\left(k_h k_k\right)\right)$ in view of the fact that $\forall x\in[a,b], z\in[c,d], 0<m_1(x,z)\, f(x,z)$. It can be seen from (5.13) that, if we have a function $P_1(x_{h-1},x_h,z_{k-1},z_k)$ is such that

$$\left(k_h k_k\right)\int_{x_{h-1}}^{x_h}\int_{z_{k-1}}^{z_k} m_1(t_1,t_2)\, f(t_1,t_2)\,\partial t_1\partial t_2 = P_1(x_{h-1},x_h,z_{k-1},z_k)\left[1+O\left(k_i k_j\right)^2\right]$$

and $P_1(x_{h-1},x_h,z_{k-1},z_k)$ is of order $O\left(\underset{((a,b),(c,d))}{Sup}\left(k_h k_j\right)\right)^3$,then the minimal equations (4.3) can ,to the same degree of approximation as involved in (5.13), take the form as

$$Q_1(x_{h-1},x_h,z_{k-1},z_k) = \text{Constant, h=1,2,..,L,k=1,2,...,M} \tag{5.15}$$

The solutions to the sets of equations (4.3) as an approximation to optimum $[x_h, z_k]$ can be obtained with the help of some iterative procedure where there approximate solutions can be taken as the starting points.

Theorem 1: If the regression of the estimation variable Y on the stratification variables X and Z, in the infinite super population, is given by $y=c(x,z)+e$ where 'e' is the error term such that

$E\left(e/_{x,z}\right)=0$ & $V\left(e/_{x,z}\right)=\eta(x,z)>0\ \forall x\in(a,b), z\in(c,d)$ with non-zero deviation of intervals, and further if the function $m_1(t_1,t_2)\, f(t_1,t_2)\in\Omega$,then the system of equations (4.5) giving strata boundaries $[x_h, z_k]$ which correspond to the minimum of $V\left(\bar{y}_{st}\right)_{opt}$ can be written as

$$\left[\left(k_hk_k\right)^2\int_{x_{h-1}}^{x_h}\int_{z_{k-1}}^{z_k} m_1(t_1,t_2)\, f\,(t_1,t_2)\partial t_1\partial t_2.\left[1+O\left(k_hk_k\right)^2\right]\right]^{\frac{2}{3}} = \left[\left(k_i k_j\right)^2\int_{x_h}^{x_{h+1}}\int_{z_k}^{z_{k+1}} m_1(t_1,t_2)\, f\,(t_1,t_2)\partial t_1\partial t_2.\left[1+O\left(k_i k_j\right)^2\right]\right]^{\frac{2}{3}}$$

while neglecting the terms of order $O\left(\underset{((a,b),(c,d))}{Sup}\left(k_hk_j\right)\right)^4$,these equations can be replaced by the approximate system of equations

$$\left(k_hk_k\right)^2\int_{x_{h-1}}^{x_h}\int_{z_{k-1}}^{z_k} m_1(t_1,t_2)\, f\,(t_1,t_2)\partial t_1\partial t_2 = \text{Constant}$$

or equivalently equal to

$$Q_1\left(x_{h-1},x_h,z_{k-1},z_k\right) = \text{constant}$$

where

$$m_1(x,z) = \left(\frac{\eta'^2 + 4\eta c'^2}{\eta^{\frac{3}{2}}}\right)_{(x,z)}$$

$$k_h = x_h - x_{h-1}$$

$$k_k = z_k - z_{k-1}$$

$$Q_1\left(x_{h-1},x_h,z_{k-1},z_k\right)\left[1+O\left(k_hk_k\right)^2\right] = \left(k_hk_k\right)^2\int_{x_{h-1}}^{x_h}\int_{z_{k-1}}^{z_k} m_1(t_1,t_2)\, f\,(t_1,t_2)\partial t_1\partial t_2$$

and

$i = h+1, h = 1,2,...,L$

$j = k+1, k = 1,2,...,M$

From Lemma 4 with x and z replaced by $\left(x_{h-1},x_h\right)$ *and* $\left(z_{k-1},z_k\right)$ respectively, so that $k = k_h, k = k_k$ *and* $I_{00}(x,z) = W_{hk}$, we obtain

$$W_{hk}\sqrt{\sigma_{hkc}^2 + \mu_{hk\eta}}$$

$$= \int_{x_{h-1}}^{x_h}\int_{z_{k-1}}^{z_k}\sqrt{\eta(t_1,t_2)}\, f\,(t_1,t_2)\partial t_1\partial t_2 + \frac{\left(k_hk_k\right)^2}{96}\int_{x_{h-1}}^{x_h}\int_{z_{k-1}}^{z_k} m_1(t_1,t_2)\, f\,(t_1,t_2)\partial t_1\partial t_2\left[1+O\left(k_hk_k\right)^2\right]$$

Now, taking summation over all strata we have

$$\sum_h\sum_k W_{hk}\sqrt{\sigma_{hkc}^2 + \mu_{hk\eta}}$$

$$= \int_a^b\int_c^d\sqrt{\eta(t_1,t_2)}\, f\,(t_1,t_2)\partial t_1\partial t_2 + \sum_h\sum_k\frac{\left(k_hk_k\right)^2}{96}\int_{x_{h-1}}^{x_h}\int_{z_{k-1}}^{z_k} m_1(t_1,t_2)\, f\,(t_1,t_2)\partial t_1\partial t_2\left[1+O\left(k_hk_k\right)^2\right]$$

Since it is obvious that $\int_a^b\int_c^d\sqrt{\eta(t_1,t_2)}\, f\,(t_1,t_2)\partial t_1\partial t_2$ is a constant because of interval defined for the variables. So, minimization of $\sum_h\sum_k W_{hk}\sqrt{\sigma_{hkc}^2 + \mu_{hk\eta}}$ is equivalent to minimization of

$d_1 = \sum_h\sum_k\frac{\left(k_hk_k\right)^2}{96}\int_{x_{h-1}}^{x_h}\int_{z_{k-1}}^{z_k} m_1(t_1,t_2)\, f\,(t_1,t_2)\partial t_1\partial t_2\left[1+O\left(k_hk_k\right)^2\right]$ as the first term is constant in the above equation. Thus, the Theorem 1 can be established alternatively to a large extent by minimizing the function d_1 as discussed by Ekman (1960).

Thus, we find that if the function $m_1(x,z)\, f(x,z)$ belongs to the class Ω then the minimum value of

$\sum_h \sum_k W_{hk} \sqrt{\sigma^2_{hkc} + \mu_{hk\eta}}$ and therefore $V(\bar{y}_{st})_{opt}$ exists and the set of strata boundaries

$[x_h, z_k]$, corresponding to the minimum, are the solutions of the system of equation (4.5) or equivalently of (5.11). These conditions in that capacity are extremely hard to understand and accordingly it is required to discover some approach to conquer this trouble .It is done by replacing these system of equations by other system of equations which are comparatively easier to solve but are only asymptotically equivalent to the exact minimal equations. The error factor is introduced because we neglect the terms of higher powers of strata widths which is of course justifiable if the number of strata is large. These option arrangement of conditions were given in (5.14) and (5.15) We have obtained these systems of equations after neglecting the terms of order $O\left(Sup\left(k_h k_k\right)\right)^4 = O\left(m^4\right)$ where

$m = \left(\underset{((a,b),(c,d))}{Sup} \left(k_h k_k\right) \right)$ on both sides of the equation (5.11).If the number of strata is large and

therefore terms of order $O\left(m^4\right)$ are quite small, the error involved in the approximate systems of equations is expected to be quite small and the set of stratification points $[x_h, z_k]$ obtained from them shall be quite near to the optimum values.

6. **CUM** $\sqrt[3]{D_1(x,z)}$ **RULE**

If the function $D_1(x,z) = m_1(x,z)\, f(x,z)$ where

$$m_1(x,z) = \frac{\eta'^2(x,z) + 4\eta(x,z)c'^2(x,z)}{\left[\eta(x,z)\right]^{\frac{3}{2}}}$$

is bounded and its first derivative exists for all x in $[a,b]$ and z in $[c,d]$,then for a given value of L and M taking equal intervals on the cumulative cube root of $D_1(x,z)$ will give AOSB $[x_h, z_k]$.

Remarks

1. Let $c(x,z) = \alpha + \beta x + \gamma z$ then $c'(x,z) = \beta \quad and \quad c'(x,z) = \gamma$ by differentiating partially w.r.t. $x \quad and \quad z$, respectively, and ultimately

$m_1(x,z) = \frac{\eta'^2(x,z) + 4\eta(x,z)c'^2(x,z)}{\left[\eta(x,z)\right]^{\frac{3}{2}}}$ =constant .Therefore, for such a case the proposed rule

reduces to the $Cum\sqrt[3]{f(x,z)}$.

2. Let $c(x,z) = \alpha + \beta x + \gamma z,$ then if either $c(x,z) = \alpha + \beta x$ or $c(x,z) = \alpha + \gamma z$, then

$$m_1(x) = \frac{\eta'^2(x) + 4\eta(x)c'^2(x)}{\left[\eta(x)\right]^{\frac{3}{2}}} \quad or \quad m_1(z) = \frac{\eta'^2(z) + 4\eta(z)c'^2(z)}{\left[\eta(z)\right]^{\frac{3}{2}}}$$

reduces to the method proposed by Singh and Sukhatme (1969) for single auxiliary variable.

3. For any distribution and given number of strata, the AOSB $\left[x_h, z_k\right]$ will remain unchanged with respect to the form of conditional variance $\eta\left(x,z\right)$, however, the efficiency of stratification as compared to no stratification will be changed with the choice of various forms of conditional variance.

7. EMPIRICAL STUDY

We shall now demonstrate empirically the effectiveness of the proposed method of findings the set of AOSB. The linear regression line Y on X and Z have been taken as $y=\alpha+\beta x+\gamma z+e$. Two forms conditional variance function $\eta\left(x,z\right)$ viz. $\eta\left(x,z\right)=\alpha$ and $\eta\left(x,z\right)=\lambda xz$ have taken, where α *and* λ are constants.

The origin is deliberately excluded from the range of the auxiliary variables X and Z otherwise $\eta\left(x,z\right)=\lambda xz$ we have $m_1\left(x,z\right)=\infty$ at x = 0, z = 0 and the function $m_1\left(x,z\right) f\left(x,z\right)$ in that case does not belong to the class Ω of functions. We could have also avoided this difficulty by taking some other suitable forms to the functions. For the empirical studies under optimum allocation, let us assume values of $\alpha=0.0214, \lambda=0.00437$ which are quite small so that the effect of taking $\eta\left(x,z\right)=\alpha$ and $\eta\left(x,z\right)=\lambda xz$ is negligibly small.

In order to obtain AOSB let us assume that the correlation coefficient (ρ) between X and Z is equal to 0.65 For this purpose the following density functions of the stratification variables X and Z have been considered.

Case I: The auxiliary variable X follows standard normal distribution with pdf as

$$f\left(x\right)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}} \quad ,x\geq 0$$

and the variable Z also follows standard normal distribution with pdf as

$$f\left(z\right)=\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}} \quad ,z\geq 0$$

In order to obtain the OSB when both the variables are standard normally distributed let us take the values of regression coefficients $\beta=0.65$ and $\gamma=0.57$ .For obtaining total 16 strata, 4 along the x-variable and 4 along the z-variable using the proposed rule Cum $\sqrt[3]{D_1\left(x,z\right)}$ by solving it in Mathematica software assuming the distribution of X and Z is truncated at x=6 and z=4 respectively, we get the stratification points as below:

Table 7.1: OSB when the auxiliary variables X and Z have both standard normal distribution and when $\eta\left(x,z\right)=\alpha$

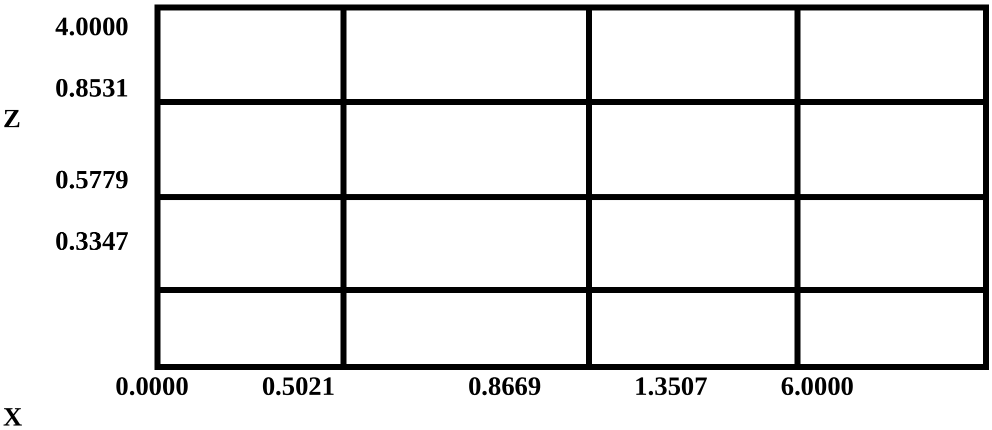

Table 7.2: OSB and Variance when the auxiliary variables are both standard normally distributed having $\eta\left(x,z\right)=\alpha$

OSB(x_h, z_k)	Variance Cum $\sqrt[3]{D_1(x,z)}$ Rule	Variance (Singh and Sukhatme 1969)	% R.E.
(0.5021,0.3347)	0.16254785	0.396418	243.8867
(0.8669,0.3347)			
(1.3507,0.3347)			
(6.0000,0.3347)			
(0.5021,0.5779)			
(0.8669,0.5779)			
(1.3507,0.5779)			
(6.0000,0.5779)			
(0.5021,0.8531)			
(0.8669,0.8531)			
(1.3507,0.8531)			
(6.0000,0.8531)			
(0.5021,4.0000)			
(0.8669,4.0000)			
(1.3507,4.0000)			
(6.0000,4.0000)			

Table 7.3: OSB when the auxiliary variables X and Z have both standard normal distribution respectively and when $\eta(x,z) = \lambda xz$

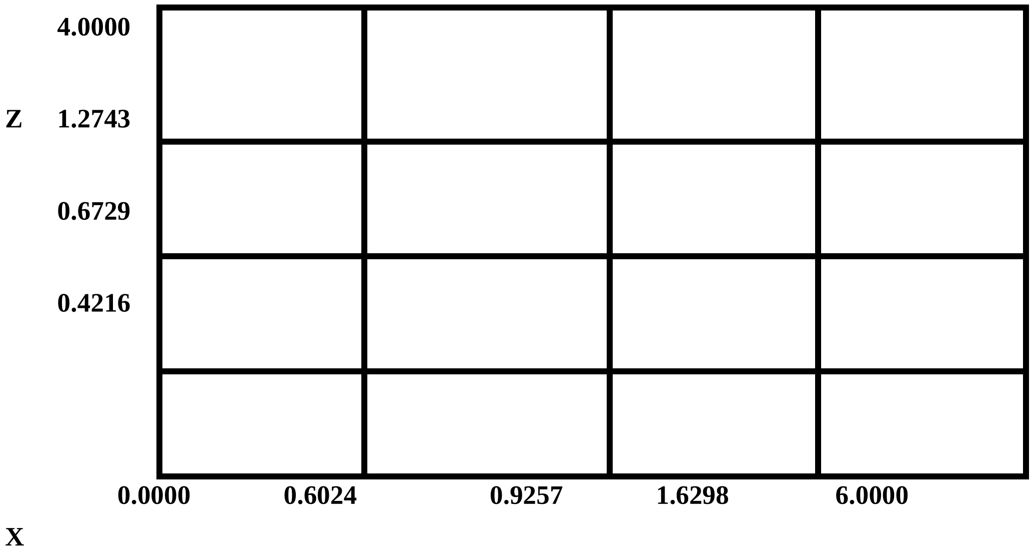

Table 7.4: OSB and Variance when the auxiliary variables are both standard normally distributed having $\eta(x,z) = \lambda xz$

OSB(x_h, z_k)	Variance (Cum $\sqrt[3]{D_1(x,z)}$ Rule)	Variance (Singh and Sukhatme 1969)	% R.E.
(0.6024,0.4216)	0.15942873	0.428619	268.84677
(0.9257,0.4216)			
(1.6298,0.4216)			
(6.0000,0.4216)			
(0.6024,0.6729)			
(0.9257,0.6729)			
(1.6298,0.6729)			
(6.0000,0.6729)			
(0.6024,1.2743)			
(0.9257,1.2743)			
(1.6298,1.2743)			
(6.0000,1.2743)			
(0.6024,4.0000)			
(0.9257,4.0000)			
(1.6298,4.0000)			
(6.0000,4.0000			

Case II: Let us consider the case when of X follows uniform distribution with pdf

$$f(x) = \frac{1}{b-a}, a \le x \le b$$

and Z follows exponential distribution with pdf

$$f(z)=e^{-z+1}, z \geq 0$$

In order to obtain OSB for the above pdf let us suppose that the variable x is defined in [1,2] and z in [1,6] and assume that values of β *and* γ be 0.56 and 0.72, respectively. Taking the above pdf's for constructing stratification points using Cum $\sqrt[3]{D_1(x,z}$ Rule for total 6 strata i,e 2 strata along x-variable and 3 along z-variable, using Mathematica software for solving the function , we get

Table 7.5: OSB and Variance when the auxiliary variables are uniformly and exponentially distributed having when $\eta(x,z)=\alpha$

OSB (x_h, z_k)	**Variance Singh and Sukhatme (1969)**	**Variance** Cum $\sqrt[3]{D_1(x,z)}$ Rule	**% R.E.**
(1.46099,1.77003) (2.00000,1.77003) (1.46099,4.07294) (2.00000,4.07294) (1.46099,6.00000) (2.00000,6.00000)	0.1553	0.096473	160.9776

Table 7.6: OSB and Variance when the auxiliary variables are uniformly and exponentially distributed having $\eta(x,z)=\lambda xz$

OSB (x_h, z_k)	**Total Variance (Singh and Sukhatme 1969)**	**Total Variance** (Cum $\sqrt[3]{D_1(x,z)}$ Rule)	**% R.E**
(1.5000,2.7856) (2.0000,2.7856) (1.5000,3.17044) (2.0000,3.17044) (1.5000,6.0000) (2.0000,6.0000)	0.173	0.0627351	275.7626

8. CONCLUSION

The proposed cumulative cube root rule can be used for determination of strata boundaries for single estimation variable using two auxiliary variables as the basis of stratification under optimum allocation. Furthermore, the percent relative efficiency of the proposed method over the method developed by Singh and Sukhatme (1969) for univariate case, shows that the proposed method is more efficient than Singh and Sukhatme (1969).The proposed rule can be applied in actual practice when the frequency distributions of the auxiliary variables are available and the form the conditional variance function is known from the previous experience.

RECEIVED: SEPTEMBER, 2019.
REVISED: DECEMBER, 2019.

REFERENCES

[1] ALLENDE-ALONSO C. M. y C. N. BOUZA-HERRERA (1987): Optimization Criteria For Multivariate Strata Construction. In "**Approximation En Optimization**", 120-138, Lecture Notes In Mathematics 1353, Springer, N. York.

[2] DALENIUS, T. (1950). The problem of optimum stratification. **Skandinavisk Aktuarietidskrift** , 33, 203-213.

[3] DANISH, F., RIZVI, S.E.H. JEELANI and M. I REASHI J.A. (2017). Obtaining Strata Boundaries under Proportional Allocation with Varying Cost of Every Unit**. Pakistan Journal Of Statistics and Operations Research,** 13, 567-574

[4] KOZAK, M. (2004): Optimal Stratification using Random search Method in Agricultural Surveys. **Statistics in Transition**, 6, 797-806.

[5] LAVALLEE, P. and HIDIRIGLOU, M. (1988). On the stratification of skewed populations. **Survey Methodology**, 14, 33-43.

[6] NIEMIRO, W. (1999). Konstrukcja optymalnej stratyfikacja metodaposzukiwan losowych. (Optimal Stratification using Random Search Method). **Wiadomosci Statystyczne**, 10, 1-9.
[7] RIZVI , S.E.H. GUPTA, J.P. and BHARGAVA, M. (2002). Optimum stratification based on auxiliary variable for compromise allocation. **Metron**, 28, 201–215.
[8] SINGH, R. and SUKHATME, B.V. (1969). Optimum stratification. **Annals of Institute of Statistical Mathematrica,** 21, 515-528.