International Journal of Modern Mathematical Sciences, 2017, 15(2): 248-260

Modern Scientific Press

*International Journal of Modern Mathematical Sciences*

Journal homepage: www.ModernScientificPress.com/Journals/ijmms.aspx

**ISSN: 2166-286X**
*Florida, USA*

*Article*

# Optimum Stratification for Bi-variate Stratification Variables with Single Study Variable

**Faizan Danish and S.E.H.Rizvi**

*Division of Statistics and Computer Science, Faculty of Basic Sciences, SKUAST-Jammu*

*Author to whom correspondence should be addressed; E-Mail: danishstat@gmail.com

*Article history:* Received 1 May 2017; Revised 26 May 2017; Accepted 31 May 2017; Published 1 June 2017.

**Abstract:** Most of the literature on survey sampling deals with obtaining optimum strata boundaries is based on only one stratifying variable. In this paper the problem of obtaining optimum strata boundaries when we have two concomitant variables with one estimation variable and the regression line between them is assumed to be linear. Neyman allocation procedure has been made for obtaining optimum strata boundaries from minimal equations. Due to complexities of minimal equations,we are supposed to have approximate to the variance of the study variable. This approximation depends only on the number of strata, the simultaneous density of stratifying variables and the correlation between the study variable and each of the stratifying variables. Numerical illustration has been done when the stratifying variables follow some particular distributions.

**Keywords**: Optimum Strata Boundaries, Minimal Equation, Neyman Allocation, Auxiliary information

**Mathematics Subject Classification 2010**: 62D05

## 1. Introduction

The principal reason for stratification in the design of sample survey is to reduce the sample variance of the estimates. In this type of sampling the whole population is partitioned into a number of strata and from each of the stratum a sample is selected by using the desired sampling design. The strata are made in such a way that the stratum is homogenous within itself and are as much as heterogeneous as possible between itself. The main factors that influence the reduction of variance to a very large extent are like choice of the variable on the basis of stratification will be one, the total number of strata that should be made from the whole population while partitioning, determination of the stratum boundaries is one the most important factor that influences the variation and the last but not the least is the design of sampling used in each stratum for selecting a sample from it. The effect of stratification using one stratifying variable has received considerable attention from research point of view. The problem of obtaining optimum strata boundaries(OSB) on a single study variable for Neyman and Proportional allocation methods was first considered by Dalenius [ 1 ] who obtained the minimal equations which then gave the OSB as their solutions. Since the equation obtained there, were implicit in nature due to which their exact solutions could not be obtained. After this several statisticians worked on it and attempted to find their approximate solutions like Dalenius and Gurney [ 2 ] , Aoyama [ 3 ], Dalenius and Hodges [ 4 ] , Durbin [ 5 ] , Mahalanobis [ 6 ] ), Singh [ 7 ].In most of these investigations of the problem of optimum stratification both the estimation variable and the variable on the basis of which stratification is made are taken to be same. Since the distribution of the estimation variable is rarely known in advance because of this it is desirable to stratify on the basis of some suitable chosen concomitant variable. Auxiliary variable was considered for stratification by Taga [ 8 ] , Singh and Sukhatme( [ 9 ] , [ 10 ], [ 11 ]), Schneeberger and Gollar [ 12 ] , Singh and Prakash [ 13 ] , Singh ([ 14 ], [15] , [16 ]) and Serfling [17] , Rizvi *et al*. [18 ]. Excluding from this, Rivest [19] suggested some iterative procedures to determine OSB. The algorithms require an initial approximation solution to strata and also there is no guarantee that the algorithm which we are used will provide the global minimum in the absence of a suitable approximate initial solution and the variance functions have more than one local minima. Gunning and Horgan [20] proposed an alternative approach to approximate stratification based on a geometric progression and the assumption of uniform stratification variable within strata. Their approach aims to equalize values of the coefficient of variation of stratification variable within strata.

In this paper, we shall take into consideration that the population values of the two stratifying variable are generated from a background of Bivariate distribution. The population values of the study variable are also assumed to be realization of a stochastic background variable and the linear regression is assumed of this variable over stratifying variable. The separate distributions of the auxiliary variables would be taken into consideration. Out of the two auxiliary variables one variable is supposed to follow exponential distribution and the other variable is assumed to be Righ-traingular distributed variable. Numerical illustration has been made of the proposed method.

## 2. Optimum Strata Points

Let there be a finite population consisting N units, about which we want to estimate the total or mean for the characteristics Y under study, using simple random sampling design. In order to have this, we divide the whole population into L × M strata such that the number of units in the (h , k)th stratum is $N_{hk}$ such that

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

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

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

Where 'n' denotes the total sample size taken from all strata. The value of population unit in the (h , k)th stratum be denoted by $y_{hki}$ (i = 1 , 2 , 3,…,.) and then the population total is

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

Since the study variable is denoted by 'Y'. Thus the unbiased estimate of population mean $\overline{y}_N$ is

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

Where $\overline{y}_{hk} = \frac{1}{n_{hk}}\sum_{i=1}^{n_{hk}} y_{hki}$ and $W_{hk} = \frac{N_{hk}}{N}$ denotes the stratum weight for the (h , k)th .

Now let us come towards the question of determining optimum strata boundaries which correspond to the minimum of variance. We assume that the finite population of size N units in a random sample from the infinite super population with same population characteristics as those of the finite population. This assumption enables us to deal with the continuous function of the density function of the estimation variable Y in super population is known. The optimum points of stratification $\left[y_{hk}\right]$ .for the case of optimum allocation method is given by the solution of the following system of equations

$$\left[\frac{\sigma^2_{hky} + \left(y_{hk} - \mu_{hky}\right)^2}{\sigma^2_{hky}} = \frac{\sigma^2_{ijy} + \left(y_{hk} - \mu_{ijy}\right)^2}{\sigma^2_{ijy}}\right]$$

Where $y_{hk}$ = the common boundary points for the (h , k)$^{th}$ and ith strata i = h+1, h+2,…,L-1 j = k+1,k+2,…, M-1 and $\mu_{hky}$ = population mean for Y in (h , k)$^{th}$ stratum

The minimal equations of this like were obtained by [1] for the case were the estimation and stratification variable are same. But in most cases the density function of Y is not known and therefore it is always desirable to obtain the equations giving optimum points of stratification for the auxiliary variables X and Z which are highly related to the study variable Y. This is possible only in most of the cases when the density functions of the auxiliary variables are known. Now we shall obtain the equations the solution to which will give the optimum points of stratification for the auxiliary variable s X and Z. The variance corresponding to these optimum strata is also minimum. For doing this, we shall assume the knowledge of the regression of Y on X and Z and also knowledge of conditional variance function $V\left( y/_{x,z} \right)$. Let the regression model of Y on X and Z be given as

$$Y = C\,(X, Z) + e \tag{2.1}$$

Where C (X, C) is a function of X and Z and 'e' is error term such that

$$E\left( e/_{x,z} \right) = 0. \qquad \forall \;\; x\varepsilon\left(a,b\right)$$

$$V\left( e/_{x,z} \right) = \eta\left(x,z\right) = 0 \qquad , \qquad z\varepsilon\left(c,d\right)$$

b-a<∞

c-d<∞

If the joint density function of (Y, X, Z) in the super population is f(y, x, z), joint marginal of X and Z is f(x, z) and the marginal density function of X and Z are f(x) and f(z) respectively.

Under regression model shown in (2.1), we have

$$W_{hk} = \int_{x_{h-1}}^{x_h} \int_{z_{k-1}}^{z_k} f\left(x,z\right) \partial x \partial z \tag{2.2}$$

is the weight of the $\left(h,k\right)^{th}$ stratum.

$$\mu_{hky} = \mu_{hkc} = \frac{1}{W_{hk}} \int_{x_{h-1}}^{x_h} \int_{z_{k-1}}^{z_k} c\left(x,z\right) f\left(x,z\right) \partial x \partial z \tag{2.3}$$

denotes the mean of the $\left(h,k\right)^{th}$ stratum, and

$$\sigma^2_{hky} = \sigma^2_{hkc} + \mu_{hk\eta} \tag{2.4}$$

Where $\left(x_{h-1}-x_h\right)$ are the boundaries $\left(h,k\right)^{th}$ and $\mu_{hk\eta}$ is the expected value of the function $\eta(x,z)$ in the $\left(h,k\right)^{th}$ stratum and $\sigma^2_{hkc}$ is given as

$$\sigma^2_{hkc}=\frac{1}{W_{hk}}\int_{x_{h-1}}^{x_h}\int_{z_{k-1}}^{z_k}c^2\left(x,z\right)f\left(x,z\right)\partial x\partial z-\left(\mu_{hkc}\right)^2 \tag{2.5}$$

## 3. Minimal Equation

Let [$x_h$ , $z_k$] denote 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 as determined by the optimum points of stratification [$x_h$ , $z_k$] the variance of the estimate $\overline{y}_{st}$ is 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 $V\left(\overline{y}_{st}\right)$ with respect to [$x_h$ , $z_k$].Before deriving the minimal equations ,let us first find out the expression for some partial derivatives which will be helpful in obtaining the equations. We have for the $\left(h,k\right)^{th}$ stratum

`

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

Differentiate it w.r.to $x_h$ and $z_k$, we get

$$W^A_{hk}=\int_{z_{k-1}}^{z_k}f\left(x_h,z\right)\partial z \tag{3.1}$$

and

$$W^B_{hk}=\int_{x_{h-1}}^{x_h}f\left(x,z_k\right)\partial x \tag{3.2}$$

Also

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

By differentiate it w.r.to $x_h$ and $z_k$, we get

$$\mu^A_{hk\eta}=\int_{z_{k-1}}^{z_k}\frac{f\left(x_h,z\right)}{w_{hk}}\left[\eta\left(x_h,z\right)-\mu_{hk\eta}\right]\partial z \tag{3.3}$$

$$\mu_{hk\eta}^{B} = \int_{x_{h-1}}^{x_h} \frac{f(x, z_k)}{w_{hk}} \left[\eta(x, z_k) - \mu_{hk\eta}\right] \partial x \tag{3.4}$$

Similarly

$$\mu_{hkc}^{A} = \int_{z_{k-1}}^{z_k} \frac{f(x_h, z)}{w_{hk}} \left[c(x_h, z) - \mu_{hkc}\right] \partial z \tag{3.5}$$

$$\mu_{hkc}^{B} = \int_{x_{h-1}}^{x_h} \frac{f(x, z_k)}{w_{hk}} \left[c(x, z_k) - \mu_{hkc}\right] \partial x \tag{3.6}$$

$$\sigma_{hkc}^{A} = \frac{1}{w_{hk}} \int_{z_{k-1}}^{z_k} \left\{ f(x_h, z)\left[c(x_h, z) - \mu_{hkc}\right]^2 - \sigma_{hkc}^2 \right\} \partial z \tag{3.7}$$

$$\sigma_{hkc}^{B} = \frac{1}{w_{hk}} \int_{x_{h-1}}^{x_h} \left\{ f(x, z_k)\left[c(x, z_k) - \mu_{hkc}\right]^2 - \sigma_{hkc}^2 \right\} \partial x \tag{3.8}$$

where A and B denotes the derivative with respect to $x_h$ and $z_k$ respectively.

Similarly for the $(i, j)^{th}$ stratum, where i=h+1 and j=k+1 the expression for the corresponding partial derivatives with respect to its lower boundaries ($x_h$ , $z_k$) are obtained as:

$$W_{ik}^{A} = -\int_{z_{k-1}}^{z_k} f(x_h, z) \partial z \tag{3.9}$$

$$W_{hj}^{B} = -\int_{x_{h-1}}^{x_h} f(x, z_k) \partial x \tag{3.10}$$

$$\mu_{ik\eta}^{A} = -\int_{z_{k-1}}^{z_k} \frac{f(x_h, z)}{w_{ik}} \left[\eta(x_h, z) - \mu_{ik\eta}\right] \partial z \tag{3.11}$$

$$\mu_{hj\eta}^{B} = -\int_{x_{h-1}}^{x_h} \frac{f(x, z_k)}{w_{hj}} \left[\eta(x, z_k) - \mu_{hj\eta}\right] \partial x \tag{3.12}$$

$$\mu_{ikc}^{A} = -\int_{z_{k-1}}^{z_k} \frac{f(x_h, z)}{w_{ik}} \left[c(x_h, z) - \mu_{ikc}\right] \partial z \tag{3.13}$$

$$\mu_{hjc}^{B} = -\int_{x_{h-1}}^{x_h} \frac{f(x, z_k)}{w_{hj}} \left[c(x, z_k) - \mu_{hjc}\right] \partial x \tag{3.14}$$

$$\sigma_{ikc}^{A} = -\frac{1}{w_{ik}} \int_{z_{k-1}}^{z_k} \left\{ f(x_h, z)\left[c(x_h, z) - \mu_{ikc}\right]^2 - \sigma_{ikc}^2 \right\} \partial z \tag{3.15}$$

$$\sigma_{hjc}^{B} = -\frac{1}{w_{hj}} \int_{x_{h-1}}^{x_h} \left\{ f\left(x, z_k\right) \left[ c\left(x, z_k\right) - \mu_{hjc} \right]^2 - \sigma_{hjc}^{2} \right\} \partial x \tag{3.16}$$

Having found the relations, we shall obtain the minimal equation for any allocation method. As we know under Neyman allocation the variance of the stratified sample mean is written as

$$V\left(\overline{y}_{st}\right) = \frac{\left[ \sum_{h=1}^{L} \sum_{k=1}^{M} W_{hk} \sigma_{hky} \right]^2}{n} \tag{3.17}$$

If the finite population correction (f.p.c) is ignored, minimizing the right side of expression of (3.17) is equalient to minimize

$$V\left(\overline{y}_{st}\right) = \sum_{h=1}^{L} \sum_{k=1}^{M} W_{hk} \sqrt{\sigma_{hky}^{2}}$$

Using (2.4) in above equation, we have

$$V\left(\overline{y}_{st}\right) = \sum_{h=1}^{L} \sum_{k=1}^{M} W_{hk} \sqrt{\sigma_{hkc}^{2} + \mu_{hk\eta}} \tag{3.18}$$

The minimization of (3.18) is obtained when differentiating to w. r. to $x_h$ and then equation to zero, as below:

$$\frac{\partial}{\partial x_h} \sum_{h=1}^{L} \sum_{k=1}^{M} W_{hk} \sqrt{\sigma_{hkc}^{2} + \mu_{hk\eta}} = 0$$

$$W_{hk} \frac{\partial}{\partial x_h} \left(\sqrt{h}\right) + \sqrt{h} \frac{\partial}{\partial x_h} W_{hk} + W_{ik} \frac{\partial}{\partial x_h} \left(\sqrt{i}\right) + \sqrt{i} \frac{\partial}{\partial x_h} W_{ik} \tag{3.19}$$

where

$$h = \sigma_{hkc}^{2} + \mu_{hk\eta} \quad \text{and} \quad i = \sigma_{ikc}^{2} + \mu_{ik\eta}$$

Now differentiating h and I w.r. to $x_h$ denoted as $h'$ and $i'$ and are given as

$$h' = \int_{z_{k-1}}^{z_k} \frac{f\left(x_h, z\right)}{W_{hk}} \left\{ \left[ c\left(x_h, z\right) - \mu_{hkc} \right]^2 - \sigma_{hkc}^2 + \eta\left(x_h, z\right) - \mu_{hk\eta} \right\} \partial z \tag{3.20}$$

$$i' = -\int_{z_{k-1}}^{z_k} \frac{f\left(x_h, z\right)}{W_{ik}} \left\{ \left[ c\left(x_h, z\right) - \mu_{ikc} \right]^2 - \sigma_{ikc}^2 + \eta\left(x_h, z\right) - \mu_{ik\eta} \right\} \partial z \tag{3.21}$$

Now using the result obtained in (3.21), we set minimal equations as:

$$\frac{\int_{z_{k-1}}^{z_k}\left\{f\left(x_h,z\right)\left[c\left(x_h,z\right)-\mu_{hk\eta}\right]^2+\sigma_{hkc}^2+\eta\left(x_h,z\right)+\mu_{hk\eta}\right\}\partial z}{\sqrt{\sigma_{hkc}^2+\mu_{hk\eta}}}$$
$$=\frac{\int_{z_{k-1}}^{z_k}\left\{f\left(x_h,z\right)\left[c\left(x_h,z\right)-\mu_{ik\eta}\right]^2+\sigma_{ikc}^2+\eta\left(x_h,z\right)+\mu_{ik\eta}\right\}\partial z}{\sqrt{\sigma_{ikc}^2+\mu_{ik\eta}}} \tag{3.22}$$

Similarly, once again in order to obtain more minimal equations for obtaining strata points, we differentiate (3.18) w.r.to $z_k$ and equate to zero, we get

$$\frac{\int_{x_{h-1}}^{x_h}\left\{f\left(x,z_k\right)\left[c\left(x,z_k\right)-\mu_{hk\eta}\right]^2+\sigma_{hkc}^2+\eta\left(x,z_k\right)+\mu_{hk\eta}\right\}\partial x}{\sqrt{\sigma_{hkc}^2+\mu_{hk\eta}}}$$
$$=\frac{\int_{x_{h-1}}^{x_h}\left\{f\left(x,z_k\right)\left[c\left(x,z_k\right)-\mu_{hj\eta}\right]^2+\sigma_{hjc}^2+\eta\left(x,z_k\right)+\mu_{hj\eta}\right\}\partial x}{\sqrt{\sigma_{hjc}^2+\mu_{hj\eta}}} \tag{3.23}$$

Where, i=h+1 and j=k+1

h=1,2,3,…,L and k= 1,2,3,…,M

It should be noted here that the system of equations that can be obtained from (3.22) and (3.23) gives the strata boundaries $\left(x_h, z_k\right)$ which corresponds to the minimum of the variance $V\left(\overline{y}_{st}\right)$ 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 $\Omega$ of function

$\forall$ x $\in[a,b]$ and z $\in[c,d]$

and if $\lambda(x,z)\in\Omega$ $\forall$ x $\in[a,b]$ and z $\in[c,d]$ then the system of equation (3.22) and (3.23)gives the strata points $\left[x_h, z_k\right]$ that minimize the variance $V\left(\overline{y}_{st}\right)$.

If the regression model defined in (2.1) is linear in two variables as

$$y=\alpha+\beta x+\gamma z+e$$

While applying variance to both the sides, we get

$$\sigma_{hky}^2 = \beta^2\sigma_{hkx}^2 + \gamma^2\sigma_{hkz}^2 + \sigma^2 \tag{3.24}$$

Where E(e) = 0 ,V(e)= $\sigma^2$ and the auxiliary variables X and Z are assumed to be uncorrelated with each other and with 'e' too. By substituting value obtained in (3.24) in (3.18), we get

$$\sum_{h=1}^{L}\sum_{k=1}^{M} W_{hk}\sqrt{\beta^2\sigma_{hkx}^2 + \gamma^2\sigma_{hkz}^2 + \sigma^2 + \mu_{hk\eta}} \tag{3.25}$$

$$= \sum_{h=1}^{L}\sum_{k=1}^{M}\sqrt{\beta^2 W_{hk}^2\sigma_{hkx}^2 + \gamma^2 W_{hk}^2\sigma_{hkz}^2 + W_{hk}^2\sigma^2 + W_{hk}^2\mu_{hk\eta}} \tag{3.26}$$

Using the approximation suggested above, we find that

$$\sum_{h=1}^{L}\sum_{k=1}^{M} W_{hk}^2\sigma_{hkx}^2 = \frac{\sum_{h=1}^{L}\sum_{k=1}^{M}\xi_{hk}^2\left[x_h - x_{h-1}\right]^4\left[z_k - z_{k-1}\right]^2}{12}$$

$$= \frac{\sum_{h=1}^{L}\sum_{k=1}^{M}\xi_{hk}^2\xi_h^{-\frac{3}{2}}\xi_k\left[x_h - x_{h-1}\right]\left[z_k - z_{k-1}\right]A_h^3 A_k}{12} \tag{3.27}$$

where

$$A_h = \xi_h^{\frac{1}{2}}\left[x_h - x_{h-1}\right]$$

$$A_h = \xi_h^{\frac{1}{2}}\left[z_k - z_{k-1}\right]$$

where $\xi_h$ and $\xi_k$ are constant value of the marginal densities with the hth stratum of X and jth stratum of Z respectively. The stratum along X and Z are constructed in such a way so that A $_h$ is equal to K ( X )/ h and $A_k$ is equal to equal to K ( z ) / k.

Now by substituting these values in (3.27), we get

$$\sum_{h=1}^{L}\sum_{k=1}^{M} W_{hk}^2\sigma_{hkx}^2 = \frac{K^3(x)K(z)}{12h^3k}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f^2(x,z) f^{\frac{-3}{2}}(x)\, f^{\frac{-1}{2}}(x)\,\partial x\partial z$$

$$= \frac{D(x,z)\sigma_x^2}{h^3k} \tag{3.28}$$

where

$$D(x,z) = \frac{K^3(x)K(z)}{12\sigma_x^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f^2(x,z) f^{\frac{-3}{2}}(x)\, f^{\frac{-1}{2}}(z)\,\partial x\partial z$$

In the similar way

$$\sum_{h=1}^{L}\sum_{k=1}^{M}W_{hk}^{2}\sigma_{hkz}^{2}=\frac{D(z,x)\sigma_{z}^{2}}{h^{3}k} \tag{3.29}$$

Once again while using the approximation systems above, we find that

$$\sum_{h=1}^{L}\sum_{k=1}^{M}W_{hk}^{2}=\sum_{h=1}^{L}\sum_{k=1}^{M}\xi_{hk}^{2}\left[x_{h}-x_{h-1}\right]^{2}\left[z_{k}-z_{k-1}\right]^{2}$$

$$=\sum_{h=1}^{L}\sum_{k=1}^{M}\xi_{hk}^{2}\xi_{h}^{\frac{-1}{2}}\xi_{k}^{\frac{-1}{2}}\left[x_{h}-x_{h-1}\right]\left[z_{k}-z_{k-1}\right]A_{h}A_{k}$$

$$=\frac{F[x,z]}{hk} \tag{3.30}$$

where $F(x,z)=\dfrac{K(x)K(z)}{hk}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f^{2}(x,z)f^{\frac{-1}{2}}(x)f^{\frac{-1}{2}}(z)\partial x\partial z$

Using (3.28),( 3.29) and (3.30) in (3.26),we get

$$\sigma_{hky}^{2}=\frac{hk}{n}\sqrt{\beta^{2}\frac{D(x,z)\sigma^{2}(x)}{h^{3}k}+\gamma^{2}\frac{D(z,x)\sigma^{2}(x)}{h^{3}k}+\sigma^{2}\frac{F(x,z)}{hk}} \tag{3.31}$$

For any pair (x, z) of stochastic variables. The correlation co-efficient between them can be obtained easily .Hence the regression co-efficient can be obtained easily.

## 4. An Application

For application of the proposed method, it has already been mentioned that the practical of the above proposed method is possible only when the probability density functions of the stratification variables are known. For instance,

Let X follows an Exponential Distribution with pdf as

$$f(x)=\begin{Bmatrix}\theta e^{-\theta x}:x>0\\0:elsewhere\end{Bmatrix} \tag{4.1}$$

and Z follows the Right Triangular distribution with pdf as

$$f(z;a,b)=\begin{Bmatrix}\dfrac{2(b-y)}{(b-a)^{2}};a\le y\le b\\0:elsewhere\end{Bmatrix} \tag{4.2}$$

where ‘a’ and ‘b’ are constants and are the lower and upper bounds of the auxiliary variable Z and the other auxiliary variable X is defined in positive real number which ranges from 0 to ∞.In order to proceed it, we need to find stratum weight and variance of both the variables in case they are independent. By substituting the density functions of both the variables given in (4.1) and (4.2) in (2.2), (2.3) and (2.5) and solving them, we get

$$W_{hk} = e^{-\theta x}\left(e^{\theta g_h}-1\right)\frac{2bV_k - 2z_{k-1}V_k - V_k^2}{(b-a)^2} \tag{4.3}$$

$$\sigma^2_{hkx} = \frac{\left(\theta^2 x_{h-1}^2 + 2\theta x_{h-1} + 2\right)e^{\theta g_h} - \theta^2\left(g_h^2 + 2x_{h-1}g_h + x_{h-1}^2\right) - 2\theta\left(x_{h-1} + g_h\right) - 2}{\theta^2\left(e^{g_h}-1\right)} - \left[\frac{\left(\theta x_{h-1}+1\right)e^{g_h} - \theta\left(x_h + g_h\right) - 1}{\theta\left(e^{g_h}-1\right)}\right]^2 \tag{4.4}$$

Similarly

$$\sigma^2_{hky} = \frac{\left(2bV_k - V_k^2\right)z_{k-1}^2 - 2z_{k-1}^3 V_k}{2bV_k - 2z_{k-1}V_k - V_k^2} - \left[\frac{\left(2bV_k - V_k^2\right)z_{k-1} - 2z_{k-1}^2 V_k}{2bV_k - 2z_{k-1}V_k - V_k^2}\right]^2 \tag{4.5}$$

By substituting values obtained in equations (4.3),(4.4) and (4.5) in (3.31),the optimum strata boundaries can be obtained.

Let us suppose we have a population of size 1000which is to be stratified into 12 strata with L = 4 and M = 3.Suppose the estimates of both stratification variable are of interest and a sample of size n = 300 is drawn from the population. By assuming the initial value in both the variables be equal to zero and the total width of variable X is 10 and of Z be 2. Solving the above equations by using these values would results in optimum strata points as shown in the below Table 1.

**Table 1:** OSB for Bi-variate auxiliary variables

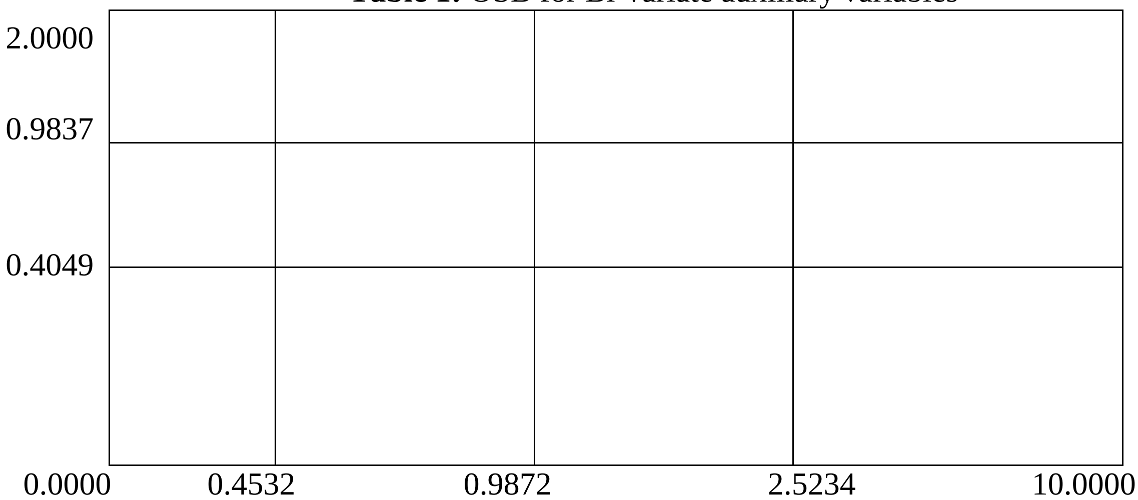

**Table 2:** Strata Boundaries and Total variance

| **OSB($x_h$ , $z_k$)** | **Total Variance** |
|---|---|
| (0.4532,0.4049) | 0.004932 |
| (0.9872,0.4049) | |
| (2.5234,0.4049) | |
| (10.0000,0.4049) | |
| (0.4532,0.9837) | |
| (0.9872,0.9837) | |
| (2.5234,0.9837) | |
| (10.0000,0.9837) | |
| (0.4532,2.0000) | |
| (0.9872,2.0000) | |
| (2.5234,2.0000) | |
| (10.0000,0.2.0000) | |

## 4. Conclusion

We propose a parametric based two-way stratification method to determine the optimum strata boundaries. The proposed technique has a wide scope of application. Since the complete data set of the study variable is often unknown which may restrict the use of many stratification techniques, while as the proposed technique requires only the values of parameter of stratification variables which may be available from the past studies. The results obtained from the numerical illustration reveals that the proposed method can produce maximum reduction in the variance of the estimator as compared to other methods. Thus, it can be safely concluded that the proposed stratification method is a better optimum for obtaining optimum strata boundaries.

## References

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

[2] Dalenius, T. and Gurney, M. The problem of optimum stratification II. *Skandinavisk Aktuarietidskrift.* 34 (1951):133-148.

[3] Aoyama, H. A study of stratified random sampling. *Ann. Ins. Stat. Math.* 6(1954): 1-36.

[4] Dalenius, T. and Hodges, J. L. Minimum variance stratification. *Journal of .American Statistics Association*. 54 (1959):88-101.

[5] Durbin, J. Review of Sampling in Sweden. *Journal of Royal Statistical Society*. 122(1959):246-248.

[6] Mahalanobis, P.C. Some aspect of design of sample surveys. *Sankhya.*12 (1953): 1-17.

[7] Singh, R. On optimum stratification fpor proportional allocation. *Sankhya*. 37(1975):109-115.

[8] Taga, Y. On optimum stratification for the objective variable based on concomitant variables using prior information. *Annals of the institute of Statistical Mathematics.* 19(1967):101-129.

[9] Singh, R., and Sukhatme, B. V. Optimum stratification. *Annals of the institute of Statistical Mathematics.* 21(3) (1969):515-528.

[10] Singh, R., and Sukhatme, B. V. Optimum stratification in sampling with varying probabilities. *Annals of the institute of Statistical Mathematics.* 24 (1972):485-494.

[11] Singh, R., and Sukhatme, B. V. Optimum stratification with ratio and regression methods of estimation. *Annals of the institute of Statistical Mathematics.* 25 (1973):627-633.

[12] Schneeberger, H. and Gollar, W. On the problem of feasibility of optimal stratifdication points according to Dalenius. *Statist. Hefte*. 20(1979): 250-256.

[13] Singh, R., and Prakash, D. Optimum stratification for equal allocation. *Annals of the institute of Statistical Mathematics.* 27 (1975):273-280.

[14] Singh, R. Approximately optimum stratification on auxiliary variable. *Journal of American Statistical Association.* 66 (1971):829-833.

[15] Singh, R. Optimum stratification for proportional allocation. *Sankhya(c),* 37 (1975a):109-115.

[16] Singh, R. An alternative method of stratification on the auxiliary variable. *Sankhya(c),* 37 (1975b):100-108.

[17] Serfling, R.J. Approximately optimal stratification. *Journal of .American Statistics Association.*63 (1968):1298-1309.

[18] Rizvi, S.E.H., Gupta, J.P. and Singh, R. Approximately optimum stratification for two study variable using auxiliary information. *Journal of Indian Society of Agricultural Statistics.* 53(3)(2002):287-298.

[19] Rivest, R. J. A generalization of Lavallee and Hidiroglou algorithm for stratification in survey. *Survey Methodology*. 28 (2002): 191-198.

[20] Gunning, P. and Horgan, J. M. A new algorithm for the construction of stratum boundaries in skewed populations. *Survey Methodology*. 30(2)(2004):159-166.