Journal of Reliability and Statistical Studies; ISSN (Print): 0974-8024, (Online): 2229-5666
Vol. 12, Issue 2 (2019): 173-185

A MATHEMATICAL PROGRAMMING APPROACH IN OPTIMUM STRATIFICATION UNDER NEYMAN ALLOCATION FOR TWO STRATIFYING VARIABLES

Faizan Danish[1], S.E.H.Rizvi[2], Manish Kumar Sharma[3], Sudhakar Dwivedi[4], Bupesh Kumar[5] and Sanjeev Kumar[6]
[1]Research Consultation Services, Doha, Qatar
[2,3]Division of Statistics and Computer science, Faculty of Basic Sciences, Sher-e-Kashmir University of Agricultural Sciences and Technology of Jammu, India
[4]Division of Agricultural Economics and ABM, Sher-e-Kashmir University of Agricultural Sciences and Technology of Jammu, India
[5]Division of Plant Breeding and Genetics, Sher-e-Kashmir University of Agricultural Sciences and Technology of Jammu, India
[6]Division of Plant Breeding and Genetics ACRA, Dhiansar, Sher-e-Kashmir University of Agricultural Sciences and Technology of Jammu, India
E Mail: [1]danishstat@gmail.com; [2]sehrizvi_stats@yahoo.co.in;
[3]manshstat@gmail.com; [4]dwivedi.sudhakar@gmail.com;
[5]bupeshsharma@gmail.com; [6]ssalgotra@gmail.com
*corresponding author: danishstat@gmail.com

Received September 20, 2018
Modified September 24, 2019
Accepted December 10, 2019

Abstract
The current study discusses the solution for obtaining stratification points under Neyman allocation having one study variable and two auxiliary variables. Using dynamic programming approach non-linear programming problem has been solved. The proposed technique has gained in precision rather than using only one auxiliary variable. Numerical illustration has been given in which each of the auxiliary variable is supposed to follow different distribution. Through the empirical study, the proposed method has been compared with the Ravindra and Sukhatme (1969) and Khan *et al*.(2005) methods with the conclusion of having its more relative efficiency.

Key Words: Optimum Stratification, Dynamic Programming, Auxiliary Information.

1. Introduction
The construction of stratification points was pioneered by Dalenius (1950) with single study stratification variable. In case of having single variable under study and its frequency distribution is known, it can be used for determining the strata boundaries. Dalenius and Gurney (1951), Mahalanobis (1952), Dalenius and Hodges (1959) have discussed regarding the study by treating frequency distribution of study variable for stratification points. Due to complexities of minimal equations approximate solutions were obtained.

However, sometimes the difficulties in study variable estimation force us to use auxiliary variable for stratification variable. Authors like Singh (1971, 1975) and Rizvi *et al*. (2002) have used auxiliary variable as a stratification variable. The most advanced and more precise method known as mathematical programming method is used by many authors like Khan et al. (2003). Khan et al. (2005), Khan*et al*. (2009) Sebnem (2011), Rao*et al*. (2012),Fonolahi and Khan (2014), Rao*et al*. (2014), Khan*et al*. (2015) etc. Using diverged cost of units in the whole strata, Danish et al. (2017a) proposed a method for calculating stratification points. Danish *et al* (2017b) discussed different contributions towards obtaining stratification points. Taking into account the problem of optimum strata width (OSW), Danish and Rizvi (2017) used MPP for uniformly study variable.

In this paper, we are taking into consideration two stratifying variables used for stratification. By using Mathematical programming approach, the OSB has been obtained. Let X and Z be two auxiliary variables having Y as study variable. Dividing the whole population into strata, we have

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

N_{hk} indicates size of (h, k)th stratum and N total population size.

$$\sum_{h}\sum_{k} n_{hk} = n$$

n_{hk} is the units taken from (h , k)th stratum and n is total sample size.
The population total can be expressed as

$$y = \sum_{h}\sum_{k}\sum_{i} y_{hki} \text{ ,(I = 1,2,3,…, } N_{hk})$$

Where Population unit y_{hki} in (h ,k)th stratum

Unbiased estimator of mean of the population $\overline{y}_{st} = \sum_{h}\sum_{k} W_{hk}\,\overline{y}_{hk}$

where , $\overline{y}_{hk} = \dfrac{1}{n_{hk}}\sum_{i} y_{hki}$ and ‘ $W_{hk} = \dfrac{N_{hk}}{N}$ is weight of (h, k)th stratum

2. Formulation of Problem under Neyman allocation

If the sampling cost per unit is same to different strata keeping the constant sample size and the allocation of sample sizes is done as

$$n_{hk} = \frac{nW_{hk}\sigma_{hky}}{\sum\sum W_{hk}\sigma_{hky}}$$

Substituting value of n_{hk} for constant n, minimum variance can be computed by

$$\sum_h \sum_k \frac{(1-f)}{n_{hk}} W_{hk}^2 \sigma_{hky}^2$$

then we get

$$V\left(\overline{y}_{st}\right) = \frac{\left(\sum_h \sum_k W_{hk}\sigma_{hky}\right)^2}{n} - \frac{\sum_h \sum_k W_{hk}\sigma_{hky}^2}{N} \tag{2.1}$$

There may be difficulty in using this as the value of σ_{hky} will usually be unknown. However, the stratum variance may be obtained from previous surveys or from a specially planned pilot survey. The other alternative is to conduct the main survey in a phased manner and utility the data collected in the first phase for ensuring better allocation in the second phase.

(2.1) can be written as (finite correction is ignored)

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

However, minimising this is equivalent to minimize (since 'n' is fixed constant)

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

If the stratification variable is not taken as study variable 'Y', we propose a model based on bi-variate stratified sampling design. The model be

$$Y = \varphi(x, z) + \varepsilon \tag{2.3}$$

where, $\varphi(x,z)$is function of two auxiliary variables and is equal to $\alpha + \beta x + \gamma z$ and 'ε' indicate experimental error and is expressed as

$$E\left(\varepsilon/(x,z)\right) = 0 \;\&\; V\left(\varepsilon/(x,z)\right) = \phi(x,z)\; \forall(x,z)$$

We can write 'μ_{hky}' and 'σ_{hky}^2'indicates stratum mean and variance respectively using Singh and Sukhatme (1969)

$$\mu_{hky} = \mu_{hk\lambda} \tag{2.4}$$

and

$$\sigma_{hky}^2 = \sigma_{hk\lambda}^2 + \mu_{hk\phi} \tag{2.5}$$

where $\mu_{hk\lambda}$ and $\mu_{hk\phi}$are the average value of $\varphi(x,z)$ and $\phi(x,z)$ respectively and $\sigma_{hk\lambda}^2$ variance of $\varphi(x,z)$

We have $\sigma_{hky}^2 = \sigma_{hk\lambda}^2 + \sigma_{hk\varepsilon}^2$ for uncorrelated 'φ' &'ε' (2.6)

Let f(x, z) is joint marginal density function of the two auxiliary variables.

Using (2.2) variance for estimating population mean, for obtaining stratification points, we have

$a = x_0 \le x_1 \le ... \le x_{L-1} \le x_L = b$ and $c = z_0 \le z_1 \le ... \le z_{M-1} \le z_M = d$ where x∈ [a,b],z∈[c,d]

If the main functions are known and also integrable then, $W_{hk}, \sigma^2_{hk\lambda}$ and $\mu_{hk\phi}$ can be written as

$$W_{hk} = \int_{x_{h-1}}^{x_h} \int_{z_{k-1}}^{z_k} f(x,z)\partial x \partial z \tag{2.7}$$

$$\sigma^2_{hk\lambda} = \frac{1}{W_{hk}} \int_{x_{h-1}}^{x_h} \int_{z_{k-1}}^{z_k} \varphi^2(x,z)\ f(x,z) dx dz - \mu^2_{hk\lambda} \tag{2.8}$$

and $$\mu_{hk\phi} = \frac{1}{W_{hk}} \int_{x_{h-1}}^{x_h} \int_{z_{k-1}}^{z_k} \phi(x,z) f(x,z) \partial x \partial z \tag{2.9}$$

Where, $$\mu_{hk\lambda} = \frac{1}{W_{hk}} \int_{x_{h-1}}^{x_h} \int_{z_{k-1}}^{z_k} \lambda(x,z) f(x,z) \partial x \partial z \tag{2.10}$$

where $(x_h - x_{h-1})$ and $(z_k - z_{k-1})$ are the extreme values of the (h,k)th stratum.

However for regression line 2.1, when 'X' and 'Z' are independent of error term 'ε'

then $$\sigma^2_{hky} = \beta^2 \sigma^2_{hkx} + \gamma^2 \sigma^2_{hkz} \tag{2.11}$$

$$W_{hk} = \int_{x_{h-1}}^{x_h} \int_{z_{k-1}}^{z_k} f(x,z)\partial x \partial z \tag{2.12}$$

$$\sigma^2_{hkx} = \frac{1}{W_{hk}} \int_{x_{h-1}}^{x_h} x^2 f(x) \partial x \int_{z_{k-1}}^{z_k} \partial z - \mu^2_{hkx} \tag{2.13}$$

$$\sigma^2_{hkz} = \frac{1}{W_{hk}} \int_{z_{k-1}}^{z_k} z^2 f(z) \partial z \int_{x_{h-1}}^{x_h} \partial x - \mu^2_{hkz} \tag{2.14}$$

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

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

Thus (2.1) can be written in boundary points

Let $$\phi(x_h, x_{h-1}, z_k, z_{k-1}) = W_{hk}\sigma_{hky}$$

$$= W_{hk}\sqrt{\left(\sigma^2{}_{hk\lambda} + \mu_{hk\phi}\right)} \tag{2.15}$$

we have already let the range

$$d_x = b - a = x_L - x_0 \tag{2.16}$$

$$t_z = d - c = z_M - z_0 \tag{2.17}$$

For stratification points, we have

Minimize $\sum_h \sum_k \phi_{hk}\left(x_h, x_{h-1}, z_k, z_{k-1}\right)$

Subject to

$$a = x_0 \le x_1 \le ... \le x_{L-1} \le x_L = b \quad , \quad c = z_0 \le z_1 \le ... \le z_{M-1} \le z_M = d \tag{2.18}$$

and

$$\sum_h \sum_k n_{hk} = n$$

For the rectangular stratification

let $V_h = x_h - x_{h-1}$ and $U_k = z_k - z_{k-1}$ denotes the total deviation from both the sides of the (h,k)th stratum. we have

$$\sum_h V_h = \sum_h \left(x_h - x_{h-1}\right) = b - a = d_x \tag{2.19}$$

$$\sum_k U_k = \sum_k \left(z_k - z_{k-1}\right) = d - c = t_z \tag{2.20}$$

Further, we can write

Minimize $\sum_h \sum_k \phi_{hk}\left(x_h, x_{h-1}, z_k, z_{k-1}\right)$ Subject to

$$\sum_h V_h = d_x \,, \sum_k U_k = t_z \tag{2.21}$$

and

$$V_h \ge 0 and U_k \ge 0$$

Initially, $\left(x_0, z_0\right)$ are the initial values of the auxiliary variables and are known. We can write

Minimize $\sum_h \sum_k \phi_{hk}\left(V_h, U_k\right)$ Subject to

$$\sum_h V_h = d_x \,, \sum_k U_k = t_z \tag{2.22}$$

and

$$V_h \ge 0 and U_k \ge 0$$

3. The solution Procedure

Since the problem expressed in (2.22) is a multistage decision problem. Using dynamic programming we have

$$\text{Minimize } \sum_{h=1}^{L_1} \sum_{k=1}^{M_1} \phi_{hk}\left(x_{h-1}, x_h, z_{k-1}, z_k\right) \tag{3.1}$$

Subject to

$\sum_{h=1}^{L-1} V_h = d_{L_1}$

$\sum_{k=1}^{M-1} U_k = t_{M_1}$,h=1,2,...,L_1 and k =1,2,...,M_1

and

$V_h \geq 0 and U_k \geq 0$

where , $\left(L_1 \times M_1\right) \leq \left(L \times M\right)$,i,e $L_1 < L, M_1 < M$, $d_{L_1} < V, t_{M_1} < M$

Note that if $d_{L_1} = V$ and $t_{M_1} = U$ then $\left(L_1 \times M_1\right) = \left(L \times M\right)$

We can transform as

$d_{L_1} = V_1 + V_2 + \ldots + V_{L_1}$

$d_{L_1-1} = V_1 + V_2 + \ldots + V_{L_1-1} = d_{L_1} - V_{L_1}$

$d_{L_1-2} = V_1 + V_2 + \ldots + V_{L_1-2} = d_{L_1-1} - V_{L_1-1}$

.

.

.

$d_2 = V_1 + V_2 = d_3 - V_3$

$d_1 = V_1 = d_2 - V_2$

Similarly, we have

$t_{M_1} = U_1 + U_2 + \ldots + U_{M_1}$

$t_{M_1-1} = U_1 + U_2 + \ldots + U_{M_1-1} = t_{M_1} - U_{M_1}$

$t_{M_1-2} = U_1 + U_2 + \ldots + U_{M_1-2} = t_{M_1-1} - U_{M_1-1}$

.

.

.

$t_2 = U_1 + U_2 = t_3 - U_3$

$t_1 = U_1 = t_2 - U_2$

Let

$$\phi_{L_1 \times M_1}\left(V_{L_1} \times U_{M_1}\right) = Min\left[\sum_{h=1}^{L_1}\sum_{k=1}^{M_1} \phi_{hk}\left(V_h, U_k\right) \Big/ \sum_{h=1}^{L_1} V_h = d_{L_1}, \sum_{k=1}^{M_1} U_k = t_{M_1}\right]$$

$$and \quad V_h \geq 0, U_k \geq 0; h = 1,2,3,\ldots,L_1 \quad ; \quad k = 1,2,3,\ldots,M_1$$

$$and \quad 1 \leq L_1 \leq L \quad , \quad 1 \leq M_1 \leq M$$

means the minimum estimation of the objective function of the condition (20), that is,

$$\phi_{L_1\times M_1}\left(d_{L_1},t_{M_1}\right)=Min\left[\sum_{h=1}^{L_1-1}\sum_{k=1}^{M_1-1}\phi_{hk}\left(V_h,U_k\right)\Bigg/\sum_{h=1}^{L_1-1}V_h=d_{L_1-1},\sum_{k=1}^{M_1-1}U_k=t_{M_1-1}\right]$$

$$and \quad V_h\geq 0,U_k\geq 0;h=1,2,3,...,L_1 \quad and \quad k=1,2,3,...,M_1$$

Thus to get recursively $\phi_{L\times M}\left(d_x,t_z\right)$ by defining $\phi_{L_1\times M_1}\left(V_{L_1},U_{M_1}\right)$

$$\phi_{L_1\times M_1}\left(d_{L_1},t_{M_1}\right)=Min\left[\begin{array}{l}\phi_{L_1\times M_1}\left(V_{L_1},U_{M_1}\right)+\\ \left[\sum_{h=1}^{L_1-1}\sum_{k=1}^{M_1-1}\phi_{hk}\left(V_h,U_k\right)\Bigg/\sum_{h=1}^{L_1-1}V_h=d_{L_1}-V_{L_1},\sum_{k=1}^{M_1-1}U_k=t_{M_1}-U_{M_1}\right]\end{array}\right]$$

$$and \quad V_h\geq 0,U_k\geq 0;h=1,2,3,...,L_1 \quad and \quad k=1,2,3,...,M_1$$

For fixed value of $\left(V_{L_1},U_{M_1}\right)$, $0\leq d_{L_1}\leq V$, $0\leq t_{M_1}\leq U$.

$$\phi_{L_1\times M_1}\left(d_{L_1},t_{M_1}\right)=\phi_{L_1\times M_1}\left(V_{L_1},U_{M_1}\right)+Min\left[\sum_{h=1}^{L_1-1}\sum_{k=1}^{M_1-1}\phi_{hk}\left(V_h,U_k\right)\Bigg/\begin{array}{l}\sum_{h=1}^{L_1-1}V_h=d_{L_1}-V_{L_1}\\ \sum_{k=1}^{M_1-1}U_k=t_{M_1}-U_{M_1}\end{array}\right]$$

and ,$V_h \geq 0$, $U_k \geq 0$

and $L_1 \varepsilon$ [1,L],$L_2\varepsilon$[1,M]

Proceeding in the same way we can get the stratification points.

4. Empirical study

Study4.1:Let X follow a distribution which is named after Augustine Cauchy, a continuous distribution with probability density function as given below:

$$f\left(x\right)=\frac{1}{\pi\tau\left(1+\left(\frac{x-x_0}{\tau}\right)^2\right)} \qquad ;-\infty<x<\infty$$

where 'x_0' and 'τ' is the location and scale parameter

However, the special case of the Cauchy distribution when the location parameter $x_0=0$ and the scale parameter $\tau=1$ is called the standard Cauchy distribution with probability density function given by

$$f(x)=\begin{cases}\dfrac{1}{\pi\left(1+x^2\right)} & ;-\infty<x<\infty \\ 0 & ; \quad otherwise\end{cases} \tag{4.1}$$

so let us assume that the X variable follows a distribution which is having pdf as given in (4.1).and let the other variable Z follows standard power distribution having pdf as

$$f(z)=\begin{cases}\dfrac{\delta z^{\delta-1}}{\theta^{\delta}} & ;z\geq\theta \\ 0 & ; \quad otherwise\end{cases} \tag{4.2}$$

where $\delta>0$ and $\theta>0$

In order to optimize (2.22) to obtain OSB when having the distribution function of X and Z as given in (4.1) and (4.2) respectively, we have to find the values of W_{hk} , σ^2_{hkx} and σ^2_{hkx} for that substitute value the pdf's in equations (2.12),(2.13) and (2.14),we get

$$W_{hk}=\frac{[I_1][I_2]}{\pi\theta^{\delta}} \tag{4.3}$$

$$\sigma^2_{hkx}=\frac{U_k}{4\pi\theta^{\delta}\left(I_1I_2\right)^2}4\left(V_h-I_1\right)\left(I_1I_2\right)^3-\pi U_k^2\left[\theta^{\delta}\log\left(\frac{1+\left(V_h+x_{h-1}\right)^2}{1+x_{h-1}^2}\right)\right]^2 \tag{4.4}$$

$$\sigma^2_{hkz}=\frac{\pi\delta}{\left(I_1I_2\right)^2}\left\{V_h\left(I_1I_2\right)\left[\left(U_k+z_{k-1}\right)^{\delta+2}-\left(z_{k-1}\right)^{\delta+2}\right]-\pi\delta\left[\left(U_k+z_{k-1}\right)^{\delta+1}-\left(z_{k-1}\right)^{\delta+1}\right]\right\} \tag{4.5}$$

where

$I_1=\tan^{-1}\left(V_h+x_{h-1}\right)-\tan^{-1}\left(x_{h-1}\right)$ and $I_2=\left[\left(U_k+z_{k-1}\right)^{\delta}-\left(z_{k-1}\right)^{\delta}\right]$

Thus (2.22) can be written as

Minimize

$$\sum_h\sum_k\frac{[I_1][I_2]}{\pi\theta^{\delta}}\sqrt{\begin{aligned}&\frac{\beta^2U_k}{4\pi\theta^{\delta}\left(I_1I_2\right)^2}4\left(V_h-I_1\right)\left(I_1I_2\right)^3-\pi U_k^2\left[\theta^{\delta}\log\left(\frac{1+\left(V_h+x_{h-1}\right)^2}{1+x_{h-1}^2}\right)\right]^2\\&+\frac{\gamma^2\pi\delta}{\left(I_1I_2\right)^2}\left\{V_h\left(I_1I_2\right)\left[\left(U_k+z_{k-1}\right)^{\delta+2}-\left(z_{k-1}\right)^{\delta+2}\right]-\pi\delta\left[\left(U_k+z_{k-1}\right)^{\delta+1}-\left(z_{k-1}\right)^{\delta+1}\right]\right\}\end{aligned}}$$

Subject to

$$\begin{aligned} &\sum_h V_h = d_x \\ &\sum_k U_k = t_z \\ &\forall V_h \geq 0, U_k \geq 0 \quad , \quad \begin{matrix} h = 1,2,...,L \\ k = 1,2,...,M \end{matrix} \end{aligned} \tag{4.6}$$

To proceed (4.6) in order to obtain OSB, let us assume that the variable X be defined in the interval [0,1] i,e. d_x= 1 and the variable Z too is defined with same interval [0,1].Also for solution of the MPP it is assumed the value of δ=3 and θ = 1.The MPP (4.6) becomes

Minimize

$$\sum_h \sum_k \frac{[I_1]U_k}{\pi} \sqrt{\begin{array}{l} \dfrac{\beta^2 U_k}{4\pi (I_1 U_k)^2} 4(V_h - I_1)(I_1 U_k)^3 - \pi U_k^2 \left[\log\left(\dfrac{1+(V_h + x_{h-1})^2}{1+x_{h-1}^2} \right) \right]^2 \\ + \dfrac{\gamma^2 \pi 3}{(I_1 U_k)^2} \left\{ V_h U_k (I_1) \left[(U_k + z_{k-1})^5 - (z_{k-1})^5 \right] - 3\pi \left[(U_k + z_{k-1})^4 - (z_{k-1})^4 \right] \right\} \end{array}}$$

Subject to

$$\begin{aligned} &\sum_h V_h = 1 \\ &\sum_k U_k = 1 \\ &\forall V_h \geq 0, U_k \geq 0 \quad , \quad \begin{matrix} h = 1,2,...,L \\ k = 1,2,...,M \end{matrix} \end{aligned} \tag{4.7}$$

Utilizing the distribution function a simulation study in R for estimating the β and γ i,e. β=0.41 and γ=1.46.while running a programming in LINGO to the MPP (4.7) for 6(2×3) strata ,the results obtained as

OSB (x_h, z_k)	**Total Variance (Proposed method)**	**Total Variance (Khan *et al.* 2008)**
(0.3516,0.5371) (1.0000,0.5371) (0.3516,0.7953) (1.0000,0.7953) (0.3516,1.0000) (1.0000,1.0000)	0.003518	0.09916

Table 4.1: Displays OSB and Variance of proposed method

Which shows the lesser variance in proposed method rather than Khan*et al.*2008 which suggests that the using of two auxiliary variable gives more precise results than using only single auxiliary variable.

Study 4.2:Let the pdf's of the two auxiliary variables are given as

$$f(x)=\begin{cases}\dfrac{1}{b-a} & ,a\le x\le b\\ o & ,otherwise\end{cases} \tag{4.8}$$

and

$$f(z)=\begin{cases}e^{-z+1} & ;\quad z\ge 0\\ 0 & ;otherwise\end{cases} \tag{4.9}$$

which means the distribution that is assumed to be followed by 'X' is Uniform distribution defined in the interval [a,b] while as exponential distribution is assumed to be followed by Z. In order to minimize (2.22) we need to obtain values of (2.12),(2.13) and (2.14) using the pdf's (4.8) and (4.9), we get

$$W_{hk}=\frac{V_h}{b-a}e^{-z_k+1}\left[e^{U_k}-1\right] \tag{4.10}$$

$$\sigma_{hkx}^2=\frac{4U_k\left(V_h^2+3x_hx_{h-1}\right)\left[e^{-z_k+1}\left[e^{U_k}-1\right]\right]^3-3U_k^2\left(V_h+2x_{h-1}\right)^2}{12\left[e^{-z_k+1}\left[e^{U_k}-1\right]\right]^2} \tag{4.11}$$

$$\sigma_{hkz}^2=\frac{b-a}{\left(e^{U_k}-1\right)^2}\left\{\begin{aligned}&z_{k-1}^2\left(e^{U_k}-1\right)^2-\left(e^{U_k}-1\right)\left[U_k^2+2e^{U_k}\left(1+z_{k-1}\right)-2\left(1+z_k\right)\right]\\&-\left[e^{U_k}\left(1+z_{k-1}\right)-U_k-z_{k-1}-1\right]^2\end{aligned}\right\} \tag{4.12}$$

Using (4.10)-(4.12) in MPP (2.22),we have

Minimize
$$\sum_h\sum_k\frac{V_h}{b-a}e^{-z_k+1}\left[e^{U_k}-1\right]\sqrt{\begin{aligned}&\beta^2\frac{4U_k\left(V_h^2+3x_hx_{h-1}\right)\left[e^{-z_k+1}\left[e^{U_k}-1\right]\right]^3-3U_k^2\left(V_h+2x_{h-1}\right)^2}{12\left[e^{-z_k+1}\left[e^{U_k}-1\right]\right]^2}\\&+\gamma^2\frac{b-a}{\left(e^{U_k}-1\right)^2}\left\{\begin{aligned}&z_{k-1}^2\left(e^{U_k}-1\right)^2-\left(e^{U_k}-1\right)\left[U_k^2+2e^{U_k}\left(1+z_{k-1}\right)-2\left(1+z_k\right)\right]\\&-\left[e^{U_k}\left(1+z_{k-1}\right)-U_k-z_{k-1}-1\right]^2\end{aligned}\right\}\end{aligned}}$$

Subject to

$$\begin{aligned}&\sum_h V_h=d_x\\&\sum_k U_k=t_z\end{aligned} \tag{4.13}$$

$$\forall V_h \geq 0, U_k \geq 0 \quad , \quad \begin{matrix} h = 1,2,...,L \\ k = 1,2,...,M \end{matrix}$$

By assuming $a = 1 = x_0, b = 2$ and $1 \leq z \leq 6$,i,e. $d_x = 1$ and having these condition over the variables a simulation study is to be made in statistical software R to obtain β = 0.56 and γ =1.25 . To obtain total 6 (2×3) strata substitute these values in MPP (4.13) we get MPP as

Minimize

$$\sum_h \sum_k V_h e^{-z_k+1}\left[e^{U_k}-1\right] \sqrt{\begin{matrix} 0.002175 \dfrac{4U_k\left(V_h^2+3x_h x_{h-1}\right)\left[e^{-z_k+1}\left[e^{U_k}-1\right]\right]^3 - 3U_k^2\left(V_h+2x_{h-1}\right)^2}{\left[e^{-z_k+1}\left[e^{U_k}-1\right]\right]^2} \\ + \dfrac{1.56}{\left(e^{U_k}-1\right)^2}\begin{Bmatrix} z_{k-1}^2\left(e^{U_k}-1\right)^2 - \left(e^{U_k}-1\right)\left[U_k^2+2e^{U_k}\left(1+z_{k-1}\right)-2\left(1+z_k\right)\right] \\ -\left[e^{U_k}\left(1+z_{k-1}\right)-U_k-z_{k-1}-1\right]^2 \end{Bmatrix} \end{matrix}}$$

Subject to

$$\sum_h V_h = 1$$

$$\sum_k U_k = 5$$

(4.14)

$$\forall V_h \geq 0, U_k \geq 0 \quad , \quad \begin{matrix} h = 1,2 \\ k = 1,2,3 \end{matrix}$$

By running a computer programme in LINGO to the optimization problem (4.14), we obtain the following results

OSB (x_h, z_k)	**Total Variance (Proposed method)**	**Total Variance (Ravindra Singh and Sukhatme 1969)**	**Total Variance (Khan *et al.* 2005)**
(1.5000,1.9825) (2.0000,1.9825) (1.5000,4.2586) (2.0000,4.2586) (1.5000,6.0000) (2.0000,6.0000)	0.0009251	0.0740	0.02081

Table: 4.2: Displays OSB and Variance of proposed method and others

The above table displays the OSB when the two auxiliary variables have distribution functions as defined in equations (4.8) and (4.9).Table reveals results have more relative efficiency than the variance obtained by Singh (1969) and Khan *et al.*(2005).Thus, the current method is more preferable than existing methods.

5. Conclusion

In this study we deal with determining optimum strata boundaries while using one dependent variable having two independent variables used as stratification variables. The problem has been formed as a mathematical programming problem which has been solved by dynamic programming approach. It is observed that making strata with the help of auxiliary information using auxiliary variable of the populations having above mentioned distribution functions, prompts considerable gains in the accuracy of the assessments while utilizing the proposed method. Numerical illustration concludes that the current method is preferable methods developed by Singh and Sukhatme (1969) and Khan *et al*.(2005).

References

1. Dalenius, T. (1950).The problem of optimum stratification-Ii, Skand. Aktuartidskr, 33, p. 203-213.
2. Dalenius, T. and Gurney, M. (1951). The problem of optimum stratification-II, Skand. Aktuartidskr, 34, p. 133-148.
3. Dalenius, T. and Hodges, J. L. (1959). Minimum variance stratification, J. Amer. Statist. Assoc., 54, p. 88-101.
4. Danish , F. and Rizvi, S.E.H. (2017). On optimum stratification using mathematical programming approach, International Research Journal of Agricultural Economics and Statistics, 8(2), p. 435-439.
5. Danish, F., Rizvi, S.E.H. Jeelani, M. I and Reashi J.A. (2017a). Obtaining strata boundaries under proportional allocation with varying cost of every unit, Pak. J. Stat. Oper. Res., 13(3), p. 567-574.
6. Danish, F., Rizvi, S.E.H. Jeelani, Sharma, M.K. and M. I.J (2017b). Optimum stratification using mathematical programming approach: a review, Stat. Appl. Prob. Lett. , 4(3), p. 123-129.
7. Fanolahi, A.V. and Khan, M.G.M. (2014).Determining the optimum strata boundaries with constant cost factor, Conference: IEEE Asia Pacific World Congress on Computer Science And Engineering (Apwc), At Plantation Island, Fiji.
8. Khan, M.G.M. , Khan, E.A. and Ahsan , M.J. (2003). An optimal multivariate stratified sampling design using dynamic programming, Aust. N. Z. J. Stat, 45(1), p.107–113.
9. Khan, M. G. M., Najmussehar and Ahsan, M. J. (2005). Optimum stratification for exponential study variable under neyman allocation, J. Indian Soc. Agricultural Statist., 59(2), p. 146-150.
10. Khan, M. G. M., Nand, N. and Ahmad,N. (2008). Determining the optimum strata boundary points using dynamic programming, Survey Methodology, 34 (2), p. 205-214.
11. Khan, M. G. M., Ahmad, N. and Khan, Sabiha (2009). Determining the optimum stratum boundaries using mathematical programming, J. Math. Model. Algorithms, Springer, Netherland, 8(4), p. 409- 423.
12. Khan, M. G. M. , Rao, D., Ansari, A.H. and Ahsan, M.J. (2015). Determining optimum strata boundaries and sample sizes for skewed population with log-normal distribution, Communication in Statistics –Simulation and Computation, 44, p.1364-1387.
13. Mahalanobis, P. C. (1952). Some aspects of the design of sample surveys, Sankhya, 12, p. 1-7.

14. Rao, D., Khan, M.G.M. and Khan, S. (2012). Mathematical programming on multivariate calibration estimation in stratified sampling, World Academy of Science, Engineering and Technology, 6, p. 12-27.
15. Rao, D.K., Khan, M.G.M. and Reddy, K.G. (2014). Optimum stratification of a skewed population, International Journal of Mathematical, Computational, Natural and Physical Engineering, 8(3), p. 497-500.
16. Rizvi, S. E. H., Gupta, J. P. and Bhargava, M. (2002). Optimum stratification based on auxiliary variable for compromise allocation, Metron, 28(1), p. 201-215.
17. Sebnem, Er.(2011). Computation methods for optmum stratification: an overview. Int.Statistical Inst.:Proc.58th World Stastistical Congress. Dulbin, (Sts 058).
18. Singh, R. (1971). Approximately optimum stratification on the auxiliary variable, J. Amer. Statist. Assoc., 66, p. 829-833.
19. Singh, R. (1975). An alternate method of stratification on the auxiliary Variable, Sankhya, 37, p. 100-108.
20. Singh, R. and Sukhatme, B. V. (1969). Optimum stratification for equal allocation, Ann. Inst. Statist. Math., 27, p. 273-280.