

Journal of Statistical Theory and Applications Vol. 20(1), March 2021, pp. 46-60 DOI: https://doi.org/10.2991/jsta.d.210107.001; ISSN 1538-7887 https://www.atlantis-press.com/journals/jsta



# Construction of Strata for a Model-Based Allocation Under a Superpopulation Model

Bhuwaneshwar Kumar Gupt<sup>1,\*</sup>, Md. Irphan Ahamed<sup>2</sup>

<sup>1</sup>Department of Statistics, North-Eastern Hill University, Shillong, 793022, India <sup>2</sup>Department of Mathematics, Umshyrpi College, Shillong, 793004, India

#### **ARTICLE INFO**

#### ABSTRACT

Article History

Received 21 Apr 2019 Accepted 07 Oct 2020

#### Keywords

Approximately optimum strata boundaries Auxiliary variable Optimum strata boundaries Superpopulation models

2000 Mathematics Subject Classification: 62D05.

This paper considers the problem of optimum stratification for a model-based allocation under a superpopulation model. The equations giving optimum points of stratification have been derived and a few methods for finding approximately optimum points of stratification have been obtained from the equations. Numerical illustrations using generated data have been worked out and the proposed methods of stratification have been compared with equal interval stratification.

© 2021 The Authors. Published by Atlantis Press B.V. This is an open access article distributed under the CC BY-NC 4.0 license (http://creativecommons.org/licenses/by-nc/4.0/).

# 1. INTRODUCTION

The classical work of Tschuprow [1] and Neyman [2] on allocation of sample size to strata opened a space for further research on allocation and stratification aspects in stratified sampling. However, initial works on problem of allocation of sample size to strata as well as optimum stratification were based on the values of study variable y itself. When the information on an auxiliary variable highly correlated with the study variable y is available, it was demonstrated by Cochran [3] that a superpopulation model could be constructed in which the finite population under study could be treated as a random sample from an infinite population (superpopulation). It could also be used for construction of strata and allocation of sample size to strata.

Hanurav [4] and Rao [5] started using auxiliary information for allocation of sample size to strata in which the following superpopulation model was considered.

$$\begin{array}{l} \text{(i)} \quad \xi \left( y_i | x_i \right) = \alpha + \beta x_i \\ \text{(ii)} \quad v \left( y_i | x_i \right) = \sigma^2 x_i^g \\ \text{(iii)} \quad \zeta \left( y_i, y_j | x_i, x_j \right) = 0 \end{array} \right\}$$

$$(1)$$

where  $\alpha$ ,  $\beta$ ,  $\sigma^2$  and g were superpopulation parameters with  $\sigma^2 > 0$  and  $g \ge 0$ . The script letters  $\xi$ ,  $\nu$ ,  $\zeta$  denoted conditional expectation, variance and covariance given x's respectively.

Hanurav [4] studied the problem of allocation and obtained the allocation as  $n_h \propto N_h \sigma_{hx}$  for simple random sampling with replacement (SRSWR) within each stratum under particular case g = 2 of the model (1). This allocation was obtained from Tschuprow–Neyman allocation when the unknown proportionate values of  $\sigma_{hy}^2$ 's were replaced by known proportionate values of  $\sigma_{hx}^2$ 's, which were the estimates of  $\sigma_{hy}^2$ 's. Rao [5] too examined analytically the justification for the assumption that the unknown proportionate values of  $\sigma_{hy}^2$ 's were not quite different from the proportionate values of known  $\sigma_{hx}^2$ 's. He proved that  $\sigma_{hy}^2$ 's could be expected to be in the same proportion as  $\sigma_{hx}^2$ 's, if

the squares of the corrected coefficients of variation of *x* character, defined by  $\frac{\sigma_{hx}^2}{X_h^2 - \frac{\delta_h}{N_h^2}}$ , where  $\delta_h = N_h \left( \sum_{j=1}^{N_h} X_{hj}^2 - \sum_{j=1}^{N_h} X_{hj}^g \right)$ , are equal in all

<sup>\*</sup>Corresponding author. Email: bhuwaneshwargupt@gmail.com

The problem of optimum allocation of sample size to strata for probability proportional to size with replacement (PPSWR) within each stratum under a particular case  $\alpha = 0$  of the model (1) was considered by Gupt and Rao [8].

On the other hand, in stratified sampling, ever since Dalenius [9] had pioneered the work on optimum stratification based on estimation variable for Tschuprow–Neyman allocation, later workers have been extending the work in various perspectives and dimensions till date. Regarding finding of approximate solutions to the equations giving optimum strata boundaries (OSBs), Dalenius and Hodges [10] were the first ones who proposed cum  $\sqrt{f}$  rule. In case of the problem of optimum stratification using auxiliary variable, the original main works, inter alia, were done by Dalenius and Gurney [11] and Taga [12] who considered it in the case of Tschuprow–Neyman and proportional allocations (PA) respectively. Subsequently, Singh and Sukhatme [13,14], Serfling [15], Singh and Parkash [16] and Singh [17–20] furthered the work on optimum stratification based on auxiliary variable for various allocation methods in which a number of methods of finding approximate solutions to all the equations giving OSB were also obtained.

Whenever an auxiliary variable highly correlated with the study variable is available, Singh and Sukhatme [13] stipulated a superpopulation model as follows — for which the form of regression of estimation variable y on the concomitant variable x and also the form of the variance function were known.

$$y = c(x) + e \operatorname{such} \operatorname{that} \xi(e|x) = 0 \operatorname{and} v(e|x) = \varphi(x)$$
(2)

where c(x) and  $\varphi(x)$  are real valued functions of x with  $\varphi(x) > 0$ , for all values of x in the range (a, b) with  $(b - a) < \infty$ .

Singh and Sukhatme [13] derived OSB and approximately optimum strata boundaries (AOSBs) based on the auxiliary variable for Tschuprow–Neyman Optimum Allocation (TNOA) and PA under the superpopulation model (2) and empirically illustrated under model (1) for the particular case g = 1.

Yadava and Singh [21] derived equations for OSB for allocation proportional to strata totals of an auxiliary variable and also developed a few methods of obtaining their approximate solutions.

Gupt [22,23] considered sample size allocation problem by modifying the above model (1) in such a way that the element of correlation among units within the same stratum was incorporated; he derived three model-based allocations by assuming some conditions for approximation.

In this paper, we consider problem of optimum stratification for the following allocation which is one of the three model-based allocations, viz.,

$$n_h \alpha N_h \sqrt{\mu_h} \left( x^g \right) \tag{3}$$

provided  $\theta_h(g) = \frac{\sigma_{h(x)}}{\sqrt{\mu_h(x^g)}}$  are equal in all strata.

The equations that give OSB for the allocation (3) for stratified SRSWR have been derived in Section 2. Moreover, these results will hold true for stratified simple random sampling without replacement (SRSWOR) design too when finite population correction is ignored in each stratum. The limiting lower bound of the variance of population mean when the number of strata tends to infinity has been shown in Section 3. The methods of obtaining AOSB to the equations that give OSB have been obtained in Section 4. Numerical illustrations by using generated populations have been worked out in Section 5. Conclusion is given in Section 6.

# 2. EQUATIONS GIVING OSB

The model-based allocation (3) can be expressed as

$$n_h = n \frac{W_h \sqrt{\mu_h \left(x^g\right)}}{\sum W_h \sqrt{\mu_h \left(x^g\right)}} \tag{4}$$

where  $W_h$  is the proportion of population units in the *h*th stratum,  $\mu_h(x^g)$  is the mean for  $x^g$  in the *h*th stratum.

For the above expression (4), the variance of the estimate of the population mean can be obtained as

$$\nu\left(\bar{y}_{st}\right) = \frac{1}{n} \sum_{h=1}^{L} \frac{W_h \left\{\beta^2 \sigma_{hx}^2 + \sigma^2 \mu_h \left(x^g\right)\right\}}{\sqrt{\mu_h \left(x^g\right)}} \sum_{h=1}^{L} W_h \sqrt{\mu_h \left(x^g\right)}.$$
(5)

If f(x) is taken as the marginal density function for the stratification variable x,

We know that

$$W_{h}\mu_{h}\left(x^{g}\right) = \int_{x_{h-1}}^{x_{h}} x^{g} f(x) \, dx.$$
(6)

$$2W_h \sigma_{hx} \frac{\delta \sigma_{hx}}{\delta x_h} = (x_h - \mu_{hx})^2 f(x_h) - \sigma_{hx}^2 f(x_h).$$
<sup>(7)</sup>

Taking into consideration of the superpopulation model (1) we get,

$$\sigma_{hy}^2 = \beta^2 \sigma_{hx}^2 + \sigma^2 \mu_h(x^g). \tag{8}$$

The variance expression in (5) is partially differentiated with respect to  $x_h$  (h = 1, 2, ..., L - 1) and equated to zero to obtain minimum variance.

On differentiating (5) partially with respect to  $x_h$ , we get

$$\sum_{h=1}^{L} \left[ \frac{W_h \left\{ \beta^2 \sigma_{hx}^2 + \sigma^2 \mu_h \left( x^g \right) \right\}}{\sqrt{\mu_h \left( x^g \right)}} \right] \frac{\delta}{\delta x_h} \sum_{h=1}^{L} W_h \sqrt{\mu_h \left( x^g \right)}$$
$$+ \sum_{h=1}^{L} W_h \sqrt{\mu_h \left( x^g \right)} \frac{\delta}{\delta x_h} \sum_{h=1}^{L} \left[ \frac{W_h \left\{ \beta^2 \sigma_{hx}^2 + \sigma^2 \mu_h \left( x^g \right) \right\}}{\sqrt{\mu_h \left( x^g \right)}} \right] = 0$$

$$\Rightarrow \sum_{h=1}^{L} \frac{W_h \{\beta^2 \sigma_{hx}^2 + \sigma^2 \mu_h \left(x^g\right)\}}{\sqrt{\mu_h \left(x^g\right)}} \left\{ W_h \frac{\delta \left(h^*\right)}{\delta x_h} + \left(h^*\right) \frac{\delta W_h}{\delta x_h} + W_i \frac{\delta \left(i^*\right)}{\delta x_h} + \left(i^*\right) \frac{\delta W_i}{\delta x_h} \right\} + \sum_{h=1}^{L} W_h \sqrt{\mu_h \left(x^g\right)} \left\{ W_h \frac{\delta \left(h\right)}{\delta x_h} + \left(h\right) \frac{\delta W_h}{\delta x_h} + W_i \frac{\delta \left(i\right)}{\delta x_h} + \left(i\right) \frac{\delta W_i}{\delta x_h} \right\} = 0$$

where 
$$i^* = h^* + 1$$
,  $(h^*) = \sqrt{\mu_h(x^g)}$ ,  $i = h + 1$ ,  $(h) = \frac{\{\beta^2 \sigma_{hx}^2 + \sigma^2 \mu_h(x^g)\}}{\sqrt{\mu_h(x^g)}}$ . (9)

By using the definitions in (6), (7) and relation (8)

$$W_{h}\frac{\delta(h^{*})}{\delta x_{h}} + (h^{*})\frac{\delta W_{h}}{\delta x_{h}} + W_{i}\frac{\delta(i^{*})}{\delta x_{h}} + (i^{*})\frac{\delta W_{i}}{\delta x_{h}}$$
$$= \frac{\mu_{h}\left(x^{g}\right) + x_{h}^{g}}{2\sqrt{\mu_{h}\left(x^{g}\right)}}f(x_{h}) + \frac{\mu_{i}\left(x^{g}\right) + x_{h}^{g}}{2\sqrt{\mu_{i}\left(x^{g}\right)}}\left(-f(x_{h})\right)$$
(10)

and

$$W_{h}\frac{\delta(h)}{\delta x_{h}} + (h)\frac{\delta W_{h}}{\delta x_{h}} + W_{i}\frac{\delta(i)}{\delta x_{h}} + (i)\frac{\delta W_{i}}{\delta x_{h}}$$

$$= \frac{\{\beta^{2}\sigma_{hx}^{2} + \sigma^{2}\mu_{h}(x^{g})\}\{\mu_{h}(x^{g}) - x_{h}^{g}\} + 2\mu_{h}(x^{g})\{\beta^{2}(x_{h} - \mu_{hx})^{2} + \sigma^{2}x_{h}^{g}\}}{2\{\mu_{i}(x^{g})\}^{3/2}}f(x_{h})$$

$$+ \frac{\{\beta^{2}\sigma_{ix}^{2} + \sigma^{2}\mu_{i}(x^{g})\}\{\mu_{i}(x^{g}) - x_{h}^{g}\} + 2\mu_{i}(x^{g})\{\beta^{2}(x_{h} - \mu_{ix})^{2} + \sigma^{2}x_{h}^{g}\}}{2\{\mu_{i}(x^{g})\}^{3/2}}\{-f(x_{h})\}}$$
(11)

Then, finally using (10) and (11), the equations giving OSB for model-based allocation (3) can be obtained as

$$\frac{\left\{\beta^{2}\sigma_{hx}^{2}+\sigma^{2}\mu_{h}\left(x^{g}\right)\right\}\left\{\mu_{h}\left(x^{g}\right)-x_{h}^{g}\right\}+2\mu_{h}\left(x^{g}\right)\left\{\beta^{2}\left(x_{h}-\mu_{hx}\right)^{2}+\sigma^{2}x_{h}^{g}\right\}}{2\left\{\mu_{h}\left(x^{g}\right)\right\}^{3/2}}\sum_{h=1}^{L}W_{h}\sqrt{\mu_{h}\left(x^{g}\right)} +\frac{\left\{\mu_{h}\left(x^{g}\right)+x_{h}^{g}\right\}}{2\sqrt{\mu_{h}\left(x^{g}\right)}}\sum_{h=1}^{L}\left[\frac{W_{h}\left\{\beta^{2}\sigma_{hx}^{2}+\sigma^{2}\mu_{h}\left(x^{g}\right)\right\}}{\sqrt{\mu_{h}\left(x^{g}\right)}}\right] \\ =\frac{\left\{\beta^{2}\sigma_{ix}^{2}+\sigma^{2}\mu_{i}\left(x^{g}\right)\right\}\left\{\mu_{i}\left(x^{g}\right)-x_{h}^{g}\right\}+2\mu_{i}\left(x^{g}\right)\left\{\beta^{2}\left(x_{h}-\mu_{ix}\right)^{2}+\sigma^{2}x_{h}^{g}\right\}}{2\left\{\mu_{i}\left(x^{g}\right)\right\}^{3/2}}\sum_{h=1}^{L}W_{h}\sqrt{\mu_{h}\left(x^{g}\right)} +\frac{\left\{\mu_{i}\left(x^{g}\right)+x_{h}^{g}\right\}}{2\sqrt{\mu_{i}\left(x^{g}\right)}}\sum_{h=1}^{L}\left[\frac{W_{h}\left\{\beta^{2}\sigma_{hx}^{2}+\sigma^{2}\mu_{h}\left(x^{g}\right)\right\}}{\sqrt{\mu_{h}\left(x^{g}\right)}}\right].$$
(12)

The above equations (12) give the OSB of the auxiliary variable x.

# 3. A FEW METHODS OF FINDING APPROXIMATE SOLUTIONS TO THE EQUATIONS GIVING OSB

In order to find the approximate solutions of the equations (12), we follow the techniques of Singh and Sukhatme [13] and Yadava and Singh [21]. For this purpose, it is required to assume the existence of partial derivatives of f(x), c(x) and  $\psi(x)$ , where we have defined  $\psi(x) = x^g$ ,  $c(x) = \alpha + \beta x$ . Then, we obtain series expansions of this system of equations (Singh and Sukhatme [13]) using the identities (Ekman [24,25]) and Taylor's series expansion, about the point  $x_h$  which is assumed as the common boundary of *h*th and (h + 1) th strata. Considering the right hand side of Equation (12) and all the derivatives used hereafter, in this paper, are evaluated at  $t = x_h$  in the interval  $t \in [x_h, x_{h+1}]$ , we have

$$\mu_{i\psi} = \psi + \frac{\psi'}{2}k_i + \frac{\psi'f' + 2f\psi''}{12f}k_i^2 + \frac{ff''\psi' + ff'\psi'' + f^2\psi''' - \psi'f'^2}{24f^2}k_i^3 + O\left(k_i^4\right).$$

$$\Rightarrow \mu_i \left( x^g \right) - x_h^g = \frac{\psi'}{2} k_i + \frac{\psi' f' + 2f \psi''}{12f} k_i^2 + \frac{f f'' \psi' + f f' \psi'' + f^2 \psi^{'''} - \psi' f'^2}{24f^2} k_i^3 + O\left(k_i^4\right).$$

Now, we can get

$$\begin{split} \beta^2 \sigma_{ix}^2 + \sigma^2 \mu_i \left( x^g \right) &= \sigma^2 \psi + \left( \frac{\sigma^2 \psi'}{2} \right) k_i + \left( \frac{\beta^2 f + 2\sigma^2 f \psi'' + \sigma^2 f' \psi'}{12f} \right) k_i^2 \\ &+ \left( \frac{\sigma^2 f f'' \psi' + \sigma^2 f f' \psi'' + \sigma^2 f^2 \psi^{'''} - \sigma^2 f^2 \psi'}{24f^2} \right) k_i^3 + O\left( k_i^4 \right) \end{split}$$

$$\begin{aligned} & \left\{\beta^{2}\sigma_{ix}^{2} + \sigma^{2}\mu_{i}\left(x^{g}\right)\right\}\left\{\mu_{i}\left(x^{g}\right) - x_{h}^{g}\right\} \\ &= \frac{\sigma^{2}\psi\psi'}{2}k_{i} + \frac{\sigma^{2}f'\psi\psi' + 2\sigma^{2}f\psi\psi'' + 3\sigma^{2}f\psi'^{2}}{12f}k_{i}^{2} \\ &+ \frac{\left(\frac{+2\sigma^{2}f^{2}\psi'\psi'' + \beta^{2}f^{2}\psi' + 2\sigma^{2}f^{2}\psi\psi'' + \sigma^{2}ff'\psi'^{2}}{\sigma^{2}ff''\psi\psi' + \sigma^{2}ff'\psi\psi'' - \sigma^{2}ff'\psi\psi'' + \sigma^{2}ff'\psi'^{2}}\right)_{24f^{2}}k_{i}^{3} + O\left(k_{i}^{4}\right). \end{aligned}$$
(13)

Also,

$$2\mu_{i} \left(x^{g}\right) \left\{\beta^{2} \left(\mu_{ix} - x_{h}\right)^{2} + \sigma^{2} x_{h}^{g}\right\}$$

$$= 2\sigma^{2} \psi^{2} + \sigma^{2} \psi \psi' k_{i} + \left(\frac{3\beta^{2} f \psi + \sigma^{2} f' \psi \psi' + 2\sigma^{2} f \psi \psi''}{6f}\right) k_{i}^{2}$$

$$+ \frac{\left(\frac{+\sigma^{2} f f' \psi \psi'' + \sigma^{2} f^{2} \psi \psi''' - \sigma^{2} f'^{2} \psi \psi'}{3\beta^{2} f^{2} \psi' + 2\beta^{2} f f' \psi + \sigma^{2} f f'' \psi \psi'}\right)}{12f^{2}} k_{i}^{3} + O\left(k_{i}^{4}\right)$$
(14)

Adding (13) and (14) we get,

$$\{\beta^{2}\sigma_{ix}^{2} + \sigma^{2}\mu_{i}\left(x^{g}\right)\}\{\mu_{i}\left(x^{g}\right) - x_{h}^{g}\} + 2\mu_{i}\left(x^{g}\right)\{\beta^{2}\left(\mu_{ix} - x_{h}\right)^{2} + \sigma^{2}x_{h}^{g}\}$$

$$= 2\sigma^{2}\psi^{2} + \left(\frac{3\sigma^{2}\psi\psi'}{2}\right)k_{i} + \frac{\left(\frac{+\sigma^{2}f\psi'^{2} + 2\beta^{2}f\psi}{\sigma^{2}f'\psi\psi' + 2\sigma^{2}f\psi\psi''}\right)}{4f}k_{i}^{2}$$

$$+ \frac{\left(\frac{+2\sigma^{2}ff'\psi'^{2} + 4\sigma^{2}f'\psi\psi'' + 7\beta^{2}f'^{2}\psi' + 4\beta^{2}ff'\psi}{3\sigma^{2}ff'\psi\psi' + 3\sigma^{2}ff'\psi\psi'' + 3\sigma^{2}f^{2}\psi\psi''' - 3\sigma^{2}f'^{2}\psi\psi'}\right)}{24f^{2}}k_{i}^{3}$$

$$+ O\left(k_{i}^{4}\right)$$

$$(15)$$

Moreover

$$\frac{1}{2 \left[\mu_{i}\left(x^{g}\right)\right]^{3/2}} = \frac{1}{2\psi^{3/2}} \left[1 - \frac{3\psi'}{4\psi}k_{i} + \frac{15f\psi'^{2} - 4f'\psi\psi' - 8f\psi\psi''}{32f\psi^{2}} + \frac{-8ff'\psi''\psi^{2} - 8f^{2}\psi'''\psi^{2} + 8f'^{2}\psi'\psi^{2})}{(20ff'\psi\psi'^{2} + 40f^{2}\psi\psi'\psi'' - 35f^{2}\psi'^{3} - 8ff''\psi'\psi^{2}}k_{i}^{3} + O\left(k_{i}^{4}\right).$$
(16)

Multiplying Equations (15) and (16), we get,

$$\frac{\left\{\beta^{2}\sigma_{ix}^{2} + \sigma^{2}\mu_{i}\left(x^{g}\right)\right\}\left\{\mu_{i}\left(x^{g}\right) - x_{h}^{g}\right\} + 2\mu_{i}\left(x^{g}\right)\left\{\beta^{2}\left(\mu_{ix} - x_{h}\right)^{2} + \sigma^{2}x_{h}^{g}\right\}}{2\left(\mu_{i}\left(x_{i}^{g}\right)\right)^{3/2}} = \sigma^{2}\sqrt{\psi} + \frac{\sigma^{2}\psi'^{2} + 8\beta^{2}\psi}{32\psi^{3/2}}k_{i}^{2} + \frac{2\sigma^{2}f'\psi\psi'^{2} + 4\sigma^{2}f\psi\psi'\psi'' - 8\beta^{2}f\psi\psi' + 16\beta^{2}f'\psi^{2} - 3\sigma^{2}f\psi'^{3}}{192f\psi^{5/2}}k_{i}^{3} + O\left(k_{i}^{4}\right).$$
(17)

Similarly, we have derived

$$\frac{\mu_i\left(x^g\right) + x_h^g}{2\sqrt{\mu_i\left(x^g\right)}} = \sqrt{\psi} + \frac{\psi'^2}{32\psi^{3/2}}k_i^2 + \frac{2f'\psi\psi'^2 + 4f\psi\psi'\psi'' - 3f\psi'^3}{192f\psi^{5/2}}k_i^3 + O\left(k_i^4\right)$$
(18)

Thus from (9), (17) and (18), we can again express the equations (12) as

$$\left[ g_{1f}k_{h}^{2} - \left(g_{1f}\right)'\frac{k_{h}^{3}}{3} + O\left(k_{h}^{4}\right) \right] \sum_{h=1}^{L} W_{h}\left(h^{*}\right) + \left[ g_{2f}k_{h}^{2} - \left(g_{2f}\right)'\frac{k_{h}^{3}}{3} + O\left(k_{i}^{4}\right) \right] \sum_{h=1}^{L} W_{h}\left(h\right)$$

$$= \left[ g_{1f}k_{i}^{2} + \left(g_{1f}\right)'\frac{k_{i}^{3}}{3} + O\left(k_{i}^{4}\right) \right] \sum_{h=1}^{L} W_{h}\left(h^{*}\right) + \left[ g_{2f}k_{i}^{2} + \left(g_{2f}\right)'\frac{k_{i}^{3}}{3} + O\left(k_{i}^{4}\right) \right] \sum_{h=1}^{L} W_{h}\left(h\right).$$

$$(19)$$

Now, for the purpose of tackling (19), we prove the following lemma:

51

**Lemma 3.1.** If  $(x_h, x_{h+1})$  are boundaries of the *i*<sup>th</sup> stratum and  $k_i = x_{h+1} - x_h$ , the expressions  $\sum_{h=1}^{L} W_h(h^*)$  and  $\sum_{h=1}^{L} W_h(h)$  can, under the conditions of AOSB, be approximated to give fixed values for a given value of *L*- no of strata, where i = h + 1.

Proof: Using series expansions in powers of interval width  $k_i$  for  $W_i$ ,  $\mu_{i\psi}$  and  $\sigma_{ix}^2$  by following certain techniques and strategies (Ekman [24,25], Singh and Sukhatme [13], Yadava and Singh [21]).

We have already got from above

$$\sqrt{\mu_{h}(x^{g})} = \sqrt{\psi} \left[ 1 + \frac{\psi'}{4\psi}k_{i} + \frac{8f\psi\psi'' + 4f'\psi\psi' - 3f\psi'^{2}}{96f\psi^{2}}k_{i}^{2} + \frac{8ff''\psi^{2}\psi' + 8ff'\psi^{2}\psi'' + 8f^{2}\psi^{2}\psi'' - 8f^{2}\psi^{2}\psi' - 8f^{2}\psi\psi'\psi'' - 4ff'\psi\psi'^{2} + 3f^{2}\psi'^{3}}{384f^{2}\psi^{3}}k_{i}^{3} + O\left(k_{i}^{4}\right) \right].$$
(20)

And 
$$W_i = k_i f + \frac{k_i^2}{2} f' + \frac{k_i^3}{6} f'' + \frac{k_i^4}{24} f''' + O\left(k_i^5\right).$$
 (21)

Multiplying (20) and (21)

$$W_{i}\sqrt{\mu_{h}(x^{g})} = \sqrt{\psi} \left[ k_{i}f + \frac{f\psi' + 2f'\psi}{4\psi}k_{i}^{2} + \frac{8f\psi\psi'' + 16f'\psi\psi' - 3f\psi'^{2} + 16f''\psi^{2}}{96\psi^{2}}k_{i}^{3} + \frac{-10f'\psi\psi'^{2} + 3f\psi'^{3} + 16f'''\psi^{3}}{24f''\psi^{2}\psi' + 24f'\psi^{2}\psi'' + 8f\psi^{2}\psi''' - 8f\psi\psi'\psi''}k_{i}^{4} + O\left(k_{i}^{5}\right) \right]$$

$$(22)$$

Considering the expansion of the following term by Taylor's series

$$\int_{x_{h}}^{x_{h+1}} \sqrt{\psi(t)} f(t) dt$$

$$= \left(\sqrt{\psi}f\right) k_{i} + \left(\sqrt{\psi}f\right)' \frac{k_{i}^{2}}{2!} + \left(\sqrt{\psi}f\right)'' \frac{k_{i}^{3}}{3!} + \left(\sqrt{\psi}f\right)^{'''} \frac{k_{i}^{4}}{4!} + O\left(k_{i}^{5}\right)$$

$$= \sqrt{\psi} \left[ k_{i}f + \frac{f\psi' + 2f'\psi}{4\psi} k_{i}^{2} + \frac{2f\psi\psi'' + 4f'\psi'\psi - f\psi'^{2} + 4f''\psi^{2}}{24\psi^{2}} k_{i}^{3} + \frac{24f'\psi^{2}\psi'' + 8f\psi^{'''}\psi^{2} + 24f''\psi^{2}\psi' - 12f'\psi\psi'^{2} + 16f^{'''}\psi^{3} - 12f\psi\psi'\psi'' + 6f\psi'^{3}}{384\psi^{3}} k_{i}^{4} + O\left(k_{i}^{5}\right) \right]$$
(23)

Subtracting (23) from (22)

$$W_{i}\sqrt{\mu_{i}(x^{g})} - \int_{x_{h}}^{x_{h+1}} \sqrt{\psi(t)}f(t) dt$$

$$= \frac{k_{i}^{2}}{96} \left[ \frac{f\psi'^{2}}{\psi\sqrt{\psi}} k_{i}^{2} + \frac{4f\psi\psi'\psi'' + 2f'\psi\psi'^{2} - 3f\psi'^{3}}{4\psi^{2}\sqrt{\psi}} k_{i}^{2} + O\left(k_{i}^{3}\right) \right]$$

$$= \frac{k_{i}^{2}}{96} \left[ g_{1}^{*}fk_{i}^{2} + \frac{1}{2} \frac{d}{dx_{h}} \left( g_{1}^{*}f \right)' k_{i}^{2} + O\left(k_{i}^{3}\right) \right], \text{ where } g_{1}^{*} = \frac{\psi'^{2}}{\psi^{3/2}}$$
(24)

The following lemma was proved by Singh and Sukhatme [13], by using Taylor's series expansion at the point t = y.

Lemma 3.2.

$$\left[\int_{y}^{x} \sqrt[\lambda]{f(t) dt}\right]^{\lambda} = k^{\lambda} \left[f(y) + \frac{f'(y)}{2}k + O\left(k^{2}\right)\right] = k^{\lambda-1} \int_{y}^{x} f(t) dt \left[1 + O\left(k^{2}\right)\right]$$

By using Lemma (3.2) in (24) and considering the fact that if we have large number of strata whose strata widths  $k_h$  are small, the higher powers of  $k_h$  in the expansion can be neglected. So, by neglecting the terms of order O ( $m^5$ ), where  $m = \sup_{(a,b)} (k_h)$ , we get

$$W_i \sqrt{\mu_i(x^g)} - \int_{x_h}^{x_{h+1}} \sqrt{\psi(t)} f(t) dt = \frac{1}{96} \left[ \int_{x_h}^{x_{h+1}} \sqrt[3]{g_1^*(t)f(t)} \right]^3.$$

By using cum  $g_1^*(t)f(t)$  or  $\sqrt[3]{g_1^*(t)f(t)}$  rule in manner in which Yadava and Singh [21] proceeded for obtaining AOSB -  $\int_{x_h}^{x_{h+1}} \sqrt[3]{g_1^*(t)f(t)} dt =$ 

 $\frac{1}{L}\int_{a}^{b}\sqrt[3]{g_1^*(t)f(t)}dt.$  We get

$$\sum_{h=1}^{L} W_h \sqrt{\mu_h(x^g)} = \int_a^b \sqrt{\psi(x)} f(x) \, dx + \frac{1}{96L^2} \left[ \int_a^b \sqrt[3]{g_1^*(x)} f(x) \, dx \right]^3.$$
(25)

From (25), we have known that under the conditions of AOSB,  $\sum_{h=1}^{L} W_h(h^*)$  in (19) can be assumed as a fixed value for a given L.

In the same way, the expression  $\sum_{h=1}^{L} W_h(h)$  in (19) can also be reduced to a fixed value, for a given L, as follows:

$$\frac{W_{i} \{\beta^{2} \sigma_{ix}^{2} + \sigma^{2} \mu_{i} (x^{g})\}}{\sqrt{\mu_{i} (x^{g})}} = \frac{1}{\sqrt{\psi}} \left[ \sigma^{2} f \psi k_{i} + \frac{\sigma^{2} f \psi' + 2\sigma^{2} f' \psi}{4} k_{i}^{2} + \frac{\left\{ \begin{array}{c} +16\sigma^{2} f f'' \psi^{2} - 3\sigma^{2} f^{2} \psi'^{2} \\ 8\beta^{2} f^{2} \psi + 8\sigma^{2} f^{2} \psi \psi'' + 16\sigma^{2} f f' \psi \psi' \right\}}{96 f \psi} k_{i}^{3} \\ + \frac{\left\{ \begin{array}{c} +8\sigma^{2} f^{3} \psi''' \psi^{2} + 3\sigma^{2} f^{3} \psi'^{3} + 16\beta^{2} f^{2} f' \psi^{2} + 16\sigma^{2} f^{2} f'' \psi^{3} - 8\beta^{2} f^{3} \psi \psi' \right\}}{96 f \psi} k_{i}^{3} \\ + \frac{\left\{ \begin{array}{c} +8\sigma^{2} f^{3} \psi''' \psi^{2} + 3\sigma^{2} f^{3} \psi \psi' \psi'' + 24\sigma^{2} f^{2} f'' \psi^{2} + 24\sigma^{2} f^{2} f'' \psi' \psi^{2} + 24\sigma^{2} f^{2} f' \psi' \psi^{2} + 24\sigma^{2} f^{2} f' \psi'' \psi^{2} \right\}}{384 f^{2} \psi^{2}} k_{i}^{4} + O \left( k_{i}^{5} \right).$$
(26)

$$\begin{split} & & \int_{x_{h}}^{x_{h+1}} \sigma^{2} \sqrt{\psi(t)} f(f) dt \\ &= \frac{1}{\sqrt{\psi}} \left[ \sigma^{2} f \psi k_{i} + \frac{\sigma^{2} f \psi' + 2\sigma^{2} f' \psi}{4} k_{i}^{2} + \frac{\{8\sigma^{2} f^{2} \psi \psi'' + 16\sigma^{2} f' f \psi \psi' - 4\sigma^{2} f^{2} \psi'^{2} + 16\sigma^{2} f f'' \psi^{2} \}}{96 f \psi} k_{i}^{3} \\ &+ \frac{\left\{ \frac{+16\sigma^{2} f^{2} f''' \psi^{3} - 12\sigma^{2} f^{3} \psi \psi' \psi'' + 6\sigma^{2} f^{3} f' \psi'^{3} - 12\sigma^{2} f^{2} f' \psi^{2} \psi'' + 8\sigma^{2} f^{3} \psi''' \psi^{2} + 24\sigma^{2} f^{2} f'' \psi' \psi^{2} - 12\sigma^{2} f^{2} f' \psi \psi'^{2} \right\}}{384 f^{2} \psi^{2}} k_{i}^{4} + O\left(k_{i}^{5}\right) \end{split}$$

$$\end{split}$$

Subtracting (27) from (26)

$$\begin{aligned} & \frac{W_i \left\{\beta^2 \sigma_{ix}^2 + \sigma^2 \mu_i \left(x^g\right)\right\}}{\sqrt{\mu_i \left(x^g\right)}} - \int_{x_h}^{x_{h+1}} \sigma^2 \sqrt{\psi \left(t\right)} f(t) dt \\ & = \frac{\left(8\beta^2 \psi + \sigma^2 \psi'^2\right) f}{96\psi^{3/2}} k_i^3 + \frac{\left(2\sigma^2 f' \psi \psi'^2 + 4\sigma^2 f \psi \psi' \psi'' - 3\sigma^2 f \psi'^3 + 16\beta^2 f' \psi^2 - 8\beta^2 f \psi \psi'\right)}{384\psi^{5/2}} k_i^4 + O\left(k_i^5\right) \end{aligned}$$

Thus, in this case too, as obtained in (25), we get

$$\sum_{h=1}^{L} \frac{W_h \left(\beta^2 \sigma_{hx}^2 + \sigma^2 \mu_h \left(x^g\right)\right)}{\sqrt{\mu_h \left(x^g\right)}} = \int_a^b \sigma^2 \sqrt{\psi(x)} f(x) \, dx + \frac{1}{96L^2} \left[ \int_a^b \sqrt[3]{g_2^*(x)} f(t) \, dt \right]^3 \tag{28}$$

where  $g_2^*(x) = \frac{8\beta^2\psi(x) + \sigma^2\psi'^2(x)}{\psi^{3/2}(x)}$ 

Under the conditions of AOSB, the expression  $\sum_{h=1}^{L} W_h(h)$  in (19) can also be assumed as a fixed value for a given value of *L*. Thus the proof of lemma is completed.

Now, for the sake of convenience, we put two constant quantities *P* and *Q* in place of  $\sum_{h=1}^{L} W_h(h^*)$  and  $\sum_{h=1}^{L} W_h(h)$  in (19), and Equation (19) can be rewritten as

$$\begin{bmatrix} g_1 f k_h^2 - (g_1 f)' \frac{k_h^3}{3} + O(k_h^4) \end{bmatrix} P + \begin{bmatrix} g_2 f k_h^2 - (g_2 f)' \frac{k_h^3}{3} + O(k_h^4) \end{bmatrix} Q$$
$$= \begin{bmatrix} g_1 f k_i^2 + (g_1 f)' \frac{k_i^3}{3} + O(k_i^4) \end{bmatrix} P + \begin{bmatrix} g_2 f k_i^2 + (g_2 f)' \frac{k_i^3}{3} + O(k_i^4) \end{bmatrix} Q$$

where  $g_1(t) = \frac{\sigma^2 \psi'^2(t) + 8\beta^2 \psi(t)}{\psi^{3/2}(t)}, g_2(t) = \frac{\psi'^2(t)}{\psi^{3/2}(t)}$ 

$$\Rightarrow g_3 f k_h^2 - (g_3 f)' \frac{k_h^3}{3} + O(k_h^4) = g_3 f k_i^2 + (g_3 f)' \frac{k_i^3}{3} + O(k_i^4),$$

where  $g_{3}(t) = Pg_{1}(t) + Qg_{2}(t)$ 

$$\Rightarrow (g_{3}f) k_{h}^{2} \left[ 1 - \frac{(g_{3}f)'}{g_{3}f} \frac{k_{h}}{3} + O(k_{h}^{2}) \right] = (g_{3}f) k_{i}^{2} \left[ 1 + \frac{(g_{3}f)'}{g_{3}f} \frac{k_{i}}{3} + O(k_{i}^{2}) \right]$$

On raising power 3/2 and then applying binomial expansion

$$\Rightarrow \left(g_{3}f\right)^{3/2}k_{h}^{3}\left[1-\frac{k_{h}}{2}\frac{\left(g_{3}f\right)'}{g_{3}f}+O\left(k_{h}^{2}\right)\right]=\left(g_{3}f\right)^{3/2}k_{i}^{3}\left[1+\frac{k_{i}}{2}\frac{\left(g_{3}f\right)'}{g_{3}f}+O\left(k_{i}^{2}\right)\right]$$
(29)

By using the Lemma 3.2 and on further simplification, the system of Equations (29) can be transformed into

$$k_{h}^{2} \int_{x_{h-1}}^{x_{h}} g_{3}(t) f(t) dt \left(1 + O\left(k_{h}^{2}\right)\right) = k_{i}^{2} \int_{x_{h}}^{x_{h+1}} g_{3}(t) f(t) dt \left(1 + O\left(k_{i}^{2}\right)\right)$$
$$\Rightarrow k_{h}^{2} \int_{x_{h}}^{x_{h}} g_{3}(t) f(t) dt = \text{constant} = c_{1}$$
(30)

 $\Rightarrow k_h^2 \int_{x_{h-1}} g_3(t) f(t) dt = \text{constant} = c_1$ 

The Equation (30) can again be easily proved to be equivalent to

$$\int_{V_{h-1}}^{x_h} \sqrt[3]{g_3(t)f(t)}dt = c_2$$
(31)

where 
$$c_2 = \frac{1}{L} \int_{a}^{b} \sqrt[3]{g_3(t)f(t)} dt$$
 (32)

For evaluating the value of constant  $c_2$ , we can use (32) and approximate solutions may be determined by fixing  $x_{h-1}$  and calculating upper boundaries.

The procedures used in this method of finding AOSB have finally yielded the following theorem.

**Theorem 3.1.** If the function  $g_3(t)f(t)$  is bounded and first two derivatives for all x in (a, b) exist, for a given number of strata taking equal intervals on cumulative  $g_3(t)f(t)$  or  $\sqrt[3]{g_3(t)f(t)}$  yields approximately OSBs on the auxiliary variable.

#### 4. LIMIT EXPRESSION FOR THE VARIANCE

The variance expression in (5) can further be reduced to a form that will give an insight into the pattern of reduction of the variance of the estimate  $\bar{y}_{st}$  with the increase in the number of strata. It is also shown that  $V(\bar{y}_{st})$  does not tend to zero when the number of strata tends to infinity. It is a sequel to the techniques used by Yadava and Singh [21].

In Lemma 3.1, it is already shown that

$$\sum_{h=1}^{L} W_h \sqrt{\mu_h(x^g)} = A + \frac{1}{L^2} B, \text{ where}$$

$$A = \int_a^b \sqrt{\psi(t)} f(t) \, dt \text{ and } B = \frac{1}{96} \left[ \int_a^b \sqrt[3]{g_1^*(t)} f(t) dt \right]^3.$$
And  $\sum \frac{W_h \{\beta^2 \sigma_{hx}^2 + \sigma^2 \mu_h(x^g)\}}{\sqrt{\mu_h(x^g)}} = C + \frac{D}{L^2}, \text{ where}$ 

$$C = \sigma^2 \int_a^b \sqrt{\psi(t)} f(t) \, dt, \, D = \frac{1}{L^2} \left[ \int_a^b \sqrt[3]{g_2^*(t)} f(t) dt \right]^3.$$

Therefore,

$$V(\bar{y}_{st}) = \frac{1}{n} \left( A + \frac{B}{L^2} \right) \left( C + \frac{D}{L^2} \right).$$
(33)

In the above expression (33), it is seen that as  $L \to \infty \Rightarrow V(\bar{y}_{st}) \to \frac{AC}{n}$ . It shows that  $V(\bar{y}_{st})$  does not tend to zero when the number of strata tends to infinity, and hence the following theorem.

**Theorem 4.1.** When the approximately OSBs are obtained by using cum  $\sqrt[3]{g_3(t)f(t)}$  rule, it is observed that  $\lim_{L\to\infty} V(\bar{y}_{st}) = \frac{AC}{n}$ .

# 5. NUMERICAL ILLUSTRATION BY USING GENERATED DATA

For numerical investigation, the following three densities of *x*, which were not only used by Singh and Sukhatme [13] but also by most of the later workers who furthered researches in the area of problem of construction of strata, are used in this paper too. We calculate the solutions of the Equation (12), corresponding solutions of the approximation methods (31) and (32), and the sampling variances of the stratified sampling for equal interval stratification for L = 2, 3, 4, 5, 6. In the case of exponential density of *x*, slight deviation from equal interval stratification is considered for L = 6 to avoid some inconveniences that came upon in using the generated data. The relative efficiencies of the equations giving OSB and their corresponding a few methods of approximation with respect to equal interval stratification are separately shown in Tables 1–9 by taking g = 1, g = 1.5 and g = 2 successively.

- i. Rectangular f(x) = 1,  $1 \le x \le 2$ .
- ii. Right triangular f(x) = 2(2 x),  $1 \le x \le 2$ .
- iii. Exponential  $f(x) = e^{-x+1}$ ,  $1 \le x < \infty$ .

The regression function c(x) is taken to be linear with the slope at 45°. The constant  $\sigma^2$  is determined in each case for the different values of *g* in such a way that 90% of the total variation is accounted for by the regression. In the case of the exponential distribution, we truncate the distribution such that the area under the curve to the right of the truncation point is 0.05. The optimum points of stratification are found by successive iterations. In solving the methods of finding AOSB (31) and (32), suitable techniques for solving numerical algebraic and transcendental equations, and numerical integrations are used.

# 6. CONCLUSION

In this paper, all the proposed methods of stratification—the equations giving OSBs and their methods of approximations—are found to be highly efficient in stratifying heteroscedastic populations. In uniform populations, equal interval stratification is considered to be efficient stratification method and all our proposed methods of stratification perform, in most cases, with nearly same with or, in a few cases, more efficiencies than that of equal interval stratification in the generated populations following uniform density function for all the strengths of heteroscedasticity, i.e., g = 1, 1.5, 2. Moreover the proposed methods of stratifications perform with relatively higher

Table 1	Uniform	distribution, $g =$	1
---------	---------	---------------------	---

No. of Strata ( <i>L</i> )	Stratification by Using Equations (12)		Equal Interval Stratification		Relative Efficiency	Stratification by Using Methods of Approximation		Relative Efficiency
	Points	$nV\left(\overline{Y}_{st}\right)$	Points	$nV\left(\overline{Y}_{st}\right)$		Points	$nV\left(\overline{Y}_{st}\right)$	
2	1.5170	0.02783	1.5	0.0277	99.68	1.5011	0.0283	97.87
3	1.3259,	0.01641	1.334,	0.0164	100.00	1.3203,	0.0163	100.61
	1.6712		1.667			1.6582		
4	1.2564,	0.01278	1.25,	0.0128	99.84	1.2388,	0.0130	98.41
	1.4789,		1.5,			1.49009		
	1.7343		1.75			1.7477		
5	1.2124,	0.01202	1.20,	0.0132	109.58	1.1904,	0.0119	110.39
	1.3695,		1.40,			1.3861,		
	1.5302,		1.60,			1.5865,		
	1.7451		1.80			1.7912		
6	1.1699,	0.00964	1.1667,	0.0088	90.87	1.1583,	0.0089	98.87
	1.3208,		1.3334,			1.32033,		
	1.5200,		1.499,			1.4857,		
	1.7233,		1.667,			1.6543,		
	1.8767		1.8334			1.8257		

#### **Table 2**Uniform distribution, g = 1.5.

No. of Strata ( <i>L</i> )	Stratification Equation	by Using s (12)	Equal Int Stratifica	erval ation	Relative Efficiency	Stratification Method Approxin	Stratification by Using Methods of Approximation	
-	Points	$nV\left(\overline{Y}_{st}\right)$	Points	$nV\left(\overline{Y}_{st}\right)$	-	Points	$nV\left(\overline{Y}_{st}\right)$	-
2	1.5135	0.0279	1.5	0.0278	99.53	1.4800	0.0283	98.20
3	1.3418,	0.0163	1.334,	0.0164	100.68	1.3169,	0.0163	100.68
	1.6792		1.667			1.6557		
4	1.2558,	0.0127	1.25,	0.0127	99.92	1.2361,	0.0129	98.53
	1.4785,		1.50,			1.4832,		
	1.7337		1.75			1.7403		
5	1.2121,	0.0120	1.20,	0.0131	109.26	1.1882,	0.0119	110.46
	1.3695,		1.40,			1.3832,		
	1.5302,		1.60,			1.5853,		
	1.7447		1.80			1.7940		
6	1.1697,	0.0095	1.1667,	0.0086	90.52	1.1571,	0.0089	96.95
	1.3205,		1.334,			1.3183,		
	1.5107,		1.499,			1.4854,		
	1.7232,		1.6667,			1.6562,		
	1.8774		1.8334			1.8305		

# **Table 3** Uniform distribution, g = 2.

No. of Strata ( <i>L</i> )	Stratification by Using Equations (12)		Stratification by Using Equations (12)		Equal Int Stratifica	Equal Interval Stratification		re Stratification by Using cy Methods of Approximation		Stratification by Using Methods of Approximation		Relative Efficiency
	Points	$nV\left(\overline{Y}_{st}\right)$	Points	$nV\left(\overline{Y}_{st}\right)$		Points	$nV\left(\overline{Y}_{st}\right)$					
2	1.5102	0.0280	1.5	0.0278	99.39	1.4734	0.0283	98.41				
3	1.3408,	0.0167	1.334,	0.0164	98.26	1.3118,	0.0165	99.45				
	1.6783		1.667			1.6536						

(continued)

No. of Strata ( <i>L</i> )	Stratification Equation	h by Using s (12)	Equal Int Stratifica	erval ation	Relative Efficiency	Stratification Method Approxin	Stratification by Using Methods of Approximation	
	Points	$nV\left(\overline{Y}_{st}\right)$	Points	$nV\left(\overline{Y}_{st}\right)$		Points	$nV\left(\overline{Y}_{st}\right)$	
4	1.2516,	0.0127	1.25,	0.0127	99.69	1.2300,	0.0128	98.60
	1.4762,		1.50,			1.4813,		
	1.7334		1.75			1.7422		
5	1.2119,	0.0119	1.20,	0.0130	109.59	1.1840,	0.0118	110.05
	1.3696,		1.40,			1.3774,		
	1.5303,		1.60,			1.5853,		
	1.7446		1.80			1.7955		
6	1.1696,	0.0094	1.1667,	0.0084	90.05	1.1543,	0.0086	97.68
	1.3204,		1.334,			1.3163,		
	1.5106,		1.50,			1.4841,		
	1.7233,		1.6667,			1.6570,		
	1.8784		1.8334			1.8323		

**Table 3** Uniform distribution, g = 2. (*Continued*)

**Table 4** Right triangular distribution, g = 1.

No. of Strata ( <i>L</i> )	Stratification Equation	n by Using Is (12)	Equal Interval Stratification		Relative Efficiency	Stratification by Using Methods of Approximation		Relative Efficiency
	Points	$nV\left(\overline{Y}_{st}\right)$	Points	$nV\left(\overline{Y}_{st}\right)$		Points	$nV\left(\overline{Y}_{st}\right)$	
2	1.3794	0.0226	1.5	0.0278	122.98	1.3926	0.0226	123.04
3	1.2754, 1.5898	0.0131	1.334, 1.667	0.0138	105.11	1.2529, 1.5504	0.0135	102.15
4	1.1915, 1.4006, 1.6349	0.0104	1.25, 1.50, 1.75	0.0121	117.29	1.1864, 1.3944, 1.6373	0.0103	117.64
5	1.1804, 1.3812, 1.5800, 1.7367	0.0099	1.20, 1.40, 1.60, 1.80	0.0101	102.33	1.1480, 1.3085, 1.4871, 1.6960	0.0093	108.73
6	1.1450, 1.2891, 1.4357, 1.5897 1.7367	0.0079	1.1667, 1.334, 1.50, 1.6667, 1.8334	0.0093	117.35	1.1507, 1.2842, 1.4300, 1.5920, 1.7831	0.0081	115.17

**Table 5**Right triangular distribution, g = 1.5.

No. of Strata ( <i>L</i> )	Stratification Equation	h by Using as (12)	Equal Int Stratifica	terval ation	Relative Efficiency	Stratification by Using Methods of Approximation		Relative Efficiency
	Points	$nV\left(\overline{Y}_{st}\right)$	Points	$nV\left(\overline{Y}_{st}\right)$		Points	$nV\left(\overline{Y}_{st}\right)$	
2	1.3759	0.0225	1.5	0.0280	124.73	1.3864	0.0225	124.73
3	1.2737,	0.0131	1.334,	0.0142	108.54	1.2477,	0.0136	104.94
	1.5889		1.667			1.5435		
4	1.1842,	0.0103	1.25,	0.0122	118.40	1.1832,	0.0101	120.52
	1.3959,		1.50,			1.3892,		
	1.6344		1.75			1.6330		
5	1.1799,	0.0099	1.20,	0.0101	102.34	1.1450,	0.0092	109.92
								(continued)

No. of Strata (L)	Stratification by Using Equations (12)		Equal Interval Stratification		Relative Efficiency	Stratification by Using Methods of Approximation		Relative Efficiency
	Points	$nV\left(\overline{Y}_{st}\right)$	Points	$nV\left(\overline{Y}_{st}\right)$		Points	$nV\left(\overline{Y}_{st}\right)$	
	1.3807,		1.40,			1.3034,		
	1.5799,		1.60,			1.4183,		
	1.7371		1.80			1.6903		
6	1.1447,	0.0078	1.1667,	0.0092	117.92	1.1477,	0.0080	115.70
	1.2888,		1.334,			1.2791,		
	1.4356,		1.50,			1.4230,		
	1.5897,		1.6667,			1.5850,		
	1.7371		1.8334			1.7776		

#### **Table 5**Right triangular distribution, g = 1.5. (Continued)

**Table 6** Right triangular distribution, g = 2.

No. of Strata ( <i>L</i> )	Stratification Equation	tratification by Using Equations (12)		Equal Interval Stratification		Stratification by Using Methods of Approximation		Relative Efficiency
	Points	$nV\left(\overline{Y}_{st}\right)$	Points	$nV\left(\overline{Y}_{st}\right)$		Points	$nV\left(\overline{Y}_{st}\right)$	
2	1.3732	0.0224	1.5	0.0283	126.40	1.3803	0.0221	127.71
3	1.2720,	0.0130	1.334,	0.0142	109.29	1.2431,	0.0129	110.39
	1.5883		1.667			1.5375		
4	1.1837,	0.0102	1.25,	0.0121	119.17	1.1790,	0.0098	123.42
	1.3957,		1.50,			1.3830,		
	1.6341		1.75			1.6273		
5	1.1794,	0.0098	1.20,	0.0101	102.97	1.1430,	0.0087	115.33
	1.3804,		1.40,			1.3002,		
	1.5800,		1.60,			1.4775,		
	1.7377		1.80			1.6884		
6	1.1445,	0.0077	1.1667,	0.0091	118.73	1.1447,	0.0072	126.70
	1.2887,		1.334,			1.2744,		
	1.4358,		1.50,			1.4174,		
	1.5900,		1.6667,			1.5791,		
	1.7377		1.8334			1.7734		

#### **Table 7**Exponential distribution, g = 1.

No. of Strata ( <i>L</i> )	Stratification by Using Equations (12)		Equal Interval Stratification		Relative Efficiency	Stratification by Using Methods of Approximation		Relative Efficiency
	Points	$nV\left(\overline{Y}_{st}\right)$	Points	$nV\left(\overline{Y}_{st}\right)$		Points	$nV\left(\overline{Y}_{st}\right)$	
2	2.2012	0.1817	2.50	0.1984	109.19	2.0632	0.1747	113.52
3	1. 6974,	0.1037	2.0, 3.0	0.1208	116.43	1.6503, 2.5605	0.1055	114.49
	2.5992							
4	1.6526,	0.0857	1.75,	0.0888	103.63	1.4686,	0.0802	110.75
	2.3296,		2.50,			2.0625,		
	3.1672		3.25			2.8591		
5	1.4058,	0.0674	1.6,	0.0807	119.77	1.3663,	0.0657	122.86
	1.9199,		2.2,			1.8066,		
								(continued)

No. of Strata ( <i>L</i> )	Stratification Equation	by Using s (12)	Equal Int Stratifica	terval ation	Relative Efficiency	Stratification by Using Methods of Approximation		Relative Efficiency
	Points	$nV\left(\overline{Y}_{st}\right)$	Points	$nV\left(\overline{Y}_{st}\right)$		Points	$nV\left(\overline{Y}_{st}\right)$	
	2.5131,		2.8,			2.3489,		
	3.2297		3.4			3.0445		
6	1.3710,	0.0575	1.5,	0.0644	111.90	1.3007,	0.0504	126.40
	1.8163,		2.0,			1.6503,		
	2.3082,		2.3,			2.0626,		
	2.7549,		2.8,			2.5602,		
	3.2835		3.5			3.1815		

**Table 7** Exponential distribution, g = 1. (*Continued*)

**Table 8**Exponential distribution, g = 1.5.

No. of Strata ( <i>L</i> )	Stratification by Using Equations (12)		Equal Interval Stratification		Relative Efficiency	Stratification by Using Methods of Approximation		Relative Efficiency
	Points	$nV\left(\overline{Y}_{st}\right)$	Points	$nV\left(\overline{Y}_{st}\right)$		Points	$nV\left(\overline{Y}_{st}\right)$	
2	2.1016	0.1715	2.50	0.1997	116.42	2.0247	0.1704	117.16
3	1. 6954,	0.1017	2.0,	0.1219	119.78	1.6264,	0.1041	117.08
	2.5915		3.0			2.5684		
4	1.6513,	0.0856	1.75,	0.0891	104.10	1.4466,	0.0762	116.98
	2.3279,		2.50,			2.0249,		
	3.1658		3.25			2.8127		
5	1.4039,	0.0661	1.6,	0.0806	121.88	1.3484,	0.0645	125.07
	1.9191,		2.2,			1.7741,		
	2.5132,		2.8,			2.3082,		
	3.2294		3.4			3.0633		
6	1.3695,	0.0606	1.5,	0.0682	112.57	1.2856,	0.0540	126.38
	1.1816,		2.0,			1.6221,		
	2.3095,		2.30,			2.0250,		
	2.7586,		2.8,			2.5191,		
	3.2748		3.5			3.1482		

#### **Table 9** Exponential distribution, g = 2.

No. of Strata ( <i>L</i> )	Stratification by Using Equations (12)		Equal Interval Stratification		Relative Efficiency	Stratification by Using Methods of Approximation		Relative Efficiency
	Points	$nV\left(\overline{Y}_{st}\right)$	Points	$nV\left(\overline{Y}_{st}\right)$		Points	$nV\left(\overline{Y}_{st}\right)$	
2	2.0613	0.1672	2.50	0.2014	120.90	1.9881	0.1683	119.71
3	1. 6939,	0.1001	2.0,	0.1228	122.73	1.5951,	0.1045	117.55
	2.5846		3.0			2.4780		
4	1.6503,	0.0854	1.75,	0.0893	104.56	1.4259,	0.0739	120.89
	2.3271,		2.50,			1.9983,		
	3.1667		3.25			2.7732		
5	1.4021,	0.0670	1.6,	0.0804	120.11	1.3315,	0.0648	124.10
	1.9186,		2.2,			1.7429,		
								(continued)

No. of Strata ( <i>L</i> )	Stratification by Using Equations (12)		Equal Interval Stratification		Relative Efficiency	Stratification by Using Methods of Approximation		Relative Efficiency
	Points	$nV\left(\overline{Y}_{st}\right)$	Points	$nV\left(\overline{Y}_{st}\right)$		Points	$nV\left(\overline{Y}_{st}\right)$	
	2.5144,		2.8,			2.2680,		
	3.2311		3.4			2.9714		
6	1.3511,	0.0585	1.5,	0.0675	115.31	1.2714,	0.0530	127.23
	1.8094,		2.0,			1.5952,		
	2.3114,		2.30,			1.9883,		
	2.7637,		2.8,			2.4780,		
	3.2886		3.5			3.1141		

**Table 9** Exponential distribution, g = 2. (*Continued*)

efficiencies than equal interval stratification in stratifying populations of right triangular and exponential probability density functions for all the considered strengths of heteroscedasticity. Hence, it is observed that all the proposed methods of stratification perform efficiently in stratifying less skewed and lower level of heteroscedastic populations as well as highly skewed and higher level of heteroscedastic populations. These methods can be used effectively in stratifying heteroscedastic populations based on auxiliary variable which is highly correlated with estimation variable.

# **CONFLICTS OF INTEREST**

There is no conflict of interest between the authors.

# **AUTHORS' CONTRIBUTIONS**

The authors are equally involved in the work and the joint effort of the authors have led to the making of the paper.

# **Funding Statement**

There is no funding for the research work from any funding agency.

# ACKNOWLEDGEMENTS

We, the authors, are grateful to the anonymous reviewers for the inputs and guidance provided to us in improving the quality of the paper.

# REFERENCES

- 1. A.A. Tschuprow, Metron. 2 (1923), 461-493.
- 2. J. Neyman, J. R. Stat. Soc. 97 (1934), 558–625.
- 3. W.G. Cochran, Ann. Math. Stat. 17 (1946), 164-177.
- 4. T.V. Hanurav, Optimum Sampling Strategies and Some Related Problems, unpublished Ph.D. Thesis, Indian Statistical Institute, 1965.
- 5. T.J. Rao, Ann. Inst. Stat. Math. 20 (1968), 159-166.
- 6. R.D. Narain, J. Indian. Soc. Agric. Stat. 3 (1951), 169-174.
- 7. D.G. Horvitz, D.J. Thompson, J. Am. Stat. Assoc. 47 (1952), 663-685.
- 8. B.K. Gupt, T.J. Rao, J. Indian. Soc. Agric. Stat. 50 (1997), 199–208.
- 9. T. Dalenius, Skand. Akt. 33 (1950), 203-213.
- 10. T. Dalenius, J.L. Hodges, Skand. Akt. 3 (1959), 198-203.
- 11. T. Dalenius, M. Gurney, Skand. Akt. 34 (1951), 133-148.
- 12. Y. Taga, Ann. Inst. Stat. Math. 19 (1967), 101-130.
- 13. R. Singh, B.V. Sukhatme, Ann. Inst. Stat. Math. 21 (1969), 515–528.
- 14. R. Singh, B.V. Sukhatme, Ann. Inst. Stat. Math. 24 (1972), 485–494.
- 15. R.J. Serfling, J. Am. Stat. Assoc. 63 (1968), 1298-1309.

- 16. R. Singh, D. Prakash, Ann. Inst. Stat. Math. 27 (1975), 273-280.
- 17. R. Singh, J. Am. Stat. Assoc. 66 (1971), 829-833.
- 18. R. Singh, Sankhya(c). 37 (1975), 109–115.
- 19. R. Singh, An Alternate Method of Stratification on the Auxiliary Variable, Sankhya(c). 37 (1975), 100-108.
- 20. R. Singh, Aust. NZ. J. Stat. 17 (1975), 12-21.
- 21. S.S. Yadava, R. Singh, Commun. Stat-Theor. Method. 13 (1984), 2793–2806.
- 22. B.K. Gupt, Metron-Int. J. Stat. LXI (2003), 35-52.
- 23. B.K. Gupt, Allocation of Sample Size in Stratified Sampling Under Superpopulation Models, LAP LAMBERT Academic Publishing AV Akademikerverlag GmbH & Co.KG, Saarbrucken, Germany, 2012.
- 24. G. Ekman, Ann. Math. Stat. 30 (1959), 219–229.
- 25. G. Ekman, Ann. Math. Stat. 30 (1959), 1131-1134.