Component of Statistics Canada
Catalogue no. 12-001-X Business Survey Methods Division

Article

Determining the optimum strata boundary points using dynamic programming

by Mohammad G. Mostafa Khan, Niraj Nand and Nesar Ahmad

December 2008

Survey Methodology, December 2008
Vol. 34, No. 2, pp. 205-214
Statistics Canada, Catalogue No. 12-001-X

Determining the optimum strata boundary points using dynamic programming

Mohammad G. Mostafa Khan, Niraj Nand and Nesar Ahmad [1]

Abstract

Optimum stratification is the method of choosing the best boundaries that make strata internally homogeneous, given some sample allocation. In order to make the strata internally homogenous, the strata should be constructed in such a way that the strata variances for the characteristic under study be as small as possible. This could be achieved effectively by having the distribution of the main study variable known and create strata by cutting the range of the distribution at suitable points. If the frequency distribution of the study variable is unknown, it may be approximated from the past experience or some prior knowledge obtained at a recent study. In this paper the problem of finding Optimum Strata Boundaries (OSB) is considered as the problem of determining Optimum Strata Widths (OSW). The problem is formulated as a Mathematical Programming Problem (MPP), which minimizes the variance of the estimated population parameter under Neyman allocation subject to the restriction that sum of the widths of all the strata is equal to the total range of the distribution. The distributions of the study variable are considered as continuous with Triangular and Standard Normal density functions. The formulated MPPs, which turn out to be multistage decision problems, can then be solved using dynamic programming technique proposed by Bühler and Deutler (1975). Numerical examples are presented to illustrate the computational details. The results obtained are also compared with the method of Dalenius and Hodges (1959) with an example of normal distribution.

Key Words: Stratified random sampling; Optimum stratification; Triangular distribution; Standard normal distribution; Mathematical programming problem; Multistage decision problem; Dynamic programming technique.

1. Introduction

The basic consideration involved in the determination of optimum strata boundaries (OSB) is that the strata should be internally as homogenous as possible, that is, the stratum variances σ_h^2 should be as small as possible, given some sample allocation. When a single characteristic is under study and the distribution of the study variable is available, the OSB can be determined by cutting the range of this distribution at suitable points. This problem of determining the OSB was first discussed by Dalenius (1950), when the study variable itself is used as stratification variable. He presented a set of minimal equations that could be solved for finding OSB. Unfortunately these equations could not usually be solved because of their implicit nature. Hence attempts have been made by several authors to obtain the approximate strata boundaries using classical methods. Given the number of strata, Dalenius and Gurney (1951) suggested that the strata boundaries be determined when $W_h\sigma_h$ remain constant, where W_h is the weight of stratum h. Mahalanobis (1952) and Hansen and Hurwitz (1953) have suggested that the strata boundaries can be determined when $W_h\mu_h$ remain constant. Aoyama (1954) suggested an approximate rule and recommended to make strata of equal width $x_h - x_{h-1}$, where x_{h-1} and x_h are the boundaries of stratum h. Ekman (1959) determined the strata boundaries with the condition that $W_h(x_h - x_{h-1}) = \text{constant}$. Dalenius and Hodges (1959) recommended to construct the strata by taking equal intervals on the cumulative of $\sqrt{f(x)}$. Sethi (1963) proposed a method to work out the boundaries given by the calculus equations

$$\frac{(x_h - \mu_h)^2 + \sigma_h^2}{\sigma_h} = \frac{(x_{h+1} - \mu_{h+1})^2 + \sigma_{h+1}^2}{\sigma_{h+1}}$$

for a standard continuous distribution resembling the study population.

In a comparison on some of the classical approximate methods, the Ekman method and the Dalenius and Hodges method are proved to work consistently well (see Cochran 1961, Hess, Sethi and Balakrishnan 1966, Murthy 1967) but the later is more convenient and easier to apply (see Nicoloni 2001).

Unnithan (1978) suggested an iterative method using Shanno's modified Newton method for determining the strata boundaries that leads to a local minimum of the variance for Neyman allocation, if a suitable initial solution is chosen. The procedure is proved to be faster than the Dalenius and Hodges iterative procedure. Later on Unnithan and Nair (1995) gave a method of selecting an appropriate starting point for modified Newton method that may lead to a global minimum of the variance.

Lavallée and Hidiroglou (1988) proposed an algorithm to construct stratum boundaries for a power allocated stratified sample of non-certainty sample units. Hidiroglou and

1. Mohammad G. Mostafa Khan, Department of Statistics, Iowa State University, Ames, Iowa 50011, U.S.A., (M.G.M. Khan was on leave from the School of Computing, Information and Mathematical Sciences, University of the South Pacific, Suva, Fiji); Niraj Nand and Nesar Ahmad, School of Computing, Information and Mathematical Sciences, University of the South Pacific, Suva, Fiji.

Srinath (1993) presented a more general form of the algorithm, which by assigning different values to operating parameters yields a power allocation, a Neyman allocation, or a combination of these allocations. Sweet and Sigman (1995) and Rivest (2002) reviewed Lavallée and Hidiroglou algorithm and proposed their modified versions of the algorithm that incorporate the different relationships between the stratification and study variables. Detlefsen and Veum (1991) investigated the Lavallée and Hidiroglou algorithm for several strata and observed that the algorithm's convergence was slow or non-existent. They also found that different starting points lead to different OSBs for the same population.

Niemiro (1999) proposed a random search method in the stratification problem but the algorithm did not guarantee that it leads to global optimum. Furthermore, it would go wrong in a case of a large population, as it requires too many iteration steps (see Kozak 2004).

Nicolini (2001) suggested a method, named *Natural Class Method* (NCM), to oppose the most utilized Dalenius and Hodges method but neither method was proved to be more efficient than other.

Lednicki and Wieczorkowski (2003) presented a method of stratification using the simplex method of Nelder and Mead (1965). Later Kozak (2004) presented the modified random search algorithm as a method of the optimal stratification. The Kozak algorithm was quite faster and efficient as compared to Rivest, and Lednicki and Wieczorkowski but it could not guarantee that the algorithm leads to the global optimum.

Bühler and Deutler (1975) formulated the problem of determining OSB as an optimization problem that can be solved by a dynamic programming technique. This approach is also used by Lavallée (1987, 1988) for determining the OSB which would divide the population domain of two stratification variables into distinct subsets such that the precision of the variables of interest is maximized.

Khan, Khan and Ahsan (2002) considered the problem of finding OSB as an equivalent problem of determining Optimum Strata Width (OSW). The authors formulated the problem of OSW as a Mathematical Programming Problem (MPP). Following the Bühler and Deutler's dynamic programming approach, they solve the MPP that gives exact solution, if the frequency distribution of the study variable is known and the number of strata is fixed in advance. Khan *et al.* (2002) applied their procedure to work out OSB to the population having uniform and right triangular distribution. Later Khan, Najmussehar and Ahsan (2005) extended this dynamic programming approach for determining the OSB for an exponential study variable also.

In this paper the problem of determining OSB for the study variables with Triangular and Standard Normal distributions are discussed. Viewing the fact that these problems are equivalent to the problems of determining OSW, we formulate the problems as MPPs and solve them by following Bühler and Deutler's dynamic programming approach. The formulated MPPs minimize the variance of the estimated population parameter under Neyman allocation subjected to a restriction that sum of the widths of all the strata is equal to the total range of the distribution of the study variable. In Section 2, a review of dynamic programming approach proposed by Bühler and Deutler (1975) is presented. In Section 3, the details of the formulation of the problems of OSW as MPPs are provided. The solution procedure using dynamic programming technique to solve the MPPs is discussed in Section 4. The computational details of the solution procedure is illustrated with numerical examples in Section 5. Finally, in Section 6, an investigation is carried out to compare the results obtained by the dynamic programming method and the cum $\sqrt{f}$ method of Dalenius and Hodges (1959) with an example from a population of normal distribution. It reveals that the proposed dynamic programming method yields a gain in efficiency over the cum $\sqrt{f}$ method.

2. Determination of OSB using dynamic programming techniques: A review of Bühler and Deutler's approach

Let X be a random study variable, discrete or continuous, with probability density function $f(x)$, $a \le x \le b$. To estimate the population mean μ by a stratified sample, X is partitioned into L strata $[a, x_1]$, $(x_1, x_2]$, ..., $(x_{L-1}, b]$ such that

$$a = x_0 \le x_1 \le x_2 \le, \ldots, \le x_{L-1} \le x_L = b. \tag{1}$$

Suppose that from stratum h $(h = 1, 2, \ldots, L)$, which contains N_h units, a sample of size n_h with units y_{hj} $(h = 1, 2, \ldots, L;\ j = 1, 2, \ldots, n_h)$ is selected. Then the stratified mean $\bar{x}_{st} = \sum_{h=1}^{L} W_h \bar{x}_h$ is an unbiased estimate of μ with variance

$$V(\bar{x}_{st}) = \sum_{h=1}^{L} W_h \sigma_h^2 \left(\frac{W_h}{n_h} - \frac{1}{N} \right), \tag{2}$$

where $W_h = N_h / N$, $\bar{x}_h = 1/n_h \sum_{j=1}^{n_h} y_{hj}$, $\sigma_h^2 = [1/(N_h - 1)] \times \sum_{j=1}^{N_h} (y_{hj} - \mu_h)^2$ and $\mu_h = 1/N_h \sum_{j=1}^{N_h} y_{hj}$.

When the frequency function $f(x)$ is known, the values of W_h and σ_h in (2) can be obtained by

$$W_h = \int_{x_{h-1}}^{x_h} f(x)\, dx, \tag{3}$$

$$\sigma_h^2 = \frac{1}{W_h}\int_{x_{h-1}}^{x_h} x^2 f(x)\,dx - \mu_h^2, \tag{4}$$

where

$$\mu_h = \frac{1}{W_h}\int_{x_{h-1}}^{x_h} x f(x)dx \tag{5}$$

is the mean and (x_{h-1}, x_h) are the boundaries of h^{th} stratum.

Then (2) reads as the function of strata boundary points and sample sizes, that is,

$$V(\bar{x}_{st}) = V(\bar{x}_{st} \mid x_1, \ldots, x_{L-1}, n_1, \ldots, n_L).$$

If n_h are fixed, the objective of the optimum stratification is to determine stratum boundary points $(x_1, \ldots, x_{L-1})$ such that $V(\bar{x}_{st})$ is minimum. Further, if the sampling ratios n_h/N_h are small or the sampling is with replacement, then the following optimization problems are obtained, depending on the type of allocation of total sample size $(n = \sum_{h=1}^{L} n_h)$ to strata.

1. Proportional allocation $(n_h = n \cdot W_h)$

$$\text{Minimize } \sum_{h=1}^{L} W_h \sigma_h^2$$
$$\text{subject to } a = x_0 \le x_1 \le x_2 \le, \ldots, \le x_{L-1} \le x_L = b \tag{6}$$

2. Equal allocation $(n_h = n/L)$

$$\text{Minimize } \sum_{h=1}^{L} W_h^2 \sigma_h^2$$
$$\text{subject to } a = x_0 \le x_1 \le x_2 \le, \ldots, \le x_{L-1} \le x_L = b \tag{7}$$

3. Neyman allocation $(n_h = n \cdot W_h \sigma_h / \sum_{h=1}^{L} W_h \sigma_h)$

$$\text{Minimize } \sum_{h=1}^{L} W_h \sigma_h$$
$$\text{subject to } a = x_0 \le x_1 \le x_2 \le, \ldots, \le x_{L-1} \le x_L = b. \tag{8}$$

The problems (6) to (8) have the following structure:

$$\text{Minimize} \sum_{h=1}^{L} \phi_h(x_{h-1}, x_h),$$
$$\text{subject to } a = x_0 \le x_1 \le x_2 \le, \ldots, \le x_{L-1} \le x_L = b. \tag{9}$$

Bühler and Deutler (1975) have suggested a recursive optimization method for solving (9) using a dynamic programming technique as follows:

Consider an optimization problem with the special structure:

$$\text{Minimize } \sum_{h=1}^{m} u_h(z_{h-1}, y_h),$$
$$\text{subject to } z_h = v_h(z_{h-1}, y_h),$$
$$z_h \in Z_h,$$
$$y_h \in S_h(z_{h-1}),$$
$$z_0 = z';\ h = 1, 2, \ldots, m, \tag{10}$$

where $m =$ number of stages, $u_h =$ stage return functions, $v_h =$ stage transformation functions, $Z_h =$ state spaces, $S_h =$ decision spaces, and $z' =$ initial state. Then a dynamic programming procedure using Bellman's principle of optimality (Bellman 1957) can be used to solve (10).

If $m = L$, $z_0 = a$, $z_L = b$, $Z_h = [a, b]$, $Z_{h-1} = [a, b - y_h]$, $S_h(z_{h-1}) = [0, b - z_{h-1}]$ with $z_{h-1} \in Z_{h-1}$, $u_h(z_{h-1}, y_h) = \phi_h(z_{h-1}, y_h + z_{h-1})$ with $y_h \in S_h(z_{h-1})$, $v_h(z_{h-1}, y_h) = y_h + z_{h-1}$, then (10) is transformed to the following problem:

$$\text{Minimize } \sum_{h=1}^{L} \phi_h(z_{h-1}, y_h + z_{h-1}),$$
$$\text{subject to } z_h = y_h + z_{h-1},$$
$$z_h \in [a, b],$$
$$y_h \in [0, b - z_{h-1}],$$
$$z_0 = a,\ z_L = b;\ h = 1, 2, \ldots, L. \tag{11}$$

The problem (11) is an equivalent problem of (9) as they hold the following results:

1. If $(x_1^*, \ldots, x_{L-1}^*)$ is an optimum solution of (9), then $y_h^* = x_h^* - x_{h-1}^*$, $z_h^* = x_h^*$ is an optimum of (11).
2. If $y_h^*(h = 1, \ldots, L)$, $z_h^*(h = 1, \ldots, L-1)$ is an optimum solution of (11), then $x_h^* = z_h^*(h = 1, \ldots, L-1)$ is an optimum solution of (9).

If $\Phi_h(z_{h-1})$ is the optimum value of objective function at stage h with the available state z_{h-1}, then the backward recursive equation to solve (11) using a dynamic programming technique is given by

$$\Phi_h(z_{h-1}) = \min[\phi_h(z_{h-1}, y_h + z_{h-1}) + \Phi_{h+1}(z_h) \mid z_h = y_h + z_{h-1}] \tag{12}$$

on $y_h \in S_h(z_{h-1})$ with initially $\Phi_{L+1} \equiv 0$.

3. Formulation of the problem of OSW as an MPP

In this section the Bühler and Deutler's approach discussed above is extended for a study variable with a continuous density function $f(x)$. The problem (11) is transformed into an equivalent problem of determining OSW by considering $y_h = z_h - z_{h-1} = x_h - x_{h-1}$ as strata widths and then the objective function and the constraints are constructed as functions of y_h. The MPP is treated as a multistage decision problem in which at each stage the value of the OSW and hence the OSB for a stratum is worked out using dynamic programming technique with a forward recursive equation.

Let $f(x)$ be the frequency function and x_0 and x_L are the smallest and largest values of x. If the population mean is estimated under Neyman allocation, then the problem of determining the strata boundaries is to cut up the range,

$$x_L - x_0 = d, \tag{13}$$

at intermediate points $x_1 \le x_2 \le, ..., \le x_{L-1}$ such that $\sum_{h=1}^{L} W_h \sigma_h$ in (8) is minimum.

Consider that $f(x)$ has n piece-wise continuous linear or non-linear functions as follows:

$$f(x) = \begin{cases} g_1(x); & x_0 = a_0 \le x \le a_1, \\ g_2(x); & a_1 < x \le a_2, \\ \vdots & \\ g_n(x); & a_{n-1} < x \le a_n = x_L. \end{cases} \tag{14}$$

Also assume that out of L strata, l_i be the number of strata to be formed under the density function $g_i(x)$; $i = 1, 2, ..., n$ and $\sum_{i=1}^{n} l_i = L$.

If $f(x)$ in (14) is integrable, using the expressions (3), (4) and (5), W_h, σ_h^2 and μ_h are obtained as a function of the boundary points x_h and x_{h-1}. Thus the objective function in (8) could be expressed as a function of boundary points x_h and x_{h-1} only. Let

$$\phi_h(x_h, x_{h-1}) = W_h \sigma_h.$$

Note that the above function has different values for different density functions in (14).

Thus, the problem (8) can be treated as an optimization problem to find $x_1, x_2, ..., x_{L-1}$ as stated in (9).

Let $y_h = x_h - x_{h-1} \ge 0$ denote the width of the h^{th} $(h = 1, 2, ..., L)$ stratum.

With the above definition of y_h, the range of the distribution given in (13) is expressed as the function of the stratum widths as:

$$\sum_{h=1}^{L} y_h = \sum_{h=1}^{L} (x_h - x_{h-1}) = x_L - x_0 = d. \tag{15}$$

The k^{th} stratification point x_k; $k = 1, 2, ..., L-1$ is then expressed as:

$$\begin{aligned} x_k &= x_0 + y_1 + y_2 + \ ... \ + y_k \\ &= x_{k-1} + y_k, \end{aligned}$$

which is a function of k^{th} stratum width and $(k-1)^{\text{th}}$ stratum boundary.

Considering $z_k = x_k$ and adding (15) as a constraint, the problem (11) can be rewritten as an equivalent problem of determining OSW as:

$$\begin{aligned} &\text{Minimize} \ \sum_{h=1}^{L} \phi_h(y_h, x_{h-1}), \\ &\text{subject to} \ \sum_{h=1}^{L} y_h = d, \\ &\text{and} \qquad y_h \ge 0; \ h = 1, 2, ..., L. \end{aligned} \tag{16}$$

Initially, x_0 is known. Therefore, the first term, that is, $\phi_1(y_1, x_0)$ in the objective function of the MPP (16) is a function of y_1 alone. Once y_1 is known, the next stratification point $x_1 = x_0 + y_1$ will be known and the second term in the objective function $\phi_2(y_2, x_1)$ will become a function of y_2 alone.

Therefore, stating the objective function as a function of y_h alone the MPP (16) is expressed as:

$$\begin{aligned} &\text{Minimize} \ \sum_{h=1}^{L} \phi_h(y_h), \\ &\text{subject to} \ \sum_{h=1}^{L} y_h = d, \\ &\text{and} \qquad y_h \ge 0; \ h = 1, 2, ..., L. \end{aligned} \tag{17}$$

The Sections 3.1 and 3.2 illustrate the formulation of the problem of determining OSW as an MPP for Triangular and Standard Normal study variables respectively.

3.1 MPP for triangular distribution

Let the study variable x be following the Triangular distribution on the interval [a, b] with the probability density function:

$$f(x) = \begin{cases} \dfrac{2(x-a)}{(b-a)(c-a)}; & a \le x \le c \\ \dfrac{2(b-x)}{(b-a)(b-c)}; & c < x \le b, \end{cases} \tag{18}$$

where a is a location parameter, b is a scale parameter and c is the shape parameter.

It has two piece-wise functions.

When $a \le x \le c$, from (18) and using (3), (5) and (4), W_h, μ_h, and σ_h^2 are obtained as:

$$W_h = \frac{y_h(y_h + 2a_h)}{(b-a)(c-a)}, \quad (19)$$

$$\mu_h = \frac{\frac{2}{3}y_h^2 + 2y_h x_{h-1} - a y_h + 2a_h x_{h-1}}{y_h + 2a_h},$$

and

$$\sigma_h^2 = \frac{y_h^2[y_h^2 + 6a_h y_h + 6a_h^2]}{18(y_h + 2a_h)^2}, \quad (20)$$

where $y_h = x_h - x_{h-1}$, $a_h = x_{h-1} - a$ and $a \le x_{h-1} \le x_h \le c.$

Thus from (19) and (20),

$$W_h \sigma_h = \frac{y_h^2\sqrt{y_h^2 + 6a_h y_h + 6a_h^2}}{3\sqrt{2}(b-a)(c-a)}. \quad (21)$$

Similarly, when $c < x \le b$, from (18) and using (3), (5) and (4), it can be demonstrated that

$$W_h = \frac{y_h(2b_h - y_h)}{(b-a)(b-c)}, \quad (22)$$

$$\mu_h = \frac{3b_h y_h - 3y_h x_{h-1} + 6b_h x_{h-1} - 2y_h^2}{3(2b_h - y_h)},$$

and

$$\sigma_h^2 = \frac{y_h^2(6b_h^2 - 6b_h y_h + y_h^2)}{18(2b_h - y_h)^2}, \quad (23)$$

where $y_h = x_h - x_{h-1}$, $b_h = b - x_{h-1}$ and $c < x_{h-1} \le x_h \le b.$

Thus, from (22) and (23),

$$W_h \sigma_h = \frac{y_h^2\sqrt{6b_h^2 - 6b_h y_h + y_h^2}}{3\sqrt{2}(b-a)(b-c)}. \quad (24)$$

Let λ_1 and λ_2 be the last and the first stratum formed under the first and second piece-wise function of (18) respectively. If any stratum (say, l) falls under both functions, then λ_1 and λ_2 are not considered to be two different strata but the fractions of the same l^{th} stratum. Then, using (21) and (24) the MPP (17) could be expressed as the problem of determining the OSW for the study variable with Triangular frequency function as:

$$\text{Minimize}\left\{\sum_{h=1}^{\lambda_1}\frac{y_h^2\sqrt{y_h^2 + 6a_h y_h + 6a_h^2}}{3\sqrt{2}(b-a)(c-a)} + \sum_{h=\lambda_2}^{L}\frac{y_h^2\sqrt{6b_h^2 - 6b_h y_h + y_h^2}}{3\sqrt{2}(b-a)(b-c)}\right\},$$

$$\text{subject to} \quad \sum_{h=1}^{L} y_h = d,$$

$$\text{and} \quad y_h \ge 0; \; h = 1, 2, \ldots, L, \quad (25)$$

where $d = b - a$.

3.2 MPP for normal distribution

The study variable x is said to have a Standard Normal distribution if its probability density function is given by

$$f(x) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right); \quad -\infty < x < \infty.$$

As in section 3.1, using the definition (3), (5) and (4), it can be seen that

$$W_h = \frac{erf\left(\frac{y_h + x_{h-1}}{\sqrt{2}}\right) - erf\left(\frac{x_{h-1}}{\sqrt{2}}\right)}{2}, \quad (26)$$

$$\mu_h = \frac{\sqrt{2}\left[\exp\left(-\frac{x_{h-1}^2}{2}\right) - \exp\left(-\frac{(y_h + x_{h-1})^2}{2}\right)\right]}{\sqrt{\pi}\left[erf\left(\frac{y_h + x_{h-1}}{\sqrt{2}}\right) - erf\left(\frac{x_{h-1}}{\sqrt{2}}\right)\right]},$$

and

$$\begin{aligned}\sigma_h^2 = \Bigg\{\sqrt{2\pi}\Bigg[& x_{h-1}\exp\left(-\frac{x_{h-1}^2}{2}\right)erf\left(\frac{y_h + x_{h-1}}{\sqrt{2}}\right) \\ & - (y_h + x_{h-1})\exp\left(-\frac{(y_h + x_{h-1})^2}{2}\right)erf\left(\frac{y_h + x_{h-1}}{\sqrt{2}}\right) \\ & - x_{h-1}\exp\left(-\frac{x_{h-1}^2}{2}\right)erf\left(\frac{x_{h-1}}{\sqrt{2}}\right) \\ & + (y_h + x_{h-1})\exp\left(-\frac{(y_h + x_{h-1})^2}{2}\right)erf\left(\frac{x_{h-1}}{\sqrt{2}}\right)\Bigg] \\ & + \pi\left[erf\left(\frac{y_h + x_{h-1}}{\sqrt{2}}\right) - erf\left(\frac{x_{h-1}}{\sqrt{2}}\right)\right]^2 \\ & - 2\left[\exp\left(-\frac{x_{h-1}^2}{2}\right) - \exp\left(-\frac{(y_h + x_{h-1})^2}{2}\right)\right]^2\Bigg\} \\ & \div \pi\left[erf\left(\frac{y_h + x_{h-1}}{\sqrt{2}}\right) - erf\left(\frac{x_{h-1}}{\sqrt{2}}\right)\right]^2 \quad (27)\end{aligned}$$

where $y_h = x_h - x_{h-1}$, $erf(x_h) - erf(x_{h-1}) = (2/\sqrt{\pi}) \times \int_{x_{h-1}}^{x_h}\exp(-u^2)\,du$ and $h = 1, 2, \ldots, L.$

Therefore, using the values in (26) and (27) the MPP (17) can be expressed as

$$\text{Minimize} \sum_{h=1}^{L} \text{Sqrt}\left\{ \frac{1}{2\sqrt{2\pi}} \left[x_{h-1} \exp\left(-\frac{x_{h-1}^2}{2}\right) erf\left(\frac{y_h + x_{h-1}}{\sqrt{2}}\right) \right. \right.$$
$$- (y_h + x_{h-1}) \exp\left(-\frac{(y_h + x_{h-1})^2}{2}\right) erf\left(\frac{y_h + x_{h-1}}{\sqrt{2}}\right)$$
$$- x_{h-1} \exp\left(-\frac{x_{h-1}^2}{2}\right) erf\left(\frac{x_{h-1}}{\sqrt{2}}\right)$$
$$\left. + (y_h + x_{h-1}) \exp\left(-\frac{(y_h + x_{h-1})^2}{2}\right) erf\left(\frac{x_{h-1}}{\sqrt{2}}\right) \right]$$
$$+ \frac{1}{4}\left[erf\left(\frac{y_h + x_{h-1}}{\sqrt{2}}\right) - erf\left(\frac{x_{h-1}}{\sqrt{2}}\right) \right]^2$$
$$\left. - \frac{1}{2\pi}\left[\exp\left(-\frac{x_{h-1}^2}{2}\right) - \exp\left(-\frac{(y_h + x_{h-1})^2}{2}\right) \right]^2 \right\}$$

$$\text{subject to} \sum_{h=1}^{L} y_h = d$$
$$\text{and} \quad y_h \geq 0; \; h = 1, 2, \ldots, L. \tag{28}$$

4. The solution procedure using dynamic programming technique

The MPP (17) is a multistage decision problem in which the objective function and the constraints are separable of y_h, which allow us to use a dynamic programming technique as illustrated by Bühler and Deutler (1975) for the problem (11).

Consider the following subproblem of (17) for first $k(< L)$ strata:

$$\text{Minimize} \sum_{h=1}^{k} \phi_h(y_h),$$
$$\text{subject to} \sum_{h=1}^{k} y_h = d_k,$$
$$\text{and} \quad y_h \geq 0; \; h = 1, 2, \ldots, k, \tag{29}$$

where $d_k < d$ is the total width available for division into k strata or the state value at stage k. Note that $d_k = d$ for $k = L$.

The transformation functions are given by

$$\begin{aligned} d_k &= y_1 + y_2 + \ldots + y_k, \\ d_{k-1} &= y_1 + y_2 + \ldots + y_{k-1} = d_k - y_k, \\ d_{k-2} &= y_1 + y_2 + \ldots + y_{k-2} = d_{k-1} - y_{k-1}, \\ &\vdots \qquad \vdots \\ d_2 &= y_1 + y_2 = d_3 - y_3, \\ d_1 &= y_1 = d_2 - y_2. \end{aligned}$$

Let $\Phi_k(d_k)$ denote the minimum value of the objective function of (29), that is,

$$\Phi_k(d_k) = \min\left[\sum_{h=1}^{k} \phi_h(y_h) \middle| \sum_{h=1}^{k} y_h = d_k, \right.$$
$$\left. \text{and} \;\; y_h \geq 0; h = 1, 2, \ldots, k \right].$$

With the above definition of $\Phi_k(d_k)$, the MPP (17) is equivalent to finding $\Phi_L(d)$ recursively by finding $\Phi_k(d_k)$ for $k = 1, 2, \ldots, L$ and $0 \leq d_k \leq d$.

We can write:

$$\Phi_k(d_k) = \min\left[\phi_k(y_k) + \sum_{h=1}^{k-1} \phi_h(y_h) \middle| \sum_{h=1}^{k-1} y_h = d_k - y_k, \right.$$
$$\left. \text{and} \;\; y_h \geq 0; \; h = 1, 2, \ldots, k-1. \right]$$

For a fixed value of y_k; $0 \leq y_k \leq d_k$,

$$\Phi_k(d_k) = \phi_k(y_k)$$
$$+ \min\left[\sum_{h=1}^{k-1} \phi_h(y_h) \middle| \sum_{h=1}^{k-1} y_h = d_k - y_k, \right.$$
$$\left. \text{and} \;\; y_h \geq 0; h = 1, 2, \ldots, k-1 \right].$$

Using the Bellman's principle of optimality, we write a forward recursive equation, instead of backward recursive equation as suggested by Bühler and Deutler in (12), for using dynamic programming technique as:

$$\Phi_k(d_k) = \min_{0 \leq y_k \leq d_k} \left[\phi_k(y_k) + \Phi_{k-1}(d_k - y_k) \right], \; k \geq 2. \tag{30}$$

For the first stage, that is, for $k = 1$:

$$\Phi_1(d_1) = \phi_1(d_1) \Rightarrow y_1^* = d_1, \tag{31}$$

where $y_1^* = d_1$ is the optimum width of the first stratum. The relations (30) and (31) are solved recursively for each $k = 1, 2, \ldots, L$ and $0 \leq d_k \leq d$, and $\Phi_L(d)$ is obtained. From $\Phi_L(d)$ the optimum width of L^{th} stratum, y_L^*, is obtained. From $\Phi_{L-1}(d - y_L^*)$ the optimum width of $(L-1)^{\text{th}}$ stratum, y_{L-1}^*, is obtained and so on until y_1^* is obtained.

Note that depending upon the piece-wise function(s) in (14) under which the stratum is formed, $\phi_k(y_k)$ in (30) will take different value for each y_k as follows:

$$y_k = x_k - x_{k-1} \le a_i - a_0, \quad \text{for some } i \ (i = 1, 2, ..., n)$$

and,

$$x_k \in [a_{i-1}, a_i], \qquad \text{for some } i \ (i = 1, 2, ..., n).$$

5. Numerical illustrations

In this section the computational details of the solution procedure discussed in section 4 for the MPPs (25) and (28) are presented.

5.1 Triangular distribution

Let us assume that $a = x_0 = 0$, $c = 1$ and $b = x_L = 2$. This implies that $d = x_L - x_0 = 2$ and the MPP (25) is expressed as:

$$\text{Minimize } \left\{ \sum_{h=1}^{\lambda_1} \frac{y_h^2 \sqrt{y_h^2 + 6a_h y_h + 6a_h^2}}{6\sqrt{2}} + \sum_{h=\lambda_2}^{L} \frac{y_h^2 \sqrt{6b_h^2 - 6b_h y_h + y_h^2}}{6\sqrt{2}} \right\},$$

$$\text{subject to } \sum_{h=1}^{L} y_h = 2,$$

$$\text{and} \qquad y_h \ge 0; \ h = 1, 2, ..., L, \tag{32}$$

where $a_h = x_{h-1}$ and $b_h = 2 - x_{h-1}$.

Using (30) and (31), the recursive equations for solving MPP (32) can be stated as:

For the first stage $(k = 1)$

$$\Phi_1(d_1) = \frac{d_1^3}{6\sqrt{2}} \qquad \text{at} \qquad y_1 = d_1, \tag{33}$$

and for the stages $(k \ge 2)$

$$\Phi_k(d_k) = \begin{cases} \min \left[\dfrac{y_k^2 \sqrt{y_k^2 + 6a_k y_k + 6a_k^2}}{6\sqrt{2}} + \Phi_{k-1}(d_k - y_k) \right] & \text{if } 0 \le d_k \le 1, \\ \min \left[\dfrac{y_k^2 \sqrt{6b_k^2 - 6b_k y_k + y_k^2}}{6\sqrt{2}} + \Phi_{k-1}(d_k - y_k) \right] & \text{if } 1 < d_k \le 2, \end{cases} \tag{34}$$

where the min function is on $0 \le y_k \le d_k$, $a_k = x_{k-1} = d_k - y_k$ and $b_k = 2 - x_{k-1} = 2 - d_k + y_k$.

Substituting this values of a_k and b_k, (34) becomes

$$\Phi_k(d_k) = \begin{cases} \min \left[\dfrac{y_k^2 \sqrt{y_k^2 + 6(d_k - y_k) d_k}}{6\sqrt{2}} + \Phi_{k-1}(d_k - y_k) \right] & \text{if } 0 \le d_k \le 1, \\ \min \left[\dfrac{y_k^2 \sqrt{y_k^2 + 6(2 - d_k + y_k)(2 - d_k)}}{6\sqrt{2}} + \Phi_{k-1}(d_k - y_k) \right] & \text{if } 1 < d_k \le 2, \end{cases} \tag{35}$$

where the min function is on $0 \le y_k \le d_k$.

Then solving the recursive equations (33) and (35) by executing a computer program developed for the solution procedure given in section 4, the OSWs are obtained. The results of optimum strata widths y_h^* and hence the optimum strata boundaries x_h^* along with the values of the objective function $\sum_{h=1}^{L} \phi_h(y_h)$ for $L = 2, 3, 4, 5$ and 6 are presented in Table 1.

Table 1
Optimum strata widths and boundaries of triangular distribution

No. of strata L	Optimum Strata Widths (OSW) (y_h^*)	Optimum Strata Boundaries (OSB) $(x_h^* = x_{h-1}^* + y_h^*)$	Optimum values of the objective function $\sum_{h=1}^{L} \phi_h(y_h) = \sum_{h=1}^{L} W_h \sigma_h$
2	$y_1^* = 1.000000$ $y_2^* = 1.000000$	$x_1^* = 1.000000$	0.2357022604
3	$y_1^* = 0.838081$ $y_2^* = 0.411608$ $y_3^* = 0.750311$	$x_1^* = 0.838081$ $x_2^* = 1.249689$	0.1655523797
4	$y_1^* = 0.645751$ $y_2^* = 0.354249$ $y_3^* = 0.354249$ $y_4^* = 0.645751$	$x_1^* = 0.645751$ $x_2^* = 1.000000$ $x_3^* = 1.354249$	0.1226262641
5	$y_1^* = 0.582819$ $y_2^* = 0.319725$ $y_3^* = 0.252176$ $y_4^* = 0.299439$ $y_5^* = 0.545841$	$x_1^* = 0.582819$ $x_2^* = 0.902544$ $x_3^* = 1.154720$ $x_4^* = 1.454159$	0.0998893913
6	$y_1^* = 0.497369$ $y_2^* = 0.272849$ $y_3^* = 0.229782$ $y_4^* = 0.229782$ $y_5^* = 0.272849$ $y_6^* = 0.497369$	$x_1^* = 0.497369$ $x_2^* = 0.770218$ $x_3^* = 1.000000$ $x_4^* = 1.229782$ $x_5^* = 1.502631$	0.0829362498

5.2 Normal distribution

Let x follow the Standard Normal distribution in the interval (x_0, x_L). For the purpose of illustration, we assume that $x_0 = -4$ and $x_L = 4$. Then $d = 8$, which gives MPP (28) as:

$$\text{Minimize } \sum_{h=1}^{L} \left\{ \text{Sqrt}\left\{ \frac{1}{2\sqrt{2\pi}} \times \left[x_{h-1} \exp\left(-\frac{x_{h-1}^2}{2}\right) erf\left(\frac{y_h + x_{h-1}}{\sqrt{2}}\right) - (y_h + x_{h-1}) \exp\left(-\frac{(y_h + x_{h-1})^2}{2}\right) erf\left(\frac{y_h + x_{h-1}}{\sqrt{2}}\right) - x_{h-1} \exp\left(-\frac{x_{h-1}^2}{2}\right) erf\left(\frac{x_{h-1}}{\sqrt{2}}\right) + (y_h + x_{h-1}) \exp\left(-\frac{(y_h + x_{h-1})^2}{2}\right) erf\left(\frac{x_{h-1}}{\sqrt{2}}\right) \right] + \frac{1}{4}\left[erf\left(\frac{y_h + x_{h-1}}{\sqrt{2}}\right) - erf\left(\frac{x_{h-1}}{\sqrt{2}}\right) \right]^2 - \frac{1}{2\pi}\left[\exp\left(-\frac{x_{h-1}^2}{2}\right) - \exp\left(-\frac{(y_h + x_{h-1})^2}{2}\right) \right]^2 \right\} \right\},$$

$$\text{subject to } \sum_{h=1}^{L} y_h = 8,$$

$$\text{and } \quad y_h \geq 0;\ h = 1, 2, \ldots, L. \tag{36}$$

We have

$$\begin{aligned} x_{k-1} &= x_0 + y_1 + y_2 + \ldots + y_{k-1} \\ &= -4 + y_1 + y_2 + \ldots + y_{k-1} \\ &= d_{k-1} - 4 \\ &= d_k - y_k - 4. \end{aligned}$$

Substituting this value of x_{k-1} in (36) and using (30) and (31), the recursive equations for solving MPP (36) are obtained as:

For first stage $(k = 1)$:

$$\Phi_1(d_1) = \left\{ \text{Sqrt}\left\{ \frac{1}{2\sqrt{2\pi}} \left[-\exp\left(-\frac{1}{2}\right) erf\left(\frac{(d_1 - 4)}{\sqrt{2}}\right) - (d_1 - 4)\exp\left(-\frac{(d_1-4)^2}{2}\right) erf\left(\frac{(d_1 - 4)}{\sqrt{2}}\right) + \exp\left(-\frac{1}{2}\right) erf\left(-\frac{1}{\sqrt{2}}\right) + (d_1 - 4)\exp\left(-\frac{(d_1-4)^2}{2}\right) erf\left(-\frac{1}{\sqrt{2}}\right) \right] + \frac{1}{4}\left[erf\left(\frac{(d_1-4)}{\sqrt{2}}\right) - erf\left(-\frac{1}{\sqrt{2}}\right) \right]^2 - \frac{1}{2\pi}\left[\exp\left(-\frac{1}{2}\right) - \exp\left(-\frac{(d_1-4)^2}{2}\right) \right]^2 \right\} \right\} \tag{37}$$

at $y_1 = d_1$,
and for the stages $k \geq 2$:

$$\Phi_k(d_k) = \min_{0 \leq y_k \leq d_k} \left\{ \text{Sqrt}\left\{ \frac{1}{2\sqrt{2\pi}} \times \left[(d_k - y_k - 4)\exp\left(-\frac{(d_k - y_k - 4)^2}{2}\right) erf\left(\frac{d_k - 4}{\sqrt{2}}\right) - (d_k - 4)\exp\left(-\frac{(d_k - 4)^2}{2}\right) erf\left(\frac{d_k - 4}{\sqrt{2}}\right) - (d_k - y_k - 4)\exp\left(-\frac{(d_k - y_k - 4)^2}{2}\right) erf\left(\frac{(d_k - y_k - 4)}{\sqrt{2}}\right) + (d_k - 4)\exp\left(-\frac{(d_k - 4)^2}{2}\right) erf\left(\frac{(d_k - y_k - 4)}{\sqrt{2}}\right) \right] + \frac{1}{4}\left[erf\left(\frac{d_k - 4}{\sqrt{2}}\right) - erf\left(\frac{(d_k - y_k - 4)}{\sqrt{2}}\right) \right]^2 - \frac{1}{2\pi}\left[\exp\left(-\frac{(d_k - y_k - 4)^2}{2}\right) - \exp\left(-\frac{(d_k - 4)^2}{2}\right) \right]^2 \right\} + \Phi_{k-1}(d_k - y_k) \right\}. \tag{38}$$

Solving the recursive equations (37) and (38), the optimum strata widths y_h^* and hence the optimum strata boundaries x_h^* are obtained. Table 2 shows these results along with the values of the objective function $\sum_{h=1}^{L} \phi_h(y_h)$ for $L = 2, 3, 4, 5$ and 6.

Table 2
Optimum strata widths and boundaries of standard normal distribution

No. of strata L	Optimum Strata Widths (OSW) (y_h^*)	Optimum Strata Boundaries (OSB) $(x_h^* = x_{h-1}^* + y_h^*)$	Optimum values of the objective function $\sum_{h=1}^{L} \phi_h(y_h) = \sum_{h=1}^{L} W_h \sigma_h$
2	$y_1^* = 4.000000$ $y_2^* = 4.000000$	$x_1^* = 0.000000$	0.6021710931
3	$y_1^* = 3.450300$ $y_2^* = 1.099400$ $y_3^* = 3.450300$	$x_1^* = -0.549700$ $x_2^* = 0.549700$	0.4265717619
4	$y_1^* = 3.124570$ $y_2^* = 0.875430$ $y_3^* = 0.875430$ $y_4^* = 3.124570$	$x_1^* = -0.875430$ $x_2^* = 0.000000$ $x_3^* = 0.875430$	0.3297899642
5	$y_1^* = 2.896360$ $y_2^* = 0.767900$ $y_3^* = 0.671480$ $y_4^* = 0.767900$ $y_5^* = 2.896360$	$x_1^* = -1.103640$ $x_2^* = -0.335740$ $x_3^* = 0.335740$ $x_4^* = 1.103640$	0.2686646379
6	$y_1^* = 2.722440$ $y_2^* = 0.702200$ $y_3^* = 0.575360$ $y_4^* = 0.575360$ $y_5^* = 0.702200$ $y_6^* = 2.722440$	$x_1^* = -1.277560$ $x_2^* = -0.575360$ $x_3^* = 0.000000$ $x_4^* = 0.575360$ $x_5^* = 1.277560$ $x_6^* = 4.000000$	0.2265979522

Table 3
Frequency distribution of x and cum $\sqrt{f(x)}$

Class	Frequency $f(x)$	Cum $\sqrt{f(x)}$
(-3.98)-(-3.58)	2	1.4
(-3.58)-(-3.18)	6	3.8
(-3.18)-(-2.78)	23	8.6
(-2.78)-(-2.38)	59	16.3
(-2.38)-(-1.98)	155	28.7
(-1.98)-(-1.58)	296	45.9
(-1.58)-(-1.18)	630	71.0
(-1.18)-(-0.783)	1,015	102.9
(-0.783)-(-0.383)	1,361	139.8
(-0.383)-0.017	1,551	179.2
0.017-0.417	1,495	217.9
0.417-0.817	1,315	254.2
0.817-1.22	1,003	285.9
1.22-1.62	613	310.7
1.62-2.02	285	327.6
2.02-2.42	128	338.9
2.42-2.82	38	345.1
2.82-3.22	18	349.3
3.22-3.62	7	351.9

6. Discussion

In this section we will undertake a numerical investigation into the effectiveness of the dynamic programming method to the Dalenius and Hodges' cum $\sqrt{f}$ method, which is the most frequently used and better known method. For this purpose, we have generated data of size $N = 10{,}000$ for a population with standard normal density function $f(x) = (1/\sqrt{2\pi})\exp(-x^2/2)$, which have been grouped into 19 equal classes. In Table 3 the class frequencies are given in column 2 while their cumulative roots are given in column 3.

For this example the smallest and the largest values of x are $x_0 = -3.98$ and $x_L = 3.62$ respectively. Therefore, the range of the distribution $d = 7.60$.

The OSB are determined for this distribution by using cum $\sqrt{f}$ method and also dynamic programming method. For each $L = 2, 3, 4, 5$ and 6 the variance $\sum_{h=1}^{L} W_h \sigma_h$ is calculated, which is used for the efficiency of the two methods of stratification. The results of this investigation are given in Table 4. From the last column of table it can be seen that the OSB obtained by dynamic programming method are more efficient for all $L = 1, 2, ..., 6$. Although, the efficiency of cum $\sqrt{f}$ method depends on the initial choice of the number of classes but there is no theory which gives the best number of classes (see Hedlin 2000).

Table 4
Relative efficiency of dynamic programming method

L	(Cum $\sqrt{f}$ method)		Dynamic programming method		Relative efficiency
	OSB	$\sum_{h=1}^{L} W_h \sigma_h$	OSB	$\sum_{h=1}^{L} W_h \sigma_h$	
2	-0.017	0.60131	-0.00034	0.60126	100.00832
3	-0.783 0.417	0.43177	-0.55015 0.54884	0.42576	101.41159
4	-0.783 -0.017 0.817	0.33067	-0.87593 -0.00081 0.87395	0.32905	100.49233
5	-1.18 -3.83 0.417 1.22	0.27066	-1.10418 -0.33656 0.33452 1.10147	0.26799	100.99631
6	-1.18 -0.783 -0.017 0.417 1.22	0.24242	-1.27813 -0.57619 -0.00115 0.57369 1.27462	0.22598	107.27498

Finally, the other methods available in the literature such as Aoyama (1954), Ekman (1959), Sethi (1963), *etc.* are mostly classical methods to obtain approximate strata boundaries. Many authors such as Unithan (1978), Lavallée and Hidiroglou (1988), Sweet and Sigman (1995), Rivest (2002), *etc.* suggested iterative procedures. These iterative procedure require initial approximate solutions. Also there is no guarantee that an iterative procedure will give the global minimum in the absence of a suitable approximate initial solution and the variance function have more than one local minimum. The advantage of the dynamic programming method is that it gives the global minimum of the objective function and it does not require any initial approximate solutions.

Acknowledgement

The authors are grateful to the Associate Editor and referees for their invaluable comments and suggestions to improve the manuscript.

References

Aoyama, H. (1954). A study of stratified random sampling. *Annals of the Institute of Statistical Mathematics*, 6, 1-36.

Bellman, R.E. (1957). *Dynamic Programming*. Princetown University Press, New Jersey.

Bühler, W., and Deutler, T. (1975). Optimal stratification and grouping by dynamic programming. *Metrika*, 22, 161-175.

Cochran, W.G. (1961). Comparison of methods for determining stratum boundaries. *Bulletin of the International Statistical Institute*, 38, 2, 345-358.

Dalenius, T. (1950). The problem of optimum stratification-II. *Skand. Aktuartidskr*, 33, 203-213.

Dalenius, T., and Gurney, M. (1951). The problem of optimum stratification. *Skand. Aktuartidskr*, 34, 133-148.

Dalenius, T., and Hodges, J.L. (1959). Minimum variance stratification. *Journal of the Americal Statistical Association*, 54, 88-101.

Detlefsen, R.E., and Veum, C.S. (1991). Design issues for the retail trade sample surveys of the U.S. Bureau of the Census. *Proceedings of the Survey Research Methods Section*, American Statistical Association, 214-219.

Ekman, G. (1959). Approximate expression for conditional mean and variance over small intervals of a continuous distribution. *Annals of the Institute of Statistical Mathematics*, 30, 1131-1134.

Hansen, M.H., and Hurwitz, W.N. (1953). On the theory of sampling from finite population. *The Annals of Mathematical Statistics*, 14, 333-362.

Hedlin, D. (2000). A procedure for stratification by an extended Ekman rule. *Journal of Official Statistics*, 16, 15-29.

Hess, I., Sethi, V.K. and Balakrishnan, T.R. (1966). Stratification: A practical investigation. *Journal of the Americal Statistical Association*, 61, 71-90.

Hidiroglou, M.A., and Srinath, K.P. (1993). Problems associated with designing subannual bussiness surveys. *Journal of Bussiness and Economic Statistics*, 11, 397-405.

Khan, E.A., Khan, M.G.M. and Ahsan, M.J. (2002). Optimum stratification: A mathematical programming approach. *Culcutta Statistical Association Bulletin*, 52 (special), 205-208.

Khan, M.G.M., Najmussehar, and Ahsan, M.J. (2005). Optimum stratification for exponential study variable under Neyman allocation. *Journal of Indian Society of Agricultural Statistics*, 59(2), 146-150.

Kozak, M. (2004). Optimal stratification using random search method in agricultural surveys. *Statistics in Transition*, 6, 5, 797-806.

Lavallée, P. (1987). Some contributions to optimal stratification. Master Thesis, Carleton University, Ottawa, Canada.

Lavallée, P. (1988). Two-way optimal stratification using dynamic programming. *Proceedings of the Section on Survey Research Methods*, American Statistical Association, Virginia, 646-651.

Lavallée, P., and Hidiroglou, M. (1988). On the stratification of skewed populations. *Survey Methodology*, 14, 33-43.

Lednicki, B., and Wieczorkowski, R. (2003). Optimal stratification and sample allocation between subpopulations and strata. *Statistics in Transition*, 6, 287-306.

Mahalanobis, P.C. (1952). Some aspects of the design of sample surveys. *Sankhyā*, 12, 1-7.

Murthy, M.N. (1967). *Sampling Theory and Methods*. Statistical Publishing Society, Calcutta.

Nelder, J.A., and Mead, R. (1965). A simplex method for function minimization. *Computer Journal*, 7, 308-313.

Nicolini, G. (2001). A method to define strata boundaries. *Departmental Working Papers 2001-01*, Department of Economics, University of Milan, Italy (www.economia.unimi.it/pubb/wp83.pdf).

Niemiro, W. (1999). Konstrukcja optymalnej stratyfikacja metoda poszukiwan losowych. (Optimal Stratification using Random Search Method). *Wiadomosci Statystyczne*, 10, 1-9.

Rivest, L.-P. (2002). A Generalization of Lavallée and Hidiroglou algorithm for stratification in business survey. *Survey Methodology*, 28, 191-198.

Sethi, V.K. (1963). A note on optimum stratification of population for estimating the population mean. *Australian Journal of Statistics*, 5, 20-33.

Sweet, E.M., and Sigman, R.S. (1995). Evaluation of model-assisted procedures for stratifying skewed populations using auxiliary data. *Proceedings of the Survey Research Methods Section*, American Statistical Association, 491-496.

Unnithan, V.K.G. (1978). The minimum variance boundry points of stratification. *Sankhyā*, 40, C, 60-72.

Unnithan, V.K.G., and Nair, N.U. (1995). Minimum variance stratification. *Communications in Statistics*, 24(1), 275-284.