On Optimum Allocation in Stratified Random Sampling

JONGHO KIM

1. Introduction

In this paper we will consider the optimum allocation for estimation of population means in stratified simple random sampling in case a multi-variate problem is treated in a normal population. In ordinary sampling survey we should consider many variables at the same time, but we treat the estimation problems of these variables independently, and the optimum allocation method by Neyman is obliged to be modified so that a certain compromised size is taken from among different sample sizes for each variate. In these circumstances we must know the correlation coefficients among the variables in each stratum, but if it were possible to get from previous experiences, we would get an optimum allocation procedure for the estimation of means. Even if we stratify the sampling units by each variate separately as used in the ordinary case, we can also get an optimum size of allocation for the multi-variate case.

2. Optimum allocation

As the purpose for stratified random sampling, we can say the enhancement of precision and the decrease of cost. According to the way how we can come to the purpose that we try to, because the sample sizes n_i in the respective strata must be determined before sampling, the problem of allocation of the sample size for each stratum is happened to us. The better method of allocation we use, the better precision of sample will be. Accordingly, in order to get the best precision of the sample, we have to determine the sample size n_i in each stratum. In this case, the best way of determination of the sample size n_i will be obtained by the standardization of a certain comparison. The standard are variance and cost: It is desirable to take the least variance as much as when we use a certain amount of cost and in case error is constant, we can decrease the cost. Moreover, after determination of the cost function, we will be able to determine the sample size n_i in each stratum for variance to become the minimum under the condition of the cost is constant. This determination of n_i is called a optimum allocation of the sample size in each stratum. In optimum allocation, we can get minimum variance when we give the constant cost. This allocation of the sample size is called optimum allocation by Deming. On the other hand, under the condition for the sample size being constant the allocation of the sample size in the stratum taking minimum variance is called optimum allocation by Neyman. And

Received by the editors September 26, 1968.

though we can get the allocation if we only know variance within stratum in optimum allocation by Neyman, we can know variance within stratum or not. If we ignore variance within stratum and put them into the same, there is a method which we make the allocation of the sample size in each stratum. We call it the proportional allocation. It is possible to apply these various allocation not only to one variate population but also to multi-variate population.

3. Allocation method for multi-variate case

Let us now consider the estimation problem for the means of a certain multi-variate normal population by stratified random sampling.

First of all we treat in case of two-variables x and y.

NOTATION;

K: the number of strata

N : the sizes of the population

n: the sizes of the sample

N_i: the sizes of the i-th stratum

n_i: the sample sizes of the i-th stratum

X and Y: the population means

$\bar{x}$ and $\bar{y}$: the sample estimates

$\sigma_{\bar{x}}^2$: the variance of $\bar{x}$

$\sigma_{\bar{y}}^2$: the variance of $\bar{y}$

$\sigma_{\bar{x}\bar{y}}$: the covariance of $\bar{x}$ and $\bar{y}$

ρ: the correlation coefficient of $\bar{x}$ and $\bar{y}$

Then the probability density function of two-variate normal distribution is given in the following:

$$(1)\qquad f(\bar{x},\bar{y})=\frac{1}{2\pi\sigma_{\bar{x}}\sigma_{\bar{y}}\sqrt{1-\rho^2}}\exp\left[-\frac{1}{2(1-\rho^2)}\left\{\left(\frac{\bar{x}-\bar{X}}{\sigma_{\bar{x}}}\right)^2-2\rho\frac{\bar{x}-\bar{X}}{\sigma_{\bar{x}}}\frac{\bar{y}-\bar{Y}}{\sigma_{\bar{y}}}+\left(\frac{\bar{y}-\bar{Y}}{\sigma_{\bar{y}}}\right)^2\right\}\right]$$
$$(-1<\rho<1,\quad \sigma_{\bar{x}}>0,\quad \sigma_{\bar{y}}>0)$$

If we put

$$-\frac{1}{2(1-\rho^2)}\left\{\left(\frac{\bar{x}-\bar{X}}{\sigma_{\bar{x}}}\right)^2-2\rho\frac{\bar{x}-\bar{X}}{\sigma_{\bar{x}}}\frac{\bar{y}-\bar{Y}}{\sigma_{\bar{y}}}+\left(\frac{\bar{y}-\bar{Y}}{\sigma_{\bar{y}}}\right)^2\right\}=-\frac{\lambda^2}{2}$$

then

$$(2)\qquad \left(\frac{\bar{x}-\bar{X}}{\sigma_{\bar{x}}}\right)^2-2\rho\frac{\bar{x}-\bar{X}}{\sigma_{\bar{x}}}\frac{\bar{y}-\bar{Y}}{\sigma_{\bar{y}}}+\left(\frac{\bar{y}-\bar{Y}}{\sigma_{\bar{y}}}\right)^2=\lambda^2(1-\rho^2).$$

Therefore we know that the function $f(\bar{x},\bar{y})$ is constant without reference to $\bar{x}$ and $\bar{y}$. We see that the locus of point $(\bar{x},\bar{y})$ satisfying above condition is a quadratic curve. By the transformation

$$\frac{\bar{x}-\bar{X}}{\sigma_{\bar{x}}}=\frac{1}{\sqrt{2}}(\xi-\eta),\quad \frac{\bar{y}-\bar{Y}}{\sigma_{\bar{y}}}=\frac{1}{\sqrt{2}}(\xi+\eta)$$

we have

$$(3)\qquad (1-\rho)\xi^2+(1+\rho)\eta^2=\lambda^2(1-\rho^2).$$

Evidently the equation (3) shows an ellipse with center $(\bar{X}, \bar{Y})$, the ellipse is called an ellipse of concentration. The area of the ellipse of concentration is proportional to $\sqrt{1-\rho^2}$ when λ, $\sigma_{\bar{x}}$ and $\sigma_{\bar{y}}$ are constant. The area decrease and points $(\bar{x}, \bar{y})$ are dense about $(\bar{X}, \bar{Y})$ when $|\rho|$ increase. Also the ellipse of concentration is given by the variance-covariance matrix, namely

$$(4) \qquad (\bar{x}-\bar{X}, \ \bar{y}-\bar{Y})\begin{pmatrix}\sigma_{\bar{x}}^2 & \sigma_{\bar{x}\bar{y}} \\ \sigma_{\bar{x}\bar{y}} & \sigma_{\bar{y}}^2\end{pmatrix}^{-1}(\bar{x}-\bar{X}, \ \bar{y}-\bar{Y})'=(\xi, \ \eta)\lambda^{-1}(\xi, \ \eta)'.$$

Making use of the expressions

$$(5) \qquad \bar{x}=\frac{1}{N}\sum_{i=1}^{K}N_i\bar{x}_i, \qquad \bar{y}=\frac{1}{N}\sum_{i=1}^{K}N_i\bar{y}_i$$

for the estimation of means $\bar{X}$ and $\bar{Y}$ in stratified random sampling with

$$(6) \qquad \bar{x}_i=\frac{1}{n_i}\sum_{l=1}^{n_i}x_{il}, \qquad \bar{y}_j=\frac{1}{n_j}\sum_{k=1}^{n_j}y_{jk}.$$

We obtain

$$(7) \qquad E(\bar{x}\,\bar{y})=\frac{1}{N^2}E\Big(\sum_{i=1}^{K}\sum_{j=1}^{K}N_iN_j\bar{x}_i\bar{y}_j\Big)=\frac{1}{N^2}\sum_{i=1}^{K}\sum_{j=1}^{K}N_iN_j\,E(\bar{x}_i\bar{y}_j)$$

$$(8) \qquad E(\bar{x}_i\bar{y}_i)=\frac{N_i-n_i}{N_i-1}\,\frac{1}{n_i}\Big(\frac{1}{N_i}\sum_{l=1}^{N_i}X_{il}Y_{il}\Big)+\frac{n_i-1}{n_i}\,\frac{N_i}{N_i-1}\,\bar{X}_i\bar{Y}_i$$

and for $i \neq j$

$$(9) \qquad E(\bar{x}_i\bar{y}_j)=\frac{1}{N_iN_j}\sum_{l=1}^{N_i}\sum_{j=1}^{N_j}X_{il}Y_{jk}.$$

Therefore, it follows from (7), (8) and (9) that

$$(10) \qquad \sigma_{\bar{x}}^2=E(\bar{x}^2)-\{E(\bar{x})\}^2=\frac{1}{N^2}\sum_{i=1}^{K}N_i^2\,\frac{N_i-n_i}{N_i-1}\,\frac{\sigma_{x_i}^2}{n_i}$$

$$(11) \qquad \sigma_{\bar{y}}^2=E(\bar{y}^2)-\{E(\bar{y})\}^2=\frac{1}{N^2}\sum_{i=1}^{K}N_i^2\,\frac{N_i-n_i}{N_i-1}\,\frac{\sigma_{y_i}^2}{n_i}$$

$$(12) \qquad \begin{aligned}\sigma_{\bar{x}\bar{y}} &= E(\bar{x}\bar{y})-E(\bar{x})E(\bar{y}) \\ &= \frac{1}{N^2}\sum_{i=1}^{K}N_i^2\Big\{\frac{N_i-n_i}{N_i-1}\,\frac{1}{n_i}\Big(\frac{1}{N_i}\sum_{l=1}^{N_i}X_{il}Y_{il}\Big)+\frac{n_i-1}{n_i}\,\frac{N_i}{N_i-1}\,\bar{X}_i\bar{Y}_i\Big\} \\ &\quad +\frac{1}{N^2}\sum_{i\neq j}^{K}\sum^{K}N_iN_j\frac{1}{N_iN_j}\sum_{l=1}^{N_i}\sum_{k=1}^{N_j}X_{il}Y_{jk}-\frac{1}{N^2}\sum_{i=1}^{K}N_i^2\bar{X}_i\bar{Y}_i-\frac{1}{N^2}\sum_{i\neq j}^{K}\sum^{K}N_iN_j\bar{X}_i\bar{Y}_j \\ &= \frac{1}{N^2}\sum_{i=1}^{K}N_i^2\,\frac{N_i-n_i}{N_i-1}\,\frac{\sigma_{x_iy_i}}{n_i}.\end{aligned}$$

Thus the equation (4) becomes

$$(13) \qquad (\xi, \ \eta)\lambda^{-1}(\xi, \ \eta)'=(\xi, \ \eta)\begin{pmatrix}\dfrac{\sigma_{\bar{x}}^2}{\sigma_{\bar{x}}^2\sigma_{\bar{y}}^2-\sigma_{\bar{x}\bar{y}}^2} & -\dfrac{\sigma_{\bar{x}\bar{y}}}{\sigma_{\bar{x}}^2\sigma_{\bar{y}}^2-\sigma_{\bar{x}\bar{y}}^2} \\ -\dfrac{\sigma_{\bar{x}\bar{y}}}{\sigma_{\bar{x}}^2\sigma_{\bar{y}}^2-\sigma_{\bar{x}\bar{y}}^2} & \dfrac{\sigma_{\bar{y}}^2}{\sigma_{\bar{x}}^2\sigma_{\bar{y}}^2-\sigma_{\bar{x}\bar{y}}^2}\end{pmatrix}(\xi, \ \eta)'$$

$$= \frac{\xi^2\sigma_{\bar{y}}^2 - 2\xi\eta\sigma_{\bar{x}}\sigma_{\bar{y}} + \eta^2\sigma_{\bar{y}}^2}{\sigma_{\bar{x}}^2\sigma_{\bar{y}}^2 - \sigma_{\bar{x}\bar{y}}^2}$$

$$= \frac{\xi^2\frac{1}{N^2}\sum_{i=1}^{K}N_i^2\frac{N_i-n_i}{N_i-1}\frac{\sigma_{y_i}^2}{n_i} - 2\xi\eta\frac{1}{N^2}\sum_{i=1}^{K}N_i^2\frac{N_i-n_i}{N_i-1}\frac{\sigma_{x_iy_i}}{n_i} + \eta^2\frac{1}{N^2}\sum_{i=1}^{K}N_i^2\frac{N_i-n_i}{N_i-1}\frac{\sigma_{x_i}^2}{n_i}}{\sigma_{\bar{x}}^2\sigma_{\bar{y}}^2 - \sigma_{\bar{x}\bar{y}}^2}$$

and the area of the ellipse of concentration is given by $\pi\sqrt{\sigma_{\bar{x}}^2\sigma_{\bar{y}}^2 - \sigma_{\bar{x}\bar{y}}^2}$.

Hence, in order to decide the optimum size n_i for the minimum error in the simultaneous estimation, we have to minimize the value $(c_{\bar{x}}^2\sigma_{\bar{y}}^2 - \sigma_{\bar{x}\bar{y}}^2)$ under the condition $n = \sum_{i=1}^{K} n_i$.

Let us consider an equation

$$(14) \quad Q = \left(\sum_{i=1}^{K} N_i^2\frac{N_i-n_i}{N_i-1}\frac{\sigma_{x_i}^2}{n_i}\right)\left(\sum_{i=1}^{K} N_i^2\frac{N_i-n_i}{N_i-1}\frac{\sigma_{y_i}^2}{n_i}\right) - \left(\sum_{i=1}^{K} N_i^2\frac{N_i-n_i}{N_i-1}\frac{\sigma_{x_iy_i}}{n_i}\right)^2 + \lambda\sum_{i=1}^{K} n_i$$

where λ is a Lagrange's multiplier. Differentiating with respect to n_i and putting zero, we have

$$(15) \qquad \frac{N_i^2}{n_i^2}(c_{x_i}^2\sigma_{\bar{y}}^2 + \sigma_{y_i}^2\sigma_{\bar{x}}^2 - 2\sigma_{x_iy_i}\sigma_{\bar{x}\bar{y}}) = \lambda.$$

From (15), we have

$$(16) \qquad n_i = N_i\sqrt{\sigma_{x_i}^2\sigma_{\bar{y}}^2 + \sigma_{y_i}^2\sigma_{\bar{x}}^2 - 2\sigma_{x_iy_i}\sigma_{\bar{x}\bar{y}}} \,/\sqrt{\lambda}$$

hence n_i is apparently proportional to N_i and it depends on the quantities σ_{x_i}, σ_{y_i} and σ_{xy_i}. Putting

$$\sigma_{x_iy_i} = \rho_i\sigma_{x_i}\sigma_{y_i} \quad \text{and} \quad N_i - n_i \fallingdotseq N_i - 1$$

according to the equation (15), we have

$$N_i^2(c_{x_i}^2\sigma_{\bar{y}}^2 + \sigma_{y_i}^2\sigma_{\bar{x}}^2 - 2\sigma_{x_i\bar{y}_i}\sigma_{\bar{x}\bar{y}}) = \lambda n_i^2$$

and substituting the expressions

$$\sigma_{\bar{x}}^2 = \frac{1}{N^2}\sum_{i=1}^{K}N_i^2\frac{\sigma_{x_i}^2}{n_i}, \quad \sigma_{\bar{y}}^2 = \frac{1}{N^2}\sum_{i=1}^{K}N_i^2\frac{\sigma_{y_i}^2}{n_i}, \quad \sigma_{\bar{x}\bar{y}} = \frac{1}{N^2}\sum_{i=1}^{K}N_i^2\frac{\sigma_{x_iy_i}}{n_i}$$

we get

$$(17) \qquad N_i^2\left(\sigma_{x_i}^2\sum_{i=1}^{K}N_i^2\frac{\sigma_{y_i}^2}{n_i} + \sigma_{y_i}^2\sum_{i=1}^{K}N_i^2\frac{\sigma_{x_i}^2}{n_i} - 2\sigma_{x_iy_i}\sum_{i=1}^{K}N_i^2\frac{\sigma_{x_iy_i}}{n_i}\right) = \lambda N^2 n_i^2.$$

Putting

$$a_{ij} = N_i^2N_j^2(\sigma_{x_i}^2\sigma_{y_j}^2 + \sigma_{y_i}^2\sigma_{x_j} - 2\sigma_{x_iy_i}\sigma_{x_jy_j}) > 0$$

the equation (17) represented by

$$(18) \qquad \frac{a_{i1}}{n_1} + \frac{a_{i2}}{n_2} + \cdots\cdots + \frac{a_{ik}}{n_k} = \lambda N^2 n_i^2, \quad i = 1,\ 2, \cdots\cdots, k.$$

Thus the optimum size of allocation is obtained by solving the simultaneous equations (18). For this purpose, we put $n_i = \frac{1}{u_i}$, then

$$a_{i1}u_1 + a_{i2}u_2 + \cdots\cdots + a_{ik}u_k = \frac{\lambda N^2}{u_i^2}, \qquad i = 1,\ 2, \cdots\cdots, k. \tag{19}$$

At first, we insert the first aproximate value $u^{(1)} = (u_1^{(1)}, u_2^{(1)}, \cdots\cdots, u_k^{(1)}) = (1, 1 \cdots\cdots, 1)$ into the left side of the equation (19) and get the second approximate value $u^{(2)} = (u_1^{(2)}, u_2^{(2)}, \cdots\cdots, u_K^{(2)})$ with

$$u_i^{(2)} = \sqrt{\lambda}\, N\Big(\sum_{j=1}^{K} a_{ij}\, u_j^{(1)}\Big)^{-\frac{1}{2}} \tag{20}$$

Inserting this $u^{(2)}$ into the left side of the equation (19), we have the third approximate value $u^{(3)}$ and so on. We arrive at a limit.

A proof of the existence of this limit is given in the following :

From the above assumption, we get

$$u_i^{(2)} = \sqrt{\lambda}N\Big(\sum_{j=1}^{K} a_{ij}\, u_j^{(1)}\Big)^{-\frac{1}{2}} = \sqrt{\lambda}N\Big(\sum_{j=1}^{K} a_{ij}\Big)^{-\frac{1}{2}}$$

$$u_i^{(3)} = \sqrt{\lambda}N\Big(\sum_{j=1}^{K} a_{ij}\, u_j^{(2)}\Big)^{-\frac{1}{2}}$$

$$u_i^{(4)} = \sqrt{\lambda}N\Big(\sum_{j=1}^{K} a_{ij}\, u_j^{(3)}\Big)^{-\frac{1}{2}}$$

$$u_i^{(4)} - u_i^{(2)} = \sqrt{\lambda}N\Big(\sum_{j=1}^{K} a_{ij}\, u_j^{(3)}\Big)^{-\frac{1}{2}} - \sqrt{\lambda}N\Big(\sum_{j=1}^{K} a_{ij}\Big)^{-\frac{1}{2}}$$

$$= \sqrt{\lambda}N\Big(\sum_{j=1}^{K} a_{ij}\, u_j^{(3)}\Big)^{-\frac{1}{2}}\Big(\sum_{j=1}^{K} a_{ij}\Big)^{-\frac{1}{2}}\Big\{\Big(\sum_{j=1}^{K} a_{ij}\Big)^{\frac{1}{2}} - \Big(\sum_{j=1}^{K} a_{ij}\, u_j^{(3)}\Big)^{\frac{1}{2}}\Big\}$$

$$= \sqrt{\lambda}N\Big(\sum_{j=1}^{K} a_{ij}\, u_j^{(3)}\Big)^{-\frac{1}{2}}\Big(\sum_{j=1}^{K} a_{ij}\Big)^{-\frac{1}{2}}\Big\{\Big(\sum_{j=1}^{K} a_{ij}\, u_j^{(3)}\Big)^{\frac{1}{2}} + \Big(\sum_{j=1}^{K} a_{ij}\Big)^{\frac{1}{2}}\Big\}^{-1}\sum_{j=1}^{K} a_{ij}(1 - u_j^{(3)}).$$

We may assume that $0 < u_j^{(3)} \leqq 1$, and obtain $u_i^{(4)} \geqq u_i^{(2)}$.

Similarly,

$$u_i^{(5)} - u_i^{(3)} = \sqrt{\lambda}\,N\Big(\sum_{j=1}^{K} a_{ij}\, u_j^{(4)}\Big)^{-\frac{1}{2}}\Big(\sum_{j=1}^{K} a_{ij}\, u_j^{(2)}\Big)^{-\frac{1}{2}}\Big\{\Big(\sum_{j=1}^{K} a_{ij}\, u_j^{(2)}\Big)^{\frac{1}{2}} - \Big(\sum_{j=1}^{K} a_{ij}\, u_j^{(4)}\Big)^{\frac{1}{2}}\Big\}$$

$$= \sqrt{\lambda}\,N\Big(\sum_{j=1}^{K} a_{ij}\, u_j^{(4)}\Big)^{-\frac{1}{2}}\Big(\sum_{j=1}^{K} a_{ij}\, u_j^{(2)}\Big)^{-\frac{1}{2}}\Big\{\Big(\sum_{j=1}^{K} a_{ij}\, u_j^{(4)}\Big)^{\frac{1}{2}} + \Big(\sum_{j=1}^{K} a_{ij}\, u_j^{(2)}\Big)^{\frac{1}{2}}\Big\}^{-1}\Big\{\sum_{j=1}^{K} a_{ij}(u_j^{(2)} - u_j^{(4)})\Big\}$$

Therefore, $u_i^{(5)} \leqq u_i^{(3)}$

Thus we have the following

$$0 < u_i^{(2)} \leqq u_i^{(4)} \leqq u_i^{(6)} \leqq \cdots\cdots \leqq 1, \text{ and} \tag{21}$$

$$0 < \cdots\cdots \leqq u_i^{(7)} \leqq u_i^{(5)} \leqq u_i^{(3)} < 1 \tag{22}$$

Hence we see that the limits $u_i^{(2k)}$ and $u_i^{(2k+1)}$ exist.

Moreover, since

(23) $$u_i^{(2k+1)} \geqq u_i^{(2k)}$$

(24) $$u_i^{(2k)} \leqq u_i^{(2k-1)},$$

we get

$$(u_i^{(2k-3)} - u_i^{(2k+2)}) - (u_i^{(2k+1)} - u_i^{(2k)}) = (u_i^{(2k+3)} - u_i^{(2k+1)}) + (u_i^{(2k)} - u_i^{(2k-2)}) \leqq 0.$$

Hence

$$|u_i^{(2k+3)} - u_i^{(2k+2)}| \leqq |u_i^{(2k+1)} - u_i^{(2k)}|,$$

and

$$\lim_{k\to\infty} |u_i^{(2k+1)} - u_i^{(2k)}| = 0.$$

This shows that the existence of a limit $u_i^{(k)}$.

4. Method of stratification

As was seen in the preceding section, we may treat the equation (14) in order to stratify the sampling units to get better estimates of means. The essential part of the method for a proper stratification is to minimize the expression such as

(25) $$G = \Big(\sum_{i=1}^{K} \frac{N_i^2}{n_i} \sigma_{x_i}^2\Big)\Big(\sum_{i=1}^{K} \frac{N_i^2}{n_i} \sigma_{y_i}^2\Big) - \Big(\sum_{i=1}^{K} \frac{N_i^2}{n_i} \sigma_{x_i y_i}\Big)^2$$
$$= \sum_{i=1}^{K} \frac{N_i^4}{n_i^2} \sigma_{x_i}^2 \sigma_{y_i}^2 (1-\rho_i^2) + \sum_{i \neq j}^{K} \sum^{K} \frac{N_i^2}{n_i} \frac{N_j^2}{n_j} \sigma_{x_i}\sigma_{y_j}(\sigma_{x_i}\sigma_{y_j} - \rho_i\rho_j\sigma_{x_j}\sigma_{y_i})$$

where ρ_i is the correlation coefficient between x and y in the i-th stratum. Let us take the second term of G in the last expression. In order to minimize this term, we have to stratify the sampling units so that $\sigma_{x_i}\sigma_{y_j} - \rho_i\rho_j\sigma_{x_j}\sigma_{y_i}$ approaches to zero. This means that the gradient of the regression line of y on x in the i-th stratum equals that of x on y in the j-th stratum for any i and j. Accordingly, the two regression lines must coincide in each stratum and have the same gradient with correlation coefficient ± 1. In this case the first term of G equals to zero and the value of G also equals to zero. In practical problems it is better to use a stratification method which makes ρ_i near to 1. Generally it is more difficult to make ρ_i near to -1 than to 1.

On the other hand, when the sampling units are stratified so that ρ_i becomes 0 in each stratum, which is used in many practical cases, the value of G becomes larger than that in the above mentioned better stratification.

5. General method for multi-variate stratification

We develope a similar argument in the case of 3 or more variables.

In the case of 3 varibles, we put

$$(26) \qquad Q = N^6 \begin{vmatrix} \sigma_{\bar{x}}^2 & \sigma_{\bar{x}\bar{y}} & \sigma_{\bar{x}\bar{z}} \\ \sigma_{\bar{x}\bar{y}} & \sigma_{\bar{y}}^2 & \sigma_{\bar{y}\bar{z}} \\ \sigma_{\bar{x}\bar{z}} & \sigma_{\bar{y}\bar{z}} & \sigma_{\bar{z}}^2 \end{vmatrix} + \lambda \sum_{i=1}^{K} n_i$$

Differentiating Q with respect to n_i and putting it to zero, we have

$$(27) \qquad \sum_{j=1}^{K} \sum_{k=1}^{K} \frac{a_{ijk}}{n_j n_k} = \lambda N^2 n_i^2, \qquad i = 1, 2, \cdots\cdots, k.$$

or putting $u_i = \frac{1}{n_i}$, we obtain

$$(28) \qquad \sum_{j=1}^{K} \sum_{k=1}^{K} a_{ijk} u_j u_k = \frac{\lambda N^2}{u_i^2}, \quad i = 1, 2, \cdots\cdots, k,$$

where

$$(29) \qquad a_{ijk} = N_i^2 N_j^2 N_k^2 \sum_{i,j,k} \begin{vmatrix} \sigma_{x_i}^2 & \sigma_{x_i y_i} & \sigma_{x_i z_i} \\ \sigma_{x_i y_j} & \sigma_{y_j}^2 & \sigma_{y_j z_j} \\ \sigma_{x_k z_k} & \sigma_{y_k z_k} & \sigma_{z_k}^2 \end{vmatrix}$$

where the summation covers all permutation (i, j, k). The method in solving this simultaneous equation (28) is carried out by the successive approximation in a similar way as in two variables.

As for a better stratification we can proceed in the same way as in two variables case.

Let us put

$$(30) \quad G = \left(\sum_{i=1}^{K} \frac{N_i^2}{n_i} \sigma_{x_i}^2\right)\left(\sum_{i=1}^{K} \frac{N_i^2}{n_i} \sigma_{y_i}^2\right)\left(\sum_{i=1}^{K} \frac{N_i^2}{n_i} \sigma_{z_i}^2\right) + 2\left(\sum_{i=1}^{K} \frac{N_i^2}{n_i} \sigma_{x_i y_i}\right)\left(\sum_{i=1}^{K} \frac{N_i^2}{n_i} \sigma_{y_i z_i}\right)\left(\sum_{i=1}^{K} \frac{N_i^2}{n_i} \sigma_{x_i z_i}\right)$$

$$- \left(\sum_{i=1}^{K} \frac{N_i^2}{n_i} \sigma_{x_i}^2\right)\left(\sum_{i=1}^{K} \frac{N_i^2}{n_i} \sigma_{y_i z_i}\right)^2 - \left(\sum_{i=1}^{K} \frac{N_i^2}{n_i} \sigma_{y_i}^2\right)\left(\sum_{i=1}^{K} \frac{N_i^2}{n_i} \sigma_{x_i z_i}\right)^2 - \left(\sum_{i=1}^{K} \frac{N_i^2}{n_i} \sigma_{z_i}^2\right)\left(\sum_{i=1}^{K} \frac{N_i^2}{n_i} \sigma_{x_i y_i}\right)^2$$

$$= \sum_{i,j,k} \frac{N_i^2 N_j^2 N_k^2}{n_i n_j n_k} \begin{vmatrix} \sigma_{x_i}^2 & \sigma_{x_i y_i} & \sigma_{x_i z_i} \\ \sigma_{x_j y_j} & \sigma_{y_j}^2 & \sigma_{y_j z_j} \\ \sigma_{x_k z_k} & \sigma_{y_k z_k} & \sigma_{z_k}^2 \end{vmatrix}$$

In order to minimize the value of this expression G, we try to minimize these values of determinants for any i, j and k. This is accomplished when the regression planes in all strata have the same direction cosines with multiple correlation coefficient 1. Because, if the value of the determinant for $j = k$ for example, is zero, then the direction cosines $R_{11}^{(j)}/\sigma_{x_j} : R_{12}^{(j)}/\sigma_{y_j} : R_{13}^{(j)}/\sigma_{z_j}$ of the normal line of regression plane of x on y and z in the i-th stratum satisfy the following equation

$$(31) \qquad \frac{R_{11}^{(j)}}{\sigma_{x_j}} \sigma + \frac{R_{12}^{(j)}}{\sigma_{y_j}} \rho_{x_i y_i} \sigma_{y_i} + \frac{R_{13}^{(j)}}{\sigma_{z_j}} \rho_{x_i z_i} \sigma_{z_i} = 0.$$

where $R_{lm}^{(j)}$ is the cofactor of ρ_{lm} in the determinant

$$R^{(j)}=\begin{vmatrix} 1 & \rho_{x_jy_j} & \rho_{x_jz_j} \\ \rho_{x_jy_j} & 1 & \rho_{x_jz_j} \\ \rho_{z_jy_j} & \rho_{y_jz_j} & 1 \end{vmatrix}$$

in the j-th stratum.

On the other hand if the multiple correlaiton coefficient equals 1 in the i-th stratum then $R^{(i)}=0$, hence

$$R_{11}^{(i)}+R_{12}^{(i)}\rho_{x_iy_i}+R_{13}^{(i)}\rho_{x_iz_i}=R^{(i)}=0$$

that is

$$\frac{R_{11}^{(i)}}{\sigma_{x_i}}\sigma_{x_i}+\frac{R_{12}^{(i)}}{\sigma_{y_i}}\rho_{x_iy_i}\sigma_{y_i}+\frac{R_{13}^{(i)}}{\sigma_{z_i}}\rho_{x_iz_i}\sigma_{z_i}=0. \quad (32)$$

From (31) and (32), we obtain two direction cosines which are proportional to each other, we can stratify the sampling units so that the regression planes in all strata become parallel to each other with multiple correlation coefficient 1.

Bibliography

1. W. G. Cochran, *Sampling Techniques.*
2. W. Edwards Deming, *Some theory of Sampling,* John Wiley, 1950.
3. H. Cramér, *Methematical Methods of Statistics.*
4. E. S. Pearson, *Sampling Problem in Industry,* Supplement to the Journal of the Royal Statistical Society, Vol. 1., No. 2, 1934.
5. S. S. Wilks, *Mathematical Statistics,* John Wiley.

Dongkook University