STRATIFIED RANDOM SAMPLING; GAIN IN PRECISION DUE TO STRATIFICATION IN THE CASE OF PROPORTIONAL ALLOCATION

TAKASHI YANAGAWA

(Received May 15, 1973; revised Aug. 30, 1973)

1. Introduction

In the current sample survey, the stratification plays an important role theoretically as well as in practical application. The main object of stratification has been thought to improve the precision in estimating the population mean and most of the studies on the stratified random sampling so far seem to have been made within its limitation. But in practice we consider the problem of estimating several population characteristics including means, variances, covariances, correlation coefficients etc. In our recent paper [4], [6] we explored the problem from a general point of view by estimating a fairly general class of functionals of the population distribution. Let $\mathcal{F}$ be a family of all continuous or all discontinuous distribution functions over p ($\geqq 1$) dimensional Euclidian space. Let us consider a functional $\theta(F)$ over $\mathcal{F}$ with given symmetric kernel $\varphi(x_1, \cdots, x_m)$ of degree m such that

$$\theta(F)=\int_{R^{pm}} \varphi(x_1, \cdots, x_m) \prod_{i=1}^{m} dF(x_i) , \qquad \text{for } \forall F \in \mathcal{F} . \tag{1.1}$$

Many of the important population characteristics can be expressed by (1.1); for example population mean of order ν is expressed by $\varphi(x)=x^\nu$, $\nu=1, 2, \cdots$ where $p=1$ and $m=1$; population variance is expressed by $\varphi(x_1, x_2)=(x_1-x_2)^2/2$ where $p=1$ and $m=2$, population covariance is expressed by $\varphi(x_1, x_2)=(x_1^{(1)}-x_2^{(1)})(x_1^{(2)}-x_2^{(2)})/2$ where $x_i=(x_i^{(1)}, x_i^{(2)})$, $i=1, 2$, $p=2$ and $m=2$ etc. In [4] we proposed a generalized U-statistics (say U_n^*) for the stratified random sample, which will be defined exactly in the next section, and in [6] U_n^* was shown to be a unique U.M.V. (uniformly minimum variance) unbiased estimator of $\theta(F)$ wrt (with respect to) the family $\mathcal{F}$ in the case of stratification where random variables in each sample taken from a stratum are independently distributed within strata. On the other hand it is well known that Hoeffding's U-statistic (say U_n) is a unique U.M.V. unbiased estimator of $\theta(F)$ wrt the family $\mathcal{F}$ in the case of simple random sampling (for example see Fraser [1], Chap. 4).

The purpose of this paper is to compare the variances of U_n^* with U_n, and then to demonstrate the gain in precision due to stratification in the case of proportional allocation. In order to explain our object in more detail we must note some results in previous work; (1) It seems to have been considered so far that a stratification always assured the gain in precision, but recently an example which opposed this fact was given by Wakimoto [3] in estimating a population variance. (2) Optimum allocation and optimum stratification have been shown to depend on the kernel φ ([4] and [6]), and then even if we could set the optimum allocation and stratification successfully for some specific population characteristic we could no more expect it to be the best but it might even be the worst way in estimating other population characteristics.

The result of this paper is roughly summarized as follows. In the case of proportional allocation gain in precision due to stratification is assured for any stratification and any underlying distribution F if we ignore the term $O(1/n^3)$, where n is a total sample size. Then the proportional allocation will be recommended in this paper in the minimax sense in estimating several population characteristics.

2. Theorems

It is shown in Isii and Taga [2] that the stratification of a population Π into L strata $\Pi_1, \cdots, \Pi_L$ may be represented by a decomposition of its distribution function F over R^p into a number L of functions $\bar{F}_1(x), \cdots, \bar{F}_L(x)$ with analogous properties to the distribution function such that the relation $\sum_{i=1}^{L} \bar{F}_i(x) = F(x)$ hold for all $x \in R^p$. Then the function $F_i(x)$ defined by $w_i^{-1}\bar{F}_i(x)$ may be considered as the distribution function corresponding to the ith stratum Π_i, where $w_i = \lim_{x\to(\infty,\cdots,\infty)} \bar{F}_i(x)$, $x=(x_1, \cdots, x_p)$, for $1 \leqq i \leqq L$, and then we get the relations

$$F(x) = \sum_{i=1}^{L} w_i F_i(x) \qquad \text{for all } x \in R^p \tag{2.1}$$

$$\sum_{i=1}^{L} w_i = 1 \quad \text{and} \quad w_i \geqq 0 \qquad \text{for all } i = 1 \sim L\,.$$

It is noted here that each w_i, the weight of the ith stratum Π_i ($1 \leqq i \leqq L$), should be known exactly to us after the stratification operation has been completed in some way.

Let us assume that number L of strata, total sample size n, sample allocation $(n_1, \cdots, n_L)$ and stratification procedure are preassigned in advance, and that each size N_i of the stratum Π_i is sufficiently large compared with n_i so that a sample from the population may be considered to be independent of each other.

Suppose that $X_{ij} \in R^p$, $j=1\sim n_i$ is a random sample of size n_i drawn from the ith stratum Π_i with the distribution function $F_i(x)$, $x \in R^p$, then our U_n^* is given as follows:

$$U_n^* = m! \sum_r \prod_{i=1}^{L} \frac{w_i^{r_i}}{r_i!} U_r \,, \tag{2.2}$$

$$U_r = \left[\binom{n_1}{r_1}\cdots\binom{n_L}{r_L}\right]^{-1} \sum_{\boldsymbol{\alpha}} \varphi(X_{1\alpha(11)},\cdots,X_{1\alpha(1r_1)},\cdots, X_{L\alpha(L1)},\cdots,X_{L\alpha(Lr_L)}) \,,$$

where the first summation in (2.2) should be taken over all combinations $(r_1,\cdots,r_L)$ of non-negative integers such that $\sum_{i=1}^{L} r_i = m$ and the second one be over all combinations $\boldsymbol{\alpha}=(\alpha(11),\cdots,\alpha(Lr_L))$ of positive integers corresponding to each $(r_1,\cdots,r_L)$ such that $1 \leqq \alpha(i1) < \cdots < \alpha(ir_i) \leqq n_i$ for $i=1\sim L$.

On the other hand let $X_i \in R^p$, $i=1\sim n$ be a simple random sample, then Hoeffding's U-statistic U_n is defined as follows.

$$U_n = \left[\binom{n}{m}\right]^{-1} \sum_{\boldsymbol{\alpha}} \varphi(X_{\alpha 1},\cdots,X_{\alpha m}) \,, \tag{2.3}$$

where the summation is extended over all combinations $(\alpha_1,\cdots,\alpha_m)$ of m integers such that $1 \leqq \alpha_1 < \cdots < \alpha_m \leqq n$.

From the U.M.V. property of U_n^* and U_n we can define the precision due to stratification by the difference of their variances. The gain in precision due to stratification for any stratifications, any population characteristics of type (1.1) and any underlying distribution F in the case of the proportional allocation will be demonstrated in the following theorem.

THEOREM 1. *Suppose* $\mathrm{E}\,\varphi^2(X_1,\cdots,X_m) < \infty$ *and the proportional allocation* $n_i = w_i n$ *for* $i=1\sim L$. *Then*
(1) *if* $m=1$, *where* m *is a degree of the kernel* φ, *we get*

$$\mathrm{Var}\,[U_n] - \mathrm{Var}\,[U_n^*] = \frac{1}{n} \sum_{i=1}^{L} w_i [\mathrm{E}\,\varphi(X_{i1}) - \theta(F)]^2 \,, \tag{2.3}$$

and then $\mathrm{Var}\,[U_n] \geqq \mathrm{Var}\,[U_n^*]$ *for any* F *and any partition* (2.1) *of* F.
(2) *if* $m \geqq 2$ *we get*

$$\begin{aligned} &\mathrm{Var}\,[U_n] - \mathrm{Var}\,[U_n^*] \\ &= \frac{m^2}{n}\left(1 - \frac{(m-1)^2}{n}\right) \sum_{i=1}^{L} w_i [\mathrm{E}\,\varphi_1(X_{i1}) - \theta(F)]^2 \end{aligned} \tag{2.4}$$

$$+\frac{m^2(m-1)^2}{2n^2}\sum_{i,j}^{L} w_i w_j \{[\mathrm{E}\,\varphi_2(X_{i1}, X_{j2})-\theta(F)]^2$$
$$+2\,\mathrm{E}\,[\varphi_2^j(X_{i1})-\varphi_1(X_{i1})-\mathrm{E}\,\varphi_2^j(X_{i1})+\mathrm{E}\,\varphi_1(X_{i1})]^2\} + O(1/n^3)\ ,$$

where

(2.5) $\varphi_c(x_1,\cdots,x_c)=\mathrm{E}\,[\varphi(X_1,\cdots,X_m)\,|\,X_i=x_i,\ i=1\sim c]$ for $c=1\sim m$

and

$$\varphi_2^j(x)=\int_{R^P}\varphi_2(x, y)dF_j(y) \qquad \text{for } j=1\sim L\ , \tag{2.6}$$

and then if we ignore the term $O(1/n^3)$ *we get* $\mathrm{Var}\,[U_n]\geqq \mathrm{Var}\,[U_n^*]$ *for any* m *and* n *which satisfy* $n\geqq(m-1)^2$, *for any distribution* F *and any partition* (2.1) *of* F.

Remark 1. If we do not ignore the term $O(1/n^3)$, the result of Theorem 1 (2) is not true. For example, consider $m=2$ and some "bad" stratification such that $F_1=\cdots=F_2=F$. Then we get from (3.2) and (3.5) in the next section

$$\begin{aligned}\mathrm{Var}\,[U_n]&-\mathrm{Var}\,[U_n^*]\\ &=-\frac{2}{n^2(n-1)}\sum_{i=1}^{L}\frac{w_i(1-w_i)}{n_i-1}\{\mathrm{Var}\,[\varphi(X_1, X_2)]\\ &\qquad -2\,\mathrm{Cov}\,[\varphi(X_1, X_2), \varphi(X_1, X_3)]\}\ ,\end{aligned}$$

where X_1, X_2, X_3 are i.i.d. random variables with distribution function F. But since it follows from Fraser [1], p. 227, Theorem 5.2 that

$$\mathrm{Var}\,[\varphi(X_1, X_2)]-2\,\mathrm{Cov}\,[\varphi(X_1, X_2), \varphi(X_1, X_3)]\geqq 0\ ,$$

then we get $\mathrm{Var}\,[U_n]\leqq \mathrm{Var}\,[U_n^*]$.

The above theorem can be extended to the two sample case. Let $X_i \in R^P$, $i=1\sim n_1$ be a simple random sample of size n_1 from a population Π^X with distribution function $F(x)$, $x\in R^P$, and $Y_j\in R^P$, $j=1\sim n_2$ be a simple random sample of size n_2 from the another population Π^Y with distribution function $G(y)$, $y\in R^P$. Let $\theta(F; G)$ be a real valued population characteristic with kernel $\varphi(x; y)$ such that

$$\theta(F; G)=\int_{R^{2P}}\varphi(x; y)dF(x)dG(y)\ . \tag{2.7}$$

A well-known generalized U-statistic which is a unique U.M.V. unbiased estimator of $\theta(F; G)$ wrt $\mathcal{F}$ is given as follows:

$$\tilde{U}=(n_1n_2)^{-1}\sum_{\alpha=1}^{n_1}\sum_{\beta=1}^{n_2}\varphi(x_\alpha; y_\beta)\ . \tag{2.8}$$

On the other hand in stratification suppose that the population Π^X with distribution function $F(x)$ and the population Π^Y with $G(y)$ are classified into L_1 and L_2 strata $\Pi_1^X, \cdots, \Pi_{L_1}^X$ and $\Pi_1^Y, \cdots, \Pi_{L_2}^Y$, respectively, in such a way the distribution functions F_i and G_j respectively corresponding to the ith stratum Π_i^X and jth stratum Π_j^Y satisfy the relations

$$F(x)=\sum_{i=1}^{L_1} v_i F_i(x) , \qquad 0\leqq v_i \leqq 1,\ \sum_{i=1}^{L_1} v_i=1 ,$$

(2.9)

$$G(y)=\sum_{j=1}^{L_2} w_j G_j(y) , \qquad 0\leqq w_j \leqq 1,\ \sum_{j=1}^{L_2} w_j=1 ,$$

for all $x \in R^P$ and $y \in R^P$.

Suppose for each $i=1\sim L_1$ and $j=1\sim L_2$ we have a sampling plan to take random samples $X_{i\alpha} \in R^P$, $\alpha=1\sim n_{1i}$ of size n_{1i} and $Y_{j\beta} \in R^P$, $\beta=1\sim n_{2j}$ of size n_{2j} from the ith stratum Π_i^X and Π_j^Y, respectively, then our corresponding estimator of $\theta(F; G)$ is proposed in [4] as follows.

$$\tilde{U}^*=\sum_{i=1}^{L_1} \sum_{j=1}^{L_2} \frac{v_i w_j}{n_{1i} n_{2j}} \sum_{\alpha=1}^{n_{1i}} \sum_{\beta=1}^{n_{2j}} \varphi(X_{i\alpha}; Y_{j\beta}) . \tag{2.10}$$

THEOREM 2. *Suppose* $\mathrm{E}\,\varphi^2(X_1; Y_1)<\infty$, *and the proportional allocation* $n_{1i}=v_i n_1$ *and* $n_{2j}=w_j n_2$ *for* $i=1\sim L_1$ *and* $j=1\sim L_2$, *then we get*

$$\begin{aligned}(2.11)\quad & \mathrm{Var}\,[\tilde{U}]-\mathrm{Var}\,[\tilde{U}^*] \\ &=\frac{(n_2-1)}{n_1 n_2} \sum_{i=1}^{L_1} v_i[\mathrm{E}\,\varphi_g(X_{i1})-\theta]^2+\frac{(n_1-1)}{n_1 n_2} \sum_{j=1}^{L_2} w_j[\mathrm{E}\,\varphi_f(Y_{j1})-\theta]^2 \\ &\quad +\frac{1}{n_1 n_2} \sum_{i=1}^{L_1} \sum_{j=1}^{L_2} v_i w_j \{[\mathrm{E}\,\varphi(X_{i1}; Y_{j1})-\theta]^2+\mathrm{E}\,[\varphi_{gj}(X_{i1})-\varphi_g(X_{i1}) \\ &\quad -\mathrm{E}\,\varphi_{gj}(X_{i1})+\mathrm{E}\,\varphi_g(X_{i1})]^2+\mathrm{E}\,[\varphi_{fi}(Y_{j1})-\varphi_f(Y_{j1}) \\ &\quad -\mathrm{E}\,\varphi_{fi}(Y_{j1})+\mathrm{E}\,\varphi_f(Y_{j1})]^2\} ,\end{aligned}$$

where $\theta=\theta(F; G)$, $\varphi_f(y)=\mathrm{E}\,[\varphi(X; Y)\,|\,Y=y]$, $\varphi_g(x)=\mathrm{E}\,[\varphi(X; Y)\,|\,X=x]$, $\varphi_{fi}(y)=\mathrm{E}\,[\varphi(X_{i1}; Y)\,|\,Y=y]$ *and* $\varphi_{gj}(x)=\mathrm{E}\,[\varphi(X; Y_j)\,|\,X=x]$, *and then we get* $\mathrm{Var}\,[\tilde{U}]\geqq \mathrm{Var}\,[\tilde{U}^*]$ *for any underlying distributions* F *and* G *and any partition* (2.9).

Remark 2. If $\theta(F; G)$ be a real valued population characteristic with kernel $\varphi(x_1, \cdots, x_{m_1}; y_1, \cdots, y_{m_2})$ such that

$$\theta(F; G)=\int_{R^{(m_1+m_2)p}} \varphi(x_1, \cdots, x_{m_1}; y_1, \cdots, y_{m_2}) \prod_{i=1}^{m_1} \prod_{j=1}^{m_2} dF(x_i)dG(y_j) ,$$

then we can give an estimator of $\theta(F; G)$ which is a generalized U-statistic based on a stratified random sample (see Yanagawa [5], p. 367). Further we can prove a corresponding result as Theorem 1 (2) if $n_1 \geqq$

$(m_1-1)^2$, $n_2 \geqq (m_2-1)^2$ and if we ignore the term $O(1/n^3)$, where $n=n_1+n_2$. The details, which can be given through similar discussions as the previous theorems, are abbreviated, since we need quite complicated notations and tedious computations. The author conjectures that a similar result holds for an estimating problem of more general population characteristics $\theta(F; G)$ whose kernel is given by $\varphi(x_1, \cdots, x_{m_1}; y_1, \cdots, y_{m_2}; \cdots; z_1, \cdots, z_{m_k})$.

3. Proof of the theorem

PROOF OF THEOREM 1. (1) is clear since for $m=1$ U_n^* and U_n result in respectively

$$U_n^* = \sum_{i=1}^{L} \frac{w_i}{n_i} \sum_{\alpha=1}^{n_i} \varphi(X_{i\alpha}) \qquad \text{and} \qquad U_n = \frac{1}{n} \sum_{i=1}^{n} \varphi(X_i) .$$

(2) First we note the following relations which are obtained easily by simple manipulation by using Stirling's formula such that $n! \sim (2\pi)^{1/2} n^{n+1/2} e^{-n}$.

$$\text{(3.1)} \qquad \frac{\binom{n-r}{s-c}}{\binom{n}{s}} = \begin{cases} 1 - rsn^{-1} + An^{-2} + O(n^{-3}) & \text{for } c=0 \text{ and } r, s \geqq 0 , \\ s[n^{-1} - (r-1)(s-1)n^{-2} + O(n^{-3})] & \text{for } c=1 \text{ and } r, s \geqq 1 , \\ s(s-1)[n^{-2} + O(n^{-3})] & \text{for } c=2 \text{ and } r, s \geqq 2 , \end{cases}$$

where A is a constant.

Let us put

$$\zeta_{crs} = \text{Cov}\,[\varphi_{cr}(X_{11}, \cdots, X_{1c_1}, \cdots, X_{L1}, \cdots, X_{Lc_L}), \\ \varphi_{cs}(X_{11}, \cdots, X_{1c_1}, \cdots, X_{L1}, \cdots, X_{Lc_L})] ,$$

where

$$\varphi_{cr}(x_{11}, \cdots, x_{1c_1}, \cdots, x_{L1}, \cdots, x_{Lc_L}) \\ = \text{E}\,[\varphi(X_{11}, \cdots, X_{1r_1}, \cdots, X_{L1}, \cdots, X_{Lr_L}) \,|\, X_{ij} = x_{ij}, \\ 1 \leqq i \leqq L,\ 1 \leqq j \leqq C_i] ,$$

$$\mathbf{r} = (r_1, \cdots, r_L) , \quad \mathbf{s} = (s_1, \cdots, s_L) \quad \text{and} \quad \mathbf{c} = (c_1, \cdots, c_L) \\ \text{for } c_i \leqq \min(r_i, s_i) \qquad \text{for } i = 1 \sim L .$$

Then the variance of U_n^* is (see [6])

$$\text{(3.2)} \qquad \text{Var}\,[U_n^*] = (m!)^2 \sum_{\mathbf{r},\mathbf{s}} \prod_{i=1}^{L} \frac{w_i^{r_i+s_i}}{r_i! s_i!} \sum_{\mathbf{c}} \prod_{j=1}^{L} \frac{\binom{r_j}{c_j}\binom{n_j-r_j}{s_j-c_j}}{\binom{n_j}{s_j}} \zeta_{\mathbf{crs}} ,$$

where the first summation of the right-hand side of (3.2) should be taken over all combinations of non-negative integers $\boldsymbol{r}=(r_1,\cdots,r_L)$ and $\boldsymbol{s}=(s_1,\cdots,s_L)$ such that $\sum_{i=1}^{L} r_i=\sum_{i=1}^{L} s_i=m$, and the second one is all combinations of non-negative integers $\boldsymbol{c}=(c_1,\cdots,c_L)$ such that $c_i \leqq \min(r_i, s_i)$ for $1\leqq i\leqq L$. We want to expand $\mathrm{Var}[U_n^*]$ such as $\mathrm{Var}[U_n^*]=A_1n^{-1}+A_2n^{-2}+O(n^{-3})$. To get A_1 and A_2 exactly it is necessary and sufficient from (3.1) to treat three different cases of $\boldsymbol{c}$ such that (1) $\boldsymbol{c}=(1,0,\cdots,0),(0,1,0,\cdots,0),\cdots,(0,\cdots,0,1)$, (2) $\boldsymbol{c}=(1,1,0,\cdots,0),\cdots,(0,\cdots,0,1,1)$ and (3) $\boldsymbol{c}=(2,0,\cdots,0),\cdots,(0,\cdots,0,2)$. In the first case $\boldsymbol{c}=(0,\cdots,0,\overset{i}{\check{1}},0,\cdots,0)$ we get from (3.1) that

$$\prod_{j=1}^{L}\binom{r_j}{c_j}\binom{n_j-r_j}{s_j-c_j}\bigg/\binom{n_j}{s_j}$$
$$=\frac{r_is_i}{n_i}-\frac{r_is_i}{n_i}\sum_{j\neq i}^{L}\frac{r_js_j}{n_j}-\frac{r_i(r_i-1)s_i(s_i-1)}{n_i^2}+O\left(\frac{1}{n^3}\right).$$

Thus the right-hand side of (3.2) reduces to

$$B_1=(m!)^2\sum_{i=1}^{L}\frac{w_i^2}{n_i}\sum_{r,s}{}'\prod_{\alpha=1}^{L}\frac{w_\alpha^{r_\alpha+s_\alpha}}{r_\alpha!s_\alpha!}\zeta_{(1)rs}^{(i)}$$
$$-(m!)^2\sum_{i\neq j}^{L}\frac{w_i^2w_j^2}{n_in_j}\sum_{r,s}{}''\prod_{\alpha=1}^{L}\frac{w_\alpha^{r_\alpha+s_\alpha}}{r_\alpha!s_\alpha!}\zeta_{(1)rs}^{(i)}$$
$$-(m!)^2\sum_{i=1}^{L}\frac{w_i^4}{n_i^2}\sum_{r,s}{}''\prod_{\alpha=1}^{L}\frac{w_\alpha^{r_\alpha+s_\alpha}}{r_\alpha!s_\alpha!}\zeta_{(1)rs}^{(i)}+O\left(\frac{1}{n^3}\right),$$

where the summations $\sum'$ and $\sum''$ should be taken over all combinations of non-negative integers $\boldsymbol{r}=(r_1,\cdots,r_L)$ and $\boldsymbol{s}=(s_1,\cdots,s_L)$ such that $\sum_{i=1}^{L} r_i=\sum_{i=1}^{L} s_i=m-1$ and $\sum_{i=1}^{L} r_i=\sum_{i=1}^{L} s_i=m-2$, respectively, $\zeta_{(1)rs}^{(i)}=\mathrm{Cov}[\varphi_{1r}^{(i)}(X_{i1}), \varphi_{1s}^{(i)}(X_{i1})]$ and $\varphi_{1r}^{(i)}(x)=\mathrm{E}[\varphi(X_{11},\cdots,X_{1r_1},\cdots,X_{L1},\cdots,X_{Lr_L})\mid X_{i1}=x]$. Then we get

$$B_1=m^2\sum_{i=1}^{L}\frac{w_i^2}{n_i}\mathrm{Var}[\varphi_1(X_{i1})]-m^2(m-1)^2\sum_{i\neq j}^{L}\frac{w_i^2w_j^2}{n_in_j}\mathrm{Cov}[\varphi_2(X_{i1},X_{j2}),$$
$$\varphi_2(X_{i1},X_{j3})]-m^2(m-1)^2\sum_{i=1}^{L}\frac{w_i^4}{n_i^2}\mathrm{Cov}[\varphi_2(X_{i1},X_{i2}),$$
$$\varphi_2(X_{i1},X_{i3})]+O\left(\frac{1}{n^3}\right),$$

where

$$\varphi_c(x_1,\cdots,x_c)=\mathrm{E}[\varphi(X_1,\cdots,X_m)\mid X_i=x_i,\ i=1\sim c]. \tag{3.3}$$

In the second case $\boldsymbol{c}=(0,\cdots,0,\overset{i}{\check{1}},0,\cdots,0,\overset{j}{\check{1}},0,\cdots,0)$ we get from (3.1) that

$$\prod_{\alpha=1}^{L}\binom{r_\alpha}{c_\alpha}\binom{n_\alpha-r_\alpha}{s_\alpha-c_\alpha}\Big/\binom{n_\alpha}{s_\alpha}=\frac{r_ir_js_is_j}{n_in_j}+O\Big(\frac{1}{n^3}\Big)\,.$$

Thus the right-hand side of (3.2) reduces to

$$B_2=(m!)^2\sum_{i<j}^{L}\frac{w_i^2w_j^2}{n_in_j}\sum_{r,s}{}''\prod_{\alpha=1}^{L}\frac{w_\alpha^{r_\alpha+s_\alpha}}{r_\alpha!s_\alpha!}\zeta_{(1)rs}^{(i,j)}\,,$$

where the summation $\sum''$ is the same as above, $\zeta_{(1)rs}^{(i,j)}=\mathrm{Cov}\,[\varphi_{(1)r}^{(i,j)}(X_{i1},X_{j1}),\ \varphi_{(1)s}^{(i,j)}(X_{i1},X_{j1})]$ and

$$\varphi_{(1)r}^{(i,j)}(x,y)=\mathrm{E}\,[\varphi(X_{11},\cdots,X_{1r_1},\cdots,X_{L1},\cdots,X_{Lr_L})\,|\,X_{i1}=x,\ X_{j1}=y]\,.$$

Then we get

$$B_2=m^2(m-1)^2\sum_{i<j}^{L}\frac{w_i^2w_j^2}{n_in_j}\mathrm{Var}\,[\varphi_2(X_{i1},X_{j1})]+O\Big(\frac{1}{n^3}\Big)\,.$$

In the last case $\boldsymbol{c}=(0,\cdots,0,\overset{i}{\check{2}},0,\cdots,0)$ we get from (3.1) that

$$\prod_{j=1}^{L}\binom{r_j}{c_i}\binom{n_i-r_j}{s_j-c_j}\Big/\binom{n_j}{s_j}=\frac{r_i(r_i-1)s_i(s_i-1)}{2n_i^2}+O\Big(\frac{1}{n^3}\Big)\,.$$

Thus the right-hand side of (3.2) reduces to

$$B_3=(m!)^2\sum_{i=1}^{L}\frac{w_i^4}{2n_i^2}\sum_{r,s}{}''\prod_{\alpha=1}^{L}\frac{w_\alpha^{r_\alpha+s_\alpha}}{r_\alpha!s_\alpha!}\zeta_{(2)rs}^{(i)}\,,$$

where the summation $\sum''$ is the same as above,

$$\zeta_{(2)rs}^{(i)}=\mathrm{Cov}\,[\varphi_{(2)r}^{(i)}(X_{i1},X_{i2}),\ \varphi_{(2)s}^{(i)}(X_{i1},X_{i2})]$$

and

$$\varphi_{(2)r}^{(i)}(x,y)=\mathrm{E}\,[\varphi(X_{11},\cdots,X_{1r_1},\cdots,X_{L1},\cdots,X_{Lr_L})\,|\,X_{i1}=x,\ X_{i2}=y]\,.$$

Then

$$B_3=\frac{m^2(m-1)^2}{2}\sum_{i=1}^{L}\frac{w_i^4}{n_i^2}\mathrm{Var}\,[\varphi_2(X_{i1},X_{i2})]+O\Big(\frac{1}{n^3}\Big)\,.$$

Summarizing the above three results we get

$$\begin{aligned}\mathrm{Var}\,[U_n^*]&=B_1+B_2+B_3+O(1/n^3)\\&=m^2\sum_{i=1}^{L}\frac{w_i^2}{n_i}\mathrm{Var}\,[\varphi_1(X_{i1})]\\&\quad-m^2(m-1)^2\sum_{i,j}^{L}\frac{w_i^2w_j^2}{n_in_j}\mathrm{Cov}\,[\varphi_2(X_{i1},X_{j2}),\ \varphi_2(X_{i1},X_{j3})]\\&\quad+\frac{m^2(m-1)^2}{2}\sum_{i,j}^{L}\frac{w_i^2w_j^2}{n_in_j}\mathrm{Var}\,[\varphi_2(X_{i1},X_{j2})]+O\Big(\frac{1}{n^3}\Big)\,.\end{aligned}$$

Especially in the case of proportional allocation, $n_i = w_i n$ for $i = 1 \sim L$, we get

$$\begin{aligned} \text{(3.4)} \qquad \operatorname{Var}[U_n^*] = & \frac{m^2}{n} \sum_{i=1}^{L} w_i \operatorname{Var}[\varphi_1(X_{i1})] \\ & + \frac{1}{n^2} \Big\{ \frac{m^2(m-1)^2}{2} \sum_{i,j}^{L} w_i w_j \operatorname{Var}[\varphi_2(X_{i1}, X_{j2})] \\ & - m^2(m-1)^2 \sum_{i,j}^{L} w_i w_j \operatorname{Cov}[\varphi_2(X_{i1}, X_{j1}), \varphi_2(X_{i1}, X_{j3})] \Big\} \\ = & O(1/n^3) . \end{aligned}$$

On the other hand the variance of U_n is given as follows (see Fraser [1], p. 225):

$$\text{(3.5)} \qquad \operatorname{Var}[U_n] = \sum_{c=1}^{m} \binom{m}{c} \binom{n-m}{m-c} \zeta_c \Big/ \binom{n}{m} ,$$

where $\zeta_c = \operatorname{Cov}\{\varphi(X_1, \cdots, X_m), \varphi(X_1, \cdots, X_c, X_{m+1}, \cdots, X_{2m-c})\}$ and $X_1, X_2, \cdots, X_{2m-c}$ are i.i.d. random variables with distribution function $F(x)$, $x \in R^p$. By using (3.1) this reduces to

$$\text{(3.6)} \qquad \operatorname{Var}[U_n] = m^2\zeta_1/n + m^2(m-1)^2(\zeta_2 - 2\zeta_1)/2n^2 + O(1/n^3) .$$

Further from (2.1)

$$\text{(3.7)} \qquad \begin{aligned} \zeta_1 & = \operatorname{E}[\varphi_1(X_1) - \theta]^2 \\ & = \sum_{i=1}^{L} w_i \operatorname{E}[\varphi_1(X_{i1}) - \theta]^2 \\ & = \sum_{i=1}^{L} w_i \operatorname{Var}[\varphi_1(X_{i1})] + \sum_{i=1}^{L} w_i [\operatorname{E}\varphi_1(X_{i1}) - \theta]^2 , \\ \zeta_2 & = \operatorname{E}[\varphi_2(X_1, X_2) - \theta]^2 \\ & = \sum_{i,j}^{L} w_i w_j \operatorname{E}[\varphi_2(X_{i1}, X_{j2}) - \theta]^2 \\ & = \sum_{i,j}^{L} w_i w_j \operatorname{Var}[\varphi_2(X_{i1}, X_{j2})] + \sum_{i,j}^{L} w_i w_j [\operatorname{E}\varphi_2(X_{i1}, X_{j2}) - \theta]^2 \end{aligned}$$

where $\varphi_c(x_1, \cdots, x_c)$ is defined in (3.3) and X_{i1} is a random variable with distribution function $F_i(x)$, $x \in R^p$, for $i = 1 \sim L$. Then from (3.4), (3.6) and (3.7)

$$\begin{aligned} \text{(3.8)} \qquad & \operatorname{Var}[U_n] - \operatorname{Var}[U_n^*] \\ & = \frac{m^2}{n} \Big(1 - \frac{(m-1)^2}{n}\Big) \sum_{i=1}^{L} w_i (\operatorname{E}\varphi_1(X_{i1}) - \theta)^2 \\ & \quad + \frac{m^2(m-1)^2}{2n^2} \Big\{ \sum_{i,j}^{L} w_i w_j [\operatorname{E}\varphi_2(X_{i1}, X_{j2}) - \theta]^2 \end{aligned}$$

$$+2\sum_{i,j}^{L} w_i w_j [\mathrm{Cov}\,[\varphi_2(X_{i1}, X_{j2}), \varphi_2(X_{i1}, X_{j3})] - \mathrm{Var}\,[\varphi_1(X_{i1})]]\Big\}$$
$$+O(1/n^3)\,.$$

Since

$$\sum_{i,j}^{L} w_i w_j \{\mathrm{Cov}\,[\varphi_2(X_{i1}, X_{j2}), \varphi_2(X_{i1}, X_{j3})] - \mathrm{Var}\,[\varphi_1(X_{i1})]\}$$
$$=\sum_{i,j}^{L} w_i w_j\, \mathrm{E}\,\{\varphi_2^j(X_{i1}) - \varphi_1(X_{i1}) - \mathrm{E}\,\varphi_2^j(X_{i1}) + \mathrm{E}\,\varphi_1(X_{i1})\}^2\,,$$

where

$$\varphi_2^j(x) = \int \varphi_2(x, y)dF_j(y)\,.$$

Then we finally get

$$\mathrm{Var}\,[U_n] - \mathrm{Var}\,[U_n^*]$$
$$= \frac{m^2}{n}\left(1 - \frac{(m-1)^2}{n}\right)\sum_{i=1}^{L} w_i[\mathrm{E}\,\varphi_1(X_{i1}) - \theta]^2$$
$$+\frac{m^2(m-1)^2}{2n^2}\sum_{i,j}^{L} w_i w_j \{[\mathrm{E}\,\varphi_2(X_{i1}, X_{j2}) - \theta]^2$$
$$+2\,\mathrm{E}\,[\varphi_2^j(X_{i1}) - \varphi_1(X_{i1}) - \mathrm{E}\,\varphi_2^j(X_{i1}) + \mathrm{E}\,\varphi_1(X_{i1})]^2\} + O(1/n^3)\,.$$

Thus we complete the proof of Theorem 1.

PROOF OF THEOREM 2. Similarly as one sample case we get in the case of proportional allocation $n_{1_i} = v_i n_1$ and $n_{2_j} = w_j n_2$ for $i = 1 \sim L_1$, $j = 1 \sim L_2$, that

$$\mathrm{Var}\,[\tilde{U}^*] = \sum_{i=1}^{L_1} v_i\, \mathrm{Var}\,[\varphi_g(X_{i1})]/n_1 + \sum_{j=1}^{L_2} w_j\, \mathrm{Var}\,[\varphi_f(Y_{j1})]/n_2 \tag{3.9}$$
$$+\sum_{i=1}^{L_1}\sum_{j=1}^{L_2} v_i w_j \{\mathrm{Var}\,[\varphi(X_{i1}; Y_{j1})] - \mathrm{Cov}\,[\varphi(X_{i1}; Y_{j1}),$$
$$\varphi(X_{i1}; Y_{j2})] - \mathrm{Cov}\,[\varphi(X_{i1}; Y_{j1}), \varphi(X_{i2}; Y_{j1})]\}/n_1 n_2\,,$$

where $\varphi_f(y) = \mathrm{E}\,[\varphi(X; Y) \mid Y = y]$ and $\varphi_g(x) = \mathrm{E}\,[\varphi(X; Y) \mid X = x]$.

On the other hand the variance of $\tilde{U}$ is given by

$$\mathrm{Var}\,[\tilde{U}] = \mathrm{E}\,[\varphi_g(X_1) - \theta]^2(n_2 - 1)/n_1 n_2 + \mathrm{E}\,[\varphi_f(Y_1) - \theta]^2(n_1 - 1)/n_1 n_2$$
$$+\mathrm{E}\,[\varphi(X_1; Y_1) - \theta]^2/n_1 n_2\,.$$

Further by (2.9)

$$\mathrm{Var}\,[\tilde{U}] = \sum_{i=1}^{L_1} v_i\, \mathrm{Var}\,[\varphi_g(X_{i1})](n_2 - 1)/n_1 n_2 \tag{3.10}$$
$$+\sum_{i=1}^{L_1} v_i[\mathrm{E}\,\varphi_g(X_{i1}) - \theta]^2(n_2 - 1)/n_1 n_2$$

$$+\sum_{j=1}^{L_2} w_j \operatorname{Var}[\varphi_f(Y_{j1})](n_1-1)/n_1n_2$$
$$+\sum_{j=1}^{L_2} w_j[\mathrm{E}\,\varphi_f(Y_{j1})-\theta]^2(n_1-1)/n_1n_2$$
$$+\sum_{i=1}^{L_1}\sum_{j=1}^{L_2} v_iw_j \operatorname{Var}[\varphi(X_{i1};Y_{j1})]/n_1n_2$$
$$+\sum_{i=1}^{L_1}\sum_{j=1}^{L_2} v_iw_j[\mathrm{E}\,\varphi(X_{i1};Y_{j1})-\theta]^2/n_1n_2\,.$$

Thus from (3.9) and (3.10)

$$\begin{aligned}
&\operatorname{Var}[\tilde{\mathcal{U}}]-\operatorname{Var}[\tilde{\mathcal{U}}^*]\\
&=\sum_{i=1}^{L_1} v_i[\mathrm{E}\,\varphi_g(X_{i1})-\theta]^2(n_2-1)/n_1n_2\\
&\quad+\sum_{j=1}^{L_2} w_j[\mathrm{E}\,\varphi_f(Y_{j1})-\theta]^2(n_1-1)/n_1n_2\\
&\quad+\sum_{i=1}^{L_1}\sum_{j=1}^{L_2} v_iw_j[\mathrm{E}\,(X_{i1};Y_{j1})-\theta]/n_1n_2\\
&\quad+\sum_{i=1}^{L_1}\sum_{j=1}^{L_2} v_iw_j\{\operatorname{Cov}[\varphi(X_{i1};Y_{j1}),\varphi(X_{i1};Y_{j2})]-\operatorname{Var}[\varphi_g(X_{i1})]\\
&\qquad+\operatorname{Cov}[\varphi(X_{i1};Y_{j1}),\varphi(X_{i2};Y_{j1})]-\operatorname{Var}[\varphi_f(Y_{j1})]\}/n_1n_2\\
&=\frac{(n_2-1)}{n_1n_2}\sum_{i=1}^{L_1} v_i[\mathrm{E}\,\varphi_g(X_{i1})-\theta]^2+\frac{(n_1-1)}{n_1n_2}\sum_{j=1}^{L_2} w_j[\mathrm{E}\,\varphi_f(Y_{j1})-\theta]^2\\
&\quad+\frac{1}{n_1n_2}\sum_{i=1}^{L_1}\sum_{j=1}^{L_2} v_iw_j\{[\mathrm{E}\,\varphi(X_{i1};Y_{j1})-\theta]^2+\mathrm{E}\,[\varphi_{gj}(X_{i1})-\varphi_g(X_{i1})\\
&\quad-\mathrm{E}\,\varphi_{gj}(X_{i1})+\mathrm{E}\,\varphi_g(X_{i1})]^2+\mathrm{E}\,[\varphi_{fi}(Y_{j1})-\varphi_f(Y_{j1})\\
&\quad-\mathrm{E}\,\varphi_{fi}(Y_{j1})+\mathrm{E}\,\varphi_f(Y_{j1})]^2\}\,,
\end{aligned}$$

where

$$\varphi_{fi}(y)=\mathrm{E}\,[\varphi(X_{i1};Y)\,|\,Y=y] \quad\text{and}\quad \varphi_{gj}(x)=\mathrm{E}\,[\varphi(X;Y_{j1})\,|\,X=x]\,.$$

Thus we get Theorem 2.

Acknowledgement

The author expresses his hearty thanks to the referee for his helpful comments on the paper.

KYUSHU UNIVERSITY

REFERENCES

[1] Fraser, D. A. S. (1957). *Nonparametric Methods in Statistics*, John Wiley & Sons, Inc.

[2] Isii, K. and Taga, Y. (1969). On optimum stratification for multivariate distributions, *Skand. Aktuartidskr.*, **52**, 24-38.

[3] Wakimoto, K. (1971). Stratified random sampling (I): Estimation of the population variance, *Ann. Inst. Statist. Math.*, **23**, 232-252.

[4] Yanagawa, T. and Wakimoto, K. (1972). Estimation of some functional of the population distribution based on a stratified random sample, *Ann. Inst. Statist. Math.*, **24**, 137-151.

[5] Yanagawa, T. (1972). Stratified random sampling; Testing hypotheses based on $\mathcal{U}$-statistics, *Memoirs of the Fac. of Sci. Kyushu Univ.*, Series A, Math., XXVI, No. 2, 361-379.

[6] Taga, Y., Wakimoto, K. and Yanagawa, T. (1973). Generalized $\mathcal{U}$-statistics for stratified random samples, *Skand. Aktuartidskr.*, 30-51.