YOKOHAMA MATHEMATICAL
JOURNAL VOL. 40, 1993

GENERALIZATION OF HT STRATEGY AND ITS APPLICATION

By

YASUSHI TAGA

(Received August 17, 1992; Revised November 17, 1992)

Summary. The Horvitz-Thompson strategy (HT strategy) can be generalized by redefining the so-called inclusion probabilities so that both strategies of with-replacement sampling and without-replacement sampling are represented in the same forms of their estimators and variance formulas. Using the above result we can find a feasible procedure which improves a given strategy of with-replacement sampling to a better strategy of without-replacement sampling under suitable conditions.

Besides necessary and sufficient conditions for the existing of a sampling design are shown which induces the first-order or second-order inclusion probabilities given in advance in case of without-replacement sampling design.

1. Introduction

Let us consider a finite population U consisting of N units which are identified with the index set $\{1, 2, \cdots, N\}$. For each unit i the real-valued variate y_i is labeled, and each subset $s=(i_1, i_2, \cdots, i_n)$ of the population U is called "sample" and the number of units in s "sample size". The set S of all possible 2^N samples is called "sample space", and any probability distribution $p(s)$ over S is called "sampling design".

Our main concern is to estimate the population total $y=y_1+y_2+\cdots+y_N$ by the linear unbiased estimator $t(s)$ with its variance as small as possible under some suitable sampling design $p(s)$. A pair (p, t) of sampling design $p(s)$ and estimator $t(s)$ is called "strategy".

In case of without-replacement sampling it is well-known that the so-called Horvitz-Thompson estimator (HT estimator) defined by

$$t(s)=\sum_{i\in s}\frac{y_i}{\pi_i}, \tag{1.1}$$

is unbiased for the population total y and its variance $V_p(t)$ is given by

$$V_p(t)=\sum_{i=1}^{N}\pi_i(1-\pi_i)\left(\frac{y_i}{\pi_i}\right)^2+\sum_{i\neq j}^{N}(\pi_{ij}-\pi_i\pi_j)\frac{y_i}{\pi_i}\frac{y_j}{\pi_j}, \tag{1.2}$$

1991 Mathematics Subject Classification: 62D05.
Key words and phrases: varying probability sampling, HT estimator, strategy.

where π_i and π_{ij} denote the first-order and second-order inclusion probabilities respectively, i.e.,

$$\pi_i=\sum_{s\ni i}p(s)\,, \tag{1.3}$$

and

$$\pi_{ij}=\sum_{s\ni i,j}p(s)\,. \tag{1.4}$$

In case of fixed sample size n (FS(n) in short) $V_p(t)$ given by (1.2) can be rewritten into the so-called "Yates-Grundy form" as

$$V_p(t)=\sum_{i<j}^{N}(\pi_i\pi_j-\pi_{ij})\left(\frac{y_i}{\pi_i}-\frac{y_j}{\pi_j}\right)^2. \tag{1.5}$$

It is easily shown that the relations

$$\sum_{i=1}^{N}\pi_i=n\,, \tag{1.6}$$

and

$$\sum_{j=1}^{N}\pi_{ij}=n\pi_i\,,\qquad (1\leqq i\leqq N) \tag{1.7}$$

hold in any FS(n) design, where $\pi_{ii}=\pi_i$. $(1\leqq i\leqq N)$

Those results shown above and some optimalities (admissibilities) or minimaxities of various strategies have been studied by many reserchers as shown in references [1]~[10].

In this paper we generalize the HT estimator (or strategy) by redefining inclusion probabilities so that the generalized HT estimator (GHT estimator) and its variance formula can be represented in the same form in both cases of with-replacement sampling and without-replacement sampling.

Then we show in case of FS(n) design that a given strategy of with-replacement sampling could be improved to some strategy of without-replacement sampling under suitable conditions as shown in *Lemma* 1 and *Theorem.*

In addition we show the necessary and sufficient conditions for existing of a sampling design $p(s)$ which induces the first-order inclusion probabilities π_i's or the second-order inclusion probabilities π_{ij}'s given in advance (see *Lemma* 2 and *Lemma* 3), and of course those conditions must be considered firstly in advance of proving *Theorem.*

2. Generalized HT (GHT) estimator

Let ν_i be the frequency of the ith element of U appeared in a sample s $(1\leqq i\leqq N)$, and $\tilde{\pi}_i$, $\tilde{\pi}_{ij}$ be expectations of ν_i and $\nu_i\nu_j$ respectively, i.e.,

$$\tilde{\pi}_i=E_p\{\nu_i\}\,, \tag{2.1}$$

and

$$\tilde{\pi}_{ij}=E_p\{\nu_i\nu_j\} \tag{2.2}$$

for $1\leqq i,\ j\leqq N$. Note that $\tilde{\pi}_i$ and $\tilde{\pi}_{ij}$ are identical with the first-order and second-order inclusion probabilities π_i and π_{ij} respectively in the case of without-replacement sampling.

Now let us define generalized HT (GHT) estimator $\tilde{t}(s)$ by

$$\tilde{t}(s)=\sum_{i=1}^{N}\nu_i\frac{y_i}{\tilde{\pi}_i}. \tag{2.3}$$

Then it is easily seen that $\tilde{t}(s)$ is a linear and unbiased estimator of the population total y and its variance $V_p(\tilde{t})$ is given by

$$V_p(\tilde{t})=\sum_{i,j=1}^{N}(\tilde{\pi}_{ij}-\tilde{\pi}_i\tilde{\pi}_j)\frac{y_i}{\tilde{\pi}_i}\frac{y_j}{\tilde{\pi}_j}. \tag{2.4}$$

In case of FS(n), $V_p(\tilde{t})$ above can be rewritten into the Yates-Grundy form

$$V_p(\tilde{t})=\sum_{i<j}^{N}(\tilde{\pi}_i\tilde{\pi}_j-\tilde{\pi}_{ij})\left(\frac{y_i}{\tilde{\pi}_i}-\frac{y_j}{\tilde{\pi}_j}\right)^2, \tag{2.5}$$

which is a similar form as in (1.5).

Note that the following relations among $\tilde{\pi}_i$ and $\tilde{\pi}_{ij}$ hold for any FS(n) design,

$$\sum_{i=1}^{N}\tilde{\pi}_i=n, \tag{2.6}$$

and

$$\sum_{j=1}^{N}\tilde{\pi}_{ij}=n\tilde{\pi}_i, \qquad (1\leqq i\leqq N), \tag{2.7}$$

since $\sum_{i=1}^{N}\nu_i=n$ holds in case of FS(n) design.

Note that the GHT estimator reduces to the usual HT estimator in case of without-replacement sampling design since $\tilde{\pi}_i$ and $\tilde{\pi}_{ij}$ are nothing but the first-order and second-order inclusion probabilities.

Example 1. (HH-strategy)

Hansen-Hurwitz strategy (HH strategy) is defined as in the following:

$$t_{HH}(s)=\sum_{i=1}^{N}\nu_i\frac{y_i}{np_i}, \tag{2.8}$$

where $p_i\geqq 0$, $\sum_{i=1}^{N}p_i=1$ and $(\nu_1, \nu_2, \cdots, \nu_N)$ represents a random vector disributed according to the multinomial distrubution MUL $(n\,;\, p_1, p_2, \cdots, p_N)$.

Then it is easily seen that the following relations

$$\tilde{\pi}_i=E_p\{\nu_i\}=np_i, \tag{2.9}$$

$$\tilde{\pi}_{ij}=E_p\{\nu_i\nu_j\}=n(n-1)p_ip_j, \tag{2.10}$$

$$\tilde{\pi}_{ii}=E_p\{\nu_i^2\}=n(n-1)p_i^2+np_i \tag{2.11}$$

hold for $1 \leqq i \leqq N$, $1 \leqq i < j \leqq N$.

3. A procedure for improving a strategy of with-replacement sampling design

Let us consider a with-replacement sampling design $\tilde{p}(s)$ of fixed sample size n from a finite population U of size N, and let $\{\tilde{\pi}_i\}$, $\{\tilde{\pi}_{ij}\}$ be expectations of ν_i and $\nu_i \nu_j$ respectively as defined by (2.1) and (2.2).

Let us put $b_i = \tilde{\pi}_{ii} - \tilde{\pi}_i = E_{\tilde{p}}\{\nu_i(\nu_i - 1)\} > 0$ for $1 \leqq i \leqq N$, and suppose that the following condition (3.1) is satisfied:

$$b_1 + b_2 + \cdots + b_N \geqq 2b_m \qquad (b_m = \max_{1 \leq i \leq N} b_i). \tag{3.1}$$

Then, using *Lemma* 1, we can find a_{ij}'s such that the relations $a_{ij} = a_{ji} \geqq 0$, $a_{ii} = 0$ and $\sum_{j=1}^{N} a_{ij} = b_i$ holds for $1 \leqq i,\ j \leqq N$.

Defining π_i^* and π_{ij}^* such that $\pi_i^* = \pi_{ii}^* = \tilde{\pi}_i$ for $1 \leqq i \leqq N$ and $\pi_{ij}^* = \tilde{\pi}_{ij} + a_{ij}$ for $1 \leqq i < j \leqq N$, then π_i^* and π_{ij}^* satisfy the following conditions:

$$\pi_i^* = \tilde{\pi}_i, \qquad \pi_{ij}^* \geqq \tilde{\pi}_{ij} \qquad \text{for } 1 \leqq i < j \leqq N, \tag{3.2}$$

and

$$\pi_{kl}^* > \tilde{\pi}_{kl}, \qquad \text{for some } (k, l)\ 1 \leqq k < l \leqq N. \tag{3.3}$$

If there exists a design $p^*(s)$ of without-replacement sampling having π_i^* and π_{ij}^* as the first-order and second-order inclusion probabilities respectively, it is easily seen that HT estimator $t^* = \sum_{i \in s} y_i / \pi_i^*$ has the smaller variance $V_{p^*}(t^*)$ than the variance $V_{\tilde{p}}(\tilde{t})$ of estimator $\tilde{t}$ given in (2.4). Because $V_{p^*}(t^*)$ and $V_{\tilde{p}}(\tilde{t})$ are expressed in the Yates-Grundy forms as (1.5) and (2.5), i.e.,

$$V_{p^*}(t) = \sum_{i<j}^{N} (\pi_i^* \pi_j^* - \pi_{ij}^*) \left(\frac{y_i}{\pi_i^*} - \frac{y_j}{\pi_j^*} \right)^2, \tag{3.4}$$

and

$$V_{\tilde{p}}(\tilde{t}) = \sum_{i<j}^{N} (\tilde{\pi}_i \tilde{\pi}_j - \tilde{\pi}_{ij}) \left(\frac{y_i}{\tilde{\pi}_i} - \frac{y_j}{\tilde{\pi}_j} \right)^2 \tag{3.5}$$

respectively where $\pi_i^* = \tilde{\pi}_i$, $\pi_{ij}^* \geqq \tilde{\pi}_{ij}$ for $1 \leqq i < j \leqq N$ and $\pi_{kl}^* > \tilde{\pi}_{kl}$ for some (k, l) from (3.2) and (3.3).

Now the conditions for existing of $\{\pi_i^*\}$, $\{\pi_{ij}^*\}$ and $p^*(s)$ stated above are examined in the following way.

First let us consider the existence of design $p^*(s)$ which induces $\{\pi_i^*\}$ or $\{\pi_{ij}^*\}$ as the first-order or second-order inclusion probabilities respectively.

Let $\boldsymbol{p} = (p(s_1), p(s_2), \cdots, p(s_M))^t$ be a vector of probability distribution of design $p(s)$, $\boldsymbol{\pi}_1^* = (\pi_1^*, \pi_2^*, \cdots, \pi_N^*)^t$ and $C = [c_{ij}]$ be $N \times M$ matrix with components c_{ij} defined by

(3.6) $$c_{ik}=\begin{cases} 1 & \text{if } s_k\ni i, \\ 0 & \text{otherwise,} \end{cases} \quad \left(\begin{matrix} 1\leqq i\leqq N \\ 1\leqq k\leqq M=\binom{N}{n} \end{matrix}\right).$$

Then the following relation between $\boldsymbol{p}$ and $\boldsymbol{\pi}_1^*$ must hold

(3.7) $$C\mathrm{p}=\boldsymbol{\pi}_1^*,$$

where $\boldsymbol{\pi}_1^*$ is the given vector above and $\mathrm{rank}(C)=N$ for $n<N$ as shown in *Lemma* 2 below.

Then we can see the condition for the existing of design $p^*(s)$ with the first-order inclusion probabilities $\{\pi_i^*\}$ is that the equation (3.7) has a non-negative solution $\boldsymbol{p}=\boldsymbol{p}^*$, i.e., $p^*(s)\geqq 0$ for any s.

Note that a necessary and sufficient conditions for existing of non-negative solusion $\boldsymbol{p}=\boldsymbol{p}^*$ can be explicitly obtained only after solving equation (3.7).

Similary we can show that a necessary and sufficient condition for existing of $p^*(s)$ with the second-order inclusion probabilities $\{\pi_{ij}^*\}$ can be obtained if the following equation (3.8) has non-negative solution $\boldsymbol{p}^*=(p^*(s_1), \cdots, p^*(s_M))^t$, i.e., $p^*(s_i)\geqq 0$ for $1\leqq k\leqq M=\binom{N}{n}$:

(3.8) $$D\boldsymbol{p}=\boldsymbol{\pi}_2^*,$$

where $\boldsymbol{p}=(p(s_1), \cdots, p(s_M))^t$, $\boldsymbol{\pi}_2^*=(\pi_{12}^*, \pi_{13}^*, \cdots, \pi_{N-1,N}^*)^t$ and $D=[d_{(ij)k}]$ is $\binom{N}{2}\times\binom{N}{n}$ matrix where $d_{(ij)k}$ defined by

(3.9) $$d_{(ij)k}=\begin{cases} 1 & \text{if } s_k\ni i,\ j\ (i\text{th and } j\text{th elements of } U) \\ 0 & \text{otherwise,} \end{cases}$$

for $1\leqq k\leqq\binom{N}{n}$ and $1\leqq i\leqq j\leqq N$.

It is shown in *Lemma* 3 below that $\mathrm{rank}(D)=\min\left\{\binom{N}{2}, \binom{N}{n}\right\}$. Note that the equation (3.8) reduces to (3.7) by multiplying a suitable matrix H to the both sides of (3.8) from the left, if the following conditions (3.10) are satisfied.

(3.10) $$\sum_{j=1}^{N}\pi_{ij}^*=n\pi_i^*, \qquad \pi_{ii}^*=\pi_i^* \qquad \text{for } 1\leqq i\leqq j\leqq N.$$

Therefore if the equation (3.7) has a unique solution $\boldsymbol{p}=\boldsymbol{p}^*$, then the equation (3.8) has the same solution $\boldsymbol{p}=\boldsymbol{p}^*$. In conclusion we state the following theorem.

Theorem. *Let $\{\tilde{\pi}_i\}$, $\{\tilde{\pi}_{ij}\}$ be the first-order, second-order inclusion probabilities respectively induced from a given with-replacement sampling design $\tilde{p}(s)$, and suppose that $0<\tilde{\pi}_i<1$ without loss of generality.*

If $b_i=\tilde{\pi}_{ii}-\tilde{\pi}_i$ $(1\leqq i\leqq N)$ *satisfy the condition* (3.1), *then we can find* $\{\pi_i^*\}$ *and* $\{\pi_{ij}^*\}$ *satisfying the conditions* (3.2) *and* (3.3).

Further if the equation (3.8) *has a non-negative solution* $\boldsymbol{p}=\boldsymbol{p}^*$ *for* $\boldsymbol{\pi}_2^*=(\pi_{12}^*, \pi_{13}^*, \cdots, \pi_{N-1,N}^*)^t$, *then we can take* $p^*(s)$ *for a design of withont-replacement sampling which induces* π_i^* *and* π_{ij}^* *as the first-order and second-order inclusion probabilities respectively.*

Thus if the conditions above are all satisfied, then a given strategy of with-replacement sampling $(\tilde{p}, \hat{t})$ *can be improved to a strategy of without-replacement sampling* (p^*, t^*), *i.e.,*

$$V_{p^*}(t^*)<V_{\tilde{p}}(\hat{t}).$$

Remark. This theorem shows that we can find a feasible procedure to improve a given strategy of with-replacement sampling to a better strategy of without-replacement sampling under suitable conditions as above though it may not be best or admissible.

4. Lemmas

Lemma 1. *Let* $b_1, b_2, \cdots, b_N$ *be given non-negative numbers satisfying the condition* $0\leqq b_1\leqq b_2\leqq \cdots \leqq b_N$.

Then the necessary and sufficient condition for existing of $N\times N$ *matrix* $A=[a_{ij}]$ *such that*

$$a_{ij}=a_{ji}\geqq 0,\quad a_{ii}=0,\qquad \text{for } 1\leqq i\leqq j\leqq N, \tag{4.1}$$

and

$$\sum_{j=1}^{N} a_{ij}=b_i \qquad \text{for } 1\leqq i\leqq N, \tag{4.2}$$

is

$$\sum_{i=1}^{N-1} b_i\geqq b_N\,. \tag{4.3}$$

Proof. It is clear that the condition (4.3) is necessary.

To prove the sufficiency of (4.3) we shall construct an $N\times N$ matrix A satisfying the conditions (4.1) and (4.2) in the following way.

Let us define $N\times N$ matrices E_i, F_i such that

$$E_1=\frac{1}{N-1}\begin{bmatrix} 0 & 1 & \cdots & \cdots & 1 \\ 1 & 0 & 1 & \cdots & 1 \\ \vdots & \ddots & \ddots & \ddots & \cdots \\ \cdots & & \ddots & \ddots & 1 \\ 1 & \cdots & \cdots & 1 & 0 \end{bmatrix},$$

$$E_2=\frac{1}{N-2}\begin{bmatrix} 0 & 0 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & 1 & \cdots & \cdots & 1 \\ \vdots & 1 & 0 & \ddots & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & & \ddots & \ddots & 1 \\ 0 & 1 & \cdots & \cdots & 1 & 0 \end{bmatrix}, \cdots$$

$$E_{N-1}=\begin{bmatrix} 0 & 0 & 0 & \cdots & \cdots & 0 \\ 0 & 0 & 0 & \cdots & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & 0 & 0 \\ \vdots & & & 0 & 0 & 1 \\ 0 & 0 & \cdots & 0 & 1 & 0 \end{bmatrix}$$

and

$$F_1=\begin{bmatrix} 0 & 0 & \cdots & \cdots & 0 & 1 \\ 0 & 0 & \cdots & \cdots & 0 & 1 \\ \vdots & \vdots & \ddots & & \vdots & \vdots \\ \vdots & \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \cdots & 0 & 1 \\ 1 & 1 & \cdots & \cdots & 1 & 0 \end{bmatrix}, \quad F_2=\begin{bmatrix} 0 & 0 & \cdots & \cdots & 0 & 0 \\ 0 & 0 & \cdots & \cdots & 0 & 1 \\ \vdots & \vdots & \ddots & & \vdots & \vdots \\ \vdots & \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \cdots & 0 & 1 \\ 0 & 1 & \cdots & \cdots & 1 & 0 \end{bmatrix}, \cdots$$

$$F_{N-1}=\begin{bmatrix} 0 & 0 & \cdots & \cdots & 0 & 0 \\ 0 & 0 & \cdots & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & & & \cdots \\ \vdots & \vdots & & \ddots & & 0 \\ 0 & 0 & \cdots & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 0 & 1 & 0 \end{bmatrix}.$$

Further define $N\times N$ matrices E and F such that

(4.4) $$E=b_1E_1+(b_2-b_1)E_2+\cdots+(b_{N-1}-b_{N-2})E_{N-1},$$

and

(4.5) $$F=b_1F_1+(b_2-b_1)F_2+\cdots+(b_{N-1}-b_{N-2})F_{N-1}.$$

Then it is easily seen that E and F are symmetric $N\times N$ matrices satisfying the conditions

$$E\cdot(1,\cdots,1)^t=(b_1,b_2,\cdots,b_{N-1},b_{N-1})^t$$

and

$$F\cdot(1,\cdots,1)^t=(b_1,b_2,\cdots,b_{N-1},\sum_{i=1}^{N-1}b_i)^t,$$

Since $b_{N-1}\leqq b_N\leqq\sum_{i=1}^{N-1}b_i$ by (4.3), there exists a non-negative number λ $(0\leqq\lambda\leqq1)$ satistying the relation

(4.6) $$\lambda b_{N-1}+(1-\lambda)\sum_{i=1}^{N-1} b_i=b_N .$$

Then the $N\times N$ matrix $A=[a_{ij}]$ satisfying (4.1) and (4.2) can be obtained by putting $A=\lambda E+(1-\lambda)F$. Note that λ is determined from (4.6) such that

(4.7) $$b=\left(\sum_{i=1}^{N-1} b_i-b_N\right)\Big/\sum_{i=1}^{N-2} b_i .$$

Lemma 2. *Let C be $N\times M$ matrix, the (i, j) element c_{ij} of which is defined by* (3.6). *Then for any n and N $(n<N)$, the* rank *of C is equal to N, i.e.,*

(4.8) $$\text{rank}\,(C)=N, \quad \textit{where } M=\binom{N}{n}.$$

Proof. For any fixed n and $N=n+1$, it is clear that $M=n+1$ and $(n+1)\times(n+1)$ matrix represented by

$$C=\begin{bmatrix} 1 & 1 & \cdots & 1 & 0 \\ \vdots & \vdots & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & 1 \\ \vdots & 1 & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & \vdots \\ 1 & 0 & \cdot^{\cdot^{\cdot}} & & 1 \\ 0 & 1 & \cdots & \cdots & 1 \end{bmatrix},$$

and that $\text{rank}(C)=n+1=N$.

Further let us show (4.8) holds for n and $N+1$ under the assumption that (4.8) holds for any n and N.

For convenience sake let us put suffix N to C such as C_N in case of N. Then it is easily seen that $C_{N+1}=\begin{bmatrix} C_N & G \\ \mathbf{0}^t & \mathbf{1}^t \end{bmatrix}$, where $\mathbf{0}=(0, 0, \cdots, 0)^t$ and $\mathbf{1}=(1, 1, \cdots, 1)^t$ are $\binom{N}{n}$ and $\binom{N}{n-1}$ dimensional vectors respectively and G is an $N\times\binom{N}{n-1}$ matrix, and that $\text{rank}\,(C_{N+1})=\text{rank}\,(C_N)+1=N+1$. Therefore (4.8) holds for any n and $N(n<N)$ by induction.

Lemma 3. *Let D be a matrix $\binom{N}{2}\times\binom{N}{n}$, the (i, j) element of which is defined by* (3.9). *Then it holds that*

(4.9) $$\text{rank}\,(D)=\min\left\{\binom{N}{2}, \binom{N}{n}\right\}$$

for any n and N $(2\leqq n\leqq N)$.

Proof. In case $N=3$ and $n=2$ the matrix D is represented as 3×3 unit matrix and so (4.9) holds.

Let us assume (4.9) holds for any n and N in case $\binom{N}{2}\leqq\binom{N}{n}$, and show that (4.9) holds for n and $N+1$. For convenience sake let us put suffix N and n to D as D_n^N in case of n and N. Then it is easily seen that D_n^{N+1} can be represented as

$$D_n^{N+1}=\begin{bmatrix} D_n^N & R \\ O & C_N \end{bmatrix},$$

where O is $N\times\binom{N}{n}$ zero matrix, R some suitable $\binom{N}{2}\times\binom{N}{n-1}$ matrix and C_N is $N\times\binom{N}{n-1}$ matrix as shown in *Lemma* 2 of rank N. Then it is easily seen that rank $(D_n^{N+1})=\binom{N}{2}+N=\binom{N+1}{2}$.

In case $\binom{N}{2}>\binom{N}{n}$ for $n=N-1$, it is easily proved that rank $(D)=\binom{N}{N-1}=N$.

5. Examples

Example 1. In case of $N=4$ and $n=3$, 4×4 matrix C in (3.7) is given by

$$C=\begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}. \tag{5.1}$$

Then the equation (3.7) is represented by

$$\begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{bmatrix}=\begin{bmatrix} \pi_1^* \\ \pi_2^* \\ \pi_3^* \\ \pi_4^* \end{bmatrix}, \tag{5.2}$$

which has the unique solution

$$\begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{bmatrix}=\frac{1}{3}\begin{bmatrix} \pi_1^*+\pi_2^*+\pi_3^*-2\pi_4^* \\ \pi_1^*+\pi_2^*+\pi_4^*-2\pi_3^* \\ \pi_1^*+\pi_3^*+\pi_4^*-2\pi_2^* \\ \pi_2^*+\pi_3^*+\pi_4^*-2\pi_1^* \end{bmatrix}. \tag{5.3}$$

Therefore the necessary and sufficient conditions for existing of non-negative solution $\boldsymbol{p}=\boldsymbol{p}^*$ is given by

$$\pi_1^*+\pi_2^*+\pi_3^*\geqq 2\pi_4^*, \quad \text{or} \quad \pi_4^*\leqq 1, \tag{5.4}$$

assuming that $0<\pi_1^*\leqq\pi_2^*\leqq\pi_3^*\leqq\pi_4^*$.

Example 2. In case of $N=4$ and $n=2$, let us improve an HH strategy to get a better HT strategy. Consider an HH strategy such as $p_1=0.1$, $p_2=0.25$,

$p_3=0.3$, $p_4=0.35$, and then $\tilde{\pi}_1=0.2$, $\tilde{\pi}_2=0.5$, $\tilde{\pi}_3=0.6$, $\tilde{\pi}_4=0.7$. By the definition of $b_i=E_{\tilde{p}}\{\nu_i(\nu_i-1)\}=n(n-1)p_i^2$, we get $b_1=0.02$, $b_2=0.125$, $b_3=0.18$, $b_4=0.245$ and so $b_1+b_2+b_3=0.325>b_4$.

Therefore we can get a matrix $A=\lambda E+(1-\lambda)F$ as shown in the proof of *Lemma* 2 such that

$$E=\begin{bmatrix} 0 & 0.00667 & 0.00667 & 0.00667 \\ 0.00667 & 0 & 0.05917 & 0.05917 \\ 0.00667 & 0.05917 & 0 & 0.11417 \\ 0.00667 & 0.05917 & 0.11417 & 0 \end{bmatrix},$$

$$F=\begin{bmatrix} 0 & 0 & 0 & 0.02 \\ 0 & 0 & 0 & 0.125 \\ 0 & 0 & 0 & 0.180 \\ 0.02 & 0.125 & 0.180 & 0 \end{bmatrix},$$

$$\lambda=0.5517, \qquad 1-\lambda=0.4483,$$

$$A=\begin{bmatrix} 0 & 0.00368 & 0.00368 & 0.01264 \\ 0.00368 & 0 & 0.03264 & 0.08868 \\ 0.00368 & 0.03264 & 0 & 0.14368 \\ 0.01264 & 0.08868 & 0.14368 & 0 \end{bmatrix},$$

$$[\pi_{ij}^*]=\begin{bmatrix} 0.22000 & 0.05368 & 0.06368 & 0.08264 \\ 0.05368 & 0.62500 & 0.18264 & 0.26368 \\ 0.06368 & 0.18264 & 0.78000 & 0.35368 \\ 0.08264 & 0.26368 & 0.35368 & 0.94500 \end{bmatrix}.$$

Since D is the unit matrix in this case, we get $\boldsymbol{p}^*=\boldsymbol{\pi}_2^*$. Further we can get an HT strategy (p^*, t^*) better than the original HH strategy $(\tilde{p}, \tilde{t})$ given above. In case $y_1=1$, $y_2=2$, $y_3=3$, and $y_4=4$, we get $V_{p^*}(t^*)=0.4803<V_{\tilde{p}}(\tilde{t})=0.5714$.

Acknowledgement

The author expresses his hearty thanks to Professor T. Mori for his helpful comments in proving *Lemma* 1 and *Lemma* 2. He also would like to thank the referee for his helpful comments.

References

[1] Cassel, C.-M. et al., *Foundations of Inference in Survey Sampling*, A Wiley-Interscience Publication, 1977.

[2] Gabler, S., *Minimax Solutions in Sampling from Finite Populations*, Springer, 1990.

[3] Hájek, J., *Sampling From Finite Population*, Dekker, 1981.

[4] Hansen, M.H., Hurwitz, W.N. and W.G. Madow, *Sample Survey Methods and Theory*, I *and* II, New York, Wiley, 1953.

[5] Hedayat, A.S. and B.K. Sinha, *Design and Inference in Finite Population Sampling*, A Wiley-Interscience Publication, 1991.
[6] Horvitz, D.G. and D.J. Thompson, A generalization of sampling without replacement from a finite universe, *J. Amer. Statist. Assoc.* 47 (1952), 663-685.
[7] Kish, L., *Survey Sampling*, New York, Wiley, 1965.
[8] Midzuno, H., An outline of the theory of sampling system, *Ann. Inst. Statist. Math.*, 1 (1950), 149-156.
[9] Särndal, C.-E., Swensson, B. and J. Wretman, *Model Assisted Survey Sampling*, Springer-Verlag, 1992.
[10] Taga, Y., *Theory of Sample Survey* (in Japanese), Saiensu-sha, 1975.

Faculty of International Studies
Bunkyo University
1100 Namegaya
Chigasaki, 253 Japan