Ann. Inst. Statist. Math.
Vol. 50, No. 1, 49–70 (1998)

DEPENDENCE BETWEEN ORDER STATISTICS IN SAMPLES FROM FINITE POPULATION AND ITS APPLICATION TO RANKED SET SAMPLING*

KOITI TAKAHASI[1] AND MASAO FUTATSUYA[2]

[1]*Department of Basic Science, School of Science and Engineering, Ishinomaki Senshu University, Ishinomaki 986-80, Japan*
[2]*Department of Mathematical System Science, Faculty of Science and Technology, Hirosaki University, Hirosaki 036, Japan*

(Received June 4, 1996; revised February 10, 1997)

Abstract. Let $X_1, X_2, \dots, X_m$, $Y_1, Y_2, \dots, Y_n$ be a simple random sample without replacement from a finite population and let $X_{(1)} \le X_{(2)} \le \cdots \le X_{(m)}$ and $Y_{(1)} \le Y_{(2)} \le \cdots \le Y_{(n)}$ be the order statistics of $X_1, X_2, \dots, X_m$ and $Y_1, Y_2, \dots, Y_n$, respectively. It is shown that the joint distribution of $X_{(i)}$ and $X_{(j)}$ is positively likelihood ratio dependent and $Y_{(j)}$ is negatively regression dependent on $X_{(i)}$. Using these results, it is shown that when samples are drawn without replacement from a finite population, the relative precision of the ranked set sampling estimator of the population mean, relative to the simple random sample estimator with the same number of units quantified, is bounded below by 1.

Key words and phrases: Ranked set sampling, finite population, order statistics, dependence.

1. Introduction

Let $X_1, X_2, \dots, X_m$ be independently distributed according to a univariate distribution, and let $X_{(1)} \le X_{(2)} \le \cdots \le X_{(m)}$ be their order statistics. In this case, Lehmann (1966) has shown that the joint distribution of two order statistics, $X_{(i)}$ and $X_{(j)}$, is positively likelihood ratio dependent. We consider the case where $X_1, X_2, \dots, X_m$ is a simple random sample without replacement from a finite population and, therefore, $X_1, X_2, \dots, X_m$ are not independent. It is shown that the joint distribution of $X_{(i)}$ and $X_{(j)}$ is positively likelihood ratio dependent also in this case. We prove this result in Section 2.

Next, let $X_1, X_2, \dots, X_m$, $Y_1, Y_2, \dots, Y_n$ be a simple random sample of size $m + n$ without replacement from a finite population, and let $X_{(1)} \le X_{(2)} \le$

* This research was partially supported by the ISM Cooperative Research Program(92-ISM-CRP-19).

$\cdots \leq X_{(m)}$ and $Y_{(1)} \leq Y_{(2)} \leq \cdots \leq Y_{(n)}$ be two sets of the order statistics of $X_1, X_2, \ldots, X_m$ and $Y_1, Y_2, \ldots, Y_n$, respectively. We consider possible dependence between $X_{(i)}$ and $Y_{(j)}$. It is shown that $Y_{(j)}$ is negatively regression dependent on $X_{(i)}$. We prove this result in Section 3. In Section 4, we show that $\mathrm{Cov}(X_{(i)}, X_{(j)}) \geq 0$ and $\mathrm{Cov}(X_{(i)}, Y_{(j)}) \leq 0$, and we give the conditions for the equality to hold.

Finally, using these results, we shall prove a theorem on ranked set sampling (RSS) in finite populations. As it was pointed out by Patil *et al.* (1995), most of the researche in RSS has been concerned with sampling from infinite (continuous) populations. Takahasi and Futatsuya (1988), Futatsuya and Takahasi (1990), and Patil *et al.* (1995) studied RSS for estimating a population mean when sampling was done without replacement from a finite population. Takahasi and Futatsuya (1988) gave an expression for the variance of the RSS estimator ($\hat{\mu}_{RSS}$), and Patil *et al.* (1995) obtained explicit expressions for the variance of $\hat{\mu}_{RSS}$ and the corresponding relative savings. Performance of the RSS estimator is generally benchmarked against that of the simple random sampling (SRS) estimator ($\hat{\mu}_{SRS}$) with the same number of quantifications. For this purpose, we use the relative precision (RP), $RP = \mathrm{Var}(\hat{\mu}_{SRS})/\mathrm{Var}(\hat{\mu}_{RSS})$. Futatsuya and Takahasi (1990) considered the extremal finite populations maximizing relative precision. In Section 5, we show that RP is never smaller than 1, and is greater than 1 unless $N-1$ elements of the population of size N have the same value.

2. Positive likelihood ratio dependence between order statistics of a sample

Let $\boldsymbol{\Omega}$ be the finite population $\{x_1, \ldots, x_1, x_2, \ldots, x_2, \ldots, x_l, \ldots, x_l\}$ ($x_1 < x_2 < \cdots < x_l$) of size N. Let ν_a be the number of x_a in $\boldsymbol{\Omega}$ ($a = 1, 2, \ldots, l$), $f_a = \nu_1 + \nu_2 + \cdots + \nu_a$ ($a = 1, 2, \ldots, l$), $f_0 = 0$ and $\tilde{f}_a = f_{a-1} + 1$ ($a = 1, 2, \ldots, l$). We assume that $\nu_a > 0$ ($a = 1, 2, \ldots, l$). Let $X_1, X_2, \ldots, X_n$ be a simple random sample of size n without replacement from $\boldsymbol{\Omega}$. Let $X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(n)}$ be the order statistics of this sample. In this section, we prove that the joint distribution $X_{(i)}$ and $X_{(j)}$ ($i \neq j$) is positively likelihood ratio dependent.

Let us first consider the case $\boldsymbol{\Omega} = \boldsymbol{\Omega}_N = \{1, 2, \ldots, N\}$. Let $Z_1, Z_2, \ldots, Z_n$ be a simple random sample of size n without replacement from $\boldsymbol{\Omega}_N$. Let $Z_{(1)} < Z_{(2)} < \cdots < Z_{(n)}$ be the order statistics of this sample. The set of all points (s, t) satisfying $\Pr\{Z_{(i)} = s, Z_{(j)} = t\} > 0$ is denoted by $\boldsymbol{S}$. For $1 \leq i < j \leq n$, we have

$$\boldsymbol{S} = \{(s, t) \mid i \leq s \leq N - n + i, j \leq t \leq N - n + j, j - i \leq t - s\}. \tag{2.1}$$

We prove the following:

LEMMA 2.1. *Let $1 \leq i < j \leq n$. Then the joint distribution of $Z_{(i)}$ and $Z_{(j)}$ is positively likelihood ratio dependent; that is, if $s < s'$ and $t < t'$, then*

$$\begin{aligned} &\Pr\{Z_{(i)} = s, Z_{(j)} = t\} \Pr\{Z_{(i)} = s', Z_{(j)} = t'\} \\ &\quad \geq \Pr\{Z_{(i)} = s, Z_{(j)} = t'\} \Pr\{Z_{(i)} = s', Z_{(j)} = t\}, \end{aligned} \tag{2.2}$$

with equality holding if and only if $(s, t) \notin \boldsymbol{S}$ or $(s', t') \notin \boldsymbol{S}$.

PROOF. From (2.1), if $(s,t) \notin \boldsymbol{S}$ or $(s',t') \notin \boldsymbol{S}$, then $(s,t') \notin \boldsymbol{S}$ or $(s',t) \notin \boldsymbol{S}$ and, therefore, equality holds in (2.2). If $(s,t) \in \boldsymbol{S}$ and $(s',t') \in \boldsymbol{S}$, and $(s,t') \notin \boldsymbol{S}$ or $(s',t) \notin \boldsymbol{S}$, then

$$\begin{aligned}&\Pr\{Z_{(i)} = s, Z_{(j)} = t\} \Pr\{Z_{(i)} = s', Z_{(j)} = t'\} \\ &\qquad > 0 = \Pr\{Z_{(i)} = s, Z_{(j)} = t'\} \Pr\{Z_{(i)} = s', Z_{(j)} = t\}.\end{aligned}$$

Finally, if $(s,t), (s',t'), (s,t'), (s',t) \in \boldsymbol{S}$, then, since

$$\Pr\{Z_{(i)} = s, Z_{(j)} = t\} = \frac{1}{\binom{N}{n}} \binom{s-1}{i-1} \binom{t-1-s}{j-1-i} \binom{N-t}{n-j},$$

we have

$$\begin{aligned}\frac{\Pr\{Z_{(i)} = s, Z_{(j)} = t\} \Pr\{Z_{(i)} = s', Z_{(j)} = t'\}}{\Pr\{Z_{(i)} = s, Z_{(j)} = t'\} \Pr\{Z_{(i)} = s', Z_{(j)} = t\}} &= \frac{(t-s-1)(t-s-2)\cdots(t-s-j+i+1)}{(t'-s-1)(t'-s-2)\cdots(t'-s-j+i+1)} \\ &\quad \times \frac{(t'-s'-1)(t'-s'-2)\cdots(t'-s'-j+i+1)}{(t-s'-1)(t-s'-2)\cdots(t-s'-j+i+1)}.\end{aligned} \tag{2.3}$$

Because $(t-s-k)(t'-s'-k) - (t'-s-k)(t-s'-k) = (t'-t)(s'-s) > 0$, we get $(2.3) > 1$. This completes the proof of the lemma.

Now we consider the general case $\boldsymbol{\Omega} = \{x_1, \ldots, x_1, x_2, \ldots, x_2, \ldots, x_l, \ldots, x_l\}$. Let

$$\boldsymbol{G}_{ab} = \{(s,t) \mid \tilde{f}_a \le s \le f_a, \tilde{f}_b \le t \le f_b, s, t \in \boldsymbol{\Omega}_N\}.$$

We have the following lemma.

LEMMA 2.2. *For $1 \le i, j \le n$, $i \ne j$ and $1 \le a, b \le l$,*

$$\Pr\{X_{(i)} \le x_a, X_{(j)} \le x_b\} = \Pr\{Z_{(i)} \le f_a, Z_{(j)} \le f_b\}$$

and

$$\Pr\{X_{(i)} = x_a, X_{(j)} = x_b\} = \sum_{(s,t) \in \boldsymbol{G}_{ab}} \Pr\{Z_{(i)} = s, Z_{(j)} = t\}.$$

PROOF. The results come from the relations $X_{(i)} \le x_a \Leftrightarrow Z_{(i)} \le f_a$, $X_{(i)} = x_a \Leftrightarrow \tilde{f}_a \le Z_{(i)} \le f_a$, and the definition of $\boldsymbol{G}_{ab}$.

The following theorem shows positively likelihood ratio dependence between two order statistics in general finite population.

THEOREM 2.1. *Let $1 \le i < j \le n$. The joint distribution of $X_{(i)}$ and $X_{(j)}$ is positively likelihood ratio dependent; that is, if $1 \le a < a' \le l$ and $1 \le b < b' \le l$, then*

$$\begin{aligned}(2.4)\qquad &\Pr\{X_{(i)} = x_a, X_{(j)} = x_b\}\Pr\{X_{(i)} = x_{a'}, X_{(j)} = x_{b'}\}\\ &\ge \Pr\{X_{(i)} = x_a, X_{(j)} = x_{b'}\}\Pr\{X_{(i)} = x_{a'}, X_{(j)} = x_b\}\end{aligned}$$

with equality holding if and only if $\boldsymbol{G}_{ab} \cap \boldsymbol{S} = \emptyset$ or $\boldsymbol{G}_{a'b'} \cap \boldsymbol{S} = \emptyset$.

PROOF. Let $(s,t) \in \boldsymbol{G}_{ab}$ and $(s',t') \in \boldsymbol{G}_{a'b'}$. Then, $s < s'$ and $t < t'$. Therefore, from Lemma 2.1, we have

$$\begin{aligned}&\Pr\{Z_{(i)} = s, Z_{(j)} = t\}\Pr\{Z_{(i)} = s', Z_{(j)} = t'\}\\ &\ge \Pr\{Z_{(i)} = s, Z_{(j)} = t'\}\Pr\{Z_{(i)} = s', Z_{(j)} = t\}.\end{aligned}$$

Summing both sides over $(s,t) \in \boldsymbol{G}_{ab}$ and $(s',t') \in \boldsymbol{G}_{a'b'}$, and applying Lemma 2.2, we get (2.4). The condition for equality in (2.4) also comes from Lemma 2.1. This completes the proof of the theorem.

3. Negative regression dependence between order statistics of two samples

Let $\mathcal{M} = \{X_1, X_2, \dots, X_m\}$ be a simple random sample of size m without replacement from $\boldsymbol{\Omega} = \{x_1, \dots, x_1, x_2, \dots, x_2, \dots, x_l, \dots, x_l\}$ given in Section 2. Put $\boldsymbol{\Phi} = \boldsymbol{\Omega} - \mathcal{M}$. Let $\mathcal{N} = \{Y_1, Y_2, \dots, Y_n\}$ be a simple random sample of size n without replacement from $\boldsymbol{\Phi}$, and let $X_{(1)} \le X_{(2)} \le \cdots \le X_{(m)}$ and $Y_{(1)} \le Y_{(2)} \le \cdots \le Y_{(n)}$ be the order statistics of these samples, respectively. In this section, we prove that $Y_{(j)}$ is negatively regression dependent on $X_{(i)}$.

Let us first consider the case $\boldsymbol{\Omega} = \boldsymbol{\Omega}_N = \{1, 2, \dots, N\}$. In this case, let us denote $\mathcal{M} = \{X_1, X_2, \dots, X_m\}$ and $\mathcal{N} = \{Y_1, Y_2, \dots, Y_n\}$ by $\mathcal{M} = \{U_1, U_2, \dots, U_m\}$ and $\mathcal{N} = \{V_1, V_2, \dots, V_n\}$, respectively. Let $U_{(1)} < U_{(2)} < \cdots < U_{(m)}$ and $V_{(1)} < V_{(2)} < \cdots < V_{(n)}$ be the order statistics of $\{U_1, U_2, \dots, U_m\}$ and $\{V_1, V_2, \dots, V_n\}$, respectively. Let us assume that $1 \le i \le m$, $1 \le j \le n$, with $N \ge m + n$. Put $\tilde{N} = N - m - n + i + j - 1$. We define several subsets of $\boldsymbol{\Omega}_N \times \boldsymbol{\Omega}_N$ as follows:

$$\begin{aligned}\boldsymbol{O} &= \{(u,v) \mid \Pr\{U_{(i)} = u\} > 0 \text{ and } \Pr\{V_{(j)} = v\} > 0\},\\ \boldsymbol{A} &= \{(u,v) \mid \Pr\{U_{(i)} = u, V_{(j)} = v\} > 0 \text{ and } u > v\},\\ \boldsymbol{C} &= \{(u,v) \mid \Pr\{U_{(i)} = u, V_{(j)} = v\} > 0 \text{ and } u < v\},\\ \boldsymbol{B} &= \{(u,v) \mid i + j \le u \le \tilde{N}, u = v\},\\ \boldsymbol{D} &= \{(u,v) \mid i \le u \le i + j - 1, j \le v \le i + j - 1\}\end{aligned}$$

and

$$\boldsymbol{E} = \{(u,v) \mid \tilde{N} + 1 \le u \le N - m + i, \tilde{N} + 1 \le v \le N - n + j\}.$$

Then we have the following lemma.

LEMMA 3.1. *For the sets defined above, we have*

$$\boldsymbol{O} = \{(u,v) \mid i \le u \le N-m+i, j \le v \le N-n+j\},$$

(3.1) $$\boldsymbol{A} = \{(u,v) \mid i+j \le u \le N-m+i, j \le v \le \tilde{N}, u > v\},$$

(3.2) $$\boldsymbol{C} = \{(u,v) \mid i \le u \le \tilde{N}, i+j \le v \le N-n+j, u < v\},$$

(3.3) $$\{(u,v) \mid \Pr\{U_{(i)} = u, V_{(j)} = v\} > 0\} = \boldsymbol{A} \cup \boldsymbol{C}$$

and

$$\boldsymbol{O} = \boldsymbol{A} \cup \boldsymbol{B} \cup \boldsymbol{C} \cup \boldsymbol{D} \cup \boldsymbol{E},$$

where $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{D}$, $\boldsymbol{E}$ *are mutually disjoint.*

PROOF. It is sufficient to prove (3.2), because the proof of (3.1) is similar to the proof of (3.2), and others are obvious. Let us consider Table 1. Under the assumption $u < v$, $\Pr\{U_{(i)} = u, V_{(j)} = v\} > 0$ if and only if there are non-negative integers x and y such that all the entries of the table are non-negative. Using this fact, it is easy to obtain (3.2).

Table 1. Conditions for $\boldsymbol{C}$.

	$1,\dots,u-1$	u	$u+1,\dots,v-1$	v	$v+1,\dots,N$	Total
$\mathcal{M}$	$i-1$	1	y	0	$m-i-y$	m
$\mathcal{N}$	x	0	$j-1-x$	1	$n-j$	n
$\boldsymbol{\Phi}-\mathcal{N}$	$u-i-x$	0	$v-u-j+x-y$	0	$N-m-n+i+j-v+y$	$N-m-n$
Total	$u-1$	1	$v-1-u$	1	$N-v$	N

Figure 1 gives an example of the partition of $\boldsymbol{O}$ into $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{D}$, $\boldsymbol{E}$, where the points of $\boldsymbol{A}$ are the lattice points in area $\boldsymbol{A}$ in Fig. 1 and so on.

The following lemma is used in proving Theorem 3.2.

LEMMA 3.2. *Let* $(u,v) \in \boldsymbol{O}$. *Then,*

$$\Pr\{U_{(i)} = u, V_{(j)} = v\} = \Pr\{U_{(m-i+1)} = N-u+1, V_{(n-j+1)} = N-v+1\},$$
$$\Pr\{U_{(i)} = u, V_{(j)} \le v\} = \Pr\{U_{(m-i+1)} = N-u+1\} - \Pr\{U_{(m-i+1)} = N-u+1, V_{(n-j+1)} \le N-v\}$$

and

$$\Pr\{V_{(j)} \le v \mid U_{(i)} = u\} = 1 - \Pr\{V_{(n-j+1)} \le N-v \mid U_{(m-i+1)} = N-u+1\}.$$

PROOF. The results come from the fact that the joint distribution of $(U_{(m-i+1)}, V_{(n-j+1)})$ is the same as that of $(N+1-U_{(i)}, N+1-V_{(j)})$.

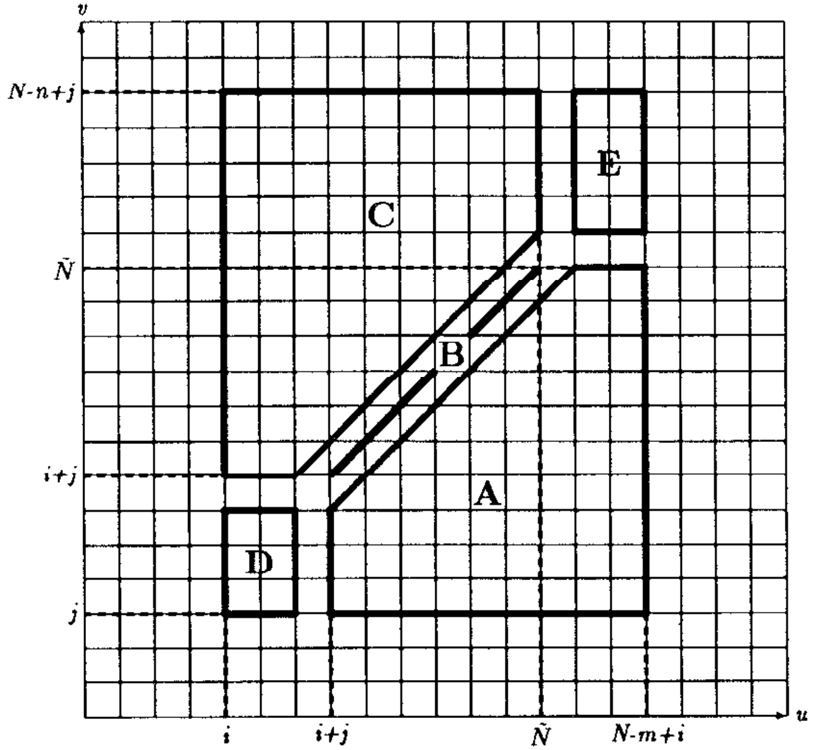

Fig. 1. Example of $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{D}$, $\boldsymbol{E}$ ($N = 20$, $m = 8$, $n = 3$, $i = 4$, $j = 3$).

Let $T_v = \sharp\{a \in \boldsymbol{\Phi} \mid a \leq v\}$ and $R_{(j)} = \sharp\{h \in \boldsymbol{\Phi} \mid h \leq V_{(j)}\}$, where $\sharp\mathcal{X}$ denotes the number of elements in the set $\mathcal{X}$. The following lemma is important for deriving the joint distribution of $U_{(i)}$ and $V_{(j)}$.

LEMMA 3.3. *Let* $(u, v) \in \boldsymbol{O}$. *Then,*

$$\begin{aligned}
&\Pr\{U_{(i)} = u, V_{(j)} \leq v\} \\
&\quad = \sum_{a=j}^{v} \Pr\{U_{(i)} = u, T_v = a\} \Pr\{R_{(j)} \leq a\} \\
&\quad = \sum_{a=j}^{v} \left[\Pr\{U_{(i)} = u, T_v = a\} \sum_{h=j}^{\min\{a, N-m-n+j\}} \Pr\{R_{(j)} = h\} \right].
\end{aligned}$$

PROOF. It is easily checked that

$$\begin{aligned}
V_{(j)} \leq v \Leftrightarrow T_v \geq j \quad &\text{and} \quad R_{(j)} \leq T_v \\
&\Leftrightarrow \bigvee_{a=j}^{v} [T_v = a, R_{(j)} \leq a],
\end{aligned}$$

where $\mathcal{X}_1 \vee \mathcal{X}_2$ means $\mathcal{X}_1$ or $\mathcal{X}_2$. Because $(U_{(i)}, T_v)$ and $R_{(j)}$ are independent, we have

$$\begin{aligned}
&\Pr\{U_{(i)} = u, V_{(j)} \le v\} \\
&= \sum_{a=j}^{v} \Pr\{U_{(i)} = u, T_v = a, R_{(j)} \le a\} \\
&= \sum_{a=j}^{v} \Pr\{U_{(i)} = u, T_v = a\} \Pr\{R_{(j)} \le a\} \\
&= \sum_{a=j}^{v} \left[\Pr\{U_{(i)} = u, T_v = a\} \sum_{h=j}^{\min\{a, N-m-n+j\}} \Pr\{R_{(j)} = h\} \right].
\end{aligned}$$

Now we can give an expression of the joint distribution of $U_{(i)}$ and $V_{(j)}$. Let

$$\begin{aligned}
&\boldsymbol{A}_1 = \{(u,v) \mid (u,v) \in \boldsymbol{A}, v = \tilde{N}\}, \\
&\boldsymbol{A}_2 = \boldsymbol{A} - \boldsymbol{A}_1 = \{(u,v) \mid i+j \le u \le N-m+i, j \le v < \tilde{N}, u > v\}, \\
&\boldsymbol{C}_1 = \{(u,v) \mid (u,v) \in \boldsymbol{C}, v = N-n+j\}
\end{aligned} \tag{3.4}$$

and

$$\boldsymbol{C}_2 = \boldsymbol{C} - \boldsymbol{C}_1.$$

THEOREM 3.1. *Let $1 \le i \le m$, $1 \le j \le n$ and $(u,v) \in \boldsymbol{O}$. Then, $\Pr\{U_{(i)} = u, V_{(j)} \le v\}$ is given as follows:*

$$\Pr\{U_{(i)} = u, V_{(j)} \le v\} = \frac{\binom{u-1}{i-1}\binom{N-u}{m-i}}{\binom{N}{m}}, \qquad (u,v) \in \boldsymbol{A}_1 \cup \boldsymbol{C}_1 \cup \boldsymbol{E}$$

$$\begin{aligned}
&\Pr\{U_{(i)} = u, V_{(j)} \le v\} \\
&= \frac{\binom{N-u}{m-i}}{\binom{N}{m}\binom{N-m}{n}} \sum_{h=j}^{\min\{u-i, v, N-m-n+j\}} \binom{h-1}{j-1}\binom{N-m-h}{n-j} \\
&\quad \times \sum_{a=\max\{h, v+1-i\}}^{\min\{u-i, v\}} \binom{v}{a}\binom{u-1-v}{u-i-a}, \qquad (u,v) \in \boldsymbol{A}_2
\end{aligned} \tag{3.5}$$

$$\begin{aligned}
&\Pr\{U_{(i)} = u, V_{(j)} \le v\} \\
&= \frac{\binom{N-u}{m-i}\binom{u-1}{u-i}}{\binom{N}{m}\binom{N-m}{n}} \sum_{h=j}^{u-i} \binom{h-1}{j-1}\binom{N-m-h}{n-j}, \qquad (u,v) \in \boldsymbol{B}
\end{aligned}$$

(3.6) $\Pr\{U_{(i)} = u, V_{(j)} \le v\}$

$$= \frac{\binom{N-u}{m-i}\binom{u-1}{i-1}}{\binom{N}{m}}$$

$$- \frac{\binom{u-1}{i-1}}{\binom{N}{m}\binom{N-m}{n}}$$

$$\times \sum_{h=n-j+1}^{\min\{N-v,N-u-m+i,N-m-j+1\}} \binom{h-1}{n-j}\binom{N-m-h}{j-1}$$

$$\times \sum_{a=\max\{h,N-v-m+i\}}^{\min\{N-v,N-u-m+i\}} \binom{N-v}{a}\binom{v-u}{N-u-m+i-a},$$

$$(u,v) \in \boldsymbol{C}_2$$

$$\Pr\{U_{(i)} = u, V_{(j)} \le v\} = 0, \qquad (u,v) \in \boldsymbol{D}.$$

PROOF. By elementary combinatorial calculations, we have the following (i), (ii) and (iii).

$$\text{(i)}\ \Pr\{R_{(j)} = h\} = \begin{cases} \dfrac{\binom{h-1}{j-1}\binom{N-m-h}{n-j}}{\binom{N-m}{n}}, & j \le h \le N-m-n+j \\ 0, & \text{otherwise.} \end{cases}$$

$$\text{(ii)}\ \Pr\{U_{(i)} = u\} = \begin{cases} \dfrac{\binom{u-1}{i-1}\binom{N-u}{m-i}}{\binom{N}{m}}, & i \le u \le N-m+i \\ 0, & \text{otherwise.} \end{cases}$$

(iii) If $v = u$ or $v = u-1$, then

$$\Pr\{U_{(i)} = u, T_v = a\} = \begin{cases} 0, & a \ne u-i \\ \Pr\{U_{(i)} = u\}, & a = u-i \end{cases}$$

If $1 \le v \le u-1$, then

$$\Pr\{U_{(i)} = u, T_v = a\}$$
$$= \begin{cases} \dfrac{1}{\binom{N}{m}}\binom{v}{a}\binom{u-1-v}{u-i-a}\binom{N-u}{m-i}, \\ \qquad \max\{0, v+1-i\} \le a \le \min\{u-i, v\} \\ 0, \qquad \text{otherwise.} \end{cases}$$

Case 1: $(u,v) \in \boldsymbol{A}_1 \cup \boldsymbol{C}_1 \cup \boldsymbol{E}$. In this case, we have $\Pr\{U_{(i)} = u, V_{(j)} \leq v\} = \Pr\{U_{(i)} = u\}$. The desired result comes from (ii).

Case 2: $(u,v) \in \boldsymbol{D}$. From Fig. 1, this is obvious.

Case 3: $(u,v) \in \boldsymbol{B}$. In this case, we have

$$\begin{aligned}
&\Pr\{U_{(i)} = u, V_{(j)} \leq v\} \\
&\quad = \Pr\{U_{(i)} = u, V_{(j)} \leq u-1\} \\
&\quad = \sum_{a=j}^{u-1} \Pr\{U_{(i)} = u, T_v = a\} \Pr\{R_{(j)} \leq a\} \qquad \text{(by Lemma 3.3)} \\
&\quad = \Pr\{U_{(i)} = u\} \Pr\{R_{(j)} \leq u-i\} \qquad (U_{(i)} = u \text{ implies } T_v = u-i) \\
&\quad = \frac{\binom{u-1}{i-1}\binom{N-u}{m-i}}{\binom{N}{m}} \sum_{h=j}^{\min\{u-i,N-m-n+j\}} \frac{\binom{h-1}{j-1}\binom{N-m-h}{n-j}}{\binom{N-m}{n}} \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{(by (ii) and (i))} \\
&\quad = \frac{\binom{N-u}{m-i}\binom{u-1}{i-1}}{\binom{N}{m}\binom{N-m}{n}} \sum_{h=j}^{\min\{u-i,N-m-n+j\}} \binom{h-1}{j-1}\binom{N-m-h}{n-j}.
\end{aligned}$$

Case 4: $(u,v) \in \boldsymbol{A}_2$. From Lemma 3.3, (i) and (ii), we have

$$\begin{aligned}
&\Pr\{U_{(i)} = u, V_{(j)} \leq v\} \\
&\quad = \sum_{a=j}^{v} \Pr\{U_{(i)} = u, T_v = a\} \sum_{h=j}^{\min\{a,N-m-n+j\}} \Pr\{R_{(j)} = h\} \\
&\quad = \sum_{a=\max\{j,v+1-i\}}^{\min\{u-i,v\}} \frac{1}{\binom{N}{m}} \binom{v}{a}\binom{u-1-v}{u-i-a}\binom{N-u}{m-i} \\
&\qquad \times \sum_{h=j}^{\min\{a,N-m-n+j\}} \frac{\binom{h-1}{j-1}\binom{N-m-h}{n-j}}{\binom{N-m}{n}} \\
&\quad = \frac{\binom{N-u}{m-i}}{\binom{N}{m}\binom{N-m}{n}} \sum_{a=\max\{j,v+1-i\}}^{\min\{u-i,v\}} \binom{v}{a}\binom{u-1-v}{u-i-a} \\
&\qquad \times \sum_{h=j}^{\min\{a,N-m-n+j\}} \binom{h-1}{j-1}\binom{N-m-h}{n-j}
\end{aligned}$$

$$= \frac{\binom{N-u}{m-i}}{\binom{N}{m}\binom{N-m}{n}} \sum_{h=j}^{\min\{u-i,v,N-m-n+j\}} \binom{h-1}{j-1}\binom{N-m-h}{n-j}$$
$$\times \sum_{a=\max\{h,v+1-i\}}^{\min\{u-i,v\}} \binom{v}{a}\binom{u-1-v}{u-i-a}.$$

Case 5: $(u,v) \in \boldsymbol{C}_2$. This case can be obtained by using Case 4. By Lemma 3.2, we have

$$\begin{aligned} (3.7) \qquad \Pr\{U_{(i)} = u, V_{(j)} \le v\} &= \Pr\{U_{(m-i+1)} = N-u+1\} \\ &\quad - \Pr\{U_{(m-i+1)} = N-u+1, V_{(n-j+1)} \le N-v\}. \end{aligned}$$

Put $i' = m-i+1$, $j' = n-j+1$, $u' = N-u+1$, $v' = N-v$ and $\tilde{N}' = N-m-n+i'+j'-1$. Then, (3.7) $= \frac{\binom{N-u}{m-i}\binom{u-1}{i-1}}{\binom{N}{m}} - \Pr\{U_{(i')} = u', V_{(j')} \le v'\}$. Since $(u,v) \in \boldsymbol{C}_2$, we have $i'+j' \le u' \le N-m+i'$, $j' \le v' < \tilde{N}'$ and $u' > v'$. From these inequalities, it can be said that (u',v') belongs to $\boldsymbol{A}_2$ in the case of $(i,j,\tilde{N}) = (i',j',\tilde{N}')$ in (3.4). Therefore, we can apply Case 4 to $\Pr\{U_{(i')} = u', V_{(j')} \le v'\}$ and obtain an expression corresponding to (3.5). Substituting $i' = m-i+1$, $j' = n-j+1$, $u' = N-u+1$ and $v' = N-v$ for this expression, we obtain (3.6).

This completes the proof.

From this theorem, we obtain the following corollary.

COROLLARY 3.1. *Suppose* $(u,v) \in \boldsymbol{O}$. *Then,*

$$\begin{aligned} &\Pr\{V_{(j)} \le v \mid U_{(i)} = u\} \\ &\quad = 1, \qquad (u,v) \in \boldsymbol{A}_1 \cup \boldsymbol{C}_1 \cup \boldsymbol{E} \\ &\quad = \sum_{h=j}^{\min\{u-i,v,N-m-n+j\}} \frac{\binom{h-1}{j-1}\binom{N-m-h}{n-j}}{\binom{N-m}{n}} \\ &\qquad \times \sum_{a=\max\{h,v+1-i\}}^{\min\{u-i,v\}} \frac{\binom{v}{a}\binom{u-1-v}{u-i-a}}{\binom{u-1}{u-i}}, \qquad (u,v) \in \boldsymbol{A}_2 \\ &\quad = \sum_{h=j}^{u-i} \frac{\binom{h-1}{j-1}\binom{N-m-h}{n-j}}{\binom{N-m}{n}}, \qquad (u,v) \in \boldsymbol{B} \end{aligned}$$

$$= 1 - \sum_{h=n-j+1}^{\min\{N-u-m+i, N-v, N-m-j+1\}} \frac{\binom{h-1}{n-j}\binom{N-m-h}{j-1}}{\binom{N-m}{n}}$$

$$\times \sum_{a=\max\{h, N-v-m+i\}}^{\min\{N-u-m+i, N-v\}} \frac{\binom{N-v}{a}\binom{v-u}{N-u-m+i-a}}{\binom{N-u}{N-u-m+i}}, \qquad (u,v) \in \boldsymbol{C}_2$$

$$= 0, \qquad (u,v) \in \boldsymbol{D}.$$

The following lemma shows a stochastic ordering between hyper geometric distributions and is essential for the proof of Theorem 3.2.

LEMMA 3.4. *Suppose that* r, M *and* L *are positive integers and* $r < M + L$. *For any* k *satisfying* $\max\{0, r-L\} < k \leq \min\{r, M\}$,

$$\sum_{a=k}^{\min\{r,M\}} \frac{\binom{M}{a}\binom{L}{r-a}}{\binom{M+L}{r}} < \sum_{a=k}^{\min\{r+1,M\}} \frac{\binom{M}{a}\binom{L+1}{r+1-a}}{\binom{M+L+1}{r+1}}.$$

PROOF. Let $c = \max\{0, n-L\}$, $d_1 = \min\{r, M\}$ and $d_2 = \min\{r+1, M\}$. Define

$$p_1(a) = \begin{cases} \dfrac{\binom{M}{a}\binom{L}{r-a}}{\binom{M+L}{r}}, & c \leq a \leq d_1 \\ 0, & \text{otherwise}, \end{cases}$$

$$p_2(a) = \begin{cases} \dfrac{\binom{M}{a}\binom{L+1}{r+1-a}}{\binom{M+L+1}{r+1}}, & c \leq a \leq d_2 \\ 0, & \text{otherwise}, \end{cases}$$

and

$$p(a) = p_2(a) - p_1(a).$$

For $c \leq a \leq d_1$,

$$p(a) = \frac{\binom{M}{a}\binom{L}{r-a}}{\binom{M+L}{r}} \frac{\{(M+L+1)a - M(r+1)\}}{(M+L+1)(r+1-a)}.$$

Put $g = \frac{M(r+1)}{M+L+1}$. It is easily checked that $c < g < d_2$. Therefore, we have

$$\begin{aligned} p(a) &< 0, &\quad \text{if} \quad & c \leq a < g, \\ p(a) &= 0, &\quad \text{if} \quad & a = g, \\ p(a) &> 0, &\quad \text{if} \quad & g < a \leq d_1 \end{aligned} \tag{3.8}$$

and particularly,

$$\begin{aligned} &p(c) < 0, \\ &p(d_2) > 0 \quad \text{if} \quad d_2 = d_1. \end{aligned}$$

If $d_2 > d_1$, then $p_2(d_2) > 0$ and $p_1(d_2) = 0$, therefore, $p(d_2) > 0$. We can now replace d_1 in (3.8) by d_2. Then, it is obvious that

$$\sum_{a=k}^{d_2} p(a) > 0, \qquad \text{if} \quad k \geq g$$

and, noting that $\sum_{a=c}^{d_2} p(a) = 0$,

$$\sum_{a=k}^{d_2} p(a) = -\sum_{a=c}^{k-1} p(a) > 0, \qquad \text{if} \quad c < k < g.$$

This completes the proof.

Before proceeding further, we define several subsets of $\boldsymbol{O}$ as follows:

$$\begin{aligned} &\tilde{\boldsymbol{O}} = \{(u,v) \in \boldsymbol{O} \mid u \neq N - m + i\}, \\ &\tilde{\boldsymbol{A}} = \boldsymbol{A} \cap \tilde{\boldsymbol{O}}, \\ &\tilde{\boldsymbol{E}} - \boldsymbol{E} \cap \tilde{\boldsymbol{O}}, \\ &\widetilde{\boldsymbol{A}_1} = \{(u,v) \in \tilde{\boldsymbol{A}} \mid v = \tilde{N}\}, \\ &\widetilde{\boldsymbol{A}_2} = \tilde{\boldsymbol{A}} - \widetilde{\boldsymbol{A}_1}, \\ &\widetilde{\boldsymbol{C}_2} = \{(u,v) \in \boldsymbol{C} - \boldsymbol{C}_1 \mid u \neq \tilde{N}\}, \\ &\boldsymbol{C}_3 = \boldsymbol{C} - \boldsymbol{C}_1 - \widetilde{\boldsymbol{C}_2}, \\ &\boldsymbol{D}_1 = \{(u,v) \in \boldsymbol{D} \mid u = i + j - 1\}, \\ &\boldsymbol{D}_2 = \boldsymbol{D} - \boldsymbol{D}_1 \end{aligned}$$

and

$$\boldsymbol{F} = \tilde{\boldsymbol{E}} \cup \boldsymbol{C}_1 \cup \widetilde{\boldsymbol{A}_1}.$$

Then,

$$\tilde{\boldsymbol{O}} = \widetilde{\boldsymbol{A}_2} \cup \boldsymbol{B} \cup \boldsymbol{C} \cup \boldsymbol{D} \cup \tilde{\boldsymbol{E}},$$

where $\widetilde{\boldsymbol{A}_2}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{D}$ and $\tilde{\boldsymbol{E}}$ are mutually disjoint.

THEOREM 3.2. *$V_{(j)}$ is negatively regression dependent on $U_{(i)}$; that is, for $(u,v) \in \tilde{\boldsymbol{O}}$,*

$$\Pr\{V_{(j)} \le v \mid U_{(i)} = u\} \le \Pr\{V_{(j)} \le v \mid U_{(i)} = u+1\}.$$

Furthermore, we obtain the following results on when equality will hold.

(i) *For $i = 1$, equality holds if and only if $(u,v) \in \boldsymbol{D}_2 \cup \boldsymbol{F} \cup \widetilde{\boldsymbol{A}_2}$.*

(ii) *For $i = m$, equality holds if and only if $(u,v) \in \boldsymbol{D}_2 \cup \boldsymbol{F} \cup \widetilde{\boldsymbol{C}_2}$.*

(iii) *For $1 < i < m$, equality holds if and only if $(u,v) \in \boldsymbol{D}_2 \cup \boldsymbol{F}$.*

PROOF. From Lemma 3.1, it follows that

$$\Pr\{V_{(j)} \le v \mid U_{(i)} = u\} = \Pr\{V_{(j)} \le v \mid U_{(i)} = u+1\} = 0 \quad \text{for} \quad (u,v) \in \boldsymbol{D}_2,$$
$$\Pr\{V_{(j)} \le v \mid U_{(i)} = u\} = \Pr\{V_{(j)} \le v \mid U_{(i)} = u+1\} = 1 \quad \text{for} \quad (u,v) \in \boldsymbol{F},$$
$$\Pr\{V_{(j)} \le v \mid U_{(i)} = u\} = 0 < \Pr\{V_{(j)} \le v \mid U_{(i)} = u+1\} \quad \text{for} \quad (u,v) \in \boldsymbol{D}_1$$

and

$$\Pr\{V_{(j)} \le v \mid U_{(i)} = u\} < 1 = \Pr\{V_{(j)} \le v \mid U_{(i)} = u+1\} \quad \text{for} \quad (u,v) \in \boldsymbol{C}_3.$$

(a) The case of $(u,v) \in \boldsymbol{B}$. Note that $u = v$ in this case. If $u = \tilde{N}$, then by Lemma 3.1, we have

$$\Pr\{V_{(j)} \le \tilde{N} \mid U_{(i)} = \tilde{N}\} < 1 = \Pr\{V_{(j)} \le \tilde{N}+1 \mid U_{(i)} = \tilde{N}+1\}.$$

Suppose that $u < \tilde{N}$. From Lemma 3.1, we have

$$\Pr\{V_{(j)} \le u \mid U_{(i)} = u+1\} = \Pr\{V_{(j)} \le u+1 \mid U_{(i)} = u+1\}.$$

From Corollary 3.1,

$$\begin{aligned}
\Pr\{V_{(j)} \le u \mid U_{(i)} = u\} &= \sum_{h=j}^{u-i} \frac{\binom{h-1}{j-1}\binom{N-m-h}{n-j}}{\binom{N-n}{n}} \\
&< \Pr\{V_{(j)} \le v \mid U_{(i)} = u+1\} \\
&= \sum_{h=j}^{u+1-i} \frac{\binom{h-1}{j-1}\binom{N-m-h}{n-j}}{\binom{N-n}{n}}.
\end{aligned}$$

(b) The case of $(u,v) \in \widetilde{\boldsymbol{A}_2}$. First, note that $(u+1,v) \in \boldsymbol{A}_2$. By Corollary 3.1, we have

(3.9) $$\Pr\{V_{(j)} \le v \mid U_{(i)} = u\}$$

$$= \sum_{h=j}^{\min\{v,u-i,N-m-n+j\}} \frac{\binom{h-1}{j-1}\binom{N-m-h}{n-j}}{\binom{N-m}{n}}$$

$$\times \sum_{a=\max\{h,v+1-i\}}^{\min\{v,u-i\}} \frac{\binom{v}{a}\binom{u-1-v}{u-i-a}}{\binom{u-1}{u-i}}$$

and

(3.10) $$\Pr\{V_{(j)} \leq v \mid U_{(i)} = u+1\}$$

$$= \sum_{h=j}^{\min\{v,u+1-i,N-m-n+j\}} \frac{\binom{h-1}{j-1}\binom{N-m-h}{n-j}}{\binom{N-m}{n}}$$

$$\times \sum_{a=\max\{h,v+1-i\}}^{\min\{v,u+1-i\}} \frac{\binom{v}{a}\binom{u+1-1-v}{u-i-a}}{\binom{u+1-1}{u+1-i}}.$$

If $i = 1$, then it is easily seen that

$$(3.9) = (3.10) = \sum_{h=j}^{\min\{v,N-m-n+j\}} \frac{\binom{h-1}{j-1}\binom{N-m-h}{n-j}}{\binom{N-m}{n}}.$$

Now, suppose that $i \geq 2$. For $v = j$, we have

$$\frac{(3.10)}{(3.9)} = \frac{(u-j)(u+1-i)}{u(u+1-i-j)} > 1.$$

Let $v \geq j+1$. First, we consider the case of $v \leq i+j-1$. In this case, we have

(3.11) $$(3.9) = \sum_{h=j}^{\min\{v,u-i,N-m-n+j\}} \frac{\binom{h-1}{j-1}\binom{N-m-h}{n-j}}{\binom{N-m}{n}}$$

$$\times \left\{ \sum_{a=h}^{\min\{v,u-i\}} \frac{\binom{v}{a}\binom{u-1-v}{u-i-a}}{\binom{u-1}{u-i}} \right\}$$

and

$$(3.12) \qquad (3.10) = \sum_{h=j}^{\min\{v,u+1-i,N-m-n+j\}} \frac{\binom{h-1}{j-1}\binom{N-m-h}{n-j}}{\binom{N-m}{n}} \times \left\{ \sum_{a=h}^{\min\{v,u+1-i\}} \frac{\binom{v}{a}\binom{u+1-1-v}{u+1-i-a}}{\binom{u+1-1}{u+1-i}} \right\}.$$

If $u = v + 1$, then we have $u = i + j$ and $v = i + j - 1$, therefore, we can show (3.10) < (3.11) by direct calculations. If $u > v + 1$, then putting $M = v$, $L = u - 1 - v$ and $r = u - i$ in Lemma 3.4, we have

$$\sum_{a=h}^{\min\{v,u-i\}} \frac{\binom{v}{a}\binom{u-1-v}{u-i-a}}{\binom{u-1}{u-i}} < \sum_{a=h}^{\min\{v,u+1-i\}} \frac{\binom{v}{a}\binom{u-1-v+1}{u-i+1-a}}{\binom{u-1+1}{u-i+1}}, \quad \text{for} \quad j \le h \le \min\{v, u-i\}.$$

Therefore, (3.12) > (3.11). Now, we consider the case of $v \ge i + j$. Then (3.9) can be decomposed into two terms;

$$(3.13) \qquad (3.9) = \sum_{h=j}^{v+1-i} \frac{\binom{h-1}{j-1}\binom{N-m-h}{n-j}}{\binom{N-m}{n}} \left\{ \sum_{a=v+1-i}^{\min\{v,u-i\}} \frac{\binom{v}{a}\binom{u-1-v}{u-i-a}}{\binom{u-1}{u-i}} \right\}$$
$$+ \sum_{h=v+1-i+1}^{\min\{v,u-i,N-m-n+j\}} \frac{\binom{h-1}{j-1}\binom{N-m-h}{n-j}}{\binom{N-m}{n}} \times \left[\sum_{a=h}^{\min\{v,u-i\}} \frac{\binom{v}{a}\binom{u-1-v}{u-i-a}}{\binom{u-1}{u-i}} \right],$$

where the second term of the right-hand side of (3.13) disappears for $v = u + 1$. Similarly, (3.10) can be written as

$$(3.14) \qquad (3.10) = \sum_{h=j}^{v+1-i} \frac{\binom{h-1}{j-1}\binom{N-m-h}{n-j}}{\binom{N-m}{n}}$$

$$\times \left\{ \sum_{a=v+1-i}^{\min\{v,u+1-i\}} \frac{\binom{v}{a}\binom{u+1-1-v}{u+1-i-a}}{\binom{u+1-1}{u+1-i}} \right\}$$

$$+ \sum_{h=v+1-i+1}^{\min\{v,u+1-i,N-m-n+j\}} \frac{\binom{h-1}{j-1}\binom{N-m-h}{n-j}}{\binom{N-m}{n}}$$

$$\times \left[\sum_{a=h}^{\min\{v,u+1-i\}} \frac{\binom{v}{a}\binom{u+1-1-v}{u+1-i-a}}{\binom{u+1-1}{u+1-i}} \right].$$

Because the quantities in { } of (3.13) and (3.14) are 1, the first terms of (3.13) and (3.14) are equal. If $v = u-1$, then the second term of (3.13) $= 0$ and the second term of (3.14) > 0. Therefore, we get (3.14) $>$ (3.13) for $v = u-1$. Now, assume that $v \leq u-2$. Put $M = v$, $L = u-1-v$, $r = u-i$ and $k = h$ for the expression in [ ] of (3.13). By Lemma 3.4, we have

$$\begin{aligned} &\text{the second term of (3.13)} \\ &< \sum_{h=v+1-i+1}^{\min\{v,u+1-i,N-m-n+j\}} \frac{\binom{h-1}{j-1}\binom{N-m-h}{n-j}}{\binom{N-m}{n}} \\ &\quad \times \sum_{a=h}^{\min\{u+1-i,v\}} \frac{\binom{v}{a}\binom{u+1-1-v}{u+1-i-a}}{\binom{u+1-1}{u+1-i}} \\ &\leq \text{the second term of (3.14)}. \end{aligned}$$

Thus, we have (3.13) $<$ (3.14).

(c) The case of $(u,v) \in \widetilde{C_2}$. Let $i' = m-i+1$, $j' = n-j+1$, $u' = N-u$, $v' = N-v$, and $\tilde{N}' = N-m-n+i'+j'-1$. Since $i \leq u < \tilde{N}$, $i+j \leq v \leq N-n+j-1$, and $u < v$, we have $i'+j' \leq u' \leq N-m+i'-1$, $j' \leq v' \leq \tilde{N}'-1$ and $v' < u'$. Therefore, we can use the result of case (b) for (i',j',u',v'). Hence, we have

$$\Pr\{V_{(j')} \leq v' \mid U_{(i')} = u'\} \leq \Pr\{V_{(j')} \leq v' \mid U_{(i')} = u'+1\}$$

with equality holding if and only if $i' = 1$; that is,

$$\begin{aligned} &\Pr\{V_{(n-j+1)} \leq N-v \mid U_{(m-i+1)} = N-u\} \\ &\quad \leq \Pr\{V_{(n-j+1)} \leq N-v \mid U_{(m-i+1)} = N-u+1\} \end{aligned}$$

with equality holding if and only if $m-i+1 = 1$. By Lemma 3.2, we have

$$\Pr\{V_{(j)} \leq v \mid U_{(i)} = u\} = \Pr\{V_{(j)} \leq v \mid U_{(i)} = u+1\} \quad \text{for} \quad i = m$$

and

$$\Pr\{V_{(j)} \le v \mid U_{(i)} = u\} < \Pr\{V_{(j)} \le v \mid U_{(i)} = u+1\} \quad \text{for} \quad i \ne m.$$

This completes the proof of the theorem.

Remark 3.1. If $(u,v) \in \boldsymbol{D}_1 \cup \boldsymbol{B} \cup \boldsymbol{C}_3$, then

$$\Pr\{V_{(j)} \le v \mid U_{(i)} = u\} < \Pr\{V_{(j)} \le v \mid U_{(i)} = u+1\} \quad \text{for any} \quad 1 \le i \le m.$$

Now let us consider the general case $\boldsymbol{\Omega} = \{x_1, \dots, x_1, x_2, \dots, x_2, \dots, x_l, \dots, x_l\}$. We easily obtain the following lemma.

LEMMA 3.5. *For* $1 \le a, b \le l$,

$$\Pr\{X_{(i)} \le x_a, Y_{(j)} \le x_b\} = \Pr\{U_{(i)} \le f_a, V_{(j)} \le f_b\}$$

and

$$\Pr\{X_{(i)} = x_a, Y_{(j)} \le x_b\} = \sum_{s=\tilde{f}_a}^{f_a} \Pr\{U_{(i)} = s, V_{(j)} \le f_b\}.$$

Let $\boldsymbol{I}_{ab} = \{(u, f_b) \mid \tilde{f}_a \le u < f_{a+1}\}$.

THEOREM 3.3. *Let* $1 \le i \le m$ *and* $1 \le j \le n$. *Then,* $Y_{(j)}$ *is negatively regression dependent on* $X_{(i)}$*, that is,*

$$\Pr\{Y_{(j)} \le x_b \mid X_{(i)} = x_a\} \le \Pr\{Y_{(j)} \le x_b \mid X_{(i)} = x_{a+1}\} \tag{3.15}$$

where $i \le f_a < f_{a+1} \le N - m + i$ *and* $j \le f_b \le N - j + 1$. *Furthermore, the following results with equality hold:*

(i) *For* $i = 1$, *equality holds if and only if* $\boldsymbol{I}_{ab} \subseteq \boldsymbol{D}_2 \cup \boldsymbol{F} \cup \widetilde{\boldsymbol{A}}_2$.
(ii) *For* $i = m$, *equality holds if and only if* $\boldsymbol{I}_{ab} \subseteq \boldsymbol{D}_2 \cup \boldsymbol{F} \cup \boldsymbol{C}_2$.
(iii) *For* $1 < i < m$, *equality holds if and only if* $\boldsymbol{I}_{ab} \subseteq \boldsymbol{D}_2 \cup \boldsymbol{F}$.

PROOF. First, note that $i \le f_a < f_{a+1} \le N - m + i \Leftrightarrow \Pr\{X_{(i)} = x_a\} > 0$ and $\Pr\{X_{(i)} = x_{a+1}\} > 0$, and $j \le f_b \le N - j + 1 \Leftrightarrow \Pr\{Y_{(j)} = x_b\} > 0$. Using Lemma 3.5, it is seen that (3.15) is equivalent to

$$\begin{aligned} &\sum_{s=\tilde{f}_a}^{f_a} \sum_{t=\tilde{f}_{a+1}}^{f_{a+1}} \Pr\{U_{(i)} = s, V_{(j)} \le f_b\} \Pr\{U_{(i)} = t\} \\ &\qquad \le \sum_{s=\tilde{f}_a}^{f_a} \sum_{t=\tilde{f}_{a+1}}^{f_{a+1}} \Pr\{U_{(i)} = t, V_{(j)} \le f_b\} \Pr\{U_{(i)} = s\}. \end{aligned} \tag{3.16}$$

From Theorem 3.2, we have

$$\begin{aligned} &\Pr\{U_{(i)} = u_1, V_{(j)} \le f_b\} \Pr\{U_{(i)} = u_2\} \\ &\qquad \le \Pr\{U_{(i)} = u_2, V_{(j)} \le f_b\} \Pr\{U_{(i)} = u_1\}, \end{aligned} \tag{3.17}$$

for $i \le u_1 < u_2 \le N - m + i$. For $1 \le u_1 \le i - 1$ or $u_2 \ge N - m + 1$, both sides of (3.17) are 0. Summing the inequalities (3.17) over $\tilde{f}_a \le u_1 \le f_a$, $\tilde{f}_{a+1} \le u_2 \le f_{a+1}$, we obtain (3.16). It remains only to check the conditions for equality. From (3.16) and Theorem 3.2, it follows that

the equality holds in (3.15)

$\Leftrightarrow$ the equality holds in (3.16) for all (s,t) such that $\tilde{f}_a \le s \le f_a$ and $\tilde{f}_{a+1} \le t \le f_{a+1}$

$\Leftrightarrow$ $(s, f_b), (s+1, f_b), \ldots, (t, f_b) \in \chi$ for all (s,t) such that $\tilde{f}_a \le s \le f_a$ and $\tilde{f}_{a+1} \le t \le f_{a+1}$, where χ denotes $\boldsymbol{D}_2 \cup \boldsymbol{F} \cup \widetilde{\boldsymbol{A}}_2$ for $i = 1$, $\boldsymbol{D}_2 \cup \boldsymbol{F} \cup \boldsymbol{C}_2$ for $i = m$, and $\boldsymbol{D}_2 \cup \boldsymbol{F}$ for $1 < i < m$.

$\Leftrightarrow$ $(s, f_b) \in \chi$ for $\tilde{f}_a \le s \le f_{a+1} - 1$

$\Leftrightarrow$ $\boldsymbol{I}_{ab} \subseteq \chi$.

This completes the proof of the theorem.

4. Covariances of two order statistics from one sample and from two samples

Let $\{Y_1, Y_2, \ldots, Y_n\}$ be a simple random sample of size n without replacement from $\boldsymbol{\Omega} = \{x_1, \ldots, x_1, x_2, \ldots, x_2, \ldots, x_l, \ldots, x_l\}$ of size N. Let $Y_{(1)} \le Y_{(2)} \le \cdots \le Y_{(n)}$ be the order statistics of $\{Y_1, Y_2, \ldots, Y_n\}$. We define $\mathcal{A}_{n:i}$ and $\mathcal{B}$ by $\{s \mid i \le s < N - n + i\}$ and $\{f_k \mid 1 \le k \le l\}$, respectively.

THEOREM 4.1. *Let $1 \le i < j \le n$. Then we have*

$$\mathrm{Cov}(Y_{(i)}, Y_{(j)}) \ge 0$$

with equality holding if and only if $\mathcal{A}_{n:i} \cap \mathcal{B} = \emptyset$ or $\mathcal{A}_{n:j} \cap \mathcal{B} = \emptyset$.

PROOF. If $\mathcal{A}_{n:i} \cap \mathcal{B} = \emptyset$, then $Y_{(i)}$ is constant, therefore, $\mathrm{Cov}(Y_{(i)}, Y_{(j)}) = 0$. Similarly, if $\mathcal{A}_{n:j} \cap \mathcal{B} = \emptyset$, then $\mathrm{Cov}(Y_{(i)}, Y_{(j)}) = 0$. Suppose that $\mathcal{A}_{n:i} \cap \mathcal{B} \ne \emptyset$ and $\mathcal{A}_{n:j} \cap \mathcal{B} \ne \emptyset$. Let $f_a = \min\{\mathcal{A}_{n:i} \cap \mathcal{B}\}$ and let $f_b = \max\{\mathcal{A}_{n:j} \cap \mathcal{B}\}$. Then, $(i, f_b) \in \boldsymbol{G}_{ab} \cap \boldsymbol{S}$ and $(f_a + 1, N - n + j) \in \boldsymbol{G}_{(a+1)(b+1)} \cap \boldsymbol{S}$. By Theorem 2.1, we get

$$\begin{aligned} &\Pr\{Y_{(i)} = x_a, Y_{(j)} = x_b\} \Pr\{Y_{(i)} = x_{a+1}, Y_{(j)} = x_{b+1}\} \\ &\quad > \Pr\{Y_{(i)} = x_a, Y_{(j)} = x_{b+1}\} \Pr\{Y_{(i)} = x_{a+1}, Y_{(j)} = x_b\}. \end{aligned}$$

Hence, $Y_{(i)}$ and $Y_{(j)}$ are not independent. Since, by Theorem 2.1, $(Y_{(i)}, Y_{(j)})$ is positive likelihood ratio dependent, $(Y_{(i)}, Y_{(j)})$ is positive quadrant dependent (See Lehmann (1959), p. 74 and Lehmann (1966), p. 1144). By Lemma 3 of Lehmann (1966), we have $\mathrm{Cov}(Y_{(i)}, Y_{(j)}) > 0$. This completes the proof of the theorem.

Let $\{X_1, X_2, \ldots, X_m, Y_1, Y_2, \ldots, Y_n\}$ be a simple random sample of size $m + n$ without replacement from $\boldsymbol{\Omega}$. Let $X_{(1)} \le X_{(2)} \le \cdots \le X_{(m)}$ be the order statistics of $X_1, X_2, \ldots, X_m$, and $Y_{(1)} \le Y_{(2)} \le \cdots \le Y_{(n)}$ the order statistics of $Y_1, Y_2, \ldots, Y_n$.

THEOREM 4.2. *Let $1 \leq i \leq m$ and $1 \leq j \leq n$. Then we have*

$$\mathrm{Cov}(X_{(i)}, Y_{(j)}) \leq 0,$$

with equality holding if and only if $\mathcal{A}_{m:i} \cap \mathcal{B} = \emptyset$ or $\mathcal{A}_{n:j} \cap \mathcal{B} = \emptyset$.

PROOF. If $\mathcal{A}_{m:i} \cap \mathcal{B} = \emptyset$ or $\mathcal{A}_{n:j} \cap \mathcal{B} = \emptyset$, then $X_{(i)}$ or $Y_{(j)}$ is constant, therefore, $\mathrm{Cov}(X_{(i)}, Y_{(j)}) = 0$. Assume that $\mathcal{A}_{m:i} \cap \mathcal{B} \neq \emptyset$ and $\mathcal{A}_{n:j} \cap \mathcal{B} \neq \emptyset$. Since, by Theorem 3.3, $Y_{(j)}$ is negatively regression dependent on $X_{(i)}$, $(X_{(i)}, Y_{(j)})$ is negatively quadrant dependent (See Lehmann (1966), p. 1144). By Lemma 3 of Lehmann (1966), we have $\mathrm{Cov}(X_{(i)}, Y_{(j)}) \leq 0$ and if $X_{(i)}$ and $Y_{(j)}$ are not independent, then we have $\mathrm{Cov}(X_{(i)}, Y_{(j)}) < 0$. Now, we shall prove that $X_{(i)}$ and $Y_{(j)}$ are not independent. Suppose that $\boldsymbol{I}_{ab} \cap (\boldsymbol{B} \cup \boldsymbol{C}_3 \cup \boldsymbol{D}_1) \neq \emptyset$. Since $(\boldsymbol{B} \cup \boldsymbol{C}_3 \cup \boldsymbol{D}_1) \cap (\boldsymbol{A}_2 \cup \boldsymbol{C}_2 \cup \boldsymbol{D}_2 \cup \boldsymbol{F}) = \emptyset$, we have $\boldsymbol{I}_{ab} \not\subseteq \boldsymbol{A}_2 \cup \boldsymbol{C}_2 \cup \boldsymbol{D}_2 \cup \boldsymbol{F}$. By Theorem 3.3, we have

$$\Pr\{Y_{(j)} \leq x_b \mid X_{(i)} = x_a\} < \Pr\{Y_{(j)} \leq x_b \mid X_{(i)} = x_{a+1}\}.$$

This implies that $X_{(i)}$ and $Y_{(j)}$ are not independent. Therefore, it suffices to show the existence of (a, b) such that $\boldsymbol{I}_{ab} \cap (\boldsymbol{B} \cup \boldsymbol{C}_3 \cup \boldsymbol{D}_1) \neq \emptyset$. Let $\mathcal{J} = \mathcal{A}_{m:i} \cap \mathcal{A}_{n:j} \cap \mathcal{B}$, $\mathcal{C}_1 = \{k \mid \max\{i, j\} \leq k < i + j\}$, $\mathcal{C}_2 = \{k \mid i + j \leq k \leq \tilde{N}\}$ and $\mathcal{C}_3 = \{k \mid \tilde{N} < k < \min\{N - m + i, N - n + j\}\}$.

Case (i): $\mathcal{J} \neq \emptyset$.

(a): $\mathcal{C}_2 \cap \mathcal{B} \neq \emptyset$. Let $f_a \in \mathcal{C}_2 \cap \mathcal{B}$. Then, $i + j \leq f_a \leq \tilde{N}$. It follows that $(f_a, f_a) \in \boldsymbol{B}$. Thus, $\boldsymbol{I}_{aa} \cap \boldsymbol{B} \neq \emptyset$.

(b): $\mathcal{C}_2 \cap \mathcal{B} = \emptyset$ and $\mathcal{C}_1 \cap \mathcal{B} \neq \emptyset$. Let $f_a = \max\{\mathcal{C}_1 \cap \mathcal{B}\}$. Then, $\max\{i, j\} < f_a < i + j$ and $\tilde{f}_a \leq i + j - 1 < f_{a+1}$. It follows that $(i + j - 1, f_a) \in \boldsymbol{I}_{aa} \cap \boldsymbol{D}_1$. Thus, $\boldsymbol{I}_{aa} \cap \boldsymbol{D}_1 \neq \emptyset$.

(c): $\mathcal{C}_1 \cap \mathcal{B} = \emptyset$, $\mathcal{C}_2 \cap \mathcal{B} = \emptyset$ and $\mathcal{C}_3 \neq \emptyset$. Let $f_a = \min\{\mathcal{C}_3 \cap \mathcal{B}\}$. We have $\tilde{N} + 1 \leq f_a \leq \min\{N - m + i, N - n + j\} - 1$ and $\tilde{f}_a \leq \tilde{N} < f_{a+1}$. It follows that $(\tilde{N}, f_a) \in \boldsymbol{I}_{aa} \cap \boldsymbol{C}_3$. Thus, $\boldsymbol{I}_{aa} \cap \boldsymbol{C}_3 \neq \emptyset$.

Case (ii): $\mathcal{J} = \emptyset$. Note that, under the assumption $\mathcal{A}_{m:i} \cap \mathcal{B} \neq \emptyset$ and $\mathcal{A}_{n:j} \cap \mathcal{B} \neq \emptyset$, $i = j$ implies $\mathcal{J} \neq \emptyset$.

(a): $i < j$. There exists $f_a \in \mathcal{B}$ such that $i \leq f_a < j$ and $N - m + i \leq f_{a+1} < N - n + j$. It is seen that $(\tilde{N}, f_a) \in \boldsymbol{I}_{a(a+1)} \cap \boldsymbol{C}_3$. Thus, $\boldsymbol{I}_{a(a+1)} \cap \boldsymbol{C}_3 \neq \emptyset$.

(b): $i > j$. There exists $f_a \in \mathcal{B}$ such that $j \leq f_a < i$ and $N - n + j \leq f_{a+1} < N - m + i$. It is seen that $(i + j - 1, f_a) \in \boldsymbol{I}_{aa} \cap \boldsymbol{D}_1$. Thus, $\boldsymbol{I}_{aa} \cap \boldsymbol{D}_1 \neq \emptyset$.

This completes the proof of the theorem.

5. Lower bound of relative precision of RSS from a finite population

We draw a simple random sample $\{X_{ijk} \mid i, j = 1, 2, \ldots, n; k = 1, 2, \ldots, r\}$ of size $n^2 r$ without replacement from a finite population $\mathbf{\Omega}$ of size N, with mean μ and variance σ^2. For each i and k, the n units $\{X_{i1k}, X_{i2k}, \ldots, X_{ink}\}$ are ranked by a visual inspection. The unit with the i-th smallest rank is quantified. This procedure involves the quantification of nr units out of the $n^2 r$ units originally drawn.

Let $X_{[i]k}$ be the i-th smallest order statistics of $\{X_{i1k}, X_{i2k}, \ldots, X_{ink}\}$. Then, the ranked set sample obtained by the above procedure can be written as $\{X_{[i]k} \mid i = 1, 2, \ldots, n; k = 1, 2, \ldots, r\}$ and the ranked set estimator $\hat{\mu}_{RSS}$ of μ is the average of $X_{[i]k}$ $(i = 1, 2, \ldots, n; k = 1, 2, \ldots, r)$:

$$\hat{\mu}_{RSS} = \bar{X}_{[n]r} = \frac{1}{nr} \sum_{k=1}^{r} \sum_{i=1}^{n} X_{[i]k}.$$

This is an unbiased estimator of μ (Takahasi and Futatsuya (1988), Patil *et al.* (1995)). Let $\bar{X}_{nr}$ be the sample mean of a simple random sample of size nr drawn without replacement from $\mathbf{\Omega}$. The definition of relative precision (RP) of $\bar{X}_{[n]r}$ (Patil *et al.* (1994)) is

$$RP = \frac{\mathrm{Var}(\bar{X}_{nr})}{\mathrm{Var}(\bar{X}_{[n]r})}.$$

In this section, we show that $RP > 1$ for almost all populations, and $RP = 1$ for very exceptional populations. The corresponding result in infinite population can be found in Takahasi and Wakimoto (1968).

Let $\{X_1, X_2, \ldots, X_n, Y_1, Y_2, \ldots, Y_n\}$ be a simple random sample of size $2n$ without replacement from $\mathbf{\Omega}$. Let $X_{(1)} \le X_{(2)} \le \cdots \le X_{(n)}$ be the order statistics of $X_1, X_2, \ldots, X_n$, and $Y_{(1)} \le Y_{(2)} \le \cdots \le Y_{(n)}$ be the order statistics of $Y_1, Y_2, \ldots, Y_n$. We shall use the following notations:

$$\mu_{n:i} = \mathrm{E}(X_{(i)}) = \mathrm{E}(Y_{(i)}),$$
$$\tilde{\alpha}_{n:ij} = \mathrm{Cov}(X_{(i)}, X_{(j)}) = \mathrm{Cov}(Y_{(i)}, Y_{(j)}),$$

and

$$\tilde{\gamma}_{n:ij} = \mathrm{Cov}(X_{(i)}, Y_{(j)}).$$

LEMMA 5.1. *Let $2n \le N$. Then,*

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \tilde{\gamma}_{n:ij} = -\frac{n^2 \sigma^2}{N-1} \tag{5.1}$$

and

$$\sum_{i=1}^{n} \tilde{\alpha}_{n:ii} = \frac{N-n}{N-1} n\sigma^2 - \sum_{i \ne j} \sum \tilde{\alpha}_{n:ij}. \tag{5.2}$$

PROOF. It is well known that

$$\operatorname{Cov}(\bar{X}, \bar{Y}) = -\frac{\sigma^2}{N-1} \tag{5.3}$$

and

$$\operatorname{Var}\bar{X} = \frac{N-n}{N-1}\frac{\sigma^2}{n}, \tag{5.4}$$

where $\bar{X} = \sum_{i=1}^{n} X_i$ and $Y = \sum_{i=1}^{n} Y_i$. On the other hand, we have

$$\operatorname{Cov}(\bar{X}, \bar{Y}) = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\tilde{\gamma}_{n:ij} \tag{5.5}$$

and

$$\operatorname{Var}\bar{X} = \frac{1}{n^2}\left[\sum_{i=1}^{n}\tilde{\alpha}_{n:ii} + \sum\sum_{i \neq j}\tilde{\alpha}_{n:ij}\right]. \tag{5.6}$$

From (5.3) and (5.5), we have (5.1), and from (5.4) and (5.6), we have (5.2).

THEOREM 5.1. *Let $n^2 r \leq N$. Then,*

$$\operatorname{Var}(\bar{X}_{[n]r}) = \operatorname{Var}(\bar{X}_{nr}) - \frac{1}{n^2 r}\left(\sum\sum_{i \neq j}\tilde{\alpha}_{n:ij} - \sum\sum_{i \neq j}\tilde{\gamma}_{n:ij}\right). \tag{5.7}$$

PROOF. By some simple calculations, we obtain

$$\operatorname{Var}(\bar{X}_{[n]r}) = \frac{1}{n^2 r^2}\left(r\sum_{i=1}^{n}\tilde{\alpha}_{n:ii} + r\sum\sum_{i \neq j}\tilde{\gamma}_{n:ij} + r(r-1)\sum_{i=1}^{n}\sum_{j=1}^{n}\tilde{\gamma}_{n:ij}\right). \tag{5.8}$$

Substituting (5.1) and (5.2) for the right-hand of (5.8), we obtain (5.7). The proof is complete.

Now, we can prove the following theorem.

THEOREM 5.2. *Let $n^2 r \leq N$. Then,*

$$RP \geq 1,$$

with equality holding if and only if one of the following conditions for populations is fulfilled:

(i) $\{x_1 < x_2 = x_3 = \cdots = x_N\}$;
(ii) $\{x_1 = x_2 = \cdots = x_{N-1} < x_N\}$;
(iii) $\{x_1 = x_2 = \cdots = x_N\}$.

PROOF. By Theorem 4.1 and Theorem 4.2, we have

$$\sum\sum_{i\neq j} \tilde{\alpha}_{n:ij} - \sum\sum_{i\neq j} \tilde{\gamma}_{n:ij} \geq 0. \tag{5.9}$$

From this and Theorem 5.1, we have $\mathrm{Var}(X_{[n]r}) \leq \mathrm{Var}(\bar{X}_{nr})$. Therefore, $RP \geq 1$. The equality in (5.9) holds if and only if $\tilde{\alpha}_{n:ij} = 0$ and $\tilde{\gamma}_{n:ij} = 0$ for all $1 \leq i, j \leq n$ $(i \neq j)$. Now, we assume that $\tilde{\alpha}_{n:ij} = 0$ and $\tilde{\gamma}_{n:ij} = 0$ for all $1 \leq i, j \leq n$ $(i \neq j)$. Then, from Theorem 4.1 or Theorem 5.1, we must have $\mathcal{A}_{n:i} \cap \mathcal{B} = \emptyset$ or $\mathcal{A}_{n:j} \cap \mathcal{B} = \emptyset$ for any (i, j), such that $1 \leq i, j \leq n$ and $i \neq j$. Suppose that, for $1 < k < n$, $\mathcal{A}_{n:k} \cap \mathcal{B} \neq \emptyset$. From this assumption, $\mathcal{A}_{n:k-1} \cap \mathcal{B} = \emptyset$ and $\mathcal{A}_{n:k+1} \cap \mathcal{B} = \emptyset$. Recall that $\mathcal{A}_{n:k} = \{s \mid k \leq s < N - n + k\}$. Because $\mathcal{A}_{n:k} \subseteq \mathcal{A}_{n:k-1} \cup \mathcal{A}_{n:k+1}$, it follows that $\mathcal{A}_{n:k} \cap \mathcal{B} = \emptyset$. This is a contradiction. Therefore, under this assumption, one of the following cases must hold:

(a) $\mathcal{A}_{n:1} \cap \mathcal{B} \neq \emptyset$ and $\mathcal{A}_{n:k} \cap \mathcal{B} = \emptyset$ for $2 \leq k \leq n$;

(b) $\mathcal{A}_{n:n} \cap \mathcal{B} \neq \emptyset$ and $\mathcal{A}_{n:k} \cap \mathcal{B} = \emptyset$ for $1 \leq k \leq n - 1$;

(c) $\mathcal{A}_{n:k} \cap \mathcal{B} = \emptyset$ for $1 \leq k \leq n$.

It is obvious that (a) implies (i), (b) implies (ii), and (c) implies (iii). Conversely, it is clear that (i), (ii) or (iii) implies $RP = 1$. This completes the proof.

From Theorem 5.2, we can say that $RP > 1$, unless $N - 1$ elements of the finite population of size N have the same value.

REFERENCES

Futatsuya, M. and Takahasi, K. (1990). Extremal finite populations maximizing relative efficiencies in ranked set sampling, *Sci. Rep. Hirosaki Univ.*, **37**, 152–159.

Lehmann, E. L. (1959). *Testing Statistical Hypotheses*, Wiley, New York.

Lehmann, E. L. (1966). Some concepts of dependence, *Ann. Math. Statist.*, **37**, 1137–1153.

Patil, G. P., Sinha, A. K. and Taillie, C. (1994). Ranked set sampling, *Handbook of Statistics, Vol. 12: Environmental Statistics* (eds. G. P. Patil and C. R. Rao), 167–200, North Holland Elsevier, New York.

Patil, G. P., Sinha, A. K. and Taillie, C. (1995). Finite population corrections for ranked set sampling, *Ann. Inst. Statist. Math.*, **47**, 621–636.

Takahasi, K. and Futatsuya, M. (1988). Ranked set sampling from a finite population, *Proc. Inst. Statist. Math.*, **36**, 55–68 (Japanese with English summary).

Takahasi, K. and Wakimoto, K. (1968). On unbiased estimates of the population mean based on the sample stratified by means of ordering, *Ann. Inst. Statist. Math.*, **20**, 1–31.