The Annals of Mathematical Statistics
1972, Vol. 43, No. 2, 373-397

ASYMPTOTIC THEORY FOR SUCCESSIVE SAMPLING WITH VARYING PROBABILITIES WITHOUT REPLACEMENT, I

BY BENGT ROSÉN

Royal Institute of Technology, Stockholm

To each of the items $1, 2, \cdots, N$ in a finite population there is associated a variate value. The population is sampled by successive drawings without replacement in the following way. At each draw the probability of drawing item s is proportional to a number $p_s > 0$ if item s remains in the population and is 0 otherwise. Let $\Delta(s; n)$ be the probability that item s is obtained in the first n draws and let Z_n be the sum of the variate values obtained in the first n draws. Asymptotic formulas, valid under general conditions when n and N both are "large", are derived for $\Delta(s; n)$, EZ_n and $\operatorname{Cov}(Z_{n_1}, Z_{n_2})$. Furthermore it is shown that, still under general conditions, the joint distribution of $Z_{n_1}, Z_{n_2}, \cdots, Z_{n_d}$ is asymptotically normal. The general results are then applied to obtain asymptotic results for a "quasi"-Horvitz-Thompson estimator of the population total.

1. Introduction, summary and notation. In recent years considerable interest has been paid to sampling with varying probabilities, a term which covers a diversity of sampling procedures. For general expositions of this subject we suggest the book [7] by Sukhatme and Sukhatme and the paper [2] by Hájek. In this paper we shall be concerned with the sampling procedure which, in accordance with the terminology in [2], is called successive sampling without replacement. With this sampling procedure items are sampled one after the other and without replacement from a finite collection, in such a way that at each draw, the probability of drawing item s is proportional to the number p_s if item s remains in the collection. Some quantities of particular interest are the following inclusion probabilities,

(1.1) $\Delta(s, n) =$ probability that item s is included in a sample of size n.

We shall assume that a variate value a_s is associated with each item s. Let $\pi = (a_1, a_2, \cdots, a_N)$ be the population of variate values. The sample sum Z_n is the sum of the variate values in a sample of size n.

Our main concern in this paper is the asymptotic behaviour of the inclusion probabilities and of the sample sum as n and N tend to infinity simultaneously. We derive approximation formulas for the mean and variance of Z_n, which are valid under general conditions when n and N both are large. Moreover we show that, still under general conditions, Z_n is approximately standard normally distributed if "standardized" with the approximations of its mean and variance.

Received June 22, 1971; revised October 21, 1971.

From a statistical point of view the "most interesting" linear statistic of the sample variates is not the sample sum, but other linear statistics which yield unbiased estimates of the population total, or equivalently of the population mean. However, once one has general results about the asymptotic behaviour of the sample sum, these results can be transformed quite easily to yield asymptotic results for other linear statistics. We illustrate this point by deriving asymptotic results for the so called Horvitz-Thompson estimator, which is defined as follows,

$$(HT)_n = \sum_{s=1}^{N} W_s \cdot \frac{a_s}{\Delta(s, n)} \tag{1.2}$$

where

$$\begin{aligned} W_s &= 1 \quad \text{if item } s \text{ is included in the sample} \\ &= 0 \quad \text{otherwise} \end{aligned} \tag{1.3}$$

and $\Delta(s, n)$ is the inclusion probability in (1.1).

This paper can be characterized as the probabilistic part of the large-sample theory for successive sampling. Problems still remain to be solved in order to "convert" the results into statistical procedures. We hope to pursue the matter in this direction later.

The key tool in the following analysis is the technique which was worked out by the author in [4]. This technique was applied in [5] to a coupon collection problem. On the basis of the results in [5] the author derived further results concerning coupon collection in [6]. Some of the results in [6] will be crucial in this paper. The course of the paper is as follows. In the next section we make precise some concepts and notation concerning successive sampling. In Section 3 we formulate the main results about the asymptotic behaviour of the inclusion probabilities and the sample sum. These results are then proved in Sections 4—13. The last section is devoted to the asymptotic behaviour of the Horvitz-Thompson estimator.

Two papers of relevance here are the earlier mentioned paper [2] by Hájek and the paper [1] also by Hájek. In [2] Hájek solves the same problems as we shall consider here for a sampling procedure which is called rejective sampling. In [1] Hájek showed, under the assumption that the sampling fraction keeps away from 0 and 1, that a necessary and sufficient condition for the validity of the central limit theorem for simple random sampling without replacement is the so-called Noether's condition (see (3.5) in [1]). Successive sampling without replacement includes as a particular case simple random sampling. When specializing our results in this paper to simple random sampling we do not obtain the central limit theorem under Noether's condition but only under the stronger assumptions (3.21) and (3.23). However, we believe that this is only a technical fallacy and we conjecture that Theorem

3.3 in this paper is true even if the assumptions (3.21) and (3.23) are replaced by Noether's condition. However, we have not been able to obtain the necessary estimates in this more general case.

We close this section by listing some notation that will be used throughout this paper. P denotes probability, E expectation, $\sigma^2(\cdot)$ variance and $\mathrm{Cov}(\cdot, \cdot)$ covariance. The superscript ^{c} indicates centering at mean, i.e. $Y^c = Y - EY$. $\mathscr{L}(X)$ denotes the distribution of the random element X and $X =_{\mathscr{L}} Y$ means that $\mathscr{L}(X) = \mathscr{L}(Y)$. $\mathscr{B}(X_1, X_2, \cdots)$ is the σ-algebra of events which is generated by the random variables $X_1, X_2, \cdots$. $E^{\mathscr{B}}$ denotes conditional expectation given the σ-algebra $\mathscr{B}$. Instead of $E^{\mathscr{B}(X)}$ we often write E^X. Weak (mass-preserving) convergence of measures is denoted by $\Rightarrow$. $N(\mu, A)$ stands for the normal distribution with mean vector μ and covariance matrix (operator) A. Matrices will usually be written as follows, $[a_{\nu\mu}; \nu, \mu = 1, 2, \cdots, d]$. $\#$ denotes number. We use the convention that when we put a non-integer λ in a place where there would naturally be an integer, then λ is interpreted as its integral part $[\lambda]$. Absolute constants will be denoted by C. If the value of a constant depends on parameter u, we write it C_u. We shall repeatedly be concerned with bounding functions $C(x, y)$ with the following properties.

(1.4) $C(x, y)$ is defined for $1 \leqq x < \infty, 0 \leqq y < \infty$, is continuous in (x, y) and is non-decreasing in each of its arguments.

Because of continuity, $C(x, y)$ is bounded on every compact part of its domain of definition.

2. On the sampling procedure. In this section we shall introduce some basic concepts, terminology and notation concerning successive sampling. Let

$$\mathbf{p} = (p_1, p_2, \cdots, p_N) \tag{2.1}$$

where

$$p_s > 0, \qquad s = 1, 2, \cdots, N \tag{2.2}$$

and

$$p_1 + p_2 + \cdots + p_N = 1. \tag{2.3}$$

We shall consider two different procedures to produce successive samples.

DRAWING PROCEDURE 1. Numbers are drawn one after the other without replacement from the collection $(1, 2, \cdots, N)$. At each draw we have

(2.4) Probability that s is drawn
= proportional to p_s if s remains in the collection
= 0 otherwise.

This drawing procedure will deliver the numbers $1, 2, \cdots, N$ in a random order, $I_1, I_2, \cdots, I_N$. The random vector $I_1, I_2, \cdots, I_N$ will be called a **p**-*permutation* of $1, 2, \cdots, N$.

More formally we define the random vector $I_1, I_2, \cdots, I_N$ as follows. Let

$$\Omega = \{(i_1, i_2, \cdots, i_N)\} \tag{2.5}$$

be the set of all permutations of $1, 2, \cdots, N$. We introduce the probability measure P on Ω by the following elementary probabilities.

$$P((i_1, i_2, \cdots, i_N)) = \frac{p_{i_1}}{1} \cdot \frac{p_{i_2}}{1 - p_{i_1}} \cdot \frac{p_{i_3}}{1 - (p_{i_1} + p_{i_2})} \cdot \cdots \cdot \frac{p_{i_N}}{1 - (p_{i_1} + \cdots + p_{i_{N-1}})} . \tag{2.6}$$

Now define the random variables $I_1, I_2, \cdots, I_N$ on (Ω, P) by

$$I_\nu = (i_1, i_2, \cdots, i_N) = i_\nu , \qquad \nu = 1, 2, \cdots, N . \tag{2.7}$$

REMARK. In (2.3) we made the normalizing assumption that the "draw probability proportionates" $p_1, p_2, \cdots, p_N$ are probabilities, i.e. that they sum up to 1 (see (2.3)). This assumption could have been dispensed with if we instead had written formula (2.6) in the following way.

$$P((i_1, i_2, \cdots, i_N)) \quad \frac{p_{i_1}}{\sum_{s=1}^N p_s} \cdot \frac{p_{i_2}}{\sum_{s=1}^N p_s - p_{i_1}} \cdot \cdots \cdot \frac{p_{i_N}}{\sum_{s=1}^N p_s - (p_{i_1} + \cdots + p_{i_{N-1}})} . \tag{2.8}$$

We shall refer to this case as the case with *genenal draw probability proportionates.* However, it will usually be convenient to have the normalizing assumption (2.3), and we adopt the following convention. Unless we explicitly state that we are dealing with general draw probability proportionates we assume that (2.3) is fulfilled.

We now regard $1, 2, \cdots, N$ as labels on the items in a finite collection and we assume that a variate value (i.e. a real number) is associated with each item. Let a_1 be the number associated with item 1, a_2 the number associated with item 2, etc. The set of numbers

$$\pi = (a_1, a_2, \cdots, a_N) \tag{2.9}$$

will be called a (finite) *population*, and $a_1, a_2, \cdots, a_N$ the *elements in the population.*

DEFINITION 2.1. Let $I_1, I_2, \cdots, I_N$ be a **p**-permutation of $1, 2, \cdots, N$ and let $\pi = (a_1, a_2, \cdots, a_N)$ be a finite population. Put

$$Y_1 = a_{I_1} , \qquad Y_2 = a_{I_2}, \cdots, Y_N = a_{I_N} . \tag{2.10}$$

The random vector $Y_1, Y_2, \cdots, Y_N$ will be called a **p**-*permutation of the elements in* π, and the random vector $Y_1, Y_2, \cdots, Y_N$, $1 \leqq n \leqq N$, a **p**-*sample of size* n *from* π.

The characteristics of a sampling situation of this type are thus $\mathbf{p} = (p_1, p_2, \cdots, p_N)$ and $\pi = (a_1, a_2, \cdots, a_N)$. We will refer to the pair $(\mathbf{p}, \pi)$ as the *sampling situation*.

As an over-all-term for $\mathbf{p}$-sampling we use the term *successive sampling (without replacement)*.

DEFINITION 2.2. The *inclusion probability* for item s in a sample of size n is

$$\Delta(s, n) = P(s \in (I_1, I_2, \cdots, I_n)), \qquad s, n = 1, 2, \cdots, N. \tag{2.11}$$

Next we consider an alternative way to obtain a $\mathbf{p}$-permutation.

DRAWING PROCEDURE 2. Numbers are drawn one after the other with replacement from the collection $(1, 2, \cdots, N)$. At each draw we have

$$p_s = \text{the probability that number } s \text{ is drawn.} \tag{2.12}$$

A number drawn is "included in the permutation" if it was not obtained previously. Let $I_1, I_2, \cdots, I_N$ be the numbers $1, 2, \cdots, N$ in the random order in which they are included in the permutation.

More formally we define the random vector $I_1, I_2, \cdots, I_N$ as follows. Let $J_1 J_2, \cdots$ be independent random variables, all having the distribution

$$P(J = s) = p_s, \qquad s = 1, 2, \cdots, N. \tag{2.13}$$

Let

$$M_n = \text{the number of different elements among } (J_1, J_2, \cdots, J_n), \qquad n = 1, 2, \cdots \tag{2.14}$$

and

$$T_\nu = \inf\{n : M_n = \nu\}, \qquad \nu = 1, 2, \cdots, N \tag{2.15}$$

i.e. T_ν is the waiting time to obtain ν different numbers. Now put

$$I_\nu' = J_{T_\nu}, \qquad \nu = 1, 2, \cdots, N. \tag{2.16}$$

The following lemma which states the equivalence of the two drawing procedures is intuitively clear. We omit a formal proof.

LEMMA 2.1.

$$(I_1', I_2', \cdots, I_N') =_{\mathscr{L}} (I_1, I_2, \cdots, I_N). \tag{2.17}$$

Because of the equivalence (2.17) we will omit the primes in (2.16) in the sequel.

We end this section by introducing some notation that will be used throughout the paper. Let $\mathbf{p}$ and π be according to (2.1) and (2.9). Put

$$L = \max_s Np_s, \tag{2.18}$$

$$l = \min_s Np_s, \tag{2.19}$$

$$\rho = L/l. \tag{2.20}$$

The *mean*, μ_π, and the *variance*, σ_π^2, *of the population* π are

(2.21) $$\mu_\pi = \frac{1}{N} \sum_{s=1}^N a_s$$

(2.22) $$\sigma_\pi^2 = \frac{1}{N-1} \sum_{s=1}^N (a_s - \mu_\pi)^2 .$$

Furthermore, we put

(2.23) $$M = \max_s |a_s| .$$

Sometimes we want to emphasize that M depends on the population. We then use the notation

(2.24) $M(\pi) =$ the maximal absolute value of the elements in π.

3. Formulation of the main results. In this section we shall formulate the main results of this paper, which give asymptotic formulas for the inclusion probabilities and for the distribution of the sample sum. We first write some of these formulas with a remainder term and then give conditions for the remainder term to be asymptotically negligible.

Let as usual $Y_1, Y_2, \cdots, Y_N$ be a $(p_1, p_2, \cdots, p_N)$-permutation of the elements in $\pi = (a_1, a_2, \cdots, a_N)$ and let

(3.1) $$Z_n = Y_1 + Y_2 + \cdots + Y_n , \qquad n = 1, 2, \cdots, N .$$

The following function will be crucial.

DEFINITION 3.1. The function $t(y)$ (depending on $\mathbf{p}$) is defined implicitly by the relation

(3.2) $$N - y = \sum_{s=1}^N e^{-p_s t(y)} , \qquad 0 \leqq y < N .$$

The above definition is always meaningful since the function

(3.3) $$\sum_{s=1}^N e^{-p_s x} , \qquad x \geqq 0$$

decreases strictly from N to 0 as x increases from 0 to ∞ (cf. (2.2)).

Our approximation formula for the inclusion probability $\Delta(s, n)$ (see (2.11)) is

(3.4) $$\Delta(s, n) = 1 - e^{-p_s t(n)} + r_\Delta(s, n) \cdot N^{-\frac{1}{2}} .$$

Put

(3.5) $$r_\Delta(n) = \max_s |r_\Delta(s, n)| .$$

We shall show that $r_\Delta(n)$ is "small" under general conditions.

The sample sum Z_n can be expressed in the following form, where W_s is defined in (1.3),

(3.6) $$Z_n = \sum_{s=1}^N W_s a_s .$$

We have

$$EW_s = \Delta(s, n) . \tag{3.7}$$

By taking expectation in (3.6) and by using (3.7) we get the well-known formula

$$EZ_n = \sum_{s=1}^N a_s \Delta(s, n) . \tag{3.8}$$

From (3.8), (3.4) and (3.5) we get the following approximation formula for EZ_n.

$$EZ_n = \sum_{s=1}^N a_s(1 - e^{-p_s t(n)}) + r_E(n) \cdot N^{\frac{1}{2}} \tag{3.9}$$

where

$$|r_E(n)| \leqq \max_s |a_s| \cdot r_\Delta(n) . \tag{3.10}$$

We introduce the following notation

$$\mu(n) = \sum_{s=1}^N a_s(1 - e^{-p_s t(n)}) , \qquad n = 1, 2, \cdots, N . \tag{3.11}$$

Then we can write (3.9) as follows

$$EZ_n = \mu(n) + r_E(n) \cdot N^{\frac{1}{2}} . \tag{3.12}$$

To formulate approximation formulas for second moments of sample sums we need some more notation. Put

$$\xi(n) = \sum_{s=1}^N p_s e^{-p_s t(n)} , \tag{3.13}$$

$$\eta(n) = \sum_{s=1}^N p_s a_s e^{-p_s t(n)} , \tag{3.14}$$

$$\begin{aligned} \sigma(m, n) &= \sum_{s=1}^N \left(a_s - \frac{\eta(m)}{\xi(m)}\right)\left(a_s - \frac{\eta(n)}{\xi(n)}\right)(1 - e^{-p_s t(m)})e^{-p_s t(n)} , \quad && 1 \leqq m \leqq n \leqq N \\ &= \sigma(n, m) , && 1 \leqq n < m \leqq N . \end{aligned} \tag{3.15}$$

As a particular case of (3.15) we put

$$\sigma^2(n) = \sigma(n, n) = \sum_{s=1}^N \left(a_s - \frac{\sum_{r=1}^N p_r a_r e^{-p_r t(n)}}{\sum_{r=1}^N p_r e^{-p_r t(n)}}\right)^2 (1 - e^{-p_s t(n)})e^{-p_s t(n)} , \qquad n = 1, 2, \cdots, N . \tag{3.16}$$

Our approximation formulas for variance and covariance of sample sums are as follows

$$\mathrm{Cov}(Z_m, Z_n) = \sigma(m, n) + r_\sigma(m, n) \cdot \sigma(m) \cdot \sigma(n) . \tag{3.17}$$

In particular we have

$$\sigma^2(Z_n) = \sigma^2(n)(1 + r_\sigma(n, n)) . \tag{3.18}$$

We shall show that $r_\sigma(m, n)$ is "small" under general conditions.

The precise formulation of the approximation results will be given in the guise of limit theorems. The appropriate limit procedure will be obtained by considering a sequence $\{(\mathbf{p}_k, \pi_k)\}_{k=1}^\infty$ of sampling situations, where

$$\mathbf{p}_k = (p_{k1}, p_{k2}, \cdots, p_{kN_k}) \tag{3.19}$$

$$\pi_k = (a_{k1}, a_{k2}, \cdots, a_{kN_k}) . \tag{3.20}$$

We adopt the following convention throughout the paper. Whenever the index k is attached to a quantity, we mean that this quantity relates to the sampling situation $(\mathbf{p}_k, \pi_k)$.

Next we list some conditions on the sequence $\{(\mathbf{p}_k, \pi_k)\}_{k=1}^{\infty}$.

$$\lim_{k\to\infty} N_k = \infty \,, \tag{3.21}$$

$$\limsup_{k\to\infty} \frac{\max_s p_{ks}}{\min_s p_{ks}} < \infty \,, \tag{3.22}$$

$$\limsup_{k\to\infty} \frac{1}{\sigma_{\pi_k}} \max |a_{ks} - \mu_{\pi_k}| < \infty \tag{3.23}$$

where σ_π and μ_π are defined in (2.21) and (2.22).

THEOREM 3.1. *If the sequence* $(\mathbf{p}_k, \pi_k)$, $k = 1, 2, \cdots$, *satisfies* (3.21) *and* (3.22), *then the corresponding sequence* $r_\Delta^{(k)}(n)$, $k = 1, 2, \cdots$, (*see* (3.5)) *satisfies*

$$\lim_{k\to\infty} \max_{\tau_1 N_k \leq n \leq \tau_2 N_k} r_\Delta^{(k)}(n) = 0 \tag{3.24}$$

for every

$$0 < \tau_1 < \tau_2 < 1 \,. \tag{3.25}$$

REMARK. According to (3.4), (3.5), (3.9) and (3.10) this theorem yields information about the asymptotic behaviour of the inclusion probabilities and the expectation of the sample sum.

THEOREM 3.2. *If the sequence* $(\mathbf{p}_k, \pi_k)$, $k = 1, 2, \cdots$, *satisfies* (3.21), (3.22) *and* (3.23), *then the corresponding sequence* $r_\sigma^{(k)}(m, n)$, $k = 1, 2, \cdots$, (*see* (3.17)) *satisfies*

$$\lim_{k\to\infty} \max_{\tau_1 N_k \leq m, n \leq \tau_2 N_k} |r_\sigma^{(k)}(m, n)| = 0 \tag{3.26}$$

for every τ_1 *and* τ_2 *which satisfy* (3.25).

Next we shall formulate results concerning asymptotic normality of sample sums. To avoid confusion we first state what we put into the term asymptotic normality.

Let $\mathbf{U}_k = (U_1^{(k)}, U_2^{(k)}, \cdots, U_d^{(k)})$, $k = 1, 2, \cdots$, be d-dimensional random vectors and let $\boldsymbol{\mu}_k = (\mu_1^{(k)}, \mu_2^{(k)}, \cdots, \mu_d^{(k)})$ and A_k, $k = 1, 2, \cdots$, be d-dimensional vectors and $d \times d$ nonnegative symmetric matrices. By the statement that $\mathbf{U}_k$ is *asymptotically* $N(\boldsymbol{\mu}_k, A_k)$*-distributed as* $k \to \infty$, we mean that

$$\mathscr{L}(A_k^{-\frac{1}{2}}(U_1^{(k)} - \mu_1^{(k)}, \cdots, U_d^{(k)} - \mu_d^{(k)})') \Rightarrow N(0, I) \qquad \text{as} \quad k \to \infty \tag{3.27}$$

where $'$ denotes transposition and I is the identity matrix.

THEOREM 3.3. *Let the sequence* $(\mathbf{p}_k, \pi_k)$, $k = 1, 2, \cdots$, *satisfy* (3.21), (3.22) *and* (3.23). *Let* d *be a* (*fixed but arbitrary*) *natural number and let* $n_k^{(1)}$, $n_k^{(2)}$, $\cdots$, $n_k^{(d)}$ *be integers such that*

$$1 \leq n_k^{(1)} < n_k^{(2)} < \cdots < n_k^{(d)} \leq N_k \,, \qquad k = 1, 2, \cdots \,. \tag{3.28}$$

Furthermore we assume that

(3.29) $$0 < \liminf_{k\to\infty} \frac{n_k^{(1)}}{N_k} \leqq \limsup_{k\to\infty} \frac{n_k^{(d)}}{N_k} < 1$$

and

(3.30) $$\liminf_{k\to\infty} \left(\frac{n_k^{(u+1)}}{N_k} - \frac{n_k^{(u)}}{N_k}\right) > 0 \,, \qquad u = 1, 2, \cdots, d-1 \,.$$

$Z_n^{(k)}$ *is defined according to* (3.1). *Then* $(Z^{(k)}_{n_k^{(1)}}, Z^{(k)}_{n_k^{(2)}}, \cdots, Z^{(k)}_{n_k^{(d)}})$ *is asymptotically*

(3.31) (a) $N((EZ^{(k)}_{n_k^{(1)}}, \cdots, EZ^{(k)}_{n_k^{(d)}}), [\mathrm{Cov}\,(Z^{(k)}_{n_k^{(\nu)}}, Z^{(k)}_{n_k^{(\mu)}}); \nu, \mu = 1, 2, \cdots, d])$-*distributed as* $k \to \infty$.

(3.32) (b) $N((\mu_k(n_k^{(1)}), \cdots, \mu_k(n_k^{(d)})), A_k(n_k^{(1)}, \cdots, n_k^{(d)}))$-*distributed as* $k \to \infty$.

where $\mu_k(n)$ *is defined in* (3.11) *and*

(3.33) $$A(n^{(1)}, n^{(2)}, \cdots, n^{(d)}) = [\sigma(n^{(\nu)}, n^{(\mu)}); \nu, \mu = 1, 2, \cdots, d]$$

with $\sigma(m, n)$ *according to* (3.15).

The following result is only a particular case of (b) in the theorem above.

COROLLARY. *If the sequence* $(\mathbf{p}_k, \pi_k)$, $k = 1, 2, \cdots$, *satisfies* (3.21), (3.22) *and* (3.23) *and if* n_k, $k = 1, 2, \cdots$, *satisfies*

(3.34) $$0 < \liminf_{k\to\infty} \frac{n_k}{N_k} \leqq \limsup_{k\to\infty} \frac{n_k}{N_k} < 1 \,,$$

then $Z^{(k)}_{n_k}$ *is asymptotically* $N(\mu_k(n_k), \sigma_k^2(n_k))$-*distributed as* $k \to \infty$, *where* $\mu_k(n)$ *and* $\sigma_k^2(n)$ *are defined in* (3.11) *and* (3.16).

4. Some relations between successive sampling and coupon collection. Let $J_1, J_2, \cdots, M_n$ and T_ν be according to (2.13)—(2.15). For fixed s, $s = 1, 2, \cdots, N$, we let $J_1(s), J_2(s), \cdots$ be independent random variables, all having the distribution.

(4.1) $$P(J(s) = r) = p_r(s) = \frac{p_r}{1 - p_s}\,, \qquad r = 1, 2, \cdots, s-1, s+1, \cdots, N\,.$$

Let $T_\nu(s)$ be defined relative to $J_1(s), J_2(s), \cdots$ in the same way as T_ν is defined relative to $J_1, J_2, \cdots$. Our analysis of the inclusion probability $\Delta(s, n)$ (see (2.11)) will be based on the following formula.

LEMMA 4.1. *For* $s, n, = 1, 2, \cdots, N$ *we have*

(4.2) $$\Delta(s, n) = 1 - E(1 - p_s)^{T_n(s)}\,.$$

In the proof of (4.2) we shall use the following conditioning result, the proof of which is straightforward and therefore omitted.

LEMMA 4.2. *Let for* $s = 1, 2, \cdots, N$, $t = 1, 2, \cdots$ $A_t(s)$ *be the event*

$$A_t(s) = \{s \notin (J_1, J_2, \cdots, J_t)\} . \tag{4.3}$$

Then, the conditional distribution of $J_1, J_2, \cdots, J_t$ *given the event* $A_t(s)$ *is the same as the distribution of* $J_1(s), J_2(s), \cdots, J_t(s)$.

PROOF OF LEMMA 4.1. According to Lemma 2.1 and the law of total probability we have, with $A_t(s)$ as in (4.3),

$$\begin{aligned} 1 - \Delta(s, n) &= P(\{s \notin (I_1, I_2, \cdots, I_n)\}) = P(\{s \notin (J_1, J_2, \cdots, J_{T_n})\}) \\ &= \textstyle\sum_{t=n}^{\infty} P(\{T_n = t\} \cap A_t(s)) \\ &= \textstyle\sum_{t=n}^{\infty} P(A_t(s)) \cdot P(\{T_n = t\} \mid A_t(s)) . \end{aligned} \tag{4.4}$$

From Lemma 4.2 we conclude that

$$P(T_n = t \mid A_t(s)) = P(T_n(s) = t) . \tag{4.5}$$

Furthermore, as is easily realized, we have

$$P(A_t(s)) = (1 - p_s)^t . \tag{4.6}$$

From (4.4)—(4.6) we get

$$1 - \Delta(s, n) = \textstyle\sum_{t=n}^{\infty} (1 - p_s)^t P(T_n(s) = t) = E(1 - p_s)^{T_n(s)} . \tag{4.7}$$

Thus the lemma is proved.

We introduce some new random variables. With $J_1, J_2, \cdots$ as in (2.13) we put

$$\begin{aligned} H_\nu &= 1 \quad \text{if} \quad J_\nu \notin (J_1, J_2, \cdots, J_{\nu-1}) \qquad \nu = 1, 2, \cdots . \\ &= 0 \quad \text{otherwise} \end{aligned} \tag{4.8}$$

Let $\pi = (d_1, d_2 \cdots, d_N)$ and put

$$Q_n = \textstyle\sum_{\nu=1}^{n} H_\nu d_{J_\nu} , \qquad n = 1, 2, \cdots . \tag{4.9}$$

We note that the distribution of $Q_1, Q_2, \cdots$ is determined by the pair $(\mathbf{p}, \pi)$.

LEMMA 4.3. *Let, for the same pair* $(\mathbf{p}, \pi)$, $D_1, D_2, \cdots, D_N$ *be* $\mathbf{p}$-*permutation of the elements in* $\pi = (d_1, d_2, \cdots, d_N)$ *and* $Q_1, Q_2, \cdots$ *be defined by* (4.9). *Then*

$$\textstyle\sum_{\nu=1}^{n} D_\nu =_{\mathscr{L}} Q_{T_n} , \qquad n = 1, 2, \cdots, N , \tag{4.10}$$

where T_n *is defined in* (2.15).

PROOF. Formula (4.10) follows from (2.14)—(2.17), (4.8) and (4.9). We omit the details.

The papers [5] and [6] are concerned with the random variables Q and T. We list some results about them which we shall need in the sequel. The following two estimates are derived in Theorems 6 and 8 in [6].

$$E|Q_n^c|^u \leqq n^{u/2} \cdot C_u \cdot (\max_s |d_s|)^u , \qquad u \geqq 0, n = 1, 2, \cdots \tag{4.11}$$

$$E|T_n - t(n)|^u \leqq n^{u/2} \cdot C_u(\rho, n/N) , \qquad u \geqq 0, 1 \leqq n < N \tag{4.12}$$

where $t(n)$ and ρ are as in Definition 3.1 and (2.20) and where $C_u(\cdot, \cdot)$ is as in (1.4).

The following estimate concerning the function $t(y)$ in Definition 3.1 is a special case of formula (3.11) in [6],

$$t(y) \leqq \frac{y}{l}\left(1 - \frac{y}{N}\right)^{-1}, \qquad 0 \leqq y < N, \tag{4.13}$$

where l is according to (2.19).

5. Proof of Theorem 3.1. Here we shall prove Theorem 3.1. The basic formula in the proof will be (4.2).

Let $t(y; s)$ be defined relative to $p_1(s), \cdots, p_{s-1}(s), p_{s+1}(s), \cdots, p_N(s)$ (see (4.1)) in the same way as $t(y)$ is defined relative to $p_1, \cdots, p_N$ (see Definition 3.1), i.e. $t(y; s)$ is defined implicitly by the following relation

$$(N-1) - y = \sum\nolimits'^{N}_{r=1} \exp\left[-p_r(s)t(y; s)\right], \qquad 0 \leqq y \leqq N-1. \tag{5.1}$$

In (5.1), and in the sequel, the prime on the summation sign indicates that the summation excludes $r = s$.

LEMMA 5.1. *With $\Delta(s, n)$ and $t(n)$ according to* (2.11) *and* (3.2) *we have for* $s, n = 1, 2, \cdots, N$,

$$|\Delta(s, n) - (1 - e^{-p_s t(n)})| \leqq C/N^{\frac{1}{2}} \sum\nolimits_{i=1}^{4} r^{(i)}(s, n), \tag{5.2}$$

where C is an absolute constant and

$$r^{(1)}(s, n) = N^{\frac{1}{2}} \frac{p_s^2}{1 - p_s} t(n; s), \tag{5.3}$$

$$r^{(2)}(s, n) = N^{\frac{1}{2}} \frac{p_s}{1 - p_s} |ET_n(s) - t(n; s)|, \tag{5.4}$$

$$r^{(3)}(s, n) = N^{\frac{1}{2}} \cdot p_s^2 \cdot E(T_n(s) - t(n; s))^2, \tag{5.5}$$

$$r^{(4)}(s, n) = N^{\frac{1}{2}} \cdot p_s \cdot |t(n; s) - t(n)|. \tag{5.6}$$

PROOF. From formula (4.2) we get

$$\begin{aligned} |\Delta(s, n) - (1 - e^{-p_s t(n)})| &\leqq |E(1 - p_s)^{T_n(s)} - Ee^{-p_s T_n(s)}| \\ &+ |Ee^{-p_s T_n(s)} - e^{-p_s t(n; s)}| + |e^{-p_s t(n; s)} - e^{-p_s t(n)}|. \end{aligned} \tag{5.7}$$

Next we list without proof some elementary inequalities which we shall use to estimate the terms to the right in (5.7).

$$|e^{-x_1} - e^{-x_2}| \leqq |x_1 - x_2|, \qquad x_1, x_2 \geqq 0 \tag{5.8}$$

$$|(e^{-x_1} - e^{-x_2}) + e^{-x_2}(x_1 - x_2)| \leqq C(x_1 - x_2)^2, \qquad x_1, x_2 \geqq 0 \tag{5.9}$$

$$0 \leqq e^{-ax} - (1 - x)^a \leqq Cax^2 e^{-ax}/(1 - x), \qquad 0 \leqq x < 1, \quad a > 0. \tag{5.10}$$

From (5.10) we get

$$(5.11)\qquad |E(1-p_s)^{T_n(s)} - Ee^{-p_sT_n(s)}| \leqq C\frac{p_s^2}{1-p_s}ET_n(s) \leqq C\frac{p_s^2}{1-p_s}(t(n;s) + |ET_n(s) - t(n;s)|)\,.$$

From (5.9) we get after some computation

$$(5.12)\qquad |Ee^{-p_sT_n(s)} - e^{-p_st(n;s)}| \leqq |e^{-p_st(n;s)}p_s(ET_n(s) - t(n;s))| + Cp_s^2E(T_n(s) - t(n;s))^2\,.$$

Finally we get from (5.8)

$$(5.13)\qquad |e^{-p_st(n;s)} - e^{-p_st(n)}| \leqq p_s|t(n;s) - t(n)|\,.$$

The lemma now follows from (5.7), (5.11)—(5.13).

Proof of Theorem 3.1. Let $r_k^{(i)}(s, n)$, $i = 1, 2, 3, 4$, be defined by (5.3)—(5.6) relative to the sampling situation $(\mathbf{p}_k, \pi_k)$, $k = 1, 2, \cdots$. The assertion in Theorem 3.1 follows from Lemma 5.1 if we show that

$$(5.14)\qquad \lim_{k\to\infty} \max_{\tau_1 N_k \leqq n \leqq \tau_2 N_k} \max_s r_k^{(i)}(s, n) = 0\,, \qquad i = 1, 2, 3, 4\,.$$

Let $l' = \min_r(N-1)p_r(s)$ and $L' = \max_r(N-1)p_r(s)$. Note that condition (3.22) can be written in the following way, with l and L as in (2.18) and (2.19).

$$(5.15)\qquad 0 < \liminf_{k\to\infty} l_k \leqq \limsup_{k\to\infty} L_k < \infty\,.$$

It is easily checked, that if (5.15) holds, then we also have

$$(5.16)\qquad 0 < \liminf_{k\to\infty} l_k' \leqq \limsup_{k\to\infty} L_k' < \infty\,.$$

From (5.3), (2.18) and (4.13) we get

$$(5.17)\qquad r_k^{(1)}(s, n) \leqq N_k^{\frac{1}{2}} \cdot \frac{L_k^2}{N_k^2} \cdot \frac{1}{1 - L_k/N_k} \cdot \frac{n}{l_k'}\left(1 - \frac{n}{N_k - 1}\right)^{-1}\,.$$

Now (5.14) for $i = 1$ follows from (5.17), (3.21), (5.15) and (5.16). Furthermore we have

$$(5.18)\qquad r_k^{(2)}(s, n) = N_k^{\frac{1}{2}}\frac{p_{ks}}{1 - p_{ks}}|ET_n^{(k)}(s) - t_k(n;s)| \leqq \frac{L_kN_k^{\frac{1}{2}}}{N_k - L_k}|ET_n^{(k)}(s) - t_k(n;s)|\,.$$

(5.14) for $i = 2$ now follows from (5.18), (5.15), (3.21) and Theorem 5 in [6]. Next, according to (5.5) and (4.12) we have

$$(5.19)\qquad r_k^{(3)}(s, n) \leqq N_k^{\frac{1}{2}} \cdot p_{ks}^2 \cdot E(T_n(s) - t(n;s))^2 \leqq N_k^{\frac{1}{2}} \cdot \frac{L_k^2}{N_k^2} \cdot n \cdot C\left(\frac{L_k'}{l_k'}, \frac{n}{N_k - 1}\right)\,.$$

Now (5.19), (5.15) and (5.16) yield that (5.14) is true for $i = 3$. To treat $r_k^{(4)}(s, n)$ we shall need the following auxiliary result, the proof of which is quite elementary and therefore omitted.

LEMMA 5.2. *Let $\varphi_1(x)$ and $\varphi_2(x)$, $0 \leqq x < \infty$, be functions which satisfy,*

(i) *$\varphi_1(x)$ and $\varphi_2(x)$ are both strictly increasing, $x \geqq 0$.*
(ii) *$\varphi_1'(x)$ and $\varphi_2'(x)$ are both non-increasing, $x \geqq 0$.*
(iii) $|\varphi_1(x) - \varphi_2(x)| \leqq A$, $x \geqq 0$.

Let $t_1(y)$ and $t_2(y)$ be the inverse functions of $\varphi_1(x)$ and $\varphi_2(x)$. Put (for y's such that $t_1(y)$ and $t_2(y)$ both are defined)

$$t^*(y) = \max(t_1(y), t_2(y)) .$$

Then,

$$|t_1(y) - t_2(y)| \leqq \frac{A}{\min(\varphi_1'(t^*(y)), \varphi_2'(t^*(y)))} . \tag{5.20}$$

To estimate $r^{(4)}(s, n)$ we apply Lemma 5.2 with

$$\varphi_1(x) = N - \sum_{r=1}^{N} e^{-p_r x} , \qquad x \geqq 0 \tag{5.21}$$

and

$$\varphi_2(x) = (N - 1) - \sum_{r=1}^{\prime N} \exp\left(- \frac{p_r}{1 - p_s} x\right) , \qquad x \geqq 0 . \tag{5.22}$$

The corresponding inverses are then $t(y)$ and $t(y; s)$ (see (3.2) and (5.1)). It is easily checked that $\varphi_1(x)$ and $\varphi_2(x)$ satisfy the conditions (i) and (ii) in Lemma 5.2. We show that condition (iii) is also satisfied. By using the mean value theorem we get, with ρ as in (2.20),

$$\begin{aligned} |\varphi_1(x) - \varphi_2(x)| &= \left|1 - e^{-p_s x} + \sum_{r=1}^{\prime N}\left(\exp\left(- \frac{p_r}{1 - p_s} x\right) - e^{-p_r x}\right)\right| \\ &\leqq 1 + \sum_{r=1}^{\prime N}\left(\frac{p_r}{1 - p_s} - p_r\right) x e^{-p_r x} \leqq 1 + p_s x e^{-(l/N)x} \\ &\leqq 1 + \frac{Lx}{N} e^{-(lx/N)} = 1 + \rho \cdot \frac{lx}{N} \cdot e^{-(lx/N)} \leqq 1 + \rho C , \end{aligned} \tag{5.23}$$

where C is a constant. Thus, condition (iii) is also fulfilled. According to (4.13) we have

$$t(y) \leqq \frac{y}{l}\left(1 - \frac{y}{N}\right)^{-1} . \tag{5.24}$$

$$t(y; s) \leqq \frac{y}{l'}\left(1 - \frac{y}{N - 1}\right)^{-1} . \tag{5.25}$$

From (5.24) and (5.25) we get

$$t^*(y) = \max(t(y) , t(y; s)) \leqq \frac{y}{l} C\left(\frac{y}{N - 1}\right) , \tag{5.26}$$

where $C(x)$ is bounded on every interval $[0, x]$, $0 \leqq x < 1$. Furthermore, we have

$$(5.27) \qquad \varphi_1'(x) = \textstyle\sum_{r=1}^{N} p_r e^{-p_r x} \geqq e^{-(L/N)x}, \qquad x \geqq 0$$

$$(5.28) \qquad \varphi_2'(x) = \textstyle\sum'^{N}_{r=1} \frac{p_r}{1-p_s} \exp\left(\frac{p_r}{1-p_s} x\right) \geqq \exp\left(-\frac{L}{1-L/N}\cdot\frac{x}{N}\right), \qquad x \geqq 0 .$$

From (5.26)—(5.28) we get

$$(5.29) \qquad \min(\varphi_1'(t^*(y)),\ \varphi_2'(t^*(y)) \geqq \exp\left(-\frac{L}{1-L/N}\cdot\frac{y}{lN} C\left(\frac{y}{N-1}\right)\right).$$

From Lemma 5.2 and (5.29) we now get

$$(5.30) \qquad |t(n) - t(n; s)| \leqq C\left(\rho, \frac{n}{N-1}\right)$$

where $C(\cdot, \cdot)$ is as in (1.4). Now (5.14) for $i = 4$ follows from (5.6) and (5.30). Thus, Theorem 3.1 is proved.

6. Two basic conditioning results. In the sequel we shall carry out a number of conditioning computations. In this section we shall consider some simple, but basic conditioning principles for successive sampling, upon which all the following conditioning computations will be based.

Let as usual $I_1, I_2, \cdots, I_N$ be a **p**-permutation of $1, 2, \cdots, N$. $\mathscr{B}_n$ will here, and in the sequel, be the following sigma-algebra of events.

$$(6.1) \qquad \mathscr{B}_n = \mathscr{B}(I_1, I_2, \cdots, I_n), \qquad n = 0, 1, 2, \cdots, N .$$

We shall need the following concept.

DEFINITION 6.1. Let $G = (s_1, s_2, \cdots, s_{N'})$ be a subset of $(1, 2, \cdots, N)$. By a *$(p_1, p_2, \cdots, p_N)$-permutation of the items in G*, we mean the random permutation of $(s_1, s_2, \cdots, s_{N'})$ which is obtained by drawing procedure 1 (see Section 2) with draw probability proportionates $p_{s_1}, p_{s_2}, \cdots, p_{s_{N'}}$ (see the remark in Section 2).

LEMMA 6.1. *The following result holds for general draw probability proportionates.*

The conditional distribution of $I_{n+1}, I_{n+2}, \cdots, I_N$ given $\mathscr{B}_n$ (i.e. given $I_1, I_2, \cdots, I_n$) is the same as the distribution of a $(p_1, p_2, \cdots, p_N)$-permutation of the items in

$$(6.2) \qquad G(\mathscr{B}_n) = (1, 2, \cdots, N)\backslash(I_1, I_2, \cdots, I_n) .$$

The proof of Lemma 6.1 is straightforward and we omit it.

LEMMA 6.2. *The following result holds for general draw probability proportionates.*

Let $I_1, I_2, \cdots, I_N$ be a $(p_1, p_2, \cdots, p_N)$-permutation of $(1, 2, \cdots, N)$ and let $D_1, D_2, \cdots, D_N$ be the corresponding **p**-*permutation of the elements in the population*

$(d_1, d_2, \cdots, d_N)$. *Then we have*

$$E^{\mathscr{B}_n} D_{n+1} = \frac{\sum_{s=1}^{N} p_s d_s - \sum_{\nu=1}^{n} p_{I_\nu} D_\nu}{\sum_{s=1}^{N} p_s - \sum_{\nu=1}^{n} p_{I_\nu}}, \quad n = 0, 1, 2, \cdots, N-1 . \tag{6.3}$$

PROOF. The following formula, which holds for general draw probability proportionates, is easily realized,

$$ED_1 = \sum_{s=1}^{N} p_s d_s / \sum_{s=1}^{N} p_s . \tag{6.4}$$

Now, from the previous lemma and (6.4) we get, with $G(\mathscr{B}_n)$ as in (6.2),

$$E^{\mathscr{B}_n} D_{n+1} = \sum_{s \in G(\mathscr{B}_n)} p_s d_s / \sum_{s \in G(\mathscr{B}_n)} p_s = \frac{\sum_{s=1}^{N} p_s d_s - \sum_{\nu=1}^{n} p_{I_\nu} D_\nu}{\sum_{s=1}^{N} p_s - \sum_{\nu=1}^{n} p_{I_\nu}} .$$

Thus, the lemma is proved.

7. Bounds on the central moments of the sample sum. The following bounds on the central moments of the sample sum will be needed repeatedly in the sequel.

LEMMA 7.1. *Let $D_1, D_2, \cdots, D_N$ be a* **p**-*permutation of the elements in $\pi = (d_1, d_2, \cdots, d_N)$. Then we have for $u \geqq 0$, $1 \leqq n \leqq N$,*

$$E|\sum_{\nu=1}^{n} D_\nu^c|^u \leqq n^{u/2} \cdot M(\pi)^u \cdot C_u(\rho, n/N) \tag{7.1}$$

where $M(\pi)$ and ρ are according to (2.24) *and* (2.20) *and $C_u(\cdot, \cdot)$ is as in* (1.4).

PROOF. According to Lemma 4.3 we have

$$\sum_{\nu=1}^{n} D_\nu =_{\mathscr{L}} Q_{T_n} = Q_{t(n)} + R_n , \tag{7.2}$$

where

$$\begin{aligned} R_n &= \sum_{\nu=t(n)+1}^{T_n} H_\nu d_{J_\nu} && \text{if} \quad T_n > t(n) , \\ &= 0 && \text{if} \quad T_n = t(n) , \\ &= \sum_{\nu=T_n+1}^{t(n)} H_\nu d_{J_\nu} && \text{if} \quad T_n < t(n) , \end{aligned} \tag{7.3}$$

and where the H-variables are defined in (4.8). From (7.3) we get

$$|R_n| \leqq |T_n - t(n)| \cdot \max_s |d_s| = M(\pi) \cdot |T_n - t(n)| . \tag{7.4}$$

By centering at means in (7.2) we get

$$\begin{aligned} E|\sum_{\nu=1}^{n} D_\nu^c|^u &= E|Q_{t(n)}^c + R_n^{c}|^u \leqq C_u \cdot E|Q_{t(n)}^c|^u + C_u \cdot E|R_n^c|^u \\ &\leqq C_u \cdot E|Q_{t(n)}^c|^u + C_u' \cdot E|R_n|^u + C_u'(E|R_n|)^u \\ &\qquad \text{according to (7.4)} \\ &\leqq C_u E|Q_{t(n)}^c|^u + C_u' \\ &\qquad \times M(\pi)^u (E|T_n - t(n)|^u + (E|T_n - t(n)|)^u) . \end{aligned} \tag{7.5}$$

By applying the estimates (4.11)—(4.13) in (7.5) we obtain (7.1) and the lemma is proved.

8. Some auxiliary results. As stated in Section 1, our basic tool in the proof of the main results will be the technique in [4]. To apply Theorem B in [4] we have to check certain conditions. In this section we shall derive results which are preparations for the verification of these conditions.

We continue to use notation which has been introduced so far. In particular L, l, ρ and M are according to (2.18)—(2.20) and (2.23).

As usual $I_1, I_2, \cdots, I_N$ is a **p**-permutation of $1, 2, \cdots, N$ and $Y_1, Y_2, \cdots, Y_N$ the corresponding **p**-permutation of the elements in $\pi = (a_1, a_2, \cdots, a_N)$. We introduce the following new quantities

$$q_{ij} = \frac{1}{N} \sum_{s=1}^{N} (Np_s)^i a_s^{\,j} , \qquad i, j = 0, 1, 2, \cdots . \tag{8.1}$$

In (8.1) and in similar formulas we use the convention that $a^0 = 1$ even if $a = 0$.

Furthermore, we define the following random variables for $i, j = 1, 2, \cdots$

$$Y_\nu(i, j) = (Np_{I_\nu})^i a_{I_\nu}^j , \qquad \nu = 1, 2, \cdots, N , \tag{8.2}$$

$$Z_n(i, j) = \sum_{\nu=1}^{n} Y_\nu(i, j) , \qquad n = 1, 2, \cdots, N , \tag{8.3}$$

$$V_n(i, j) = q_{ij} - \frac{1}{N} Z_n(i, j) , \qquad n = 1, 2, \cdots, N . \tag{8.4}$$

Note that the definitions (8.2), (8.3), (2.10) and (3.1) are related in the following way

$$Y_\nu = Y_\nu(0, 1) \quad \text{and} \quad Z_n = Z_n(0, 1) . \tag{8.5}$$

LEMMA 8.1. *The random variables $Y_1(i, j), Y_2(i, j), \cdots, Y_N(i, j)$ constitute a* **p**-*permutation of the elements in the population*

$$\pi(i, j) = ((Np_1)^i a_1^{\,j}, (Np_2)^i a_2^{\,j}, \cdots, (Np_N)^i a_N^{\,j}) . \tag{8.6}$$

PROOF. Obvious.

LEMMA 8.2. *For $i, j = 0, 1, 2, \cdots$ we have*

$$E^{\mathscr{B}_n} Y_{n+1}(i, j) = \frac{V_n(i+1, j)}{V_n(1, 0)} , \qquad n = 0, 1, 2, \cdots, N-1. \tag{8.7}$$

PROOF. According to Lemma 6.2 and the previous lemma we have

$$E^{\mathscr{B}_n} Y_{n+1}(i, j) = \frac{\sum_{s=1}^{N} p_s (Np_s)^i a_s^{\,j} - \sum_{\nu=1}^{n} p_{I_\nu} (Np_{I_\nu})^i a_{I_\nu}^j}{1 - \sum_{\nu=1}^{n} p_{I_\nu}}$$
$$= \frac{q_{i+1,j} - (1/N) Z_n(i+1, j)}{q_{1,0} - (1/N) Z_n(1, 0)} = \frac{V_n(i+1, j)}{V_n(1, 0)} .$$

Thus the lemma is proved.

We denote the expected values of the random variables Y, Z and V by the corresponding small letters, i.e.,

$$y_\nu(i, j) = EY_\nu(i, j)\,, \tag{8.8}$$

$$z_n(i, j) = EZ_n(i, j)\,, \tag{8.9}$$

$$v_n(i, j) = EV_n(i, j)\,. \tag{8.10}$$

LEMMA 8.3. *With $t(n)$ according to Definition* 3.1 *we have for*

$$i, j = 0, 1, 2, \cdots, n = 1, 2, \cdots, N\,,$$

$$v_n(i, j) = \frac{1}{N} \sum_{s=1}^{N} (Np_s)^i a_s{}^j e^{-p_s t(n)} + r(n, i, j)/N^{\frac{1}{2}} \tag{8.11}$$

where

$$|r(n, i, j)| \leqq L^i M^j r_\Delta(n) \tag{8.12}$$

and $r_\Delta(n)$ is defined in (3.5).

PROOF. For the population $\pi(i, j)$ in (8.6) we have

$$M(\pi(i, j)) \leqq L^i M^j\,, \qquad i, j = 0, 1, 2, \cdots\,. \tag{8.13}$$

Now (8.11) and (8.12) follow by some simple computations from Lemma 8.1, (8.3), (8.4), (3.9) and (8.13).

LEMMA 8.4. *For $i, j = 0, 1, 2, \cdots, n = 1, 2, \cdots, N$ we have*

$$|V_n(i, j)| \leqq L^i M^j (1 - n/N)\,, \tag{8.14}$$

$$|v_n(i, j)| \leqq L^i M^j (1 - n/N)\,, \tag{8.15}$$

$$V_n(i, 0) \geqq l(1 - n/N)\,, \tag{8.16}$$

$$v_n(1, 0) \geqq l(1 - n/N)\,. \tag{8.17}$$

PROOF. Let $G(\mathscr{B}_n)$ be according to (6.2). From (8.1)—(8.4) we get

$$V_n(i, j) = \frac{1}{N} \sum_{s \in G(\mathscr{B}_n)} (Np_s)^i a_s{}^j\,. \tag{8.18}$$

As the number of elements in $G(\mathscr{B}_n)$ is $N - n$ we get from (8.18)

$$|V_n(i j)| \leqq \frac{1}{N} \sum_{s \in G(\mathscr{B}_n)} (Np_s)^i |a_s|^j \leqq L^i M^j \frac{N - n}{N}\,.$$

Thus (8.14) is proved. (8.15) follows from (8.14) and the relation $|v_n(i, j)| = |EV_n(i, j)| \leqq E|V_n(i, j)|$. From (8.18) we get

$$V_n(1, 0) = \frac{1}{N} \sum_{s \in G(\mathscr{B}_n)} Np_s \geqq l \cdot \frac{N - n}{N}\,, \tag{8.19}$$

and (8.16) is proved. By taking expectation in (8.16) we obtain (8.17).

LEMMA 8.5. *For $i, j = 0, 1, 2, \cdots, n = 1, 2, \cdots, N, u \geqq 0$ we have*

$$E|Z_n^c(i, j)|^u \leqq n^{u/2} L^{iu} M^{ju} C_u\left(\rho, \frac{n}{N}\right), \tag{8.20}$$

$$E|V_n^c(i, j)|^u \leqq \frac{n^{u/2}}{N^u} L^{iu} M^{ju} C_u\left(\rho, \frac{n}{N}\right) \tag{8.21}$$

where $C_u(\cdot, \cdot)$ is as in (1.4).

PROOF. (8.20) follows from Lemmas 7.1 and 8.1 and (8.13). (8.21) is a consequence of (8.20) and the following formula which is easily derived from (8.4),

$$V_n^c(i, j) = -\frac{1}{N} Z_n^c(i, j) . \tag{8.22}$$

LEMMA 8.6. *For $i, j = 0, 1, 2, \cdots, n = 0, 1, 2, \cdots, N - 1$ we have*

$$\text{(a)} \quad E^{\mathscr{B}_n} Y_{n+1}(i, j) = \frac{v_n(i + 1, j)}{v_n(1, 0)} + R^{(1)}(n, i, j) \tag{8.23}$$

where

$$E|R^{(1)}(n, i, j)|^u \leqq \left(\frac{n^{\frac{1}{2}}}{N} L^i M^j\right)^u C_u\left(\rho, \frac{n}{N}\right), \tag{8.24}$$

$$\text{(b)} \quad y_{n+1}(i, j) = \frac{v_n(i + 1, j)}{v_n(1, 0)} + r(n, i, j) \tag{8.25}$$

where

$$|r(n, i, j)|^u \leqq \left(\frac{n^{\frac{1}{2}}}{N} L^i M^j\right)^u C_u\left(\rho, \frac{n}{N}\right), \tag{8.26}$$

and $C_u(\cdot, \cdot)$ is as in (1.4).

PROOF. According to (8.7) and (8.10) we have

$$E^{\mathscr{B}_n} Y_{n+1}(i, j) = \frac{V_n(i + 1, j)}{V_n(1, 0)} = \frac{v_n(i + 1, j) + V_n^c(i + 1, j)}{v_n(1, 0) + V_n^c(1, 0)} . \tag{8.27}$$

By using the identity

$$\frac{a + \alpha}{b + \beta} = \frac{a}{b} + \frac{\alpha b - a\beta}{b(b + \beta)} \tag{8.28}$$

with

$$\begin{aligned} a &= v_n(i + 1, j) , \qquad & \alpha &= V_n^c(i + 1, j) , \\ b &= v_n(1, 0) , & \beta &= V_n^c(1, 0) \end{aligned} \tag{8.29}$$

we obtain from (8.27) the formula (8.23) with

$$R^{(1)}(n, i, j) = \frac{V_n^c(i + 1, j)v_n(1, 0) - V_n^c(1, 0)v_n(i + 1, j)}{v_n(1, 0)V_n(1, 0)} . \tag{8.30}$$

From (8.30), (8.15)—(8.17), (8.21) and the formula $|a+b|^u \leqq C_u(|a|^u + |b|^u)$ we get

$$E|R^{(1)}(n,i,j)|^u \leqq \left(\frac{1}{l^2(1-n/N)^2}\right)^u \cdot C_u(L^u \cdot E|V_n{}^c(i+1,j)|^u + L^{(i+1)u}M^{ju} \cdot E|V_n{}^c(1,0)|^u) \leqq \left(\frac{n^{\frac{1}{2}}}{N} L^i M^j\right)^u C_u\left(p, \frac{n}{N}\right).$$

Thus, (a) is proved. (b) is readily obtained from (a) by taking expectation in (8.23).

LEMMA 8.7. *For $i, j = 0, 1, 2, \cdots, n = 0, 1, \cdots, N-1$ we have*

$$(8.31)\quad E^{\mathscr{B}_n}Y^c_{n+1}(i,j) = \frac{1}{N}\left(\frac{v_n(i+1,j)}{v_n(1,0)^2} Z_n{}^c(1,0) - \frac{1}{v_n(1,0)} Z_n{}^c(i+1,j)\right) + R^{(2)}(n,i,j)$$

where

$$(8.32)\quad E|R^{(2)}(n,i,j)|^u \leqq \left(\frac{n}{N^2} L^i M^j\right)^u \cdot C_u\left(\rho, \frac{n}{N}\right), \qquad u \geqq 0$$

where $C_u(\cdot, \cdot)$ is as in (1.4).

PROOF. By using the identity

$$(8.33)\quad \frac{a+\alpha}{b+\beta} = \frac{a}{b} + \frac{\alpha b - \beta a}{b^2} - \frac{\beta(\alpha b - \beta a)}{b^2(b+\beta)}$$

with a, α, b and β as in (8.29) we get from (8.27)

$$(8.34)\quad E^{\mathscr{B}_n}Y_{n+1}(i,j) = \frac{v_n(i+1,j)}{v_n(1,0)} + \frac{V_n{}^c(i+1,j)v_n(1,0) - V_n{}^c(1,0)v_n(i+1,j)}{v_n(1,0)^2} + Q(n,i,j)$$

where

$$(8.35)\quad Q(n,i,j) = -\frac{V_n{}^c(1,0)(V_n{}^c(i+1,j)v_n(1,0) - V_n{}^c(1,0)v_n(i+1,j))}{v_n(1,0)^2 \cdot V_n(1,0)}.$$

By centering in (8.34) and by using (8.22), we obtain (8.31) with

$$(8.36)\quad R^{(2)}(n,i,j) = Q(n,i,j)^c.$$

The estimate (8.32) can now be deduced in a straightforward way from (8.36), (8.35) and the estimates in Lemmas 8.4 and 8.5. We leave the details to the reader and we regard the lemma proved.

LEMMA 8.8. *For $i_1, i_2, j_1, j_2 = 0, 1, 2, \cdots, n = 0, 1, \cdots, N-1$ we have*

$$(8.37)\quad E^{\mathscr{B}_n}Y^c_{n+1}(i_1,j_1)Y^c_{n+1}(i_2,j_2) = \frac{v_n(i_1+i_2+1, j_1+j_2)}{v_n(1,0)} - \frac{v_n(i_1+1,j_1)v_n(i_2+1,j_2)}{v_n(1,0)^2} + R(n,i_1,i_2,j_1,j_2)$$

where

$$(8.38)\qquad E|R(n, i_1, i_2, j_1, j_2)|^u \leqq \left(\frac{n^{\frac{1}{2}}}{N} L^{i_1+i_2} M^{j_1+j_2}\right)^u \cdot C_u\left(\rho, \frac{n}{N}\right), \qquad u \geqq 0$$

where $C_u(\cdot, \cdot)$ *is as in* (1.4).

PROOF. We have

$$(8.39)\qquad \begin{aligned} &E^{\mathscr{B}_n} Y^c_{n+1}(i_1, j_1) Y^c_{n+1}(i_2, j_2) \\ &= E^{\mathscr{B}_n}(Y_{n+1}(i_1, j_1) - y_{n+1}(i_1, j_1))(Y_{n+1}(i_2, j_2) - y_{n+1}(i_2, j_2)) \\ &= E^{\mathscr{B}_n} Y_{n+1}(i_1, j_1) Y_{n+1}(i_2, j_2) - y_{n+1}(i_1, j_1) E^{\mathscr{B}_n} Y_{n+1}(i_2, j_2) \\ &\quad - y_{n+1}(i_2, j_2) E^{\mathscr{B}_n} Y_{n+1}(i_1, j_1) + y_{n+1}(i_1, j_1) y_{n+1}(i_1, j_2) . \end{aligned}$$

The following formula is easily realized (cf. (8.2)),

$$(8.40)\qquad Y_{n+1}(i_1, j_1) Y_{n+1}(i_2, j_2) = Y_{n+1}(i_1 + i_2, j_1 + j_2) .$$

By using (8.40), (8.23) and (8.25) in (8.39), Lemma 8.8 is obtained after some computations using the estimates (8.24) and (8.26). We omit the details and regard the lemma proved.

LEMMA 8.9. *Let* $\pi_1 = (d_1, d_2, \cdots, d_N)$ *and* $\pi_2 = (h_1, h_2, \cdots, h_N)$ *and let* $I_1, I_2, \cdots, I_N$ *be a* **p**-*permutation of* $1, 2, \cdots, N$. *Put*

$$(8.41)\qquad D_\nu = d_{I_\nu} , \qquad \nu = 1, 2, \cdots, N$$

$$(8.42)\qquad H_\nu = h_{I_\nu} , \qquad \nu = 1, 2, \cdots, N .$$

Then we have for $1 \leqq \nu_1 \neq \nu_2 < N$,

$$(8.43)\qquad |ED^c_{\nu_1} H^c_{\nu_2}| \leqq \frac{1}{N} M(\pi_1) \cdot M(\pi_2) \cdot C\left(\rho, \frac{\max(\nu_1, \nu_2)}{N}\right),$$

where $M(\pi)$ *and* ρ *are according to* (2.24) *and* (2.20) *and* $C(\cdot, \cdot)$ *is as in* (1.4).

To prove this lemma we shall need estimates of the following probabilities. Let

$$(8.44)\qquad p(n, s) = P(I_n = s) , \qquad n, s = 1, 2, \cdots, N .$$

For $s, i_1, i_2, \cdots, i_k$ being different elements from $(1, 2, \cdots, N)$ we let

(8.45) $p(n, s; i_1, i_2, \cdots, i_k) = P(I_n{}^* = s)$, where $I_1{}^*, I_2{}^*, \cdots, I^*_{N-k}$ is a $(p_1, p_2, \cdots, p_N)$-permutation of the items in $(1, 2, \cdots, N)\backslash(i_1, i_2, \cdots, i_k)$ (see Definition 6.1), $n = 1, 2, \cdots, N - k$.

LEMMA 8.10. *There is a function* $C(\cdot, \cdot)$ *which satisfies* (1.4) *such that*

$$(8.46)\qquad \text{(a)}\quad |p(n, s) - p(n + 1, s)| \leqq (1/N) p(n, s) C(\rho, n/N)$$

$$(8.47)\qquad \text{(b)}\quad p(n, s) \leqq (1/N) C(\rho, n/N)$$

$$(8.48)\qquad \text{(c)}\quad |p(n, s; i) - p(n, s)| \leqq (1/N^2) C(\rho, n/N) .$$

Although the estimates (8.46)—(8.48) look quite innocent they will be complicated to prove. We therefore postpone the proof and we first show that Lemma 8.9 follows from Lemma 8.10.

PROOF OF LEMMA 8.9. For $1 < n < N$ we have,

$$ED_1 H_n = Ed_{I_1} E^{I_1} H_n . \tag{8.49}$$

From Lemma 6.1 we get (cf. (8.45)),

$$\begin{aligned} E^{I_1} H_n = \textstyle\sum_{s \neq I_1} p(n-1, s; I_1) h_s &= \textstyle\sum_{s=1}^{N} p(n, s) h_s \\ &+ \textstyle\sum_{s \neq I_1} (p(n-1, s) - p(n, s)) h_s \\ &+ \textstyle\sum_{s \neq I_1} (p(n-1, s; I_1) - p(n-1, s)) h_s - p(n, I_1) h_{I_1} . \end{aligned} \tag{8.50}$$

By using the estimates (8.46)—(8.48) we readily get that the last three terms in (8.50) all are dominated by $N^{-1} \max_s |h_s| \cdot C(\rho, n/N)$. Furthermore, the first sum to the right in (8.50) is EH_n. Thus we get

$$E^{I_1} H_n = EH_n + \psi(I_1) , \tag{8.51}$$

where

$$|\psi(I_1)| \leqq \frac{1}{N} \max_s |h_s| \cdot C\left(\rho, \frac{n}{N}\right) . \tag{8.52}$$

Now (8.49), (8.51) and (8.52) easily yield

$$|E(D_1 - ED_1)(H_n - EH_n)| \leqq \frac{1}{N} M(\pi_1) M(\pi_2) C\left(\rho, \frac{n}{N}\right) . \tag{8.53}$$

Thereby we have proved (8.43) for $\nu_1 = 1$. Next we show that the general case can be reduced to the case $\nu_1 = 1$. Let $1 < \nu_1 < \nu_2 < N$. Then we have

$$\begin{aligned} |ED_{\nu_1}^c H_{\nu_2}^c| = |EE^{\mathscr{B}_{\nu_1 - 1}} D_{\nu_1}^c H_{\nu_2}^c| &\leqq E|E^{\mathscr{B}_{\nu_1 - 1}} D_{\nu_1}^c H_{\nu_2}^c| \\ = E|E^{\mathscr{B}_{\nu_1 - 1}} (D_{\nu_1} &- E^{\mathscr{B}_{\nu_1 - 1}} D_{\nu_1})(H_{\nu_2} - E^{\mathscr{B}_{\nu_1 - 1}} H_{\nu_2}) \\ &+ (E^{\mathscr{B}_{\nu_1 - 1}} D_{\nu_1} - ED_{\nu_1})(E^{\mathscr{B}_{\nu_1 - 1}} H_{\nu_2} - EH_{\nu_2})| . \end{aligned} \tag{8.54}$$

Let

$$\pi_1(\mathscr{B}_n) = \{d_s : s \in G(\mathscr{B}_n)\} \quad \text{and} \quad \pi_2(\mathscr{B}_n) = \{h_s : s \in G(\mathscr{B}_n)\} , \tag{8.55}$$

where $G(\mathscr{B}_n)$ is defined in (6.2). From Lemma 6.1 and (8.53) we conclude,

$$\begin{aligned} &|E^{\mathscr{B}_{\nu_1 - 1}} (D_{\nu_1} - E^{\mathscr{B}_{\nu_1 - 1}} D_{\nu_1})(H_{\nu_2} - E^{\mathscr{B}_{\nu_1 - 1}} H_{\nu_2})| \\ &\leqq \frac{1}{N - (\nu_1 - 1)} M(\pi_1(\mathscr{B}_{\nu_1 - 1})) \cdot M(\pi_2(\mathscr{B}_{\nu_1 - 1})) \cdot C\left(\rho, \frac{\nu_2 - \nu_1}{N - (\nu_1 - 1)}\right) \\ &\leqq \frac{1}{N} M(\pi_1) \cdot M(\pi_2) \cdot C\left(\rho, \frac{\nu_2}{N}\right) . \end{aligned} \tag{8.56}$$

The following estimate is easily derived from Lemma 8.6,

$$E|E^{\mathscr{B}_{\nu - 1}} D_\nu - ED_\nu|^u \leqq \left(\frac{\nu^{\frac{1}{2}}}{N} \cdot M\right)^u \cdot C_u\left(\rho, \frac{\nu}{N}\right) , \qquad u \geqq 0 . \tag{8.57}$$

By using (8.57) and the fact $\mathscr{B}_{\nu_2} \supset \mathscr{B}_{\nu_1}$, we get

$$E|(E^{\mathscr{B}_{\nu_1-1}}D_{\nu_1} - ED_{\nu_1})(E^{\mathscr{B}_{\nu_1-1}}H_{\nu_2} - EH_{\nu_2})|$$

(8.58)
$$\leqq (E(E^{\mathscr{B}_{\nu_1-1}}D_{\nu_1} - ED_{\nu_1})^2)^{\frac{1}{2}} \cdot (E(E^{\mathscr{B}_{\nu_2-1}}H_{\nu_2} - EH_{\nu_2})^2)^{\frac{1}{2}}$$
$$\leqq \frac{\nu_2}{N^2} \cdot M(\pi_1) \cdot M(\pi_2) \cdot C\left(\rho, \frac{\nu_2}{N}\right).$$

(8.43) now follows from (8.54), (8.56) and (8.58). Thus, Lemma 8.9 is proved and we turn to the proof of Lemma 8.10.

PROOF OF LEMMA 8.10. For $i_1, i_2, \cdots, i_n$ being different numbers among $(1, 2, \cdots, N)$ we put

(8.59)
$$a(i_1, i_2, \cdots, i_n) = P(I_1 = i_1, I_2 = i_2, \cdots, I_n = i_n).$$

(2.6) yields,

(8.60)
$$a(i_1, i_2, \cdots, i_n) = \frac{p_{i_1}}{1} \cdot \frac{p_{i_2}}{1 - p_{i_1}} \cdot \cdots \cdot \frac{p_{i_n}}{1 - (p_{i_1} + \cdots + p_{i_{n-1}})}.$$

We have

$$p(n, s) - p(n+1, s)$$

(8.61)
$$= \textstyle\sum_{i_n = s} a(i_1, i_2, \cdots, i_n) - \sum_{i_{n+1} = s} a(i_1, i_2, \cdots, i_{n+1})$$
$$= \textstyle\sum_{s \notin (i_1, \cdots, i_{n-1})} a(i_1, i_2, \cdots, i_{n-1}) \dfrac{p_s}{1 - (p_{i_1} + \cdots p_{i_{n-1}})}$$
$$\times \left[1 - \textstyle\sum_{i_n \notin (i_1, \cdots, i_{n-1}, s)} \dfrac{p_{i_n}}{1 - (p_i + \cdots + p_{i_n})}\right].$$

Furthermore we have,

$$\left|1 - \textstyle\sum_{i_n \notin (i_1, \cdots, i_{n-1}, s)} \dfrac{p_{i_n}}{1 - (p_{i_1} + \cdots + p_{i_n})}\right|$$
$$= \left|1 - \textstyle\sum_{i_n \notin (i_1, \cdots, i_{n-1}, s)} \dfrac{p_{i_n}}{(1 - (p_{i_1} + \cdots + p_{i_{n-1}})}\right.$$
$$\left. - \textstyle\sum_{i_n \notin (i_1, \cdots, i_{n-1}, s)} \dfrac{p_{i_n}^2}{(1 - (p_{i_1} + \cdots + p_{i_{n-1}}))(1 - (p_{i_1} + \cdots + p_{i_n}))}\right|$$
$$\leqq \left|1 - \frac{1 - (p_{i_1} + \cdots + p_{i_{n-1}} + p_s)}{1 - (p_{i_1} + \cdots + p_{i_{n-1}})}\right|$$
$$+ \left|\textstyle\sum_{i_n \notin (i_1, \cdots, i_{n-1}, s)} \dfrac{p_{i_n}^2}{(1 - (p_{i_1} + \cdots + p_{i_{n-1}}))(1 - (p_{i_1} + \cdots + p_{i_n}))}\right|$$

(8.62)
$$\text{as} \quad 1 - (p_{i_1} + \cdots + p_{i_n}) \geqq (N - n)\frac{l}{N},$$

(8.63)
$$\leqq \frac{L}{N} \cdot \frac{N}{l(N-n)} + \left(\frac{L}{N}\right)^2 \left(\frac{N}{l(N-n)}\right)^2 (N - n) \leqq \frac{1}{N} C\left(\rho, \frac{n}{N}\right),$$

where $C(\cdot, \cdot)$ is as in (1.4). By inserting (8.63) into (8.61) we obtain

$$|p(n, s) - p(n+1, s)| \leqq \frac{1}{N} C\left(\rho, \frac{n}{N}\right) \sum_{i_n = s} a(i_1, i_2, \cdots, i_n)$$
$$= \frac{1}{N} p(n, s) C\left(\rho, \frac{n}{N}\right).$$

Thus (8.46) is proved. From (8.46) we get,

$$p(n+1, s) \leqq p(n, s)\left(1 + \frac{1}{N} C\left(\rho, \frac{n}{N}\right)\right), \qquad n = 1, 2, \cdots, N-1. \tag{8.64}$$

By iterating the inequality (8.64) we obtain,

$$p(n, s) \leqq p(1, s) \prod_{\nu=1}^{n-1}\left(1 + \frac{1}{N} C\left(\rho, \frac{\nu}{N}\right)\right) \leqq p_s \cdot \exp\left(\frac{n-1}{N} C\left(\rho, \frac{n}{N}\right)\right)$$
$$\leqq \frac{1}{N} C'\left(\rho, \frac{n}{N}\right).$$

Thereby also (8.47) is proved, and we continue with the proof of (8.48). We start by deducing the following identity, where $p(n, s; i_1, \cdots, i_n)$ is defined in (8.45),

$$\begin{aligned} & p(n, s; i_1, i_2, \cdots, i_k) - p(n, s; i_0, i_1, i_2, \cdots, i_k) \\ & = \sum_{i_{k+1} \notin (s, i_0, i_1, \cdots, i_k)} \frac{p_{i_{k+1}}}{1 - (p_{i_1} + \cdots + p_{i_k})} [p(n-1, s; i_1, \cdots, i_{k+1}) \\ & \qquad - p(n-1, s; i_0, i_1, \cdots, i_{k+1})] \\ & \qquad + \frac{p_{i_0}}{1 - (p_{i_1} + \cdots + p_{i_k})} [p(n-1, s; i_0, i_1, \cdots, i_k) \\ & \qquad - p(n, s; i_0, i_1, \cdots, i_k)]. \end{aligned} \tag{8.65}$$

From Lemma 6.1 we get,

$$\begin{aligned} & p(n, s; i_1, i_2, \cdots, i_k) \\ & = \sum_{i_{k+1} \notin (s, i_1, \cdots, i_k)} \frac{p_{i_{k+1}}}{1 - (p_{i_1} + \cdots + p_{i_k})} p(n-1, s; i_1 \cdots, i_k, i_{k+1}) \end{aligned} \tag{8.66}$$

and

$$\begin{aligned} p(n, s; i_0, i_1, \cdots, i_k) = & \sum_{i_{k+1} \notin (s, i_0, i_1, \cdots, i_k)} \frac{p_{i_{k+1}}}{1 - (p_{i_0} + p_{i_1} + \cdots + p_{i_k})} \\ & \times p(n-1, s; i_0, i_1, \cdots, i_{k+1}). \end{aligned} \tag{8.67}$$

By using the identity

$$\begin{aligned} & \frac{p_{i_{k+1}}}{1 - (p_{i_0} + p_{i_1} + \cdots + p_{i_k})} \\ & = \frac{p_{i_{k+1}}}{1 - (p_{i_1} + p_{i_2} + \cdots + p_{i_k})} \\ & \qquad + \frac{p_{i_0} \cdot p_{i_{k+1}}}{(1 - (p_{i_1} + \cdots + p_{i_k}))(1 - (p_{i_0} + \cdots + p_{i_k}))} \end{aligned} \tag{8.68}$$

we obtain from (8.67),

(8.69) $p(n, s; i_0, i_1, \cdots, i_k)$

$$= \sum_{i_{k+1}\notin(s,i_0,\cdots,i_k)} \frac{p_{i_{k+1}}}{1-(p_{i_1}+\cdots+p_{i_k})} p(n-1, s; i_0, i_1, \cdots, i_{k+1})$$

$$+ \frac{p_{i_0}}{1-(p_{i_1}+\cdots+p_{i_k})} \sum_{i_{k+1}\notin(s,i_0,i_1,\cdots,i_k)} \frac{p_{i_{k+1}}}{1-(p_{i_0}+p_{i_1}+\cdots+p_{i_k})}$$

$$\times\, p(n-1, s; i_0, i_1, \cdots, i_{k+1})$$

$$= \sum_{i_{k+1}\notin(s,i_0,\cdots,i_k)} \frac{p_{i_{k+1}}}{1-(p_{i_1}+\cdots+p_{i_k})} p(n-1, s; i_0, i_1, \cdots, i_{k+1})$$

$$+ \frac{p_{i_0}}{1-(p_{i_1}+\cdots+p_{i_k})} p(n, s; i_0, i_1, \cdots, i_k)\,.$$

By substracting (8.69) from (8.66) we get (8.65), which is thereby proved. The following inequality is a consequence of Lemma 6.1, (8.46) and (8.47).

$$(8.70) \qquad |p(n-1, s; i_0, i_1, \cdots, i_k) - p(n, s; i_0, i_1, \cdots, i_k)| \leqq \frac{1}{N^2} C\left(\rho, \frac{n}{N-k}\right),$$

where $C(\cdot, \cdot)$ is as in (1.4).

Now define for $k = 0, 1, 2, \cdots, N-1$, $n = 1, 2, \cdots, N-k$,

$$(8.71) \quad \mu(n, k) = \max_{s,i_0,i_1,\cdots,i_k} |p(n, s; i_1, i_2, \cdots, i_k) - p(n, s; i_0, i_1, i_2, \cdots, i_k)|\,.$$

By taking absolute values in (8.65) and by taking maximum over $s, i_0, i_1, \cdots, i_k$, we get from (8.71), (8.70), (8.47) and (8.62),

$$(8.72) \qquad \mu(n, k) \leqq \mu(n-1, k+1) + \frac{1}{N^3} C\left(\rho, \frac{n}{N-k}\right).$$

By iterating the inequality (8.72) we obtain

$$(8.73) \qquad \mu(n, 0) \leqq \mu(1, n-1) + \frac{1}{N^2} C\left(\rho, \frac{n}{N}\right), \qquad 1 \leqq n < N\,.$$

Furthermore, as is easily checked,

$$(8.74) \qquad \mu(1, n-1) \leqq \frac{1}{N^2} C\left(\rho, \frac{n}{N}\right).$$

Now (8.73) and (8.74) yield (8.48) and Lemma 8.10 is completely proved.

REFERENCES

[1] HÁJEK, J. (1960). Limiting distributions in simple random sampling from a finite population. *Publ. Math. Inst. Hungar. Acad. Sci.* **5** 361–374.

[2] HÁJEK, J. (1964). Asymptotic theory of rejective sampling with varying probabilities from a finite population. *Ann. Math. Statist.* **35** 1491–1523.

[3] KARLIN, S. (1968). *Total Positivity*. Stanford Univ. Press.

[4] ROSÉN, B. (1969a). Asymptotic normality of sums of random elements with values in a real separable Hilbert space. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **13** 221-255.

[5] ROSÉN, B. (1969b). Asymptotic normality in a coupon collector's problem. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **13** 256-279.

[6] ROSÉN, B. (1970). On the coupon collector's waiting time. *Ann. Math. Statist.* **41** 1952-1969.

[7] SUKHATME, P. V. and SUKHATME, B. V. (1970). *Sampling Theory of Surveys with Applications*, 2nd revised ed. Asia Publishing House, London.