LIMIT THEOREMS FOR AN EXTENDED COUPON COLLECTOR'S PROBLEM AND FOR SUCCESSIVE SUBSAMPLING WITH VARYING PROBABILITIES*

by

Pranab Kumar Sen

Department of Biostatistics
University of North Carolina, Chapel Hill

Institute of Statistics Mimeo Seires No. 1235

June 1979

LIMIT THEOREMS FOR AN EXTENDED COUPON COLLECTOR'S PROBLEM AND FOR SUCCESSIVE SUBSAMPLING WITH VARYING PROBABILITIES*

PRANAB KUMAR SEN

University of North Carolina, Chapel Hill, N.C.

ABSTRACT

Asymptotic normality as well as some weak invariance principles for bonus sums and waiting times in an extended coupon collector's problem are considered and incorporated in the study of the asymptotic distribution theory of estimators of (finite) population totals in successive sub-sampling (or multistage sampling) with varying probabilities (without replacements). Some applications of these theorems are also considered.

AMS Subject Classification Nos: 60F05, 62E20.

Key Words & Phrases: Asymptotic normality, bonus sum, extended coupon collector's problem, invariance principles, successive sub-sampling, tightness, varying probabilities, waiting times.

*Work partially supported by the National Heart, Lung and Blood Institute, Contract No. NIH-NHLBI-71-2243 from the National Institutes of Health.

1. INTRODUCTION

Let $a_1, \ldots, a_N$ be the variate values of N items of a finite population and let $p_1, \ldots, p_N$ be positive numbers such that $\sum_{s=1}^N p_s = 1$. Consider a successive sampling scheme where items are sampled one after the other (without replacement) in such a way that at each draw the probability of drawing item s is proportional to p_s if item s has not already appeared in the earlier drawals. According to the usual terminology [viz., Hájek (1964) and Sukhatme and Sukhatme (1970)], we term this *successive sampling with varying probabilities (without replacement)* (SSVPWR). Let $\Delta(s, n)$ be the probability that item s is included in a sample of size n in this SSVPWR and let $S(n) = \{s: \text{item } s \in \text{sample of size } n\}$. Then, for the estimation of the population total $\tau = \sum_{s=1}^N a_s$, the *Horvitz-Thompson estimator* is

$$\hat{\tau}_n = \sum_{s=1}^N \omega_{ns} a_s / \Delta(s, n) = \sum_{S(n)} a_s / \Delta(s, n), \tag{1.1}$$

where

$$\omega_{ns} = \begin{cases} 1, & \text{items } s \in \text{sample of size } n; \\ 0, & \text{otherwise, for } 1 \le s \le N. \end{cases} \tag{1.2}$$

The varying probability structure introduces certain complications in the study of the asymptotic distribution theory of $\hat{\tau}_n$. Rosén (1970, 1972) considered an alternative approach (through the coupon collector's problem) and provided some deeper results in this context. Consider a sequence $\{J_k, k \ge 1\}$ of independent and identically distributed random variables (i.i.d.r.v.), where

$$P\{J_k = s\} = p_s \quad \text{for} \quad s = 1, \ldots, N \quad \text{and} \quad k \ge 1. \tag{1.3}$$

Let then

$$Y_{nk} = \begin{cases} a_{J_k}/\Delta(J_k, n), & \text{if } J_k \notin \{J_1,\ldots,J_{k-1}\} \\ 0, & \text{otherwise, for } k \geq 1; \end{cases} \tag{1.4}$$

$$\nu_k = \inf\{m: \text{number of distinct } J_1,\ldots,J_m = k\},\ k \geq 1. \tag{1.5}$$

Then, Rosén (1970, 1972) has shown that

$$\hat{\tau}_n \overset{\mathcal{D}}{=} \sum_{k=1}^{N} Y_{n\nu_k} = \sum_{k=1}^{\nu_n} Y_{nk} = B_{n\nu_n}, \text{ say,} \tag{1.6}$$

where $\overset{\mathcal{D}}{=}$ stands for the equality of distributions and, for every $m \geq 1$, B_{nm} is the *bonus sum* at the mth stage in a *coupon collector's problem* with the set $\{a^*_{ns}, p_s;\ 1 \leq s \leq N\}$ where $a^*_{ns} = a_s/\Delta(s, n)$, $1 \leq s \leq N$. Thus, the asymptotic normality of (randomly stopped) bonus sums provides the same for $\hat{\tau}_n$. A similar treatment holds for many other related estimators. Recently, Sen (1969) has formulated a *martingale approach* to this problem and established some invariance principles under no extra regularity conditions.

Sub-sampling (or *multi-stage sampling*) is often adapted in practice and has a great variety of applications [viz., Cochran (1963, Ch. 10)]. Here, each of the N items of the population (called the *primary units*) is composed of a number of smaller units *(sub-units)* and a common practice is to select first a sample of n primary units and then to use subsamples of sub-units in each of the selected primary units. Suppose that the sth primary unit has M_s sub-units with variate values b_{sj}, $j = 1,\ldots,M_s$, so that

$$a_s = b_{s1} + \cdots + b_{sM_s}, \quad \text{for } s = 1,\ldots,N. \tag{1.7}$$

For each s, we conceive of a set $\{p_{sj},\ 1 \leq j \leq M_s\}$ of positive numbers (such that $\sum_{j=1}^{M_s} p_{sj} = 1$) and consider a SSVPWR scheme, where m_s (out of M_s) sub-units are chosen. Then, an analogous estimator of a_s is

$$\hat{a}_s = \sum_{j=1}^{M_s} \omega^*_{sj} b_{sj} / \Delta^*_s(j, m_s). \tag{1.8}$$

where

$$\omega^*_{sj} = \begin{cases} 1, \text{ jth sub-unit } \varepsilon \text{ sub-sample of } m_s \text{ from items,} \\ 0, \text{ otherwise, for } j = 1,\ldots,M_s \end{cases} \tag{1.9}$$

and $\Delta^*_s(j, m_s)$ is the probability that the jth sub-unit belongs to the sub-sample of m_s sub-units from items s, for $1 \le j \le M_s$ and $1 \le s \le N$. From (1.1) and (1.8), one may consider the following estimator

$$\hat{\hat{\tau}}_n = \sum_{S(n)} \hat{a}_s / \Delta(s, n) = \sum_{S(n)} \left\{ \sum_{j=1}^{M_s} b_{sj} \omega^*_{sj} / \Delta^*_s(s, m_s) \right\} / \Delta(s, n)$$

$$= \sum_{s=1}^{N} \sum_{j=1}^{M_s} b_{sj} \omega_{ns} \omega^*_{sj} / [\Delta(s, n) \Delta^*_s(j, m_s)]. \tag{1.10}$$

The procedure can be extended to the multi-stage case in a natural way.

The object of the present investigation is to study the asymptotic distribution theory of estimators of τ in successive sub-sampling (SSS)VPWR schemes. Note that for each s, analogous to (1.5) and (1.6), one can define a stopping number $\nu^{(s)}_{m_s}$ and bonus sums $B^*_{s,m}$, $m \ge 1$, such that $\hat{a}_s \overset{\mathcal{D}}{=} B^*_{s,\nu_{m_s}}(s)$, for every $s(=1,\ldots,N)$. But, the multitude of the stopping numbers $\{\nu_n, \nu^{(s)}_{m_s}, s \,\varepsilon\, S(n)\}$ introduces complications in a direct extension of the Rosén approach to SSSVPWR. On the other hand, the martingale approach of Sen (1979) can more readily be extended to subsampling schemes, and this will be systematically explored here.

In Section 2, we start with an extended coupon collector's problem and present some invariance principles relating to *bonus sums* and *waiting times;* the proofs of these theorems are considered in

Section 3. Section 4 deals with the limit theorems for SSSVPWR. The concluding section is devoted to some general remarks and specific applications.

2. INVARIANCE PRINCIPLES FOR AN EXTENDED COUPON COLLECTOR'S PROBLEM

Let us consider a sequence $\{\Omega_N, N \geq 1\}$ of coupon collector's situations

$$\Omega_N = \{(a_N(1), p_N(1))', \ldots, (a_N(N), p_N(N))'\}, N \geq 1, \quad (2.1)$$

where the $a_N(s)$ are real numbers and the $p_N(s)$ are positive numbers with $\sum_{s=1}^{N} p_N(s) = 1$, $\forall N \geq 1$. Let us now consider a triangular array $\{X_N(s), 1 \leq s \leq N; N \geq 1\}$ of row-wise independent random variables where $X_N(s)$ is defined on a probability space $(\mathcal{X}_N(s), \mathcal{A}_{Ns}, \Pi_{Ns})$, for $s = 1, \ldots, N$. [Typically, $X_N(s)$ is an estimator of $a_N(s)$ with a distribution Π_{Ns}.] Then, an extended coupon collector's situation $\{\Omega_N^*\}$ is defined by letting

$$\Omega_N^* = \{((\mathcal{X}_N(s), \mathcal{A}_{Ns}, \Pi_{Ns}), p_N(s))', 1 \leq s \leq N\}, \quad N \geq 1. \quad (2.2)$$

Also, let $\{J_{nk}, k \geq 1\}$ be a (double) sequence of (row-wise) i.i.d.r.v., where

$$P\{J_{Nk} = s\} = p_N(s), \quad \text{for } 1 \leq s \leq N \text{ and } k \geq 1. \quad (2.3)$$

Let then

$$Y_{nk}^* = \begin{cases} X_N(J_{Nk}), & \text{if } J_{Nk} \notin \{J_{Nr}, r < k\} \\ 0, & \text{otherwise, for } k \geq 1; \end{cases} \quad (2.4)$$

$$Z_{Nn}^* = \sum_{k=1}^{n} Y_{Nk}^*, \quad n \geq 1; \quad Z_{N0}^* = Y_{N0}^* = 0. \quad (2.5)$$

Also, if the $X_N(s)$ are nonnegative r.v., then Z_{Nm}^* is $\nearrow$ in n, and, we let

$$U_N^*(t) = \min\{n: Z_{Nm}^* \geq t\}, \quad t \in R^+ = [0, \infty) \quad . \quad (2.6)$$

Z^*_{Nn} is termed the *bonus sum* after n coupons in the situation Ω^*_N and $U^*_N(t)$ is the *waiting time* to obtain the bonus sum t.

Let us denote by

$$a^o_N(s) = EX_N(s),\ a^*_N(s) = EX^2_N(s) \quad \text{and} \quad \sigma^2_{Ns} = V(X_N(s)) = a^*_N(s) - \{a^o_N(s)\}^2, \tag{2.7}$$

for $s = 1,\cdots,N$. As in Rosén (1969), we assume that

$$\sup_N\{\max_{1\le s\le N} Np_N(s)\} \le M_1 < \infty , \tag{2.8}$$

$$\lim_{N\to\infty}\left\{\max_{1\le s\le N} |a^o_N(s)| \Big/ \left[\sum_{s=1}^N a^*_N(s)\right]^{\frac{1}{2}}\right\} = 0 , \tag{2.9}$$

$$\liminf_{N\to\infty}\left\{\left(\sum_{s=1}^N a^*_N(s)p_N(s)\right)\Big/\left(\sum_{s=1}^N a^*_N(s)/N\right)\right\} \ge M_2 > 0. \tag{2.10}$$

Further, without any loss of generality, we may set

$$N^{-1}\sum_{s=1}^N a^*_N(s) \sim 1 \text{ as } N\to\infty . \tag{2.11}$$

Then, we assume that for every $\varepsilon > 0$,

$$\limsup_{N\to\infty}\left\{N^{-1}\sum_{s=1}^N E[(X_N(s) - a^o_N(s))^2 I(|X_N(s) - a^o_N(s)| > \varepsilon\sqrt{N})]\right\} = 0 \tag{2.12}$$

Note that by (2.9), (2.11) and (2.12), for every $\varepsilon > 0$,

$$\limsup_{N\to\infty}\left\{N^{-1}\sum_{s=1}^N E[|X^2_N(s) - a^*_N(s)|\, I(|X^2_N(s) - a^*_N(s)| > \varepsilon N)]\right\} = 0 . \tag{2.13}$$

Further, for (2.12) or (2.13) to hold, a slightly more stringent (but, more easily verifiable) condition is that for some $\delta > 0$,

$$\sup_N\left\{N^{-1}\sum_{s=1}^N E|X_N(s) - a^o_N(s)|^{2+\delta}\right\} \le M^*_1 < \infty . \tag{2.14}$$

Let us also denote by

$$\phi^o_{Nn} = \sum_{s=1}^N a^o_N(s)[1 - e^{-np_N(s)}],\quad n \ge 0, \tag{2.15}$$

$$d_{Nn}^2 = \sum_{s=1}^N a_N^*(s) e^{-np_N(s)} [1 - e^{-np_N(s)}] - n\left(\sum_{s=1}^N a_N^o(s) p_N(s) e^{-np_N(s)}\right)^2, \ n \geq 0 , \tag{2.16}$$

$$e_{Nn}^2 = \sum_{s=1}^N \sigma_{Ns}^2 \left(1 - e^{-np_N(s)}\right)^2, \ n \geq 0, \tag{2.17}$$

$$d_{Nn}^{*2} = d_{Nn}^2 + e_{Nn}^2, \quad n \geq 0. \tag{2.18}$$

Note that by (2.11) and (2.12),

$$\lim_{N\to\infty} \sup \left\{N^{-1} e_{Nn}^2\right\} < \infty, \quad \text{for every } n. \tag{2.19}$$

Also, it follows from Sen (1979) [viz., his (2.12) through (2.14)] that if

$$0 < \lim_{N\to\infty} \inf \left(N^{-1} n\right) \leq \lim_{N\to\infty} \sup \left(N^{-1} n\right) < \infty , \tag{2.20}$$

then

$$0 < \lim_{N\to\infty} \inf n^{-1} d_{Nn}^2 \leq \lim_{N\to\infty} \sup n^{-1} d_{Nn}^2 < \infty . \tag{2.21}$$

Then, we have the following

Theorem 2.1. Under (2.8), (2.9), (2.10), (2.12) *and* (2.20), *as* $N \to \infty$, $(Z_{Nn}^* - \phi_{Nn}^o)/d_{Nn}^*$ *has asymptotically a standard normal distribution.*

Actually, Theorem 2.1 can be extended to an invariance principle for $\{Z_{Nn}^* - \phi_{Nn}^o\}$ as follows. For an arbitrary T $(0 < T < \infty)$, let $A = [0, T]$ and for every N, consider a sample process $W_N = \{W_N(x), x \in A\}$ by letting

$$W_N(x) = N^{-\frac{1}{2}}\left\{Z_{N[Nx]}^* - \phi_{N[Nx]}^o\right\}, \quad x \in A, \tag{2.22}$$

(where $[s]$ denotes the largest integer $\leq s$), so that W_N belongs to the space $D[A]$, endowed with the Skorokhod J_1-topology.

Theorem 2.2. Under (2.8), (2.9), (2.10) *and* (2.12), *the finite dimensional distributions (f.d.d.) of* $\{W_N\}$ *are all asymptotically Gaussian and* $\{W_N\}$ *is tight.*

Let us now consider the case of the waiting time $\{U_N^*(t)\}$ in (2.6). Note that, by definition, for all $x, t > 0$,

$$P\{U_N^*(t) > x\} = P\{Z^*_{N[x]} < t\} \tag{2.23}$$

and hence, Theorem 2.1 yields the asymptotic normality of the standardized form of $\{U_N^*(t)\}$. Similar results hold for the f.d.d.'s. Further, $U_N^*(t)$ is monotone in t, and hence, the tightness part also follows by some routine steps. Hence, in Section 3, we proceed to prove only Theorems 2.1 and 2.2.

3. PROOFS OF THEOREMS 2.1 AND 2.2

Before we proceed to prove these theorems, we consider some preliminary results which will be needed in the sequel. Let

$$\phi_{Nn}^* = \sum_{s=1}^N X_N(s)\left(1 - e^{-np_N(s)}\right), \quad n \geq 0; \tag{3.1}$$

$$c_{Nn}^2 = \sum_{s=1}^N X_N^2(s) e^{-np_N(s)}\left(1 - e^{-np_N(s)}\right)$$

$$- n\left(\sum_{s=1}^N X_N(s) p_N(s) e^{-np_N(s)}\right)^2, \quad n \geq 0. \tag{3.2}$$

Then, we have the following

Lemma 3.1. Under the hypothesis of Theorem 2.1,

$$c_{Nn}^2/d_{Nn}^2 \to 1, \textit{ in probability, as } N \to \infty . \tag{3.3}$$

<u>*Proof:*</u> By virtue of (2.16) (2.20), (2.21) and (3.2), it suffices to show that

$$\sum_{s=1}^N (X_N(s) - a_N^o(s)) p_N(s) e^{-np_N(s)} \xrightarrow{p} 0, \text{ as } N \to \infty , \tag{3.4}$$

$$N^{-1}\sum_{s=1}^N (X_N^2(s) - a_N^*(s)) e^{-np_N(s)}\left(1 - e^{-np_N(s)}\right) \xrightarrow{p} 0, \text{ as } N \to \infty. \tag{3.5}$$

We shall only prove (3.5) as the proof of (3.4) follows on parallel lines. Let

$$\xi_{N_s} = N^{-1}(X_N^2(s) - a_N^*(s))e^{-np_N(s)}\left(1 - e^{-np_N(s)}\right), \quad s = 1,\dots,N. \tag{3.6}$$

Then, by (2.7), $E\xi_{N_s} = 0$, $1 \le s \le N$, while, by (2.12) [or (2.13)], for every $\varepsilon > 0$,

$$\sum_{s=1}^N E\left\{|\xi_{N_s}| I(|\xi_{N_s}| > \varepsilon)\right\} \longrightarrow 0, \text{ as } N \to \infty, \tag{3.7}$$

and hence,

$$\sum_{s=1}^N E\left\{\xi_{N_s} I(|\xi_{N_s}| \le \varepsilon)\right\} \longrightarrow 0, \text{ as } N \to \infty. \tag{3.8}$$

Further, (3.7) insures that for every $\varepsilon > 0$,

$$\sum_{s=1}^N P\left\{|\xi_{Ns}| > \varepsilon\right\} \longrightarrow 0 \text{ as } N \to \infty. \tag{3.9}$$

Finally, using (2.11) (at the ultimate step), for every $\varepsilon > 0$,

$$\sum_{s=1}^N \left[E\left\{\xi_{Ns}^2 I(|\xi_{Ns}| \le \varepsilon)\right\} - \left(E\left\{\xi_{Ns} I(|\xi_{Ns}| \le \varepsilon)\right\}\right)^2\right]$$

$$\le \sum_{s=1}^N E\left\{\xi_{Ns}^2 I(|\xi_{Ns}| \le \varepsilon)\right\} \le \varepsilon \sum_{s=1}^N E\left\{|\xi_{Ns}| I(|\xi_{Ns}| \le \varepsilon)\right\}$$

$$\le \varepsilon \sum_{s=1}^N E|\xi_{Ns}| \le 2\varepsilon N^{-1} \sum_{s=1}^N a_N^*(s) e^{-np_N(s)}\left(1 - p^{-np_N(s)}\right)$$

$$\le \frac{1}{2}\varepsilon N^{-1} \sum_{s=1}^N a_N^*(s) \sim \frac{1}{2}\varepsilon, \tag{3.10}$$

and this can be made arbitrarily small by choosing ε so. Hence, by an appeal to the Degenerate Convergence Theorem [viz., Loéve (1963, p. 317)], we conclude that $\sum_{s=1}^N \xi_{Ns} \xrightarrow{P} 0$ as $N \to \infty$, which proves (3.5). Q.E.D.

Let $\underset{\sim}{X}_N = (X_N(1),\dots,X_N(N))'$ and F_N be the σ-field generated by $\underset{\sim}{X}_N$.

Lemma 3.2. Under the hypothesis of Theorem 2.1, *the conditional distribution of* $(Z_{Nn}^* - \phi_{Nn}^*)/c_{Nn}$ (given F_N) *is asymptotically (in probability) normal with mean* 0 *and unit variance.*

Proof: By (2.4), (2.5) and (3.1), given F_N, we have a coupon collector's situation $\{(X_N(s),\ p_N(s)),\ 1 \le s \le N\}$ and $(Z^*_{Nn} - \phi^*_{Nn})/c_{Nn}$ corresponds to the standardized form of the bonus sum. The martingale approach developed in the proof of Theorem 3.1 of Sen (1979) works out equally well (given F_N) when the $a_N(s)$ are replaced by $X_N(s)$, while Lemma 3.1 insures that $n^{-1}C_{Nn}$ is bounded away from 0 and ∞, in probability. Hence, the proof follows directly along the lines of the proof of Theorem 3.1 of Sen (1979) and therefore the details are omitted.

By virtue of Lemma 3.1 and 3.2, we have for every real t,

$$E\left\{\exp(it(Z^*_{Nn} - \phi^*_{Nn})/d_{Nn} \mid F_N\right\} = \exp(-\tfrac{1}{2}t^2) + o_p(1), \quad \text{as} \quad N \to \infty. \tag{3.11}$$

Lemma 3.3. Under the hypothesis of Theorem 2.1, $N^{-\frac{1}{2}}(\phi^*_{Nn} - \phi^o_{Nn})$ *has asymptotically a normal distribution.*

Proof: By (2.15) and (3.1), we obtain that

$$N^{-\frac{1}{2}}\left(\phi^*_{Nn} - \phi^o_{Nn}\right) = N^{-\frac{1}{2}}\sum_{s=1}^{N}\left\{X_N(s) - a^o_N(s)\right\}\left\{1 - e^{-np_N(s)}\right\} = \sum_{s=1}^{N}\xi^*_{Nn}, \text{ say,} \tag{3.12}$$

where the ξ^*_{Ns} are independent rv's with $E\xi^*_{Ns} = 0$, $1 \le s \le N$. Also, by (2.12),

$$\sum_{s=1}^{N} E\left\{\xi^{*2}_{Ns} I(|\xi^*_{Ns}| > \varepsilon)\right\} \to 0, \text{ as } N \to \infty \ (\forall\ \varepsilon > 0). \tag{3.13}$$

Now, (3.13) insures that for every $\varepsilon > 0$, as $N \to \infty$,

$$\sum_{s=1}^{N} P\left\{|\xi^*_{Ns}| > \varepsilon\right\} \longrightarrow 0 \quad \text{and} \quad \sum_{s=1}^{N} E\left\{|\xi^*_{Ns}| I(|\xi^*_{Ns}| > \varepsilon)\right\}| \to 0, \tag{3.14}$$

and hence,

$$\left|\sum_{s=1}^{N} E\left\{\xi^*_{Ns} I(|\xi^*_{Ns}| < \varepsilon)\right\}\right| = \left|\sum_{s=1}^{N} E\left\{\xi^*_{Ns} I(|(\xi^*_{Ns}| > \varepsilon)\right\}\right|$$

$$\le \sum_{s=1}^{N} E\left\{|\xi^*_{Ns}| I(|\xi^*_{Ns}| > \varepsilon)\right\} \longrightarrow 0 \text{ as } N \to \infty. \tag{3.15}$$

Similarly,

$$\sum_{s=1}^{N}\left[E\left\{\xi^*_{Ns}I(|\xi^*_{Ns}|\le\varepsilon)\right\}\right]^2 \le \sum_{s=1}^{N}\left[E\left\{|\xi^*_{Ns}|I(|\xi^*_{Ns}|>\varepsilon)\right\}\right]^2$$

$$\le \sum_{s=1}^{N}E\left\{\xi^{*2}_{Ns}\right\}P\left\{|\xi^*_{Ns}|>\varepsilon\right\} \le \left\{\max_{1\le s\le N} P\left\{|\xi^*_{Ns}|>\varepsilon\right\}\right\}\sum_{s=1}^{N}E\left\{\xi^{*2}_{Ns}\right\}$$

$$\le \left\{\sum_{s=1}^{N}P\left\{|\xi^*_{Ns}|>\varepsilon\right\}\right\}N^{-1}e^2_{Nn} \longrightarrow 0, \text{ by (2.19) and (3.14).} \qquad (3.16)$$

Finally, by (2.17), (2.19), (3.2) and (3.3),

$$\left[\sum_{s=1}^{N}E\left\{\xi^{*2}_{Ns}I(|\xi^*_{Ns}|\le\varepsilon)\right\} - N^{-1}e^2_{Nn}\right] \longrightarrow 0, \text{ as } N\to\infty . \qquad (3.17)$$

Hence, by an appeal to the Normal Convergence Theorem [viz., Loéve (1963, p. 316)], the desired result follows. Q.E.D.

We are now in a position to prove Theorems 2.1 and 2.2. Note that for any real t,

$$E\left\{\exp\left[itN^{-\frac{1}{2}}\left\{Z^*_{Nn}-\phi^o_{Nn}\right\}\right]\right\}$$

$$= E\left\{\left[\exp\left[itN^{-\frac{1}{2}}\left(\phi^*_{Nn}-\phi^o_{Nn}\right)\right]\right]E\left[\exp\left[itN^{-\frac{1}{2}}\left(Z^*_{Nn}-\phi^*_{Nn}\right)\right]\Big|\mathcal{F}_N\right]\right\} . \quad (3.18)$$

Hence, Theorem 2.1 follows directly from (3.11), Lemma 3.3 and (3.18) along with (2.18) and (2.19) - (2.21).

Let us proceed on to the proof of Theorem 2.2. For any $m(\ge 1)$ and $\{n_{1N},\dots,n_{mN}\}$ satisfying (2.21), given $\mathcal{F}_N$, the conditional distribution of $\{N^{-\frac{1}{2}}(Z_{Nn_{jN}}-\phi^*_{Nn_{jN}}),\ 1\le j\le m\}$ can be shown to be asymptotically (in probability) multinormal by using the proof of Lemma 3.2 and the results in Section 4 of Sen (1979). Hence, to study the convergence of f.d.d.'s of $\{W_n\}$, it remains to study the asymptotic distribution of $\{N^{-\frac{1}{2}}(\phi_{Nn_{jN}}-\phi^o_{Nn_{jN}}),\ 1\le j\le m\}$. Note that for any $\underset{\sim}{\lambda}=(\lambda_1,\dots,\lambda_m)(\ne \underset{\sim}{0})$, by (2.15) and (3.1),

$$\sum_{j=1}^{m}\lambda_j N^{-\frac{1}{2}}\left(\phi^*_{Nn_{jN}} - \phi^o_{Nn_{jN}}\right)$$

$$= N^{-\frac{1}{2}}\sum_{s=1}^{N}\left\{X_N(s) - a_N^o(s)\right\}\left\{\sum_{j=1}^{m}\lambda_j\left[1 - e^{-n_{jN}p_N(s)}\right]\right\}, \qquad (3.19)$$

and hence, the proof of Lemma 3.3 may be virtually repeated to establish the asymptotic normality of the left-hand side of (3.19). This yields the asymptotic multinormality of $\{N^{-\frac{1}{2}}(\phi^*_{Nn_{jN}} - \phi^o_{Nn_{jN}}),\ 1 \le j \le m\}$ and by a direct extension of (3.18) to the vector case, the desired multinormality of $\{N^{-\frac{1}{2}}(Z^*_{Nn_{jN}} - \phi^o_{Nn_{jN}}),\ 1 \le j \le m\}$ follows. Thus, it remains only to study the tightness of $\{W_N\}$. Towards this, we write

$$W_N(s) = N^{-\frac{1}{2}}\left\{Z^*_{N[Nx]} - \phi^*_{N[Nx]}\right\} + N^{-\frac{1}{2}}\left\{\phi^*_{N[Nx]} - \phi^o_{N[Nx]}\right\}$$
$$= W_N^*(x) + W_N^o(x), \text{ say, for } x \in A. \qquad (3.20)$$

Also, we define the *modulus of continuity* $\omega_\delta(x)$, $\delta > 0$, by

$$\omega_\delta(x) = \sup\{|x(t) - x(s)|:\ |t - s| \le \delta;\ s, t \in A\}. \qquad (3.21)$$

Then, using our Lemma 3.1 and adapting the conditional probability law, given F_N, we obtain on proceeding as in Section 4 of Sen (1979) that for every $\varepsilon > 0$ and $\eta > 0$, there exist a $\delta > 0$, a positive integer N_0 and a sequence $\{E_{(N)} \subset E^N\}$ of subsets of the sample spaces of $\{\underset{\sim}{X}_N\}$ such that for $N \ge N_0$,

$$P\{\omega_\delta(W_N^*) > \varepsilon \mid F_N\} < \tfrac{1}{2}\eta,\ \forall\ \underset{\sim}{X}_N \in E_{(N)}, \qquad (3.22)$$

$$P\{\underset{\sim}{X}_N \notin E_{(N)}\} < \tfrac{1}{2}\eta. \qquad (3.23)$$

Then, by (3.22) and (3.23), for $N \ge N_0$,

$$P\{\omega_\delta(W_N^*) > \varepsilon\} < \tfrac{1}{2}\eta(1 - \tfrac{1}{2}\eta) + \tfrac{1}{2}\eta < \eta. \qquad (3.24)$$

Thus, it remains only to show that

$$P\{\omega_\delta(W_N^o) > \varepsilon\} < \eta, \quad \forall\ N \ge N_0. \qquad (3.25)$$

For this, we let

$$\widetilde{W}_N^o(x) = N^{-\frac{1}{2}}\left\{\phi^*_{N(Nx)} - \phi^o_{N(Nx)}\right\}, \; x \in A, \tag{3.26}$$

where we note that (2.15) and (3.1) are properly defined for any real n, not necessarily an integer. Then, by (3.20) and (3.26),

$$\sup\left\{|W_N^o(x) - \widetilde{W}_N^o(x)| : x \in A\right\}$$

$$= \sup\left\{N^{-\frac{1}{2}}\left|\sum_{s=1}^N \left(X_N(s) - a_N^o(s)\right)\left(e^{-NxP_N(s)} - e^{-[Nx]P_n(s)}\right)\right| : x \in A\right\}$$

$$\leq \sup_{0\leq u\leq 1}\left\{N^{-\frac{1}{2}}\left[\max_{1\leq s\leq N}\left(1-e^{-up_N(s)}\right)\right]\right\}\cdot\sup_{x\in A}\left\{N^{-1}\sum_{s=1}^N |X_N(s) - a_N^o(s)|e^{-[Nx]p_N(s)}\right\}$$

$$\leq \left(M_1 N^{-\frac{1}{2}}\right)\left(N^{-1}\sum_{s=1}^N |X_N(s) - a_N^o(s)|\right), \tag{3.27}$$

as $1 - e^{-\alpha} \leq \alpha$, $\forall\ \alpha \geq 0$ and by (2.8), $p_N(s) \leq N^{-1}M_1$, $\forall\ 1 \leq s \leq N$. On the other hand, repeating the proof of Lemma 3.1, it can be shown that $N^{-1}\sum_{s=1}^N |X_N(s) - a_N^o(s)|$ is $0_p(1)$ (as $N \to \infty$), so that the right-hand side of (3.27) is $0_p(N^{-\frac{1}{2}})$. Hence, to prove (3.25), it suffices to show that

$$P\{\omega_\delta(\widetilde{W}_N^o) > \varepsilon\} < \eta, \; \forall\ N \geq N_0. \tag{3.28}$$

Note that by definition of $\widetilde{W}_N^o$ in (3.26) and the fact that $1 - e^{-\alpha} \leq \alpha$, $\forall\ \alpha \geq 0$, we have for every $y \geq x$,

$$E[\widetilde{W}_N^o(y) - \widetilde{W}_N^o(x)]^2$$

$$= N^{-1}\sum_{s=1}^N \sigma_{Ns}^2 e^{-2Nxp_N(s)}\left(1 - e^{-N(y-x)p_N(s)}\right)^2$$

$$\leq (y - x)^2\left\{\max_{1\leq s\leq N} Np_N(s)\right\}^2\left\{N^{-1}\sum_{s=1}^N \sigma_{Ns}^2\right\}$$

$$\leq M_1^2(y - x)^2\left\{N^{-1}\sum_{s=1}^N a_N^*(s)\right\} \text{ [by (2.7) - (2.8)]}$$

$$\sim M_1(y - x)^2, \text{ by (2.11).} \tag{3.29}$$

Therefore, (3.28) follows from (3.29) and Theorem 2.3 of Billingsley (1968, p. 95). Hence, the proof of Theorem 2.2 is complete.

4. LIMIT THEOREMS FOR SSSVPWR

We are primarily interested here in the limiting distributions of $\hat{\hat{\tau}}_n$ in (1.10) and other related estimators. In the asymptotic situation (where we let $N \to \infty$), we assume that n, the number of primary units in the sample, also increases satisfying (2.10), while m_s, $s \in S(n)$ may or may not be large. In this context, we may also allow the sampling scheme for the sub-units to be rather arbitrary (not necessarily a SSVPWR), while we assume that the primary units are sampled in accordance to a SSVPWR.

To incorporate the asymptotic situation, we introduce the following modifications of the notations introduced in Section 1. For every N, we denote by $\{(a_N(1), p_N(1)),\ldots,(a_N(N), p_N(N))\}$ the scheme relating to the SSVPWR of the primary units. Similarly, we rewrite (1.7) as

$$a_N(s) = b_{Ns1}+\cdots+b_{NsM_s}, \quad s=1,\ldots,N \tag{4.1}$$

(where the M_s may also depend on N) and in (1.1) replacing the p_s by $p_N(s)$, we define $\Delta_N(s, n)$ and $S_N(n)$ in an analogous way. Based on the sub-sample of m_s sub-units from the sth primary unit, let $\tilde{a}_N(s)$ be a suitable estimator of $a_N(s)$ [not necessarily the one in (1.8)], for $s \in S_N(n)$, and consider the estimator (of $\tau_N = \sum_{s=1}^N a_N(s)$)

$$\tau^*_{Nn} = \sum_{S_N(n)} \tilde{a}_N(s)/\Delta_N(s, n) . \tag{4.2}$$

We intend to study the limiting behavior of τ^*_{Nn} [when $N \to \infty$ and n satisfies (2.20)]. We let

$$a^o_N(s) = E\tilde{a}_N(s), \quad \sigma^2_{Ns} = V(\tilde{a}_N(s)), \quad a^*_N(s) = a^o_N(s) + \sigma^2_{Ns}, \tag{4.3}$$

for $s=1,\ldots,N$, and assume that (2.8) - (2.10) hold. In addition, as in Rosén (1972), we assume that

$$\limsup_{N\to\infty}\left\{\left[\max_{1\le s\le N} p_N(s)\right] \Big/ \left[\min_{1\le s\le N} p_N(s)\right]\right\} < \infty \ . \tag{4.4}$$

Also, for every N, we consider a nondecreasing function $t_N = \{t_N(x):$ $0 \le x < \infty\}$ by letting

$$N - x = \sum_{s=1}^{N} e^{-p_N(s)t_N(x)}, \quad 0 \le x < \infty \ . \tag{4.5}$$

Let then $\tau_N^o = \sum_{s=1}^{N} a_N^o(s)$ and let

$$\delta_{Nn}^2 = \sum_{s=1}^{N} [a_N^o(s)]^2 e^{-p_N(s)t_N(n)} \Big/ \left(1 - e^{-p_N(s)t_N(n)}\right) + \sum_{s=1}^{N} \sigma_{Ns}^2 \Big/ \left(1 - e^{-p_N(s)t_N(n)}\right)$$
$$- t_N(n)\left[\sum_{s=1}^{N} a_N^o(s) p_N(s) e^{-p_N(s)t_N(n)} \quad \left(1 - e^{-p_N(s)t_N(n)}\right)\right]^2, \tag{4.6}$$

for $n \ge N$. Finally, we assume that for every $\varepsilon > 0$,

$$\overline{\lim}\left\{\max_{1\le s\le N} E([\tilde{a}_N(s) - a_N^o(s)]^2 I(|\tilde{a}_N(s) - a_N^o(s)| > \varepsilon\sqrt{N}))\right\} = 0. \tag{4.7}$$

Then, we have the following

Theorem 4.1. Under (2.8) - (2.10), (2.20), (4.4) *and* (4.7),

$$(\tau_{Nn}^* - \tau_N^o)/\delta_{Nn} \xrightarrow[\mathcal{D}]{} N(0, 1) \ . \tag{4.8}$$

If, in addition, $N^{-\frac{1}{2}}(\tau_N - \tau_N^o) \longrightarrow 0$, *then* $(\tau_{Nn}^* - \tau_N)/\delta_{Nn} \xrightarrow[\mathcal{D}]{} N(0, 1)$.

Before we proceed to prove the theorem, we note that by (4.5),

$$\frac{\partial t_N(x)}{\partial x} = \left\{\sum_{s=1}^{N} p_N(s) e^{-p_N(s)t_N(x)}\right\}^{-1} \ge 1, \tag{4.9}$$

$$\frac{\partial^2 t_N(x)}{\partial x^2} = \left\{\sum_{s=1}^{N} p_N^2(s) e^{-p_N(s)t_N(s)}\right\} \Big/ \left\{\sum_{s=1}^{N} p_N(s) e^{-p_N(s)t_N(x)}\right\}^3 > 0, \tag{4.10}$$

for every $0 \le x < \infty$, so that for every N, $t_N(x)$ is convex in $x \in [0, \infty)$. Also, define $\{\nu_{Nk}, k \ge 1\}$ as in (1.5), where for the J_s, in (1.3), we replace p_s by $p_N(s)$, $1 \le s \le N$. Then, if we let

$$g_{Nn}^2 = \left(\sum_{s=1}^N e^{-p_N(s)t_N(n)}\left(1 - e^{-p_N(s)t_N(n)}\right)\right) \Big/ \left(\sum_{s=1}^N p_N(s) e^{-p_N(s)t_N(n)}\right)^2 - t_N(n), \tag{4.11}$$

for $n \le N$, by Theorem 4 of Rosén (1970), under the hypothesis of Theorem 4.1 [namely (2.8), (2.20) and (4.4)],

$$(\nu_{Nn} - t_N(n))/g_{Nn} \xrightarrow[\mathcal{D}]{} N(0, 1). \tag{4.12}$$

Since, $t_N(n) \ge n$, $\forall\, n \ge 1$ while by (4.11), $g_{Nn}^2 = O(N)$, we obtain that

$$\nu_{Nn}/t_N(n) \xrightarrow{P} 1, \quad \text{as} \quad N \to \infty \quad \text{when (2.20) holds.} \tag{4.13}$$

Finally, it follows from Rosén (1972) that under the hypothesis of Theorem 4.1,

$$N^{\frac{1}{2}}\left[\max_{1 \le s \le N} \left|\Delta_N(s, n) - (1 - e^{-p_N(s)t_N(n)})\right|\right] \longrightarrow 0, \quad \text{as} \quad N \to \infty, \tag{4.14}$$

while, under (2.8), (2.20) and (4.4),

$$0 < \underline{\lim}\left\{\min_{1 \le s \le N} p_N(s)t_N(n)\right\} \le \overline{\lim}\left\{\min_{1 \le s \le N} p_N(s)t_N(n)\right\} < \infty, \tag{4.15}$$

so that by (4.14) and (4.15), under (2.8), (2.20) and (4.4),

$$\liminf_{N \to \infty}\left\{\min_{1 \le s \le N} \Delta_N(s, n)\right\} > 0. \tag{4.16}$$

Note that (4.7) and (4.16) insure that for every $\varepsilon > 0$,

$$\overline{\lim}\left\{\max_{1 \le s \le N} E([\tilde{a}_N(s) - a_N^o(s)]^2 I(|\tilde{a}_N(s) - a_N^o(s)| > \varepsilon\sqrt{N}\, \Delta_N(s, n)) \big/ \Delta_N^2(s, n)\right\} = 0. \tag{4.17}$$

With these preliminary results, let us now consider the proof of Theorem 4.1. We now write [parallel to (2.4)], for a given n, satisfying (2.20),

$$X_N(s) = X_N^{(n)}(s) = \tilde{a}_N(s)/\Delta_N(s, n), \quad 1 \le s \le N. \tag{4.18}$$

Then, by virtue of (4.2), $\tau^*_{Nn} = \sum_{S_N(n)} X_N^{(n)}(s)$ and parallel to (1.6), we claim that

$$\tau^*_{Nn} \overset{\mathcal{D}}{=} Z^{(n)}_{N\nu_{Nn}} , \tag{4.19}$$

where the $Z^{(n)}_{Nk}$ are defined by (2.4) - (2.5) [$X_N(s)$ being replaced by $X_N^{(n)}(s)$, $1 \le s \le N$] and ν_{Nn} is defined as in above. Hence, to prove (4.8), it suffices to show that

$$(Z^{(n)}_{N\nu_{Nn}} - \tau^o_N)/\delta_{Nn} \xrightarrow{\mathcal{D}} N(0, 1) . \tag{4.20}$$

At this stage, we use (4.13), (4.17) and Theorem 2.1, and conclude that for $\{n^*\}$ satisfying (2.20),

$$(Z^{(n)}_{Nn^*} - \phi^{(n)}_{Nn^*})/d^{(n)}_{Nn^*} \xrightarrow{\mathcal{D}} N(0, 1) , \tag{4.21}$$

where [by (2.15), (2.16) and (4.18),

$$\phi^{(n)}_{Nn^*} = \sum_{s=1}^N a^o_N(s) \left[1 - e^{-n^* p_N(s)}\right] / \Delta_N(s, n) , \tag{4.22}$$

$$\left(d^{(n)}_{Nn^*}\right)^2 = \sum_{s=1}^N \left\{ \left[a^o_N(s)\right]^2 + \sigma^2_{Ns} \right\} e^{-n^* p_N(s)} \left[1 - e^{-n^* p_N(s)}\right] / \Delta^2_N(s, n)$$
$$- n^* \left[\sum_{s=1}^N a^o_N(s) p_N(s) e^{-n^* p_N(s)} / \Delta_N(s, n)\right]^2 + \sum_{s=1}^N \sigma^2_{Ns} \left[1 - e^{-n^* p_N(s)}\right] / \Delta^2_N(s, n) . \tag{4.23}$$

By (4.6), (4.14), (4.22) and (4.23), it follows by some standard steps that

$$\phi^{(n)}_{Nt_N(n)} = \tau^o_N + o(N^{\frac{1}{2}}), \quad d^{(n)}_{Nt_N(n)}/\delta_{Nn} \longrightarrow 1, \quad \text{as } N \to \infty. \tag{4.24}$$

Hence, by (4.13), (4.21) and (4.24), to prove (4.20), it suffices to show that for every $\varepsilon > 0$ and $\eta > 0$, there exist a $\delta > 0$ and an N_0, such that

$$P\left\{ \max_{m: |m - t_N(n)| < \delta N} N^{-\frac{1}{2}} \left| Z^{(n)}_{Nm} - \phi^{(n)}_{Nm} - Z^{(n)}_{Nt_N(n)} + \phi^{(n)}_{Nt_N(n)} \right| > \varepsilon \right\} < \eta, \tag{4.25}$$

for every $N \ge N_0$, and this follows directly from the tightness part of Theorem 2.2. Since $\delta_{Nn} = 0(N^{\frac{1}{2}})$, $N^{-\frac{1}{2}}(\tau_N - \tau^o_N) \longrightarrow 0$ insures that $(\tau_N - \tau^o_N)/d^{(n)}_{Nt_N(n)} \longrightarrow 0$ as $N \to \infty$. Hence, the proof of the theorem is complete.

Next, we proceed to extend Theorem 4.1 to an invariance principle for $\{\tau^*_{Nn}\}$. For an arbitrary $c(0 < c < 1)$, let $I_c = [c, 1]$ and, for every N, consider a sample process $\zeta_N = \{\zeta_N(t), t \in I_c\}$ by letting

$$\zeta_N(t) = N^{-\frac{1}{2}}(\tau^*_{N[nt]} - \tau^o_N), \quad t \in I_c. \tag{4.26}$$

Then, ζ_N belongs to the space $D[I_c]$ and we have the following

Theorem 4.2. Under the hypothesis of Theorem 4.1, ζ_N *converges weakly to a Gaussian function on* I_c, *for every* $0 < c \le 1$.

<u>*Proof:*</u> We need to show that (i) the f.d.d.'s of $\{\zeta_N\}$ are asymptotically Gaussian and (ii) $\{\zeta_N\}$ is tight. Note that parallel to (4.19), we have

$$\{\tau^*_{Nn}, \quad n \le N\} \stackrel{\mathcal{D}}{=} \{Z^{(n)}_{N\nu_{Nn}}, \quad n \le N\}. \tag{4.27}$$

Also, for finitely many n, say $n_1, \ldots, n_b$, $b \ge 1$, satisfying (2.20), (4.13) holds as does (4.25) for $n = n_j$, $1 \le j \le b$. Hence, the proof of Theorem 4.1 can be directly extended to cover this case. Therefore, the details of (i) are omitted. To establish the tightness, we again make use of (4.27) and, thereby, it suffices to show that (a)

$$|N^{-\frac{1}{2}}(Z^{[nc]}_{N\nu_{N[N\ ]}} - \tau^o_N)| \quad \text{is} \quad O_p(1), \tag{4.28}$$

and (b) that for every $\varepsilon > 0$ and $\eta > 0$, there exist a $\delta(0 < \delta < 1 - c)$ and an N_o, such that

$$P\left\{\sup N^{-\frac{1}{2}}|Z^{(n)}_{N\nu_{Nn}} - Z^{(n')}_{N\nu_{Nn'}}| : c \le \frac{n}{N} < \frac{n'}{N} \le \frac{n}{N} + \delta \le 1\right\} > \varepsilon\Big\} < \eta, \tag{4.29}$$

for every $N \ge N_0$. Now, for every $c > 0$, $[Nc]$ satisfies (2.20) and the proof of (4.28) follows from the asymptotic normality in Theorem 4.1 along with the fact that by (4.6), $\overline{\lim}\,\delta_{N[nc]} < \infty$, $\forall\ c > 0$. Hence, we need to verify (4.29) only.

Note that for every N, ν_{Nk} is $\nearrow$ in $k(\le N)$, while, by (4.24),

$\tau_N^o = \phi_{Nt_N(n)}^{(n)} + o(N^{\frac{1}{2}})$, for every n satisfying (2.20). Hence, to prove (4.29), it suffices to show that for every n: $cN \le n < N$,

$$P\Big\{\max[|Z_{Nk}^{(n)} - \phi_{Nk}^{(n)} - Z_{Nq}^{(n')} + \phi_{Nq}^{(n')}|/\sqrt{N}: n \le n' \le (n+\delta N)\wedge N,$$
$$|k - t_N(n)| < \delta N, \ |q - t_N(n')| < \delta N] > \varepsilon\Big\} < \eta\delta\ , \tag{4.30}$$

for every $N \ge N_0$, where N_o and δ are defined as in (4.29). For this purpose, we consider a two-dimensional time-parameter stochastic process $\zeta_N^* = \{\zeta_N^*(t_1, t_2)$: $(t_1, t_2) \varepsilon I_c^2\}$, by letting

$$\zeta_N^*(t_1, t_2) = N^{-\frac{1}{2}}\left(Z_{Nt_N([Nt_2])}^{[Nt_1]} - \phi_{Nt_N([Nt_2])}^{[Nt_1]}\right), \ (t_1, t_2) \varepsilon I_c^2. \tag{4.31}$$

Note that the tightness of ζ_N^* would insure (4.30), and hence, we complete the proof of the theorem by establishing the tightness of ζ_N^*. By letting $n = [Nt_1]$, $m = [Nt_1']$, $k = t_N([Nt_2])$ and $q = t_N([Nt_2'])$, where $t_1 > t_1'$ and $t_2 > t_2'$ (so that $n \ge m$ and $k \ge q$), we obtain that

$$B_N([t_1', t_2'], [t_1, t_2]) = \zeta_N^*(t_1, t_2) - \zeta_N^*(t_1, t_2') - \zeta_N^*(t_1', t_2) + \zeta_N^*(t_1', t_2')$$
$$= N^{-\frac{1}{2}}\left\{\left(Z_{Nk}^{(n)} - Z_{Nq}^{(n)} - Z_{Nk}^{(n)} + Z_{Nq}^{(m)}\right) - \left(\phi_{Nk}^{(n)} - \phi_{Nq}^{(n)} - \phi_{Nk}^{(m)} + \phi_{Nq}^{(m)}\right)\right\}, \tag{4.32}$$

where the $Z_{Nk}^{(n)}$ are defined by (2.4) - (2.5) [and (4.18) modifying the $X_N(s)$]. Note that for $n \ge m$, $k \ge q$,

$$\left(Z_{NK}^{(n)} - Z_{Nq}^{(n)} - Z_{Nk}^{(m)} + Z_{nq}^{(m)}\right)$$
$$= \sum_{j=q+1}^{k}\left(Y_{Nj}^{(n)} - Y_{Nj}^{(m)}\right), \tag{4.33}$$

where

$$Y_{Nj}^{(n)} = \begin{cases} X^{(n)}(J_{Nj}) & \text{if } J_{Nj} \notin \{J_{N1}, \ldots, J_{Nj-1}\} \\ 0, & \text{otherwise.} \end{cases} \tag{4.34}$$

A similar equation holds for the $\phi_{Nk}^{(n)}$. As such, using (4.18), (4.33) and (4.34), we obtain by some standard steps that

$$EB_N^2([t_1', t_2'], \ [t_1, t_2]) \le M^*[(n-m)(k-q)^2/N^3], \tag{4.35}$$

where M^* is a finite, positive quantity. Thus, if we let

$t_1 - t_1' = t_2 - t_2' = \delta > 0$, the right-hand side of (4.35) is $O(\delta^3)$ [as by (4.9) - (4.10), $(k-q)/N = O(t_2 - t_1)$] and the tightness of ζ_N^* follows by an appeal to Theorem 3 of Bickel and Wichura (1971) after noting that the Lebesgue measure $[\lambda(B_N)]$ of the block B_N is δ^2, so that $\delta^3 = [\lambda(B_N)]^{3/2}$. Q.E.D.

5. SOME CONCLUDING REMARKS

In the developments in Section 4, we have treated the estimators $\tilde{a}_N(s)$ in a rather arbitrary manner. In particular, if these are unbiased estimators (of the $a_N(s)$), as is usually the case, then in (4.3), $a_N^o(s) = a_N(s)$, so that $\tau_N^o = \tau_N$. Moreover, we have not imposed any restriction on the sub-sample sizes $\{m_s\}$, apart from the uniform integrability condition in (4.7). If these m_s are all large, then the σ_{Ns}^2 will be small, so that the second term on the right-hand side of (4.6) can be neglected. In this limiting case, the variance δ_{Nn}^2 in (4.6) corresponds to the variance in the usual SSVPWR of the primary units, where one observes the $a_N(s)$. Thus, if n, the number of primary units in a SSVPWR is large and if the m_s are also so, then sub-sampling does not lead to any asymptotic increase in the relative variance of the estimators ($\hat{\tau}_{Nn}$ and τ_{Nn}^*), though in practical problems, τ_N^* is more adaptable, because it does not presuppose the knowledge of the values of $\{a_N(s)\}$.

Theorem 4.1 provides a theoretical justification of the asymptotic normality of τ_{Nn}^*, which has been occasionally used in sample survey theory without any proper justification. Theorem 4.2 can be used to extend this asymptotic normality when the sample size (n) is itself a random variable. Further, it may also be used to repeated significance testing for τ_N or in some sequential inference procedures for the same.

REFERENCES

[1] BICKEL, P.J. and WICHURA, M.J. (1971). Convergence criteria for multiparameter stochastic processes and some applications. *Ann. Math. Statist.* **42**, 1656-1670.

[2] COCHRAN, W.G. (1963). *Sampling Techniques*. (Second Edition). John Wiley, New York.

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

[4] LOÉVE, M. (1963). *Probability Theory*, Van Nostrand, Princeton, N.J.

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

[6] ROSÉN, B. (1972a). Asymptotic theory for successive sampling with varying probabilities without replacement, I. *Ann. Math. Statist.* **43**, 401-425.

[7] ROSÉN, B. (1972b). Asymptotic theory for successive sampling with varying probabilities without replacement, II. *Ann. Math. Statist.* **43**, 748-776.

[8] SEN, P.K. (1979). Invariance principles for the coupon collector's problem: a martingale approach. *Ann. Statist.* **7**, 372-380.

[9] SUKHATME, P.V. and SUKHATME, B.V. (1970). *Sampling Theory of Surveys with Applications* (Second Edition). *Asia Publ. House*, London.