The Annals of Statistics
1979, Vol. 7, No. 1, 46–59

SEQUENTIAL TESTS FOR HYPERGEOMETRIC DISTRIBUTIONS AND FINITE POPULATIONS[1]

BY TZE LEUNG LAI

Columbia University

While usual sequential analysis deals with i.i.d. observations, this paper studies sequential tests for the dependent case of sampling without replacement from a finite population. A general weak convergence theorem is obtained and it is applied to the asymptotic analysis of the tests. Motivated by such applications as election predictions and acceptance sampling, the case of hypergeometric populations is studied in detail and a simple test with a triangular continuation region is proposed and is shown to have many nice properties. The paper concludes with a general heuristic principle of "finite-population correction" which is applicable to both sequential testing and fixed-width interval estimation problems.

1. Introduction and summary. While usual sequential analysis deals with i.i.d. observations, this paper studies sequential tests for the dependent case of sampling without replacement from finite populations. In particular, we shall first consider the important case of dichotomous populations, i.e., Np items of the population are of one kind (which we designate as 1's) and $N(1-p)$ items are of another kind (which we designate as 0's). As an illustration, suppose there are two candidates A and B in an election and it is desired to project the winner when only a proportion of the votes has been counted. Unless the outcome turns out to be extremely close, the forecaster would want his prediction to be correct with a high probability and he would like to make such a prediction as early as possible. If we formulate this (two-action) problem in terms of hypothesis testing, then the forecaster has an indifference zone of size θ_N in testing the hypothesis $H: p \leqslant \frac{1}{2}(1-\theta_N)$ versus the alternative $K: p \geqslant \frac{1}{2}(1+\theta_N)$, where N is the total number of votes cast and p is the proportion in favor of A. Assume that the votes are counted in a random order and let X_n be the number of votes in favor of A at stage n. Then X_n has the hypergeometric density

$$(1.1)\qquad f(x;p,n) = P_p[X_n = x] = \binom{Np}{x}\binom{N(1-p)}{n-x}\Big/\binom{N}{n}, \qquad x = 0, 1, \ldots, n.$$

Subject to a preassigned error probability $\alpha(0 < \alpha < \frac{1}{2})$ of wrong decision, i.e.,

$$(1.2)\qquad \begin{aligned} &P_p[\text{Reject } H] \leqslant \alpha \quad \text{if} \quad p \leqslant \tfrac{1}{2}(1-\theta_N),\\ &P_p[\text{Reject } K] \leqslant \alpha \quad \text{if} \quad p \geqslant \tfrac{1}{2}(1+\theta_N), \end{aligned}$$

Received March 1976; revised July 1977.
[1]Research supported by National Science Foundation Grant No. NSF-GP-31505 X.
AMS 1970 *subject classifications.* Primary 62L10, Secondary 62L15, 60F05, 60F10.
Key words and phrases. Hypergeometric distribution, finite population, sequential analysis, Brownian motion, Brownian bridge, weak convergence.

the forecaster's question is when he can be reasonably sure that enough votes have been counted for him to decide who would be the ultimate winner. This therefore reduces to the problem of finding a sequential test of H versus K subject to (1.2). Such sequential tests of H versus K for the parameter p of finite dichotomous populations clearly have many other applications, such as in process inspection schemes, acceptance sampling, etc.

Instead of testing H versus K, consider even the simpler problem of testing $H_0 : p = \frac{1}{2}(1 - \theta_N)$ versus $H_1 : p = \frac{1}{2}(1 + \theta_N)$. In analogy with Wald's SPRT's, one may want to form the likelihood ratio

$$R_n = f\big(X_n; \tfrac{1}{2}(1 + \theta_N), n\big)/f\big(X_n; \tfrac{1}{2}(1 - \theta_N), n\big) \tag{1.3}$$

and to stop sampling at stage

$$T = T(A, B) = \inf\{n(\leqslant N): R_n \geqslant A \text{ or } R_n \leqslant B\}$$

with $0 < B < 1 < A$, accepting H_1 or H_0 according as $R_n \geqslant A$ or $R_n \leqslant B$ (cf. [7], pages 112–113 and 136–137). However, the well-known optimum properties of Wald's SPRT's do not extend to the present dependent case. If one considers the auxiliary Bayes problem with the usual 0–1 loss as the loss function, cost c per observation and prior probability λ in favor of H_1, then the posterior probability in favor of H_1 at stage n is

$$\lambda_n = \lambda R_n / \{\lambda R_n + (1 - \lambda)\}, \tag{1.4}$$

and in this Bayesian setting the sequence $(X_n)_{1 \leqslant n \leqslant N}$ is a *nonstationary* Markov chain. The sequential Bayes test accepts H_0 or H_1 according as $\lambda_{T^*} < \frac{1}{2}$ or $\lambda_{T^*} \geqslant \frac{1}{2}$, and the stopping rule T^* of the test can be found by using backward induction and the Markovian structure of $(X_n)_{1 \leqslant n \leqslant N}$ to solve the optimal stopping problem for the loss sequence $U_n = cn + \min\{\lambda_n, 1 - \lambda_n\}$, $n = 1, \ldots, N$ (cf. Chapters 3 and 5 of [4]). It can be shown that T^* takes the form

$$T^* = \inf\{n(\leqslant N): R_n \geqslant A_n \text{ or } R_n \leqslant B_n\}.$$

The quantities A_n and B_n may in general vary with n, as can be easily seen by constructing examples where this is indeed the case. Thus Bayes sequential tests of H_0 versus H_1 need no longer be of the form of SPRT's in the present dependent case.

When N is large, the computation of the likelihood ratio R_n in the SPRT becomes rather laborious for large n, and the continuation region in the Bayes test would require a great deal of computational effort. In this paper we shall consider a much simpler test. Let $S_n = 2X_n - n$. Thus in the voting application, S_n is the number of votes by which A leads B at stage n. When $N \to \infty$ and $\theta_N \to \theta$ with $0 < \theta < 1$, X_n has a limiting binomial distribution under H_0 or H_1 for every fixed n. Let $Z_n = X_{n+1} - X_n$. If $Z_1, Z_2, \cdots$ were actually i.i.d. Bernoulli random variables such that $P_p[Z_1 = 1] = p = 1 - P_p[Z_1 = 0]$, then to test $H_0' : p = \frac{1}{2}(1 - \theta)$ versus $H_1' : p = \frac{1}{2}(1 + \theta)$, the optimum test is Wald's SPRT with stopping rule $\tau' = \inf\{n \geqslant 1 : |S_n| \geqslant b\}$, where b is a positive integer. The error probabilities and the expected sample size of this test are given by:

$$\alpha = P_{H_0'}[\text{Reject } H_0'] = P_{H_1'}[\text{Reject } H_1'] = (1-\theta)^b / \{(1+\theta)^b + (1-\theta)^b\}; \tag{1.5}$$

$$E_{H_0'}\tau' = E_{H_1'}\tau' = (b/\theta)(1-2\alpha). \tag{1.6}$$

From (1.5), it follows that $b = \{\log((1-\alpha)/\alpha)\}/\{\log((1+\theta)/(1-\theta))\}$.

The above binomial test suggests the following analogous sequential test (τ, δ) of $H_0 : p = \frac{1}{2}(1-\theta_N)$ versus $H_1 : p = \frac{1}{2}(1+\theta_N)$ for the parameter p of the hypergeometric density (1.1). Without loss of generality, we shall assume that

$$0 < \theta_N < 1 \quad \text{and} \quad \tfrac{1}{2}N(1-\theta_N) \text{ is an integer.} \tag{1.7}$$

We observe that to test H_0 versus H_1, there is no need to take more than $N(1-\theta_N)+1$ observations. In fact by (1.7), $N(1-\theta_N)+1$ is odd and so $|S_{N(1-\theta_N)+1}| \geq 1$. If H_1 holds, there are $N\theta_N$ more items equal to 1 than equal to 0, so $S_{N(1-\theta_N)+1} \geq 1$. Likewise if H_0 holds, then $S_{N(1-\theta_N)+1} \leq -1$. Obviously we should also stop sampling when $|S_{N(1-\theta_N)}| \geq 2$, or when $|S_{N(1-\theta_N)-1}| \geq 3$, etc. We therefore modify the stopping rule τ' of the binomial test as follows. Given $0 < \alpha < \frac{1}{2}$, let

$$c_N = \min\{\max[\{\log((1-\alpha)/\alpha)\}/\{\log((1+\theta_N)/(1-\theta_N))\}, 1], N(1-\theta_N)+1\}. \tag{1.8}$$

To test H_0 versus H_1, we propose the stopping rule

$$\tau = \inf\{n \geq 1 : |S_n| \geq c_N - (n-1)(c_N-1)/(N(1-\theta_N))\}. \tag{1.9}$$

It is easy to see that $\tau \leq N(1-\theta_N)+1$. Clearly we can also use the stopping rule τ to test the composite hypothesis $H : p \leq \frac{1}{2}(1-\theta_N)$ versus $K : p \geq \frac{1}{2}(1+\theta_N)$. The terminal decision rule δ of the test is to accept H_0 (or H) if $S_\tau \leq -1$ and to accept H_1 (or K) if $S_\tau \geq 1$.

In most applications, N is a large number. If θ_N is not too small, then under H_0 and H_1, $Z_1, Z_2, \cdots$ are in some sense approximately independent Bernoulli and the following theorem holds in view of (1.5) and (1.6).

THEOREM 1. *Let b_N be the smallest positive integer $\geq c_N$, where c_N is as defined in* (1.8). *Suppose* $\liminf_{N\to\infty}\theta_N > 0$ *and* $\lim_{N\to\infty} N(1-\theta_N) = \infty$. *Then* $\limsup_{N\to\infty} b_N < \infty$ *and*

$$P_{H_0}[(\tau,\delta) \text{ rejects } H_0] = P_{H_1}[(\tau,\delta) \text{ rejects } H_1] = \alpha_N + o(1), \quad \text{where} \quad \alpha_N = (1-\theta_N)^{b_N}/\{(1+\theta_N)^{b_N} + (1-\theta_N)^{b_N}\}; \tag{1.10}$$

$$E_{H_0}\tau = E_{H_1}\tau = (b_N/\theta_N)(1-2\alpha_N) + o(1) \quad \text{as} \quad N \to \infty. \tag{1.11}$$

The proof of Theorem 1 is straightforward and will be omitted. In view of (1.8), $\alpha_N \leq \alpha$ for all large N under the assumptions of Theorem 1. By (1.10) and monotonicity, the error constraints (1.2) are therefore approximately satisfied if N is large and θ_N is not too small. The following theorem, which is much deeper than

Theorem 1, says that the error constraints (1.2) are still approximately satisfied by the test (τ, δ) even when θ_N is very small.

THEOREM 2. *As* $N \to \infty$ *and* $\theta_N \to 0$ *such that* $N\theta_N \to \infty$,

$$P_{H_0}[(\tau, \delta) \text{ rejects } H_0] = P_{H_1}[(\tau, \delta) \text{ rejects } H_1] \to \alpha. \tag{1.12}$$

The asymptotic relation (1.12) turns out to be quite good as a numerical approximation to the error probabilities even when N is not very large and $N\theta_N$ is quite small (say $N = 200$ and $N\theta_N = 10$). For $N = 200, 400, 800$, we have evaluated the exact error probabilities for certain special cases of θ_N and α to check the accuracy of the asymptotic approximation given by (1.10) and (1.12). Some of our results are given in Table 1 of Section 3.

In view of Theorem 2 (and Theorem 1 as well), the stopping boundaries defined by (1.9) can be regarded as a "finite-population correction" to the horizontal boundaries in the binomial test. While this correction is negligible if $\liminf_{N\to\infty}\theta_N > 0$ (or more generally if $\lim_{N\to\infty}N^{\frac{1}{2}}\theta_N \to \infty$ as shown in Section 2), it has a very significant effect if $\theta_N = O(N^{-\frac{1}{2}})$. The proof of Theorem 2 is based on a general weak convergence theorem in Section 2 for sampling without replacement from finite populations. There are three different modes of weak convergence corresponding to the three cases $\theta_N N^{\frac{1}{2}} \to \infty$, or $\theta_N N^{\frac{1}{2}} \to \beta$ (with $0 < \beta < \infty$), or $\theta_N N^{\frac{1}{2}} \to 0$ (but $N\theta_N \to \infty$). The last case is the most delicate and our analysis of the first exit time τ in this case involves the last exit time of Brownian motion and a time inversion argument.

The following theorem, whose proof will also be presented in Section 2, gives the asymptotic behavior of the expected sample size of the test (τ, δ) and shows that there is a substantial saving in the sample size (under H_0 or H_1) over the corresponding fixed sample size Neyman-Pearson test of H_0 versus H_1 (see Section 3).

THEOREM 3. *For* $0 < \alpha < \frac{1}{2}$ *and* $\beta > 0$, *let* $\rho(\alpha) = \frac{1}{2}\log((1-\alpha)/\alpha)$ *and*

$$u_k(t; \alpha, \beta) = \Phi\big((2k+1)\beta^{-1}\rho(\alpha)t^{-\frac{1}{2}} - \beta t^{\frac{1}{2}}\big) - \Phi\big((2k-1)\beta^{-1}\rho(\alpha)t^{-\frac{1}{2}} - \beta t^{\frac{1}{2}}\big),$$

where $k = 0, \pm 1, \pm 2, \cdots$ *and* Φ *is the distribution function of the standard normal distribution. Define*

$$A^*(\alpha, \beta) = \textstyle\sum_{k=-\infty}^{\infty}(-1)^k(\alpha/(1-\alpha))^k \int_0^\infty (1+t)^{-2}u_k(t; \alpha, \beta)\,dt; \tag{1.13}$$

$$A_\alpha = \textstyle\int_0^\infty\{1 - u_0(t; \alpha, 1)\}t^{-2}\,dt - \sum_{|k|=1}^{\infty}(-1)^k(\alpha/(1-\alpha))^k \int_0^\infty t^{-2}u_k(t; \alpha, 1)\,dt. \tag{1.14}$$

Then

$$A^*(\alpha, \beta) < (1-2\alpha)\rho(\alpha)/\{\beta^2 + (1-2\alpha)\rho(\alpha)\}, \tag{1.15}$$

$$A_\alpha > \{(1-2\alpha)\rho(\alpha)\}^{-1}. \tag{1.16}$$

Furthermore, as $N \to \infty$ *and* $\theta_N \to 0$,

$$(1.17) \quad E_{H_0}\tau = E_{H_1}\tau = \{(1-2\alpha)\rho(\alpha) + o(1)\}\theta_N^{-2} \quad \text{if} \quad N^{\frac{1}{2}}\theta_N \to \infty,$$

$$= \{A^*(\alpha, \beta) + o(1)\}N \quad \text{if} \quad N^{\frac{1}{2}}\theta_N \to \beta,$$

$$= N - (A_\alpha + o(1))(N\theta_N)^2 \quad \text{if} \quad N^{\frac{1}{2}}\theta_N \to 0 \quad \text{but} \quad N\theta_N \to \infty.$$

The inequalities (1.15) *and* (1.16) *are asymptotically sharp as* $\alpha \downarrow 0$; *in fact as* $\alpha \downarrow 0$ (β *fixed*),

$$(1.18) \quad A_\alpha \sim (\rho(\alpha))^{-1} \quad \text{and} \quad A^*(\alpha, \beta) = 1 - (\beta^2 + o(1))(\rho(\alpha))^{-1}.$$

The ideas used in the construction of the test (τ, δ) can be readily extended to construct sequential tests for the means of more general finite populations. In Section 4, by making use of the general weak convergence theorem of Section 2, we shall study these kinds of sequential tests.

2. The three modes of weak convergence and the proof of Theorems 2 and 3. Throughout this section we shall let $\{W(t), t \geqslant 0\}$ denote the standard Wiener process and $\{W^0(t), 0 \leqslant t \leqslant 1\}$ the Brownian bridge. For $\rho > 0$ and $\beta > 0$, define

$$(2.1) \quad T_1(\rho, \beta) = \inf\{t \geqslant 0: W(t) \geqslant \beta t + \rho\beta^{-1}\} \quad (\inf \varnothing = \infty),$$

$$(2.2) \quad T_2(\rho, \beta) = \inf\{t \geqslant 0: W(t) \leqslant \beta t - \rho\beta^{-1}\};$$

$$(2.3) \quad T_1^*(\rho, \beta) = \inf\{t \in [0, 1]: W^0(t) \geqslant \beta t + \rho\beta^{-1}(1-t)\} \quad (\inf \varnothing = 1),$$

$$(2.4) \quad T_2^*(\rho, \beta) = \inf\{t \in [0, 1]: W^0(t) \leqslant \beta t - \rho\beta^{-1}(1-t)\};$$

$$(2.5) \quad L_1(\rho, \beta) = \sup\{t \geqslant 0: W(t) \geqslant \beta + \rho\beta^{-1}t\} \quad (\sup \varnothing = 0),$$

$$(2.6) \quad L_2(\rho, \beta) = \sup\{t \geqslant 0: W(t) \leqslant \beta - \rho\beta^{-1}t\}.$$

Theorems 2 and 3 follow from Lemma 1 and Theorem 4 below.

LEMMA 1. *Let* $0 < \alpha < \frac{1}{2}$ *and* $\beta > 0$. *Set* $\rho(\alpha) = \frac{1}{2}\log((1-\alpha)/\alpha)$. *For* $i = 1, 2$, *write* $T_i = T_i(\rho(\alpha), \beta)$, $T_i^* = T_i^*(\rho(\alpha), \beta)$, $L_i = L_i(\rho(\alpha), \beta)$, *and let*

$$(2.7) \quad T = \min(T_1, T_2), \quad T^* = \min(T_1^*, T_2^*), \quad L = \max(L_1, L_2).$$

(i) $P[T_1 < T_2] = P[T_1^* < T_2^*] = P[L_1 > L_2] = \alpha$.

(ii) *Define* $A^*(\alpha, \beta)$ *and* A_α *as in* (1.13), (1.14). *Then* (1.15), (1.16) *and* (1.18) *hold, and*

$$(2.8) \quad ET = \beta^{-2}(1-2\alpha)\rho(\alpha), \quad ET^* = A^*(\alpha, \beta), \quad EL = \beta^2 A_\alpha.$$

PROOF. It is well known that $\alpha = P[T_1 < T_2]$ and ET satisfies (2.8). Since $\{(1+t)W^0(t/(1+t)), t \geqslant 0\}$ defines a standard Wiener process (cf. [5], page 402), we can write $T_1^* = T_1/(1+T_1)$, $T_2^* = T_2/(1+T_2)$. Therefore $P[T_1^* < T_2^*]$

$= P[T_1 < T_2] = \alpha$. Furthermore $T^* = T/(1 + T)$ and so $ET^* < ET/(1 + ET)$ by Jensen's inequality. We note that

$$ET^* = 1 - E(1 + T)^{-1} = \int_0^\infty (1 + t)^{-2}\, dt - \int_0^\infty (1 + t)^{-2} P[T \leqslant t]\, dt.$$

Using (4.59) of [1], we obtain the exact expression $A^*(\alpha, \beta)$ as given by (1.13) for ET^*, thus proving (1.15).

Since the processes $\{W(t), t > 0\}$ and $\{tW(1/t), t > 0\}$ have the same distribution, it is easy to see that $(1/L_1, 1/L_2)$ has the same distribution as (T_1, T_2). Therefore $P[L_1 > L_2] = P[T_1 < T_2] = \alpha$. Since L and $1/T$ have the same distribution, $EL > 1/ET$ by Jensen's inequality. Using (4.59) of [1], we obtain the exact expression for $EL = \int_0^\infty t^{-2} P[T \leqslant t]\, dt$ as given by $\beta^2 A_\alpha$, thus proving (1.16). □

The following general weak convergence theorem for finite populations shows that although there are three possible modes of weak convergence, the triangular continuation region of the type considered in (1.9) has the remarkable property that the probability of exiting the region from the upper boundary has a limit which is the same for all the three different modes of weak convergence (see (2.15) below).

THEOREM 4. *For each* N, *let* $u_{N1}, \cdots, u_{NN}$ *be real numbers* (*not necessarily distinct*) *such that*

(2.9a) $$\Sigma_{i=1}^N u_{Ni} = 0, \qquad \lim_{N\to\infty} N^{-1} \Sigma_{i=1}^N u_{Ni}^2 = 1;$$

(2.9b) $$\limsup_{N\to\infty} N^{-1} \Sigma_{i=1}^N |u_{Ni}|^r < \infty \qquad \text{for some} \quad r > 2.$$

Let $U_{N1}, \cdots, U_{NN}$ *be a random permutation of these numbers, each of the* $N!$ *permutations having probability* $1/N!$. *Let* ρ *be a positive constant and let* θ_N, a_N *and* k_N *be positive numbers and* ξ_N, ζ_N *be real numbers such that as* $N \to \infty$,

(2.10) $$k_N \sim N, \quad \theta_N \to 0 \quad \text{but} \quad N\theta_N \to \infty, \qquad a_N \sim \rho\theta_N^{-1},$$
$$|\xi_N| + |\zeta_N| = o\left(\min\{N, (N\theta_N)^2\}\right).$$

Define

(2.11) $$S_{Nn} = \Sigma_{i=1}^n (U_{Ni} - \theta_N), \qquad n = 1, \cdots, N;$$

(2.12) $$\tau_{N1} = \inf\{n : S_{Nn} \geqslant a_N(N - n + \xi_N)/k_N\} \qquad (\inf \varnothing = N),$$

(2.13) $$\tau_{N2} = \inf\{n : S_{Nn} \leqslant -a_N(N - n + \zeta_N)/k_N\};$$

(2.14) $$\tau_N = \min\{\tau_{N1}, \tau_{N2}\}.$$

(i) *Suppose* $N^{\frac{1}{2}}\theta_N \to \infty$ *as* $N \to \infty$. *Then for every* $h > 0$, *the process* $\{\theta_N \Sigma_{i=1}^{[t\theta_N^{-2}]} U_{Ni}, 0 \leqslant t \leqslant h\}$ *converges weakly to* $\{W(t), 0 \leqslant t \leqslant h\}$. *Furthermore* $(\theta_N^2 \tau_{N1}, \theta_N^2 \tau_{N2})$ *converges in distribution to* $(T_1(\rho, 1), T_2(\rho, 1))$ *as* $N \to \infty$, *and the family* $\{\theta_N^2 \tau_N, N \geqslant 1\}$ *is uniformly integrable.*

(ii) *Suppose* $N^{\frac{1}{2}}\theta_N \to \beta$ *for some* $\beta > 0$. *Then the process* $\{N^{-\frac{1}{2}} \Sigma_{i=1}^{[Nt]} U_{Ni}, 0 \leqslant t \leqslant 1\}$ *converges weakly to the Brownian bridge* $\{W^0(t), 0 \leqslant t \leqslant 1\}$ *and* $(N^{-1}\tau_{N1}, N^{-1}\tau_{N2})$ *converges in distribution to* $(T_1^*(\rho, \beta), T_2^*(\rho, \beta))$.

(iii) *Suppose* $N^{\frac{1}{2}}\theta_N \to 0$ *but* $N\theta_N \to \infty$. *Then for every* $h > 0$, *the process* $\{(N\theta_N)^{-1}\Sigma_{i=1}^{N-[(N\theta_N)^2 t]} U_{Ni},\ 0 \leqslant t \leqslant h\}$ *converges weakly to* $\{W(t), 0 \leqslant t \leqslant h\}$. *Furthermore* $((N\theta_N)^{-2}(N - \tau_{N1}),\ (N\theta_N)^{-2}(N - \tau_{N2}))$ *converges in distribution to* $(L_1(\rho, 1), L_2(\rho, 1))$, *and the family* $\{(N\theta_N)^{-2}(N - \tau_N), N \geqq 1\}$ *is uniformly integrable.*

COROLLARY. *With the same notations and assumptions as in Theorem* 4, *as* $N \to \infty$ (*and so* $\theta_N \to 0$ *but* $N\theta_N \to \infty$ *by* (2.10)),

$$\lim{}_{\to\infty} P[\tau_{N1} < \tau_{N2}] = \alpha, \qquad \textit{where} \quad \alpha = 1/(1 + e^{2\rho}). \tag{2.15}$$

To prove Theorem 4, we shall use the following two lemmas.

LEMMA 2. *Let* u_{Ni} $(i = 1, \cdots, N)$ *be real numbers satisfying* (2.9a) *and let* $U_{N1}, \cdots, U_{NN}$ *be as defined in Theorem* 4.

(i) *Suppose that as* $N \to \infty$,

$$\max{}_{1 \leqslant i \leqslant N} |u_{Ni}| = o(N^{\frac{1}{2}}). \tag{2.16}$$

Then the process $\{N^{-}\ \Sigma_{i=1}^{[Nt]} U_{Ni},\ 0 \leqslant t \leqslant 1\}$ *converges weakly to the Brownian bridge* $\{W^0(t), 0 \leqslant t \leqslant 1\}$.

(ii) *Let* (d_N) *be a sequence of positive numbers such that*

$$\lim{}_{N\to\infty} d_N = \infty, \qquad \lim{}_{N\to\infty} d_N / N = 0. \tag{2.17}$$

Assume that

$$\lim{}_{N\to\infty} N^{-1} \Sigma_{|u_{Ni}| \geqslant \varepsilon d_N^{\frac{1}{2}}}\ u_{Ni}^2 = 0 \qquad \textit{for every} \quad \varepsilon > 0. \tag{2.18}$$

Let $X_N(0) = Y_N(0) = 0$, *and for* $N/d_N > t > 0$, *define*

$$X_N(t) = d_N^{-\frac{1}{2}} \Sigma_{i=1}^{[d_N t]} U_{Ni}, \qquad Y_N(t) = d_N^{-\frac{1}{2}} \Sigma_{i=1}^{N-[d_N t]} U_{Ni}. \tag{2.19}$$

Then for every $h > 0$, *the processes* $\{X_N(t), 0 \leqslant t \leqslant h\}$ *and* $\{Y_N(t), 0 \leqslant t \leqslant h\}$ *converge weakly to* $\{W(t), 0 \leqslant t \leqslant h\}$ *as* $N \to \infty$.

Parts (i) and (ii) (for the process $X_N(t)$) of the above lemma have been established by Rosén [9]. As to the process $Y_N(t)$, we note that $Y_N(t) = -X_N^*(t)$, where $X_N^*(t) = d_N^{-\frac{1}{2}} \Sigma_{i=N-[d_N t]+1}^{N} U_{Ni}$. Since the random variables $U_{N1}, \ldots, U_{NN}$ are exchangeable, the processes $\{X_N^*(t), 0 \leqslant t \leqslant h\}$ and $\{X_N(t), 0 \leqslant t \leqslant h\}$ have the same distribution and so the weak convergence of the process $\{Y_N(t), 0 \leqslant t \leqslant h\}$ in the above lemma follows.

LEMMA 3. *With the same notations and assumptions as in Lemma* 2(ii), *assume further that* (2.9b) *holds* (*for some* $r > 2$). *Then there exists a constant* $\delta > 0$ *such that*

$$E|\Sigma_{i=1}^{n} U_{Ni}|^r \leqslant \delta n^{r/2} \qquad \textit{for all} \quad N = 1, 2, \cdots \quad \textit{and} \quad 1 \leqslant n \leqslant N. \tag{2.20}$$

Moreover, given any $\varepsilon > 0$, *there exists* $\lambda > 0$ *such that for all* $N = 1, 2, \cdots$ *and*

$N/d_N > t \geq 1,$

$$P\Big[|\Sigma_{i=1}^{[d_N s]} U_{Ni}| > \varepsilon\, d_N^{\frac{1}{2}} s \quad \text{for some } s \geq t\Big] \leq \lambda t^{-r/2}; \tag{2.21}$$

$$P\Big[|\Sigma_{i=1}^{N-[d_N s]} U_{Ni}| > \varepsilon\, d_N^{\frac{1}{2}} s \quad \text{for some } \ s \geq t\Big] \leq \lambda t^{-r/2}. \tag{2.22}$$

PROOF. By a theorem of Hoeffding ([8], page 28), for all $N = 1, 2, \cdots,$

$$E|\Sigma_{i=1}^{n} U_{Ni}|^r \leq E|\Sigma_{i=1}^{n} V_{Ni}|^r, \qquad n = 1, 2, \cdots, N, \tag{2.23}$$

where $V_{N1}, \cdots, V_{Nn}$ are a random sample of size n drawn with replacement from the population $\{u_{N1}, \cdots, u_{NN}\}$. Since $V_{N1}, V_{N2}, \cdots$ are i.i.d. with mean 0 and variance $\sigma_N^2 = N^{-1}\Sigma_{i=1}^{N} u_{Ni}^2 \to 1$ and $\sup_{N \geq 1} E|V_{N1}|^p < \infty$ for all $0 < p \leq r$ by (2.9b), there exists a constant $C > 0$ such that

$$E|(n\sigma_N)^{-\frac{1}{2}} \Sigma_{i=1}^{n} V_{Ni}|^r \leq C \quad \text{for all} \quad N = 1, 2, \cdots \quad \text{and} \quad 1 \leq n \leq N, \tag{2.24}$$

(cf. [12], Theorem 2). From (2.23) and (2.24), (2.20) follows.

As noted by Sen [10] (see also [11]), $\{n^{-1}\Sigma_{i=1}^{n} U_{Ni}, n \geq 1\}$ is a reverse martingale; hence (2.21) follows easily from the submartingale inequality and (2.20). The inequality (2.22) follows from (2.21) since $U_{N1}, \cdots, U_{NN}$ are exchangeable and $\Sigma_{i=1}^{N} U_{Ni} = 0$. □

PROOF OF THEOREM 4. It is easy to see that condition (2.9b) implies (2.16). Also (2.9b) implies that (2.18) holds for every sequence d_N of positive numbers such that $\lim_{N \to \infty} d_N = \infty$.

The case $N^{\frac{1}{2}}\theta_N \to 0$ but $N\theta_N \to \infty$ is the most delicate and we shall first consider it. Let $d_N = (N\theta_N)^2$. Then $d_N \to \infty$ and $d_N = o(N)$. In view of (2.10), for all $N \geq N_0$ (sufficiently large) and $y \geq 1$,

$$\begin{aligned} P\big[\Sigma_{i=1}^{N-[d_N t]} U_{Ni} &\geq (N - [d_N t])\theta_N \\ &\quad + a_N([d_N t] + \xi_N)/k_N \quad \text{for some } \ N/d_N > t \geq y\big] \\ &\leq P\Big[\Sigma_{i=1}^{N-[d_N t]} U_{Ni} \geq \tfrac{1}{2}\rho d_N^{\frac{1}{2}} t \quad \text{for some } \ t \geq y\Big] \leq \lambda y^{-r/2}. \end{aligned} \tag{2.25}$$

The last inequality above follows from Lemma 3. Likewise for $N \geq N_0$ and $y \geq 1$,

$$\begin{aligned} P\big[\Sigma_{i=1}^{N-[d_N t]} U_{Ni} &\leq (N - [d_N t])\theta_N \\ - a_N([d_N t] + \zeta_N)/k_N &\quad \text{for some } \ N/d_N > t \geq y\big] \leq \lambda y^{-r/2}. \end{aligned} \tag{2.26}$$

From (2.25) and (2.26), it follows that for $N \geq N_0$ and $y \geq 1$,

$$P[\tau_{N1} \leq N - d_N y] + P[\tau_{N2} \leq N - d_N y] \leq 2\lambda y^{-r/2}. \tag{2.27}$$

Since $r > 2$, (2.27) implies the uniform integrability of $d_N^{-1}(N - \tau_N)$.

Define $Y_N(t)$ as in (2.19). By Lemma 2 (ii), $\{Y_N(t), 0 \leq t \leq h\}$ converges weakly to $\{W(t), 0 \leq t \leq h\}$ for every $h > 0$. Using this fact, together with (2.10), (2.25) and (2.26), we obtain that for all fixed positive constants t_1, t_2,

$$\begin{aligned}
P[\tau_{N1} > N - d_N t_1, \tau_{N2} > N - d_N t_2] \\
= P\Big[\textstyle\sum_{i=1}^{N-[d_N t]} U_{Ni} < (N - [d_N t])\theta_N \\
+ a_N([d_N t] + \xi_N)/k_N \quad \text{for all} \quad N/d_N > t \geq t_1, \\
\textstyle\sum_{i=1}^{N-[d_N t]} U_{Ni} > (N - [d_N t])\theta_N \\
- a_N([d_N t] + \zeta_N)/k_N \quad \text{for all} \quad N/d_N > t \geq t_2\Big] \\
\to P[W(t) < 1 + \rho t \quad \text{for all} \quad t \geq t_1, \\
W(t) > 1 - \rho t \quad \text{for all} \quad t \geq t_2] \\
= P[L_1(\rho, 1) < t_1, L_2(\rho, 1) < t_2].
\end{aligned}$$

Hence $(d_N^{-1}(N - \tau_{N1}), d_N^{-1}(N - \tau_{N2}))$ converges in distribution to $(L_1(\rho, 1), L_2(\rho, 1))$.

Next consider the case $N^{\frac{1}{2}}\theta_N \to \beta$ for some $\beta > 0$. By Lemma 2(i), $\{N^{-\frac{1}{2}}\sum_{i=1}^{[Nt]} U_{Ni}, 0 \leq t \leq 1\}$ converges weakly to $\{W^0(t), 0 \leq t \leq 1\}$. Therefore in view of (2.10), for $0 < t_1, t_2 < 1$,

$$\begin{aligned}
(2.28) \quad & P[\tau_{N1} > Nt_1, \tau_{N2} > Nt_2] \\
&= P\Big[\textstyle\sum_{i=1}^{[Nt]} U_{Ni} < [Nt]\theta_N \\
&\quad + a_N(N - [Nt] + \xi_N)/k_N \quad \text{for all} \quad 0 \leq t \leq t_1, \\
&\textstyle\sum_{i=1}^{[Nt]} U_{Ni} > [Nt]\theta_N - a_N(N - [Nt] + \zeta_N)/k_N \quad \text{for all} \quad 0 \leq t \leq t_2\Big] \\
&\to P[W^0(t) < \beta t + \rho\beta^{-1}(1 - t)\forall 0 \leq t \leq t_1, W^0(t) > \beta t \\
&\qquad - \rho\beta^{-1}(1 - t)\forall 0 \leq t \leq t_2].
\end{aligned}$$

Finally for the case $N^{\frac{1}{2}}\theta_N \to \infty$, we set $d_N = \theta_N^{-2}$ and define $X_N(t)$ as in (2.19). By Lemma 2 (ii), $\{X_N(t), 0 \leq t \leq h\}$ converges weakly to $\{W(t), 0 \leq t \leq h\}$ for every $h > 0$. Using this fact and (2.10), we obtain by an argument similar to (2.28) that $(d_N^{-1}\tau_{N1}, d_N^{-1}\tau_{N2})$ converges in distribution to $(T_1(\rho, 1), T_2(\rho, 1))$. To prove the uniform integrability of $d_N^{-1}\tau_N$, we note that for all large N and $N/d_N > t > 2\rho$,

$$\begin{aligned}
P[\tau_{N2} > d_N t] &\leq P\Big[\textstyle\sum_{i=1}^{[d_N t]} U_{Ni} > [d_N t]\theta_N - a_N(N - [d_N t] + \zeta_N)/k_N\Big] \\
&\leq P\Big[\textstyle\sum_{i=1}^{[d_N t]} U_{Ni} > d_N^{\frac{1}{2}} t/4\Big] \leq \lambda t^{-r/2} \quad \text{by (2.21).} \qquad \square
\end{aligned}$$

Proof of Corollary. Let $\alpha = 1/(1 + e^{2\rho})$. Then $\rho = \frac{1}{2}\log((1 - \alpha)/\alpha)$. If $N(j)$ is a subsequence such that $N^{\frac{1}{2}}(j)\theta_{N(j)} \to \infty$, then by Lemma 1(i) and Theorem 4(i), (2.15) holds along the subsequence $N(j)$. The same conclusion is still true if $N^{\frac{1}{2}}(j)\theta_{N(j)} \to \beta$ for some $\beta > 0$ (respectively $\beta = 0$), using part (ii) (respectively (iii)) instead of part (i) of Theorem 4. $\square$

PROOF OF THEOREMS 2 AND 3. Note that under H_0, τ as defined by (1.9) is of the form (2.14) with $k_N = N(1 - \theta_N)$, $a_N = c_N - 1$, $\xi_N = \zeta_N = \{N(1 - c_N\theta_N)/(c_N - 1)\} + 1$. Clearly condition (2.10) is satisfied. Hence Theorem 2 follows from the corollary to Theorem 4, and Theorem 3 follows from Lemma 1(ii) and Theorem 4. □

3. Comparison with some exact numerical results and improvement over fixed sample size tests. When N is not too large, the error probability $P_{H_0}[(\tau, \delta)$ rejects $H_0]$ (which we denote by P) and the expected sample size $E_{H_0}\tau$ (which we denote by E) of the test (τ, δ) introduced in Section 1 can be computed by exact numerical methods on a computer, as is done in Table 1 below for certain values of N, θ_N and α. Table 1 shows that α approximates P quite well when θ_N is small. Moreover, in all cases considered, $\alpha \geqslant P$, and hence the error constraints (1.2) are satisfied.

TABLE 1

N	θ_N	α	P	E
200	.03	.2	.1735	130.2
400	.025	.2	.1842	228.8
800	.025	.2	.1907	330.5
200	.05	.1	.08915	117.2
400	.05	.1	.09228	171.2
800	.05	.1	.09418	227.0
200	.2	.1	.08120	21.4
400	.2	.1	.07796	23.6
800	.2	.1	.07662	24.8
400	.4	.05	.03019	9.30
800	.4	.05	.03139	9.33

In Table 1, for $\theta_N = .2$ or $\theta_N = .4$, the approximation of α to P is not as good as for the smaller values of θ_N. A better approximation for P in this case is provided by α_N in (1.10); moreover, (1.11) also provides a good approximation for E in this case. For example, if $\theta_N = .4$ and $\alpha = .05$, then $\alpha_N = .0326$, in good agreement with the corresponding values .03019 and .03139 of P in Table 1.

Table 2 compares the exact value of E given by Table 1 with the corresponding asymptotic approximation $A^*(\alpha, \beta)N$ given by Theorem 3 when θ_N is small but

TABLE 2

N	θ_N	$N^{\frac{1}{2}}\theta_N = \beta$	α	$\tilde{E}$	E
200	.03	.424	.2	139.6	130.2
400	.025	.5	.2	249.8	228.8
800	.025	.707	.2	363.3	330.5
200	.05	.707	.1	127.5	117.2
400	.05	1.	.1	187.1	171.2
800	.05	1.414	.1	244.2	227.0

$N^{\frac{1}{2}}\theta_N$ is not small. Since $A^*(\alpha, \beta)N$ as defined by (1.13) is rather complicated, we shall replace it by its upper bound $\tilde{E} = N(1 - 2\alpha)\rho(\alpha)/\{\beta^2 + (1 - 2\alpha)\rho(\alpha)\}$ (see the inequality (1.15)), where $\rho(\alpha) = \frac{1}{2}\log((1 - \alpha)/\alpha)$.

Let $0 < \alpha < \frac{1}{2}$ and let $z_\alpha = \Phi^{-1}(1 - \alpha)$, where Φ is the distribution function of the standard normal distribution. Let

$$(3.1) \qquad m_N = \text{positive odd integer closest to } Nz_\alpha^2/\left(N\theta_N^2 + z_\alpha^2\right).$$

(There are two such odd integers if $Nz_\alpha^2/(N\theta_N^2 + z_\alpha^2)$ is an even integer; in this case, we take m_N to be the larger one.) To test H_0 versus H_1, the fixed sample size Neyman-Pearson test which takes m_N observations and rejects H_0 if and only if $S_{m_N} > 0$ would have the following error probabilities:

$$(3.2) \qquad P^* = P_{H_0}[\text{Reject } H_0] = P_{H_1}[\text{Reject } H_1]$$
$$\to \alpha \text{ as } N \to \infty \text{ and } \theta_N \to 0 \qquad \text{such that} \quad N\theta_N \to \infty.$$

The last relation above follows from Theorem 4 and (3.3) below. The following table compares the exact error probability P^* and the sample size m_N of this Neyman-Pearson test with the exact error probability P and the expected sample size E of the test (τ, δ).

TABLE 3
$N = 800$

θ_N	α	E	P	m_N	P^*
.025	.2	330.5	.1907	469	.1998
.05	.1	227.0	.09418	361	.09948
.4	.05	9.33	.03139	17	.03867
.2	.1	24.8	.07662	39	.09646

Table 3 shows that P^* is larger than P and m_N is substantially larger than E. Also P^* is quite close to P except in the last row; here $\alpha(= .1)$ exceeds P by about .02 and P^* is close to α. If in (3.1) we replace α by P, then we obtain $m_N = 49$ (instead of 39), and with this new value of m_N, $P^* = .07116$, which is close to $P = .07662$.

We note that as $N \to \infty$ and $\theta_N \to 0$ such that $N\theta_N \to \infty$,

$$(3.3) \qquad m_N = \left(z_\alpha^2 + o(1)\right)\theta_N^{-2} \qquad \text{if } N^{\frac{1}{2}}\theta_N \to \infty,$$
$$= \left\{\left(\beta^2 + z_\alpha^2\right)^{-1} z_\alpha^2 + o(1)\right\}N \qquad \text{if } N^{\frac{1}{2}}\theta_N \to \beta\ (> 0),$$
$$= N - \left(z_\alpha^{-2} + o(1)\right)(N\theta_N)^2 \qquad \text{if } N^{\frac{1}{2}}\theta_N \to 0.$$

Let us compare (3.3) with the asymptotic formula (1.17) for E. Using (1.18), it is easy to see that for fixed $\beta > 0$, as $\alpha \downarrow 0$,

$$(3.4a) \qquad z_\alpha^2 \sim 2|\log \alpha| = (4 + o(1))(1 - 2\alpha)\rho(\alpha),$$

$$(3.4b) \qquad \left\{1 - \left(\beta^2 + z_\alpha^2\right)^{-1} z_\alpha^2\right\} / \left\{1 - A^*(\alpha, \beta)\right\} \to \frac{1}{4},$$

$$(3.4c) \qquad z_\alpha^{-2} \sim A_\alpha/4,$$

where $\rho(\alpha)$, A_α and $A^*(\alpha, \beta)$ are as in (1.17). The common factor 4 in the above three expressions can be interpreted as the "asymptotic efficiency" of the test (τ, δ) relative to the corresponding fixed sample size Neyman-Pearson test.

4. Extensions to general finite populations. Let Π be a finite population consisting of N real numbers $y_{N1}, \cdots, y_{NN}$ (not necessarily distinct) with mean $\mu_N = N^{-1}\Sigma_{i=1}^N y_{Ni}$ and variance $\sigma_N^2 = N^{-1}\Sigma_{i=1}^N (y_{Ni} - \mu_N)^2 > 0$. Suppose that the contents of the population are unknown and so are μ_N and σ_N, while the population size N is large and known. Let $Y_1, Y_2, \cdots$ be the successive observations drawn at random without replacement from the population Π and let

$$S_n = Y_1 + \cdots + Y_n, \qquad \overline{Y}_n = n^{-1}S_n, \tag{4.1}$$
$$v_n^2 = \max\{n^{-1}\Sigma_{i=1}^n (Y_i - \overline{Y}_n)^2, n^{-1}\}.$$

Using Theorem 4, the ideas of Section 1 can be easily extended to construct a sequential test of $H_0^* : \mu_N = -\theta_N\sigma_N$ versus $H_1^* : \mu_N = \theta_N\sigma_N$, where θ_N is a given small positive number. Let n_0 be a fixed positive integer. For $0 < \alpha < \frac{1}{2}$, we let $\rho(\alpha) = \frac{1}{2}\log((1-\alpha)/\alpha)$ and use the stopping rule

$$\tau^* = \inf\{n \,;\, n_0 \leqslant n \leqslant N, |S_n| \geqslant \rho(\alpha)\theta_N^{-1}v_n(1 - n/N)\}, \tag{4.2}$$

in analogy with the stopping rule τ in (1.9). The terminal decision rule is to reject H_0^* if $S_{\tau^*} \geqslant 0$ and accept H_0^* if otherwise. The following theorem shows that the error probabilities of this test are asymptotically equal to α as $N \to \infty$ and $\theta_N \to 0$ such that $N\theta_N \to \infty$.

THEOREM 5. *Assume that*

$$\liminf_{N\to\infty}\sigma_N > 0, \tag{4.3a}$$

$$\limsup_{N\to\infty} N^{-1}\Sigma_{i=1}^N |y_{Ni}|^r < \infty \qquad \textit{for some } r > 2. \tag{4.3b}$$

Then as $N \to \infty$ and $\theta_N \to 0$ such that $N\theta_N \to \infty$,

$$\begin{aligned} P[S_{\tau^*} \geqslant 0] &\to \alpha \qquad \textit{if} \quad \mu_N = -\theta_N\sigma_N, \\ P[S_{\tau^*} < 0] &\to \alpha \qquad \textit{if} \quad \mu_N = \theta_N\sigma_N. \end{aligned} \tag{4.4}$$

PROOF. We shall only consider the case $\mu_N = -\theta_N\sigma_N$. Without loss of generality, we can assume that (4.3b) holds with $2 < r \leqslant 4$. Using an argument similar to the proof of (2.20) and the Esseen-von Bahr inequality (cf. [6], noting that $1 < r/2 \leqslant 2$), it can be shown that (4.3b) implies the existence of a positive constant C such that

$$E|\Sigma_{i=1}^n\{(Y_i - \mu_N)^2 - \sigma_N^2\}|^{r/2} \leqslant Cn \qquad \text{for all} \quad N = 1, 2, \cdots \quad \text{and} \quad 1 \leqslant n \leqslant N. \tag{4.5}$$

Since $\{n^{-1}\Sigma_{i=1}^n (Y_i - \mu_N)^2, 1 \leqslant n \leqslant N\}$ and $\{\overline{Y}_n - \mu_N, 1 \leqslant n \leqslant N\}$ are reverse martingales, it follows from (2.20), (4.5) and the submartingale inequality that for every $\varepsilon > 0$, there exists $\lambda > 0$ such that

$$(4.6) \qquad P\left[\max_{n \leqslant j \leqslant N} |v_j^2 - \sigma_N^2| \geqslant \varepsilon\right] \leqslant \lambda n^{-(r/2-1)} \quad \text{for all} \quad N = 1, 2, \cdots \quad \text{and} \quad 1 \leqslant n \leqslant N.$$

Using (2.20) and the Markov inequality and noting tht $v_j^2 \geqslant j^{-1}$, it is easy to see that

$$(4.7) \qquad \lim_{N \to \infty} P[\tau^* \leqslant n] = 0 \text{ for every fixed positive integer } n.$$

Set $u_{Ni} = (y_{Ni} - \mu_N)/\sigma_N$. From (4.6), (4.7) and Theorem 4, (4.4) follows easily. □

Suppose $Y_1, Y_2, \cdots$ were in fact i.i.d. normally distributed with mean μ_N and variance σ_N^2 and suppose that σ_N^2 were known. Then to test H_0^* versus H_1^*, we would use Wald's SPRT with stopping rule

$$\tilde{\tau} = \inf\{n \geqslant n_0 : |S_n| \geqslant \rho(\alpha)\theta_N^{-1}\sigma_N\}.$$

Thus the stopping rule τ^* can be regarded as a "finite-population correction" to $\tilde{\tau}$ and the factor $1 - n/N$ in (4.2) (which leads to a triangular continuation region) as a "finite-population correction factor" to modify $\tilde{\tau}$.

The finite-population correction factor, which we have used above in the sequential testing problem, appears again in the stopping rule for the related problem of fixed-width interval estimation for the mean μ_N of the finite population. Suppose σ_N were known. Given $\theta_N > 0$ and $0 < \alpha < 1$, let $w_\alpha = \Phi^{-1}(1 - \alpha/2)$ and define

$$(4.8) \qquad m_N^* = \text{smallest integer} \geqslant Nw_\alpha^2/\left(N\sigma_N^{-2}\theta_N^2 + w_\alpha^2\right).$$

Assume that (4.3a) and (4.3b) both hold. Then as $N \to \infty$ and $\theta_N \to 0$ such that $N\theta_N \to \infty$,

$$(4.9) \qquad P\left[|\bar{Y}_{m_N^*} - \mu_N| \leqslant \theta_N\right] \to 1 - \alpha,$$

and therefore an approximate $(1 - \alpha)$-level confidence interval with prescribed width $2\theta_N$ for μ_N would be $(\bar{Y}_{m_N^*} - \theta_N, \bar{Y}_{m_N^*} + \theta_N)$ in this case. Since σ_N is actually unknown, this suggests using the fixed-width interval $(\bar{Y}_{M_N} - \theta_N, \bar{Y}_{M_N} + \theta_N)$, where M_N is the stopping rule

$$(4.10) \qquad \begin{aligned} M_N &= \inf\{n \geqslant n_0 : n \geqslant Nw_\alpha^2/\left(Nv_n^{-2}\theta_N^2 + w_\alpha^2\right)\} \\ &= \inf\{n \geqslant n_0 : nw_\alpha^{-2}\theta_N^2 \geqslant v_n^2(1 - n/N)\}. \end{aligned}$$

The term $1 - n/N$ in the last expression of (4.10) can be regarded as a finite-population correction factor to the classical Chow-Robbins rule for the i.i.d. case (cf. [3]). From (4.6) and (4.8), it follows that as $N \to \infty$ and $\theta_N \to 0$ such that $N\theta_N \to \infty$,

$$(4.11\text{a}) \qquad m_N^* = \left(w_\alpha^2 + o(1)\right)\sigma_N^{-2}\theta_N^2 \quad \text{and} \quad M_N/m_N^* \to_P 1 \quad \text{if} \quad N^{\frac{1}{2}}\theta_N \to \infty,$$

(4.11b) $m_N^* = \{(\beta^2\sigma_N^{-2} + w_\alpha^2)^{-1}w_\alpha^2 + o(1)\}N$ and

$M_N/m_N^* \to_P 1$ if $N^{\frac{1}{2}}\theta_N \to \beta(> 0)$,

(4.11c) $m_N^* = N - \{(\sigma_N w_\alpha)^{-2} + o(1)\}(N\theta_N)^2$ and

$(N - M_N)/(N - m_N^*) \to_P 1$ if $N^{\frac{1}{2}}\theta_N \to 0$.

Making use of this fact and Lemma 2, it is not hard to show that (4.9) still holds with the fixed sample size m_N^* replaced by the stopping rule M_N. For the cases $N^{\frac{1}{2}}\theta_N \to \infty$ and $N^{\frac{1}{2}}\theta_N \to \beta > 0$, these extensions of the classical fixed-width interval theory to finite populations have recently been obtained by Carroll [2] in a somewhat different context.

Acknowledgment. I wish to express my gratitude to Professor Herbert Robbins for his valuable suggestions and to Dr. Neng Rong Lee for his assistance in the numerical computations. Thanks are also due to the referees for their helpful comments.

REFERENCES

[1] ANDERSON, T. W. (1960). Modification of the sequential probability ratio test to reduce the sample size. *Ann. Math. Statist.* **31** 165–197.

[2] CARROLL, R. J. (1975). Sequential confidence intervals for the mean of a subpopulation of a finite population. Inst. Statist. Mimeo Series, Univ. of North Carolina at Chapel Hill.

[3] CHOW, Y. S. and ROBBINS, H. (1965). On the asymptotic theory of fixed-width sequential confidence intervals for the mean. *Ann. Math. Statist.* **36** 457–462.

[4] CHOW, Y. S., ROBBINS, H. and SIEGMUND, D. (1971). *Great Expectations: The Theory of Optimal Stopping*. Houghton-Mifflin, Boston.

[5] DOOB, J. L. (1949). Heuristic approach to the Kolmogorov-Smirnov theorems. *Ann. Math. Statist.* **20** 393–403.

[6] ESSEEN, C. G. and VON BAHR, B. (1965). Inequalities for the rth absolute moment of a sum of random variables, $1 < r < 2$. *Ann. Math. Statist.* **36** 299–303.

[7] GHOSH, B. K. (1970). *Sequential Tests of Statistical Hypotheses*. Addison-Wesley, Reading.

[8] HOEFFDING, W. (1963). Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.* **58** 13–30.

[9] ROSÉN, B. (1964). Limit theorems for sampling from finite populations. *Ark. Mat.* **5** 384–424.

[10] SEN, P. K. (1970). The Hájek-Rényi inequality for sampling from a finite population. *Sankhyā Ser. A* **32** 181–188.

[11] SERFLING, F. J. (1974). Pobability inequalities for the sum in sampling without replacement. *Ann. Statist.* **2** 39–48.

[12] VON BAHR, B. (1965). On the convergence of moments in the central limit theorem. *Ann. Math. Statist.* **36** 808–815.

DEPARTMENT OF MATHEMATICAL STATISTICS
COLUMBIA UNIVERSITY
NEW YORK, NEW YORK 10027