# BOUNDED LENGTH CONFIDENCE INTERVALS FOR FINITE POPULATIONS

by

Raymond J. Carroll

*Department of Statistics*
*University of North Carolina at Chapel Hill*

Institute of Statistics Mimeo Series No. 976
January, 1975

# BOUNDED LENGTH CONFIDENCE INTERVALS FOR FINITE POPULATIONS

by

Raymond J. Carroll
*University of North Carolina at Chapel Hill*

## SUMMARY

Consider an urn which has $N_K$ balls, of which $M_K$ are white and labeled by numbers. The first part of this paper deals with sequential bounded length confidence intervals for the mean value of the white balls. The stopping rules are similar to those of Chow and Robbins (1965), but since almost sure convergence is ruled out, the observations are dependent, and $M_K$, $N_K$ may not be known, different techniques are required. Theoretical considerations and a Monte-Carlo experiment show that many observations can be saved by taking the finiteness of the population into account. The second part of the paper deals with the median of the white balls; the sample median is "linearized" along the lines of Bahadur (1966). This result not only solves the problem at hand but provides a quick proof of the asymptotic normality of a sample quantile from a finite population. Many extensions of both parts are mentioned. Examples involving banking and accounting are given.

---

AMS 1970 Subject Classifications: Primary 62L10, 62F25; Secondary 62E25, 62D05, 62G15

Key Words and Phrases: Fixed-Length Confidence Intervals, Finite Populations, Order Statistics

## 1. INTRODUCTION

This paper is concerned with sequential confidence intervals of fixed width when sampling from a finite population. Two important examples have actually occurred in the business world. The first involves a large bank which wants to build a branch in the suburbs; government regulations (among other things) require the bank to determine the average deposit and total amount of deposits of *its* customers who live within a certain distance of the proposed location. The second example involves a hospital which for a number of years has been undercharging some of its Medicare customers for thermometers; the hospital asks its accountant to determine how many of its customers have been erroneously billed so it can collect from the Medicare authorities.

In both of these examples, what is required is a confidence interval for a parameter of a subset (of unknown size) of a finite population. The confidence interval should be of bounded length, because "too long" an interval is clearly undesirable. In the first example, too long an interval might not convince the government (or the board of directors) of the need for a branch bank; in the second, the hospital loses money as the lower bound of the confidence interval gets smaller.

This paper will first concentrate on the mean of the population, deferring until Section 5 a discussion of the above examples. The literature on sequentia confidence intervals of fixed length includes work by Chow and Robbins (1965), Gleser (1965), Starr (1966), Sen and Ghosh (1970), Geertsema (1970), Srivastava (1967) and Robbins, Simons and Starr (1967). This theory has been applied to ranking and selection by Robbins, Sobel and Starr (1968), Srivastava (1966), Geertsema (1972), and Carroll (1974a), (1974b). In all of these papers, the interest has been in population parameters and the samples have been i.i.d. The dependence of the random variables in this paper and the interest not in the population mean but the mean associated with a particular finite set of entities requires different stopping rules and techniques. It will be shown that explicitly taking the finiteness of the population into account results in a large savings in sample size.

To fix notation, assume that at each stage $K = 1,2, \ldots$ there are $N_K$ balls, of which $M_K$ are white and $N_K - M_K$ are red. These balls have been assigned numbers $x_{K1}, \ldots, x_{KN_K}$ with the red balls being assigned the number zero (0). Let $\xi_{K1}, \ldots, \xi_{KN_K}$ be a random permutation of $\{x_{Ki}: i = 1, \ldots, N_K\}$ and let $\eta_{K1}, \ldots, \eta_{KM_K}$ be the white balls as arranged by this permutation. Define

$$\mu_K = M_K^{-1} \sum_{i=1}^{M_K} \eta_{Ki} \qquad \sigma_K^2 = M_K^{-1} \sum_{i=1}^{M_K} (\eta_{Ki} - \mu_K)^2 .$$

The goal is to find sequential fixed width confidence intervals for $\mu_K$ using stopping rules which take less than $N_K$ observations. Suppose that from a sample of size $n$ exactly $m$ white balls $\eta_{K1}, \ldots, \eta_{Km}$ are chosen; for a given width $d_K$, the intervals will take the form

$$I_{Km} = \left(m^{-1} \sum_{i=1}^{m} \eta_{Ki} - d_K , \quad m^{-1} \sum_{i=1}^{m} \eta_{Ki} + d_K\right) .$$

Assuming $d_K \to 0$ as $K \to \infty$, the goal is to find stopping rules $M(K)$ which for a given $\alpha > 0$ satisfy

$$(1.1) \qquad \lim_{K\to\infty} P\{\mu_K \in I_{KM(K)}\} = 1 - \alpha .$$

Certainly some structure must be placed on the values given the white balls; in fact, assume that

(1.2) $\mu_K \to \mu$ , $\sigma_K^2 \to \sigma^2 (> 0)$ , $M_K/N_K \to p$ $(0 < p < 1)$ .

At the end of Section 2, the case $p = 1$ is discussed.

In Section 2, the major results for the mean are presented. The case $N_K$ known (which both examples usually satisfy) is worked out in detail. Of special interest throughout the section is the use of the notion of strong convergence in probability which corresponds (for finite populations) to almost sure convergence. An elementary proof of asymptotic efficiency is given which basically relies on two useful integral convergence results (contrast with Chow and Robbins (1965)). It is also shown that it is better to ignore the red balls except in the estimation of $M_K$ . If both $M_K$ and $N_K$ are unknown, sequential estimates of $M_K$ are given and shown to still solve the problem. When $N_K = M_K$ , i.e., one is interested in the whole population, the results can be applied. It is shown that the rules result in considerable savings in sample size when compared to rules that ignore $N_K$ when it is large.

In Section 3, the interest shifts to the median. By tying together a number of results, the median is "linearized" along the lines of Bahadur (1966). This result yields a new, much simpler proof of a result of Rosen (1964) concerning the asymptotic normality of sample quantiles. Thus, some of the applications of Bahadur's result (1966) for the i.i.d. case may also be made for the finite case.

The results of a Monte-Carlo study are given in Section 4. It seems that the rules presented here perform quite well for moderate population sizes and moderate values of the interval length $d_K$ . In Section 5 a discussion of the examples will be given.

## 2. CONFIDENCE INTERVALS FOR THE MEAN

The first steps of this section will be to present results analagous to Theorem 1 of Anscombe (1952) using simple weak convergence techniques. These will be used to show (1.1). Then two integral convergence results will be given for use in showing that the stopping rules are asymptotically efficient. Next the stopping rules will be discussed in detail when $N_K$ is known. Finally, knowledge of $N_K$ will be dispensed with. The following Lemmas are central to the results of this paper, and do not seem to have appeared previously.

Lemma 2.1. Let $r_K$ be an integer valued random variable based on $\xi_{K1}, \ldots, \xi_{KN_K}$ for which

(i) $r_K/M_K \to q$ in probability $(0 < q < 1)$

(ii) $\max\left\{M_K^{-1/2}|x_{Ki}|: i = 1, \ldots, N_K\right\} \to 0$ .

Then

$$\sigma_K^{-1}\left(r_K^{-1} - M_K^{-1}\right)^{-1/2} \left\{r_K^{-1} \sum_{i=1}^{r_K} \eta_{Ki} - \mu_K\right\} \xrightarrow{L} \Phi ,$$

where $\Phi$ is the standard normal distribution.

Proof: If $r_K$ is a sequence of constants the Lemma follows from Theorem 24.1 of Billingsley (1968) (see also Rosen (1964)). If $r_K$ is a random variable, one applies to this a modification of Theorem 17.1 of Billingsley (1968), replacing Brownian motion by the tied-down Wiener process.

Lemma 2.2. Let $n_K$ be a sequence of integers such that $n_K \to \infty$ and $n_K/M_K \to 0$. Suppose $r_K$ is as in Lemma 2.1 with $r_K/n_K \to 1$ in probability. Letting $[\circ]$ denote the greatest integer function, if for $t > 0$

$$(2.1) \qquad \sigma_K^{-1}\left([tn_K]^{-1} - M_K^{-1}\right)^{-1/2} \left\{[tn_K]^{-1} \sum_{i=1}^{[tn_K]} \eta_{Ki} - \mu_K\right\} \xrightarrow{\mathcal{L}} \Phi ,$$

then

$$\sigma_K^{-1}\left(r_K^{-1} - M_K^{-1}\right)^{-1/2} \left\{r_K^{-1} \sum_{i=1}^{r_K} \eta_{Ki} - \mu_K\right\} \xrightarrow{\mathcal{L}} \Phi .$$

Proof: The proof is similar to that of Lemma 2.1 but uses results of Kallenberg (1974) rather than Theorem 24.1 of Billingsley (1968).

Conditions under which equation (2.1) holds may be found in Hajek (1960), Madow (1948) or Taga (1964). It is interesting to note that both of the above Lemmas may be shown using Theorem 3.1 below (which modifies Anscombe's (1952) Theorem 1).

The following Lemmas will be useful in discussing the asymptotic efficiency of the stopping rules. Their proofs are straightforward and are omitted.

Lemma 2.3. Let $\{X_K\}$ be random variables with probability measures $\{P_K\}$. If $|X_K| \le c < \infty$ and $X_K \to 1$ in probability,

$$\int X_K dP_K \to 1 .$$

Lemma 2.4. Let $\{X_K, Y_K\}$ be random variables with probability measures $\{P_K\}$. If $0 \le X_K \le Y_K$, $Y_K \to d$ and $X_K \to c$ both in probability $(d > 0,\ c > 0)$, then

$$\int X_K dP_K \to c \quad \text{if} \quad \int Y_K dP_K \to d .$$

For notational convenience define $\bar{\eta}_{Km} = m^{-1} \sum_{i=1}^{m} \eta_{Ki}$ and $s_{mK}^2 = m^{-1} \sum_{i=1}^{m} (\eta_{Ki} - \bar{\eta}_{Km})^2 + m^{-1}$.

If $X_1, X_2, \ldots$ are i.i.d. random variables with mean $\mu$ and variance $\sigma^2$, Chow and Robbins (1965) proposed the stopping rule (with $t_n^2 = n^{-1} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2 + n^{-1}$ )

(2.2) $$N(d) = \text{first integer } n \ge (bt_n/d)^2 ,$$

where $\Phi(b) - \Phi(-b) = 1 - \alpha$. In analogy with (2.2) if $M_K$ were known one might use

(2.3) $M_1(K)$ = first time $m(\ge 5)$ white balls have been drawn and $m(1 - m/M_K)^{-1} \ge (bs_{mK}/d_K)^2$.

Naturally $M_1(K) \le M_K$ and one stops if $M_K$ white balls have been drawn. Since $M_K$ is known, $M_1(K)$ ignores red balls.

If $M_K$ is unknown (which is the usual case) and if, when taking a sample of size $n$ from $\xi_{K1}, \ldots, \xi_{KN_K}$ exactly $m$ white balls are drawn, a convenient estimate of $M_K$ (closely related to the maximum likelihood estimate) is $\hat{M}_{Kn} = mN_K/n$. In this case, the stopping rule becomes

(2.4) $M_2(K)$ = first time $m(\ge 5)$ white balls have been drawn and $m(1 - m/\hat{M}_{Kn})^{-1} \ge (bs_{mK}/d_K)^2$, where $n$ is the total number of balls drawn.

The statement of the main result, Theorem 2.1, is natural in light of the results of Chow and Robbins (1965). The proof depends on a number of Lemmas.

<u>Theorem 2.1.</u> Suppose that for some $\delta > 0$ , $d_K M_K^{1-\delta} \to \infty$ and that

(i) $M_K^{-1} \sum_{i=1}^{M_K} (\eta_{Ki} - \mu_K)^p \to \mu_p$ (finite), $p = 3,4$

(ii) $\left(1 + M_K d_K^2/(b\sigma_K)^2\right)^{-1} \to \psi \quad (0 \le \psi < 1)$ .

Then for $i = 1,2$

$$P\{\mu_K \in I_{KM_i(K)}\} \to 1 - \alpha \qquad (\text{"consistency"}) \tag{2.5}$$

$$M_i(K)\left(1 - M_i(K)/M_K\right)^{-1}/(b\sigma_K/d_K)^2 \to 1 \quad \text{in probability, i.e.,} \tag{2.6}$$

$$M_i(K)/M_K \to \psi \quad \text{in probability}$$

$$\left(M_K^{-1} + (b\sigma/d_K)^{-2}\right) EM_i(K) \to 1 \qquad (\text{"efficiency"}) . \tag{2.7}$$

Chow and Robbins (1965) were able to obtain results analogous to (2.5) and (2.6) by showing that $t_n^2 \to \sigma^2$ almost surely. Since the probability space change for each $K$ here, almost sure convergence of $s_{mK}^2$ to $\sigma^2$ is ruled out. Also, convergence in probability is not strong enough to get (2.6) even. Some type of convergence in between is needed; the following seems to work:

<u>Definition 2.1.</u> (Rosen (1964)) A sequence $X_{K1}, \ldots, X_{KN_K}$ of random variables *converges strongly to* $\mu$ if for any $\varepsilon > 0$ , there are $K_o, n_o$ such that

$$\sup_{K \ge K_o} \Pr\{|X_{Kn} - \mu| > \varepsilon \text{ for some } n \ge n_o\} < \varepsilon . \tag{2.8}$$

Denote this by $X_{Kn} \to \mu$ (s.p.). Note that it implies convergence in probabilit

Lemma 2.5. If the conditions (1.2) hold and if (i) of Theorem 2.1 holds, then

$$\bar{\eta}_{Km} \to \mu \text{ (s.p.)}, \ s_{mK}^2 \to \sigma^2 \text{ (s.p.)} \ \text{ and } \ M_K^{-1}\hat{M}_{Kn} \to 1 \text{ (s.p.)} .$$

Proof: From Lemma 4.2 of Rosen (1964), for any $\varepsilon > 0$ and any $m = 1,2, \ldots ,M_K$ ,

(2.9) $$\Pr\{\max_{j\geq m} |\bar{\eta}_{Kj} - \mu_K| \geq \varepsilon\} \leq 2\sigma_K^2(m^{-1} - M_K^{-1})/\varepsilon^2 .$$

Since $\mu_K \to \mu$ , this gives $\bar{\eta}_{Km} \to \mu$ (s.p.) . Now, $s_{mK}^2 = m^{-1} \sum_{i=1}^{m} (\eta_{Ki} - \mu_K)^2 + (\bar{\eta}_{Km} - \mu_K)^2 + m^{-1}$ ; applying (2.9) to the first term and using the fact that $(\bar{\eta}_{Km} - \mu_K)^2 \to 0$ (s.p,) yields $s_{mK}^2 \to \sigma_K^2$ (s.p.) . Finally, $M_K^{-1}\hat{M}_{Kn} = (N_K/M_K)n^{-1} \sum_{j=1}^{n} z_{Kj}$ , where $z_{Kj} = 1$ if $\xi_{Ki}$ is white and 0 otherwise. Since $EM_K^{-1}\hat{M}_{Kn} = 1$ , applying (2.9) completes the proof.

The next step in proving Theorem 2.1 is to show that $M_i(K) \to \infty$ in probability $(i = 1,2)$ .

Lemma 2.6. Under the conditions of Theorem 2.1, $M_i(K) \to \infty$ in probability $(i = 1,2)$ .

Proof: Since $s_{mK}^2 \geq m^{-1}$ and $d_K \to 0$ , the result is obvious for $M_1(K)$ . For $M_2(K)$ , it suffices to show that

$$P\left\{\hat{M}_{Kn} \leq M_K^{1/2} \text{ for some } m \geq 5\right\} = P\left\{\hat{M}_{Kn} \leq M_K^{1/2} \text{ some } 5 \leq m \leq M_K^{1/2}\right\} \to 0 .$$

Since $\hat{M}_{Kn} = mN_K/n \leq M_K^{1/2}$ implies there are at most $m$ successes in $mN_K/M_K^{1/2}$ trials, from (2.9),

$$P\left\{\hat{M}_{Kn} \le M_K^{1/2} \text{ some } 5 \le m \le M_K^{1/2}\right\} \le P\left\{\left(5N_K/M_K^{1/2}\right)^{-1} \sum_{i=1}^{5N_K/M_K^{1/2}} z_{Ki} \le (5N_K/M_K)^{-1}\right\} \to 0.$$

Proof of Equations (2,5) and (2.6): By Lemma 2.6, $s^2_{M_i(K)K} \to \sigma^2$ in probability. Thus, since $d_K M_K^{1-\delta} \to \infty$, $M_K - M_i(K) \to \infty$ in probability. A proof analogous to that of Chow and Robbins (1965) yields (2.6) using strong convergence in probability. Since by (2.6) $(b\sigma_K/d_K)\left(M_i(K)^{-1} - M_K^{-1}\right)^{1/2} \to 1$ in probability, applying Lemma 2.1 (if $\psi > 0$) or Lemma 2.2 (if $\psi = 0$) completes the proof.

Proof of Equation (2.7): Define the stopping rule

$$N_K^* = \text{first integer } m \ge \min\{(bs_{mK}/d_K)^2, M_K\}$$

(of course, set $N_K^* \ge 5$). Then $M_i(K) \le N_K^*$, so that

$$M_i(K)\left(M_K^{-1} + (b\sigma/d_K)^{-2}\right) \le 2(N_K^*/2)\left(M_K^{-1} + (b\sigma/d_K)^{-2}\right)$$
$$\le 4(b/d_K)^2\left(M_K^{-1} + (b\sigma/d_K)^{-2}\right)s^2_{LK},$$

where $L = N_K^*/2$. Since $0 \le \psi < 1$ implies that $M_K(b\sigma/d_K)^{-2}$ does not converge to $0$, the last term is bounded by $cs^2_{LK}$ for some constant $c > 0$. Now, $s^2_{LK} \to \sigma^2$ in probability and

$$Es^2_{LK} \le E \sup_{m\ge 1}\left\{\left|m^{-1}\sum_{i=1}^{m}(\eta_{Ki} - \mu_K)^2 - \sigma^2\right| + \sigma^2 - (\bar{\eta}_{Km} - \mu_K)^2\right\} + EL^{-1}$$
$$\le \sigma^2 + E\sup_{m\ge 1}\left|m^{-1}\sum_{i=1}^{m}(\eta_{Ki} - \mu_K)^2 - \sigma^2\right| + 1/2 .$$

Now, in general, $E|X| \le 1 + \sum_{j=1}^{\infty} P\{|X| \ge j\}$, so that by (2.9),

$$Es^2_{LK} \le c_* \text{ for some } c_* .$$

By using Lemma 2.4 and equation (2.6), one can show that any subsequence $K_1, K_2, \ldots$ must satisfy (2.7), which completes the proof.

As mentioned above, if $M_K$ is known the rule (2.3) completely ignores red balls, while if $M_K$ is unknown, the red balls are needed only to form the estimators $\hat{M}_{Kn}$. It is tempting to incorporate the red balls more explicitly into the stopping rules by treating them as variables taking on the value 0. The results below show that such rules take at least as many *total* observations (asymptotically) as (2.3) or (2.4) and are more computationally difficult because one must check for stopping after *each* ball is drawn, rather than after a white ball is drawn.

Lemma 2.7. Let $r_K$ be an integer-valued random variable as in Lemma 2.1 or 2.2 (with $N_K$ replacing $M_K$). Define

$$\sigma_{K2}^2 = N_K^{-1} \sum_{i=1}^{N_K} (\xi_{Ki} - M_K\mu_K/N_K)^2 = (M_K/N_K)\left\{\sigma_K^2 + (1 - M_K/N_K)\mu_K^2\right\}$$

$$\sigma_{K3}^2 = N_K^{-1} \sum_{i=1}^{N_K} (\xi_{Ki} - \mu_K z_{Ki})^2 = (M_K/N_K)\sigma_K^2 .$$

Then

$$\frac{(N_K/M_K)\bar{\xi}_{Kr_K} - \mu_K}{(N_K/M_K)\sigma_{K2}\left(r_K^{-1} - N_K^{-1}\right)^{1/2}} \xrightarrow{\mathcal{L}} \Phi$$

and

$$\frac{(N_K/\hat{M}_{Kr_K})\bar{\xi}_{Kr_K} - \mu_K}{(N_K/M_K)\sigma_{K3}\left(r_K^{-1} - N_K^{-1}\right)^{1/2}} \xrightarrow{\mathcal{L}} \Phi .$$

Note the surprising result that, although $\sigma_{K3}^2$ is obtained from estimating $M_K$, $\sigma_{K3}^2 \le \sigma_{K2}^2$.

Proof: The first result is immediate from Lemmas 2.1 and 2.2. For the second,

$$(N_K/\hat{M}_{Kr_K})\bar{\xi}_{Kr_K} - \mu_K = \left((N_K/M_K)\bar{\xi}_{Kr_K} - \mu_K\right) + \mu_K(1 - \hat{M}_{Kr_K}/M_K)$$
$$+ (1 - \hat{M}_{Kr_K}/M_K)^2(N_K\bar{\xi}_{Kr_K}/\hat{M}_{Kr_K})$$
$$+ (1 - \hat{M}_{Kr_K}/M_K)\left((N_K/M_K)\bar{\xi}_{Kr_K} - \mu_K\right) .$$

The last two terms converge to $0$ in probability when multiplied by $(r_K^{-1} - N_K^{-1})^{-1/2}$, and the first two terms equal

$$(N_K/M_K)\ r_K^{-1} \sum_{i=1}^{r_K} (\xi_{Ki} - \mu_K z_{Ki}) .$$

Because $\sigma_{K3}^2 \le \sigma_{K2}^2$, it suffices to consider the case $M_K$ unknown. Letting $t_{nK}^2 = n^{-1} \sum_{i=1}^{n} (\xi_{Ki} - \bar{\xi}_{Kn} z_{Ki})^2$, the natural rule is

(2.10) $N_{1K}$ = first integer $n$ for which at least $5$ white balls have been drawn and $n(1 - n/N_K)^{-1} \ge (bt_{nK}N_K/d_K M_K)^2$ .

Similar to Theorem 2.1, if $\left(1 + N_K/(b\sigma_{K3}N_K/d_K M_K)^2\right)^{-1} \to \psi_2$ $(0 \le \psi_2 < 1)$, and if $I_{Kn} = \left((N_K/\hat{M}_{Kn})\bar{\xi}_{Kn} - d_K,\ (N_K/\hat{M}_{Kn})\bar{\xi}_{Kn} + d_K\right)$, then (2.5) holds for $N_{1K}$ and $N_{1K}/M_K \to \psi_2$ in probability. Then, if $N_{2K}$ is the total number of balls chosen using $M_2(K)$,

$$\lim_{K\to\infty} N_{2K}/N_{1K} = \lim_{K\to\infty} \left(M_2(K)/p\right)/N_{1K} = \psi/\psi_2 = 1 \qquad \text{in probability} .$$

Thus one does no worse by using $M_2(K)$. However $M_2(K)$, unlike $N_{1K}$, is susceptible to the danger that one might exhaust all the white balls without stopping, in which case all $N_K$ balls will be sampled. On the other hand, a reasonable choice of $d_K$ would make this a rather rare event; even if it occurred, the above result indicates that $N_{1K}$ will be large and may in fact be more time consuming.

Now consider the situation where both $N_K$ and $M_K$ are unknown. The rule (2.4) will be used with a new estimate $\overline{M}_{Kn}$ of $M_K$ ; call this $M_3(K)$ . The main step is proving $M_K^{-1}\overline{M}_{Kn} \to 1$ (s.p.) . The usual sequential methods of estimating the size of a population are reviewed in Freeman (1973) and Samuel (1968), the latter presenting five sequential rules. This author has been unable to show that use of the maximum likelihood estimate with any of the rules B - E of Samuel (1968) satisfies $M_K^{-1}\overline{M}_{Kn} \to 1$ (s.p.) . The idea used here is to first take an initial sample, color the white balls so drawn black and replace them (estimates for $N_K$ based on coloring any ball black when chosen may also be developed). One then continues sampling, estimating $M_K$ at each step by an estimator like that in Feller (1968), pages 45 - 46. The problem is to find the initial sample size. Clearly, if a lower bound $M_{oK}$ on $M_K$ is known, the initial sample size might be obtained from (2.3) using a quantity slightly less than $M_{oK}$ .

Lemma 2.8. Assume that $M_K \geq M_{oK}$ (known) and choose $0 < a < 1$ . Let $A_{oK}$ be the stopping rule (2.3) with $aM_{oK}$ used and let $A_{1K}$ be the stopping rule (2.3) with $aM_{oK}$ used and $s_{mK}^2$ replaced by $\sigma_K^2$ . Mark the $A_{oK}$ white balls so drawn black and return them. Continue sampling without replacement, defining for $n = 1,2, \ldots$

(2.11) $\overline{M}_{Kn} = \max\{M_{oK}, A_{oK}(WB)_{Kn}/B_{Kn}\}$ , where

$(WB)_{Kn}$ = # of white or black balls drawn in n tries since $A_{oK}$

$B_{Kn}$ = # of black balls drawn in n tries since $A_{oK}$ .

Then, if $\overline{M}_{Kn}$ is used in rule (2.4) to obtain a stopping rule $M_3(K)$ , Theorem 2.1 still holds.

<u>Proof:</u> Since $\overline{M}_{Kn} \geq M_{oK}$ it follows that $M_3(K) \to \infty$ in probability. That $M_3(K) - A_{oK} \to \infty$ in probability follows easily. Defining $L(K) = M_3(K) - 1$, it follows that $M_K - M_3(K) \to \infty$ in probability since

$$\overline{M}_{KL(K)} - L(K) \geq \overline{M}_{KL(K)} L(K)/(bs_{L(K)K}/d_K)^2$$

$$\geq \overline{M}_{KL(K)}(A_{oK} - 1)/(bs_{L(K)K}/d_K)^2$$

and since $A_{oK}/A_{1K} \to 1$, $s^2_{L(K)K}/\sigma^2_K \to 1$ , both in probability. Now the theorem goes through as before.

The value a = 1 for the constant in Lemma 2.8 seems to be impermissible; if a = 1 then it is possible that $M_3(K) - A_{oK} \to \infty$ in probability may not be true. In this case, $M_K^{-1}\overline{M}_{KL(K)} \to 1$ in probability may not hold, so that (2.6) might fail.

As for the constant a in Lemma 2.8, a reasonable rule is to choose a = 1/2 if $M_{oK}$ seems to be a good lower bound (in the sense of being close to $M_K$ ), while a = 3/4 would seem to be reasonable if $M_{oK}$ is a conservative lower bound.

If no reasonable lower bound on $M_K$ is known (a case which should be somewhat rare) no intuitively obvious initial sample size presents itself. However, if one replaces $A_{oK}$ in Lemma 2.8 by $A^*_{oK}$ = first integer $m \geq (a_K b s_{mK}/d_K)^2$ (where $a_K \to 0$ ), then Lemma 2.8 will still hold.

Whether $A_{oK}$ is a good initial sample size is unknown, as is an algorithm for choosing $\{a_K\}$ if $M_{oK}$ is unknown. The procedures given here for estimating $M_K$ when $N_K$ is unknown are in some sense two stage procedures akin to the Stein two sample procedure for the mean (1945), while the sequential methods in Freeman (1973) and Samuel (1968) are analogous to purely sequential estimation procedures. Presumably, if the latter methods yield estimates which converge (s.p.) , they will be found to be more efficient.

Multivariate generalizations along the lines of Srivastava (1967) are possible by extending Theorem 3.1 below to multidimensions.

If $M_K = N_K$ , i.e., one is interested in the total population rather than a subpopulation, one can use Lemma 2.8.

Finally, all the results of this section and Section 3 below are true if at stage K the white balls are labeled by $x_1, \ldots, x_{M_K}$ which are i.i.d. with mean $\mu$ and variance $\sigma^2$ (naturally, convergence (s.p.) is replaced by convergence almost surely).

## 3. THE MEDIAN AND OTHER ROBUST ESTIMATORS

The main points of interest in this section include an analogue to Theorem 2.1 and the presentation of a Bahadur type (1966) linearization technique for sample quantiles. The latter gives a quick proof of Rosen's (1964) Theorem

15.1 concerning the asymptotic normality of a sample quantile from a finite population. Only one particular quantile will be considered although the results extend to any quantile. The $M_K$ white balls will have a $[M_K/2]$ order statistic $X_{KM_K}(M_K/2)$ , while the $[n/2]$ order statistic based on a sample of size $n$ will be $X_{Kn}(n/2)$ . The goal is to find a fixed width confidence interval for the "median" $X_{KM_K}(M_K/2)$ . The first step of this section (Theorem 3.1) will be to generalize Theorem 1 of Anscombe (1952); this result, which could have been used to prove Lemmas 2.1 and 2.2, should prove useful in extending the results of this section to other robust estimators. Once Theorem 3.1 is obtained, a generalization of Bahadur's theorem (1966) is given; this representation of $X_{Km}(m/2)$ together with Theorem 3.1 will yield results similar to Lemmas 2.1 and 2.2. The condition (3.1) below is essentially Anscombe's uniform continuity in probability (1952).

The median and other robust estimators of "location" have an obvious appeal as indicators of characteristics of a finite population in that they are not sensitive to a few wild observations. For example, in the banking situation of Section 1 one deposit might be very large and seriously effect the mean. Another appeal of these estimators is that they often have smaller variances than the sample mean; since the rules considered here are monotone functions of the sample variance, this means that the use of robust estimators will in general result in a decrease in the number of observations taken.

Theorem 3.1. Suppose that $T_{Kn}$ is a statistic based on $\eta_{K1}, \ldots, \eta_{Kn}$ and that there are constants $\mu_K, \sigma_{Kn}$ such that if $n_K \to \infty$ as $K \to \infty$ ,

$$P\left\{\sigma_{Kn_K}^{-1}(T_{Kn_K} - \mu_K) \le z\right\} \to \Phi(z) \quad \text{for all } z .$$

If $r_K/n_K \to 1$ in probability and satisfies the conditions of Lemmas 2.1 or 2.2, and if for $\varepsilon > 0,\ \gamma > 0$ there is $c > 0,\ K_o,\ n_o$ such that for $n \ge n_o$

$$\sup_{K \ge K_o} P\left\{\sigma_{Kn}^{-1}|T_{Kn} - T_{Km}| > \varepsilon \text{ some } |m - n| < cn\right\} < \gamma , \tag{3.1}$$

then

$$P\left\{\sigma_{Kn_K}^{-1}(T_{Kr_K} - \mu_K) \le z\right\} \to \Phi(z) \quad \text{for all } z .$$

The proof is much like Theorem 1 of Anscombe (1952); note that the sample mean satisfies (3.1) (by following Anscombe's proof and using (2.9)).

As in (1.2), assume that $\mu_K = X_{KM_K}(M_K/2) \to \mu$ as $K \to \infty$. Let $F_K$ be the distribution function of $\eta_{K1}, \ldots ,\eta_{KM_K}$ and let $F_{Kn}$ be the empirical distribution based on $\eta_{K1}, \ldots ,\eta_{Kn}$ (see Rosen (1964)). The following result "linearizes" $X_{Kn}(n/2)$ .

Lemma 3.1. Assume there is a distribution $F_o$ which has derivative $f_o$ (where it exists) and which satisfies the conditions given in Bahadur (1966). Assume further that

$$n^{1/2} \sup_t |F_K(t) - F_o(t)| \to 0 \quad \text{as } n,k \to \infty .$$

Then, if $\ell_{Kn} = n/2 + o(n^{1/2+\delta})$ for some $0 < \delta < 1/4$ , the following decomposition holds:

$$n^{1/2}\left|\left(X_{Kn}(n/2) - \mu\right) - f_o^{-1}(\mu)\left(\ell_{Kn}/n - F_{Kn}(\mu)\right)\right| \to 0 \text{ (s.p.)} .$$

Proof: The proof is only a slight modification of the original proof of Bahadur (1966). His equation (6) should be replaced by $a_n \sim n^{-1/2+\delta}$ for some $0<\delta<1/4$. This will mean that, for example, $n^{1/2}\alpha_{r,n,k} \to 0$ in his proof. His Lemma 1 becomes $n^{1/2}H_{Kn}(\omega) \to 0$ (s.p.); this is true since his equation (11) still holds when sampling from a finite population (Hoeffding, (1963), Theorem 4).

Lemma 3.1 provides an easy proof of Theorem 15.1 of Rosen (1964) concerning the asymptotic normality of sample quantiles from a finite population. It also shows (by means of Theorem 3.1) that if $n_K/M_K \to c < 1$ and $r_K/n_K \to 1$ in probability then

$$(3.2) \qquad 2f_o(\mu)\left(r_K^{-1} - M_K^{-1}\right)^{-1/2}\left(X_{Kr_K}(r_K/2) - X_{KM_K}(M_K/2)\right) \xrightarrow{\mathcal{L}} \Phi \ .$$

Let $c > 0$ . Geertsema (1972) considers the following estimate:

$$(3.3) \quad S_{Kn}^2 = (n/4c^2)\{X_{Kn}(n/2 + cn^{1/2}/2) - X_{Kn}(n/2 - cn^{1/2}/2)\}^2 + n^{-1} \ .$$

By Lemma 3.1, $s_{Kn}^2 \to 4^{-1}f_o^{-2}(\mu)$ (s.p.) . Using the rule $M_2(K)$ given in (2.4) with $s_{Kn}^2$ defined by (3.3), one gets:

Theorem 3.2. For the case of the median, Theorem 2.1 still holds, with $\sigma^2 = \left(4f_o^2(\mu)\right)^{-1}$ .

Proof: Equations (2.5) and (2.6) follow immediately. For (2.7), one uses Lemma 2.4 with the help of Lemma 3.3 of Geertsema (1970) and Theorem 4 of Hoeffding (1963).

This particular technique of "linearizing" an estimator can probably be applied to other robust estimators, such as M-estimators (Carroll (1974b)) and R-estimators (Antille (1974)).

## 4. A MONTE-CARLO STUDY

The results given here have been large sample results as $d_K \to 0$ and $M_K, N_K \to \infty$. As in Starr (1966) it is of interest to evaluate the performance of the rule $M_2(K)$ for moderate values of $d_K, M_K$, and $N_K$. Exact results seem computationally difficult so Monte-Carlo results will be given. It was decided to use $M_K = 100$, $N_K = 1000$ and $\alpha = .05$. The random numbers were generated by the system routine on the IBM 360 machine at the Triangle Universities Computation Center. The $100 = M_K$ white balls were labeled by Uniform (0,1) random variables, Exponential r.v.'s with parameter $\lambda = 1$, standard Normal r.v.'s, Binomial r.v.'s with $n = 10$ and $p = 1/2$, and finally by Poisson r.v.'s with parameter $\lambda = 1$. All these used the North Carolina Educational Computing Service program RANSAM. There were 100 repetitions of the experiment.

The results of these Monte-Carlo experiments are listed in Tables 1 - 5, with the following notation:

P(CD): Proportion of correct coverages

$M_2(K)$: Average number of white balls sampled

$\hat{M}_K$: Average estimate of $M_K (= 100)$ at stopping

Ratio Ex: The ratio $M_2(K)\left(1 - M_2(K)/M_K\right)^{-1}/(1.96\sigma/d_K)^2$, which should be close to one (see (2.6))

Ratio C-R: The ratio of $M_2(K)$ to the expected number of white balls using the Chow-Robbins rule

$N_2(K)$: The average *total* number of balls sampled.

The results of this preliminary Monte-Carlo study are quite encouraging. As Starr (1966) found for the normal case, as $d_K$ decreases, $P(CD)$ first decreases and then increases. In all of the cases except for the Exponential and one value of $d_K$ for the Uniform, the procedures give probabilities which are acceptably close to .95 . The results for the Exponential may be affected by the non-robustness of the sample mean. For the applications given in Section 5, $M_K$ and $N_K$ are considerably larger than in this study and the rules should become even better. For practical usage in the normal case for example, the performance will be a function of $\sigma d_K^{-1}$ rather than only $d_K^{-1}$ .

In all the cases studied, the estimates of $M_K$ obtained by $\hat{M}_{Kn}$ at stopping are quite good and seem to improve as $d_K$ decreases. The ratio $M_2(K)/N_2(K)$ is always quite close to .10 as it should be. The column "Ratio Ex" indicates that (2.7) is almost true.

The "Chow and Robbins" type procedure is basically one which sets $M_K = \infty$ as an approximation to saying $M_K$ is large. As expected, the savings in the number of white balls sampled is considerable when compared to the "Chow and Robbins" type procedure. In all the cases studied, if $d_K^{-1}$ is large enough so that $M_2(K) \geq 8$ , significant savings are obtained (see the "Ratio C-R" column). The increase in this ratio for smaller values of $d_K$ is simply due to the fact that the "Chow and Robbins" type rule by that time already exhausted all the white balls.

## TABLE 1

Monte-Carlo study for Uniform (0,1) random variables

| $d_K^{-1}$ | P(CD) | $M_2(K)$ | $\hat{M}_K$ | Ratio Ex | Ratio C-R | $N_2(K)$ |
|---|---|---|---|---|---|---|
| 1 | 1.00 | 5.00 | 123.04 | 15.84 | 15.79 | 49.26 |
| 2 | 1.00 | 5.00 | 123.04 | 3.99 | 3.95 | 49.26 |
| 3 | .99 | 5.05 | 120.87 | 1.82 | 1.77 | 49.52 |
| 4 | .93 | 5.67 | 116.15 | 1.17 | 1.12 | 57.45 |
| 5 | .87 | 7.33 | 111.79 | 1.00 | .93 | 74.85 |
| 6 | .82 | 9.45 | 109.09 | .923 | .83 | 97.12 |
| 7 | .88 | 12.46 | 105.77 | .928 | .80 | 126.34 |

## TABLE 2

Monte-Carlo study for Exponential random variables with $\lambda = 1$

| $d_K^{-1}$ | P(CD) | $M_2(K)$ | $\hat{M}_K$ | Ratio Ex | Ratio C-R | $N_2(K)$ |
|---|---|---|---|---|---|---|
| 1 | 1.00 | 5.46 | 117.94 | 1.66 | 1.61 | 55.53 |
| 2 | .92 | 8.70 | 111.85 | .73 | .64 | 88.69 |
| 3 | .78 | 16.68 | 109.34 | .71 | .55 | 166.12 |
| 4 | .76 | 25.86 | 106.45 | .73 | .48 | 256.91 |
| 5 | .77 | 38.12 | 103.99 | .83 | .45 | 377.25 |
| 6 | .81 | 49.40 | 103.56 | .90 | .49 | 485.81 |
| 7 | .87 | 58.94 | 102.94 | .94 | .59 | 582.74 |

TABLE 3

Monte-Carlo study for standard Normal random variables

| $d_K^{-1}$ | P(CD) | $M_2(K)$ | $\hat{M}_K$ | Ratio Ex | Ratio C-R | $N_2(K)$ |
|---|---|---|---|---|---|---|
| 1 | 1.00 | 5.45 | 118.81 | 1.34 | 1.29 | 54.21 |
| 2 | .93 | 12.69 | 107.98 | .88 | .75 | 130.38 |
| 3 | .87 | 25.16 | 105.07 | .91 | .66 | 251.63 |
| 4 | .90 | 38.40 | 103.50 | .95 | .57 | 382.43 |
| 5 | .92 | 51.13 | 102.63 | .99 | .51 | 507.86 |
| 6 | .93 | 61.08 | 101.90 | 1.01 | .61 | 607.06 |
| 7 | .93 | 68.89 | 102.98 | 1.02 | .69 | 677.60 |

TABLE 4

Monte-Carlo study for Binomial random variables with $n = 10$, $p = 1/2$

| $d_K^{-1}$ | P(CD) | $M_2(K)$ | $\hat{M}_K$ | Ratio Ex | Ratio C-R | $N_2(K)$ |
|---|---|---|---|---|---|---|
| 1 | .88 | 7.35 | 114.66 | .97 | .90 | 73.93 |
| 2 | .86 | 22.55 | 102.02 | .92 | .69 | 227.60 |
| 3 | .91 | 41.98 | 102.30 | .99 | .57 | 418.03 |
| 4 | .93 | 57.49 | 101.13 | 1.01 | .57 | 573.20 |
| 5 | .94 | 68.98 | 101.96 | 1.02 | .69 | 680.27 |
| 6 | .96 | 76.92 | 101.81 | 1.03 | .77 | 758.31 |
| 7 | .97 | 82.51 | 101.36 | 1.03 | .83 | 816.02 |

TABLE 5

Monte-Carlo study for Poisson random variables with $\lambda = 1$

| $d_K^{-1}$ | P(CD) | $M_2(K)$ | $\hat{M}_K$ | Ratio Ex | Ratio C-R | $N_2(K)$ |
|---|---|---|---|---|---|---|
| 1 | .97 | 5.27 | 117.13 | 1.54 | 1.49 | 53.11 |
| 2 | .87 | 11.32 | 107.23 | .91 | .80 | 116.51 |
| 3 | .87 | 22.58 | 101.63 | .93 | .71 | 230.69 |
| 4 | .90 | 35.24 | 102.22 | .97 | .62 | 351.88 |
| 5 | .90 | 46.98 | 101.52 | 1.00 | .53 | 469.29 |
| 6 | .90 | 56.29 | 101.75 | 1.00 | .56 | 560.27 |
| 7 | .91 | 64.09 | 102.14 | 1.01 | .64 | 633.52 |

## 5. EXAMPLES

For the banking problem, confidence intervals for the mean and the median have been discussed. However, the bank sometimes needs a confidence interval for the total amount of deposits $M_K\mu_K$. As will be seen in Lemma 5.2, the length $d_K$ of the confidence interval can no longer converge to 0 without exhausting the white balls.

The first step will be to prove the asymptotic results necessary. Only the case $N_K$ known will be considered since this is true in the applications.

Lemma 5.1. a) Under the conditions of Lemma 2.7,

$$\left\{N_K\sigma_{K2}(r_K^{-1} - N_K^{-1})^{1/2}\right\}^{-1} (N_K\bar{\xi}_{Kr_K} - M_K\mu_K) \xrightarrow{L} \Phi .$$

b) Under the conditions of Lemma 2.1 or 2.2, if $\hat{M}_{Kr_K}$ is the estimate of $M_K$ based on $r_K$ white balls,

$$\left\{N_K\sigma_{K2}(r_K^{-1} - N_K^{-1})^{1/2}\right\}^{-1} (\hat{M}_{Kr_K}\overline{\eta}_{Kr_K} - M_K\mu_K) \xrightarrow{\mathcal{L}} \Phi \;:$$

Proof: Part a) is simply Lemma 2.7. For part b), Lemmas 2.1 and 2.2 will be needed. Suppose that in $n$ total observations, $m$ white balls are drawn. Let $\rho_K = M_K/N_K$, $\beta_K^{-1} = \{M_K^{-1}(r_K^{-1} - N_K^{-1})^{1/2}\}$, and $z_{Ki} = 1$ if $\xi_{Ki}$ is white and $0$ otherwise. Then,

$$\beta_K\{\hat{M}_{Km}\overline{\eta}_{K_m} - M_K\mu_K\} = \beta_K M_K(\overline{\eta}_{Km} - \mu_K) + \beta_K\mu_K(\hat{M}_{Km} - M_K) + o_\rho(1) \;,$$

where $o_\rho(1)$ means a term converging to zero (0) in probability. Then

$$\beta_K\mu_K(\hat{M}_{Km} - M_K) = \beta_K\mu_K N_K\{\overline{z}_{Kn} - \rho_K\}$$

$$\begin{aligned}\beta_K M_K(\overline{\eta}_{Km} - \mu_K) &= \beta_K N_K(\overline{\xi}_{Kn} - \rho_K\mu_K) + \beta_K N_K\overline{\xi}_{Kn}(n\rho_K/m - 1)\\ &= \beta_K N_K(\overline{\xi}_{Kn} - \rho_K\mu_K) + \beta_K N_K\rho_K\mu_K(n/m)(\rho_K - m/n) + o_\rho(1)\\ &= \beta_K N_K(\overline{\xi}_{Kn} - \rho_K\mu_K) + \beta_K N_K\{\mu_K(\rho_K - \overline{z}_{Kn})\} + o_\rho(1) \;.\end{aligned}$$

These two facts, together with Lemmas 2.1 and 2.2, yield part b).

In Lemma 5.1, since both numerators have the same (asymptotic) variance, it is better to use something like rule (2.10) rather than (2.4); this is opposite to the conclusion reached as regards the mean.

Lemma 5.2. Suppose that as $K \to \infty$,

(i) $d_K N_K^{-\delta} \to \infty$ for some $\delta > 0$.

(ii) $\left(1 + N_K/(bN_K^2\sigma_{K2}/(M_K d_K))^2\right)^{-1} \to \psi_3 \qquad (0 \le \psi_3 < 1)$.

Then, if the rule (2.10) is used with $t_{nK}$ replaced by $N_k t_{nK}$, Theorem 2.1 (when appropriately modified in (2.6) and (2.7)) holds for the population total $M_K\mu_K$.

Note that as $K \to \infty$, $d_K \to 0$ is not allowed; in fact, (ii) requires that $d_K/N_K^{1/2} \to c$ $(0 < c \leq \infty)$.

For the hospital example, if $\xi_{Ki} = z_{Ki}$, then $\mu_K = 1$ and $M_K\mu_K = M_K$, so that Lemmas 5.1 and 5.2 may be used.

ACKNOWLEDGEMENTS: I wish to thank Mr. R. N. Holt of the University of North Carolina Graduate School of Business for the banking example which eventually led to this paper. Mr. Ralph Tower provided the hospital example. I also wish to thank N. L. Johnson and Gordon Simons for helpful discussions.

## REFERENCES


[1] ANSCOMBE, F. J., (1952). Large sample theory of sequential estimation. *Proc. Camb. Phil. Soc.* (48) 600-617.

[2] ANTILLE, A. (1974). A linearized version of the Hodges-Lehmann estimator. *Ann. Statist.* (2) 1308-1313.

[3] BAHADUR, R. R., (1966). A note on quantiles in large samples. *Ann. Math. Statist.* (37) 577-580.

[4] BILLINGSLEY, P. (1968). *Convergence of Probability Measures*, Wiley, New York.

[5] CARROLL, R. J. (1974a). Asymptotically nonparametric sequential selection procedures. *Inst. of Stat. Mimeo Series* #944, Univ. of North Carolina.

[6] CARROLL, R. J. (1974b). Asymptotically nonparametric sequential selection procedures II - robust estimators. *Inst. of Stat. Mimeo Series* #953, Univ. of North Carolina.

[7] CHOW, Y.S., and ROBBINS, H., (1965). On the asymptotic theory of fixed-width sequential confidence intervals for the mean. *Ann. Math. Statist.* (36) 463-467.

[8] FELLER, W. (1968). *An Introduction to Probability Theory and its Applications.* (1) (Third Edition), Wiley, New York.

[9] FREEMAN, P. R. (1973). A numerical comparison between sequential tagging and sequential recapture. *Biometrika* (60) 499-508.

[10] GEERTSEMA, J. C. (1970). Sequential confidence intervals based on rank tests. *Ann. Math. Statist.* (41) 1016-1026.

[11] GEERTSEMA, J. C. (1972). Nonparametric sequential procedures for selecting the best of k populations. *J. Am. Statist. Assoc.* (67) 614-616.

[12] GROVES, J. E., and PERNG, S. K. (1973). A class of sequential procedures for partitioning a set of populations with respect to a control. *IMS Bull.* (2) 96.

[13] GLESER, L. J. (1965). On the asymptotic theory of fixed-size sequential confidence bounds for linear regression parameters. *Ann. Math. Statist.* (36) 463-467.

[14] HÁJEK, J. (1960). Limiting distributions in simple random sampling from a finite population. *Magyar Tud. Akad. Mat. Kutató Int. Közl.* (5) 361-374.

[15] HOEFFDING, W. (1963). Probability inequalities for sums of bounded random variables. *J. Am. Stat. Assoc.* (58) 13-30.

[16] KALLENBERG, O. (1974). A note on the asymptotic equivalence of sampling with and without replacement. *Ann. Statist.* (2) 819-821.

[17] MADOW, W. G. (1948). On the limiting distributions of estimates based on samples from finite universes. *Ann. Math. Statist.* (14) 535-545.

[18] PERNG, S. K., and TONG, Y. L. (1974). A sequential solution to the inverse linear regression problem. *Ann. Statist.* (2) 535-539.

[19] ROBBINS, H., SIMONS, G., and STARR, N. (1967). A sequential analogue of the Behrens-Fisher problem. *Ann. Math. Statist.* (38) 1384-1391.

[20] ROBBINS, H., SOBEL, M., and STARR, N., (1968). A sequential procedure for selecting the largest of k means. *Ann. Math. Statist.* (39) 88-92.

[21] ROSEN, B. (1964). Limit theorems for sampling from a finite population. *Ark. Mat.* (5) 383-424.

[22] SAMUEL, E. (1968). Sequential maximum likelihood estimation of the size of a population. *Ann. Math. Statist.* (39) 1057-1068.

[23] SRIVASTAVA, M. S. (1966). Some asymptotically efficient sequential procedures for ranking and slippage problems. *J. Roy. Statist. Soc. Ser. B.* (28) 370-380.

[24] SRIVASTAVA, M. S. (1967). On fixed-width confidence bounds for regression parameters and mean vector. *J. Roy. Statist. Soc. Ser. B* (29) 132-140.

[25] STARR, N. (1966). The performance of a sequential procedure for the fixed-width interval estimation of the mean. *Ann. Math. Statist.* (37) 36-50.

[26] STEIN, C. (1945). A two-sample test for a linear hypothesis whose power is independent of the variance. *Ann. Math. Statist.* (16) 243-258.

[27] STEIN, C. (1949). Some problems in sequential estimation. *Econometrics* (17) 77-78.

[28] TAGA, Y. (1964). A note on the degree of normal approximation to the distribution function of the mean of samples from finite populations. *Ann. Inst. Stat. Math.* (16) 427-430.