AD-757 106

SEQUENTIAL ESTIMATION OF THE LARGEST
NORMAL MEAN WHEN THE VARIANCE IS UNKNOWN

Saul Blumenthal

Cornell University

DEPARTMENT
OF
OPERATIONS RESEARCH

DDC
RECEIVED
MAR 16 1973
D

COLLEGE OF ENGINEERING
CORNELL UNIVERSITY
ITHACA, NEW YORK 14850

DEPARTMENT OF OPERATIONS RESEARCH
COLLEGE OF ENGINEERING
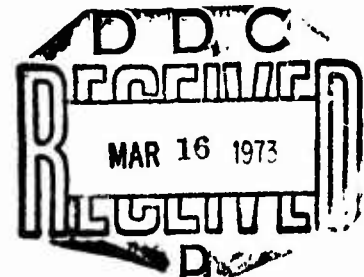CORNELL UNIVERSITY
ITHACA, NEW YORK

TECHNICAL REPORT NO. 164

January 1973

SEQUENTIAL ESTIMATION OF THE LARGEST
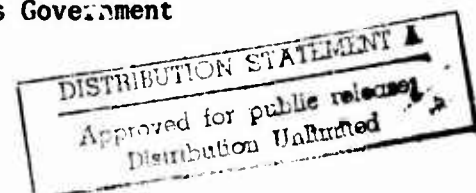NORMAL MEAN WHEN THE VARIANCE IS UNKNOWN

by

Saul Blumenthal*

*Cornell University and New York University

DOCUMENT CONTROL DATA · R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)*	2a. REPORT SECURITY CLASSIFICATION
Cornell University Operations Research Department Ithaca, New York 14850	UNCLASSIFIED
	2b. GROUP

3. REPORT TITLE

SEQUENTIAL ESTIMATION OF THE LARGEST NORMAL MEAN WHEN THE VARIANCE IS UNKNOWN

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

Technical Report

5. AUTHOR(S) *(First name, middle initial, last name)*

Saul Blumenthal

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
January, 1973		10

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
N00014-67-A-0077-0020	
b. PROJECT NO.	Technical Report No. 164
c.	9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*
d.	

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Mathematics Program Office of Naval Research Arlington, Virginia 22217

13. ABSTRACT

Given n observations from each of k populations whose distributions differ by a location parameter, the value of the largest parameter is to be estimated using the largest value of the k sample means. It is desired to design a sampling rule which guarantees that the Mean Squared Error (M.S.E.) of the estimate does not exceed a given bound when the distributions have a common but unknown scale parameter.

A sequential sampling scheme is devised based on an estimate of the scale parameter and a "least favorable" configuration of the location parameters. The sample size characteristics of the sampling plan studied under mild restrictions on the distributions involved. The M.S.E. of the resulting estimator is studied under the additional assumption of normality. A brief discussion is given of an alternate sequential plan which uses information in the sample regarding the configuration of the location parameters.

14	KEY WORDS	LINK A		LINK B		LINK C	
		ROLE	WT	ROLE	WT	ROLL	WT
	mathematical statistics						
	ranking and selection						
	sequential estimation						
	ordered parameters						

DD FORM 1473 (BACK)
1 NOV 65

II

S/N 0101-807-6821

UNCLASSIFIED
Security Classification

A-31409

1. INTRODUCTION AND SUMMARY

Given observations from k populations whose distributions are the same except for the value of a location parameter, it is desired to estimate the value of the largest of the k location parameters. A natural estimator for this problem is the largest of the k sample means. Measuring the performance of the estimator by its mean squared error (M.S.E.), it is possible to choose a sample size so that the M.S.E. is bounded by a given constant regardless of the values of the location parameters. If the common distribution function involves an unknown scale parameter, however, then no such choice of sample size is possible, but a sequential rule of the type studied by Chow and Robbins [4] can be constructed. Its limiting properties will be studied here.

Point estimation of the largest parameter for known variances has been studied before by Blumenthal and Cohen [2], [3] and Dudewicz [5], and interval estimation by Saxena and Tong [7], and Dudewicz and Tong [6]. Interval estimation with unknown variance has been examined by Tong [9], [10]. The use of the largest observed mean has no optimality properties and in fact for $k = 2$ is known to be inadmissible [2] for M.S.E. However, since the analytic form of the competing estimators is rather involved, it is suspected that this natural estimator would be widely used and its properties would therefore be of interest. In the next section, the sample size for the sequential stopping rule is shown to behave well under mild restrictions on the distributions. The behavior of the M.S.E. of the estimator formed from the sequential stopping rule is studied only for normal distributions.

1

The risk function depends on the differences between the largest parameter and the other parameters. In Section 2 we derive a conservative procedure based on the maximum risk as a function of these differences. In [1], a sequential procedure has been studied which attempts to capitalize on the information in the sample about these differences, when the variance is known. In Section 3, this procedure is extended to the case of unknown variance, and a few of its properties are outlined.

2. CONSERVATIVE PROCEDURES

Let X_i $(1 \leq i \leq k)$ be an observation from the ith population having c.d.f. $F(x, \theta_i)$ (θ_i a real number). We assume that the family of c.d.f.'s $F(x, \theta)$ satisfy

(2.1)   (i) $F(x, \theta) = F(x-\theta)$, where $F(x)$ is a c.d.f.

      (ii) $F(x) = 1 - F(x)$ or $f(x) = f(-x)$ for all x,

      (iii) The family of density functions $\{f(x-\theta)\}$ has monotone likelihood ratio.

      (iv) $\int_{-\infty}^{\infty} xf(x)dx = 0$.

Define

(2.2) $X^* = \max(X_1, \ldots, X_k); \quad \theta^* = \max(\theta_1, \ldots, \theta_k)$.

The ordered θ's will be denoted $\theta_{[1]} \leq \cdots \leq \theta_{[k]}$. The M.S.E. or risk function of X^* is easily seen to be

$$(2.3) \qquad R(X^*; \theta_1, \ldots, \theta_k) = E(X^* - \theta^*)^2 = \int_{-\infty}^{\infty} (x - \theta^*)^2 d \prod_{i=1}^{k} F(x - \theta_i) \ .$$

To choose an appropriate sample size, this risk must be studied in further detail which is done below.

Lemma 1. Let $F(x, \theta)$ satisfy condition (2.1). Then

$$(2.4) \qquad \sup_{(\theta_1, \ldots, \theta_k)} R(X^*; \theta_1, \ldots, \theta_k) = R(X^*; \theta^*, \theta^*, \ldots, \theta^*) = \int_{-\infty}^{\infty} x^2 dF^k(x)$$

and

$$(2.5) \qquad \lim_{\theta_{[k]} - \theta_{[k-1]} \to \infty} R(X^*; \theta_1, \ldots, \theta_k) = \int_{-\infty}^{\infty} x^2 dF(x) \ .$$

Proof: Write

$$R(X^*; \theta_1, \ldots, \theta_k) = \int_{-\infty}^{\theta^*} (x - \theta^*)^2 d\prod F(x - \theta_i) + \int_{\theta^*}^{\infty} (x - \theta^*)^2 d\prod F(x - \theta_i)$$

$$(2.6)$$

$$= \int_{0}^{\infty} x^2 d[\prod F(\Delta_i + x) - \prod F(\Delta_i - x)]$$

where $\Delta_i = \theta^* - \theta_i \geq 0$. Since x^2 is monotone increasing over $(0, \infty)$, the expectation will be dominated by

$$\int_{0}^{\infty} x^2 dH(x)$$

if $H(x) \leq [\prod F(\Delta_i + x) - \prod F(\Delta_i - x)]$ (see, for example, Lemma 4.1.2 of Dudewicz [5], or integrate by parts). That $H(x)$ can be taken as $[F^k(x) - F^k(-x)]$

is a consequence of Theorem 1 of Saxena and Tong [7]. Equation (2.5) follows by letting all $\Delta_i \to \infty$ (except one which is fixed at zero).

Note that for $k > 2$, the risk when all the θ's are equal, (2.4), exceeds the risk when the difference between θ^* and the others becomes infinite, (2.5). (See Lemma 4.2.3 of [5].) Regardless of k, the latter risk is just the variance of a single X_i as would be expected. In the case $k = 2$, the two risks (2.4) and (2.5) are equal as was shown by Blumenthal and Cohen [2] without the assumption of monotone likelihood ratio.

Suppose that each X_i is in fact $X_{i,n}$ and is the average of n independent identically distributed observations say Y_{ij} $(1 \le j \le n)$, so that $F(x)$ is really $F_n(x)$. A conservative procedure which will guarantee that the risk of estimating θ^* by X^*, is no greater than r (given) for any parameter configuration is to choose n as the smallest integer such that

$$\int_{-\infty}^{\infty} x^2 dF_n^k(x) \le r .$$

Suppose in addition that for each n, $F_n(x)$ is of the form $G_n(x\sqrt{n}/\sigma)$ where σ^2 is the variance of a single observation. Assume that in addition to (2.1),

$$(2.7) \qquad \int_{-\infty}^{\infty} x^2 g(x)dx = 1 .$$

Then the risk function is

$$(2.8) \quad R(X^*; \theta_1,\ldots,\theta_k) = (\sigma^2/n)\int y^2 d\Pi G_n(y + (\Delta_i \sqrt{n}/\sigma)) \le (\sigma^2/n)\int y^2 dG_n^k(y) = A_n\sigma^2/n.$$

The conservative choice of n will then be the smallest $n \ge (A_n\sigma^2/r)$.

If the value of σ^2 is not known, a reasonable way to proceed is to form an estimate of σ^2 by defining

$$(2.9a) \qquad S^2_{ni} = (1/n-1) \sum_{j=1}^{n} (Y_{ij} - X_{i,n})^2 \qquad 1 \le i \le k$$

and

$$(2.9b) \qquad S^2_n = (1/k) \sum_{i=1}^{k} S^2_{ni} .$$

Clearly S^2_n is a consistent estimator of σ^2, and a sequential stopping rule can be formulated as follows: Let N be the first n such that

$$(2.10) \qquad n \ge (A_n S^2_n/r) .$$

(If desired, n may be restricted to exceed a given value m, and $(\frac{1}{n} + S^2_n)$ should be substituted for S^2_n if $G_n(x)$ is discrete.) Then θ^* will be estimated by $X^*_N = \max_{1 \le i \le k} X_{i,n}$.

The following theorem giving the behavior of N is a restatement of the main Theorem of Chow and Robbins [4] which applies here directly. (In [4], S^2_n is defined by (2.9b) with $k = 1$. It is easy to check that their proofs are not affected by having $k > 1$.)

<u>Theorem 1.</u> Suppose $\lim_{n \to \infty} A_n = A < \infty$. Let $\lambda = (\sigma^2/r)$.

$$(2.11a) \qquad P_{\mu, \sigma^2}[N < \infty] = 1 \quad \text{for every } \mu \text{ and } \sigma^2$$

$$(2.11b) \qquad \lim_{\lambda \to \infty} N = \infty, \quad \text{a.s.}$$

(2.11c)
$$\lim_{\lambda \to \infty} (N/A\lambda) = 1 \quad \text{a.s.}$$

(2.11d)
$$\lim_{\lambda \to \infty} (E(N)/A\lambda) = 1.$$

Note: Since $\lim_{n \to \infty} G_n(x) = \phi(x)$, generally (though not necessarily always),

(2.11e)
$$A = \int_{-\infty}^{\infty} x^2 d\phi^k(x) = A^* .$$

For general distributions, it is not possible to obtain much information about the risk function, $E(X_N^* - \theta^*)^2$. Since it is not true that X_n^* forms a martingale, theorems relating to the moments of randomly stopped martingales cannot be applied. A lemma of Tong [10] shows that X_N^* converges stochastically to θ^* as $r \to 0$, but little else can be said in the general case.

Hereafter, assume that all distributions are normal. In this case the distribution of $\sqrt{N}(X_N^* - \theta^*)/\sigma$ is independent of N. In particular

$$E\{ (X_N^* - \theta^*)^2 | N \} = \{ (\sigma^2/N) \int y^2 d\Pi\phi(y + (\Delta_i \sqrt{N}/\sigma)) \}$$

(2.12)
$$\leq \{1/N\}A^* \sigma^2 .$$

The risk of X_N^* is thus related to $E(N^{-1})$. From (2.12), we see that $(1/r)E\{ (X_N^* - \theta^*)^2 | N \} \leq \{1/N\}A^*\lambda$. Thus from (2.11c) we can conclude that

(2.13)
$$\lim_{\lambda \to \infty} (1/r)E\{ (X_N^* - \theta^*)^2 | N \} \leq 1 \quad \text{a.s.}$$

This limit is conservative as can be seen by the following argument. Consider

$$\int y^2 d\Pi \Phi(y + (\Delta_i \sqrt{n}/\sigma)) \quad .$$

Applying the bounded convergence theorem to this integral we see that (recalling that $\Delta_i = 0$ for exactly one i if $\theta_{[k]}$ is unique)

$$(2.14) \qquad \lim_{n \to \infty} \int y^2 d\Pi \Phi(y + (\Delta_i \sqrt{n}/\sigma)) = \int y^2 d\Phi(y) = 1 \ .$$

By (2.11b), it is seen that (2.14) implies that

$$(2.15) \qquad \lim_{r \to 0} \int y^2 d\Pi \Phi(y + (\Delta_i \sqrt{N}/\sigma)) = 1, \quad \text{a.s.}$$

Thus, using (2.15) and (2.11c), we conclude from (2.12) that

$$(2.16) \qquad \lim_{r \to 0} (1/r) E\{(X_N^* - \theta^*)^2 | N\} = (1/A^*) \quad \text{a.s.} \quad \text{(when } \theta_{[k]} \text{ is unique) .}$$

When $\theta_{[k]}$ is not unique, the above can be modified, as follows. Suppose $\theta^* = \theta_{[k]} = \theta_{[k-1]} = \cdots = \theta_{[k-s+1]}$. Then

$$(2.17) \qquad \lim_{r \to 0} (1/r) E\{(X_N^* - \theta^*)^2 | N\} = (1/A^*) \int y^2 d\Phi^s(y), \quad \text{a.s.}$$

By Corollary 4.2.4 [5], the right sides of (2.16) and (2.17) are strictly less than one when $k > 2$, $s < k$, and are equal to one for $k = 2$, or $s = k$. Note that (2.17) depends only on the multiplicity of θ^* and otherwise is true for any parameter configuration.

If on the other hand, we regard r as fixed and examine large σ values, from (2.11c), it is seen that

$$(2.18) \quad \lim_{\sigma \to \infty} \int y^2 d\Pi\phi(y + (\Delta_i \sqrt{N}/\sigma)) = \int y^2 d\Pi\phi(y + (\Delta_i \sqrt{A^*/r})), \quad \text{a.s.},$$

and therefore using (2.11c) again in (2.12)

$$(2.19) \quad \lim_{\sigma \to \infty} (1/r)E\{(X_N^* - \theta^*)^2 | N\} = (1/A^*)\int y^2 d\Pi\phi(y + (\Delta_i \sqrt{A^*}/\sqrt{r})) \leq 1, \quad \text{a.s.}$$

Note that different limits are obtained for the conditional risk in (2.16) and (2.19) as $\lambda \to \infty$ depending on whether $\sigma \to \infty$ for fixed r or $r \to 0$ for fixed σ.

Next, the unconditional risk will be investigated, namely $(1/r)EE\{(X_N^* - \theta^*)^2 | N\}$. From (2.12) it is seen that this risk is related to $E(N^{-1})$. The following represents a slight generalization of Theorem 3_* of Starr [8].

Theorem 2. Let the minimum sample size in (2.10) be denoted as m. Let A_n in (2.10) be given by A^* (see (2.11e)) for each n. Let $\omega > 0$ be given.

$$(2.20) \quad \lim_{\lambda \to \infty} (A^*\lambda)^\omega E(N^{-\omega}) = 1 \qquad \text{for} \quad m > 1 + (2\omega/k)$$

$$= 1 + (\omega^{\omega-1}/\Gamma(\omega)) \qquad \text{for} \quad m = 1 + (2\omega/k)$$

$$= \infty \qquad \text{for} \quad m < 1 + (2\omega/k)$$

The proof is outlined in the Appendix.

The following lemma is needed to obtain the limit of the unconditional risk from (2.20) and either (2.15) or (2.18).

Lemma 2.   Let M be a positive random variable depending on a parameter t
with the properties

(2.21a) $$\lim_{t\to\infty} M = 1 \quad a.s.$$

(2.21b) $$\lim_{t\to\infty} EM = 1 \quad .$$

Let G(x) be a bounded function satisfying either

(2.22a) $$\lim_{x\to 1} G(x) = G^*$$

or

(2.22b) $$\lim_{x\to\infty} G(x) = G^* \quad .$$

Then in case (2.22a)

(2.23a) $$\lim_{t\to\infty} E(MG(M)) = G^*$$

or in case (2.22b)

(2.23b) $$\lim_{t\to\infty} E(MG(t/M)) = G^* \quad .$$

Proof: Consider (2.23a). We shall show that

(2.24) $$\lim_{t\to\infty} E[M(G(M) - G^*)] = 0 \quad .$$

By (2.22a), given $\epsilon > 0$, there is a $\delta = \delta(t)[\delta(t) \to 0$ as $\epsilon \to 0]$ s.t.

$$|G(M) - G^*| < \varepsilon \quad \text{for} \quad 1-\delta < M < 1+\delta,$$

and

$$(2.25) \qquad \left| \int_{1-\delta}^{1+\delta} M(G(M) - G^*) \right| < (1+\delta)\varepsilon .$$

By the boundedness of $G(x)$, $|G(x) - G^*| < B$ (say) and we obtain

$$(2.26) \qquad \left| \int_{0}^{1-\delta} M(G(M) - G^*) \right| < BP[M < 1-\delta]$$

as well as

$$(2.27) \qquad \left| \int_{1+\delta}^{\infty} M(G(M) - G^*) \right| < B \int_{1+\delta}^{\infty} M .$$

By (2.21a), for ε (and δ) fixed there is a $t_1(\varepsilon)$ s.t.

$$(2.28a) \qquad P[M < 1-\delta] < \varepsilon \quad \text{for} \quad t > t_1(\varepsilon)$$

$$(2.28b) \qquad P[M > 1+\delta] < \varepsilon \quad \text{for} \quad t > t_1(\varepsilon) .$$

By (2.21b), for ε fixed there is a $t_2(\varepsilon)$ s.t.

$$(2.29) \qquad \int_{0}^{\infty} M < 1+\varepsilon \quad \text{for} \quad t > t_2(\varepsilon) .$$

Thus for $t > t_3(\varepsilon) = \max(t_1(\varepsilon), t_2(\varepsilon))$,

$$(2.30) \quad \int_{1+\delta}^{\infty} M < 1+\varepsilon - \int_{1-\delta}^{1+\delta} M - \int_{0}^{1-\delta} M < 1+\varepsilon - (1-\delta)(1-2\varepsilon) = 3\varepsilon + \delta - 2\varepsilon\delta.$$

Combining (2.25), (2.26), (2.28a), (2.27) and (2.30)

(2.31) $\left|\int_0^\infty M(G(M) - G^*)\right| < B[4\epsilon + \delta - 2\epsilon\delta] + \epsilon(1+\delta)$ whenever $t > t_3(\epsilon)$,

thus proving (2.23a).

To prove (2.23b), note from (2.22b) that there is a $t_4(\epsilon)$ such that (for ϵ given)

$$|G(t/M) - G^*| < \epsilon \quad \text{if} \quad M < 1+\delta \quad \text{and} \quad t > t_4(\epsilon),$$

so that

(2.32) $\left|\int_0^{1+\delta} M(G(t/M) - G^*)\right| < \epsilon(1+\delta) \quad \text{for} \quad t > t_4(\epsilon)$.

Also, as before

(2.33) $\left|\int_{1+\delta}^\infty M(G(t/M) - G^*)\right| < B \int_{1+\delta}^\infty M$.

Using (2.33) and (2.32) with (2.30) gives the desired result and completes the proof. As a consequence of the above we have

Corollary 1:   Let m in (2.10) exceed $1 + (2/k)$. Then

(2.34) $\lim_{\sigma \to \infty} (1/r)E(X_N^* - \theta^*)^2 = (1/A^*) \int_{-\infty}^\infty x^2 d\Pi\Phi(x + \Delta_i\sqrt{A^*/r}) < 1$

and

(2.35) $\lim_{r \to 0} (1/r)E(X_N^* - \theta^*)^2 = (1/A^*) \leq 1$.

<u>Proof</u>: Let $M = (A*\sigma^2/rN)$ and apply (2.11c), (2.11d), and (2.20). For

(2.34), let $t = \sigma^2$ and use (2.12), writing $(\Delta_i\sqrt{N}/\sigma)$ as $(\Delta_i\sqrt{A*/rM})$ and

use (2.18), with (2.23a). For (2.35), let $t = (1/r)$ and use (2.12), writing

$(\Delta_i\sqrt{N}/\sigma)$ as $(\Delta_i\sqrt{A*t/M})$, and use (2.14) with (2.23b).

3. ELIMINATING ALL NUISANCE PARAMETERS

For normal distributions, a natural sequential procedure which attempts
to use all of the sample information regarding nuisance parameters would be
a combination of the rule studied in Section 2 with the rule considered in
Blumenthal [1] for the case of known σ^2. In particular, such a procedure
would use the stopping rule: stop for the first n such that

$$(3.1) \qquad r \geq (S_n^2/n) \int_{-\infty}^{\infty} y^2 d\Pi\phi[y + (\hat{\Delta}_{i,n}\sqrt{n}/S_n)]$$

where S_n^2 is given by (2.9) and $\hat{\Delta}_{i,n}$ is a consistent estimate of Δ_i, such
as $(X* - X_{i,n})$. When σ^2 is not known, it is not possible to construct a
two sample procedure whose sample size will be very nearly the same as that
of the sequential procedure, as was done in [1]. Only the sequential proce-
dure will be studied here.

As in [1], an alternative stopping rule will be defined in order to
avoid difficulties which may be caused by the behavior of the function

$$(3.2) \qquad H(x_1,\ldots,x_{k-1}) = \int_{-\infty}^{\infty} y^2 d\prod_{i=1}^{k} \phi(y + x_i), \qquad (x_k \equiv 0) .$$

Namely, let $\tilde{n}$ be the solution of

(3.3) $$\tilde{n}^2 = H(\hat{\Delta}_{1,n}\tilde{n}\sqrt{A^*}/\sqrt{r}, \ldots, \hat{\Delta}_{k-1,n}\tilde{n}\sqrt{A^*}/\sqrt{r})$$

(where we label so that $\hat{\Delta}_{k,n} \equiv 0$), and let

(3.4) $$\tilde{n} = \lceil \tilde{n}^2 S_n^2 A^*/r \rceil$$

where $\lceil x \rceil$ is the smallest integer not less than x. Sampling is stopped at the first n such that $n \geq \tilde{n}$. For a discussion of the relation of this stopping rule to that of (3.1), see [1].

Since S_n^2 is independent of the $X_{i,n}$, hence of the $\hat{\Delta}_{i,n}$, the limiting processes of Theorems 1 of this paper and [1] can be combined to give:

Theorem 3: If all $\Delta_i > 0$, $1 \leq i \leq k-1$, then for the stopping rule described after (3.4), under the assumptions of Theorem 1 of this paper, and Theorem 1 of [1],

(3.5) $$\lim_{r \to 0} (N/\lambda) = \lim_{r \to 0} E(N/\lambda) = 1 .$$

Proof: The stopping rule is of the form (2.10), with

$$A_n = H(\hat{\Delta}_{1,n}\sqrt{n}/S_n, \ldots, \hat{\Delta}_{k-1,n}\sqrt{n}/S_n) .$$

At stopping, $(\sqrt{N}/S_N) > \sqrt{A_N/r} > \sqrt{R^*/r}$, so that $P\{\hat{\Delta}_{i,N}\sqrt{N}/S_N > T, \ 1 \leq i \leq k-1\}$ can be made arbitrarily close to unity for fixed T by choosing r sufficiently small. Thus A_N converges with probability one to unity as r decreases, and (3.5) can be derived as though A_n converged deterministically to unity. This completes the proof.

If some $\Delta_i = 0$, or if the Δ_i are proportional to $\sqrt{r}$, or if the Δ_i are fixed and positive, but the limit is taken as $\sigma \to \infty$ then as in Theorem 2 of [1], there will be no probability one limit for (N/λ) as $\lambda \to \infty$, but it can be seen that (N/λ) will have some limiting distribution whose mass is concentrated on (R^*, A^*) (in the limit, it will behave as though σ were known), where $R^* = \inf H(x_1, \ldots, x_{k-1})$.

References

[1] Blumenthal, S. "Sequential Estimation of the Largest Normal Mean When the Variance is Known," Technical Report No. $/7/$, Dept. of Operations Research, Cornell University, (1973).

[2] Blumenthal, S. and Cohen, A. "Estimation of the Larger Translation Parameter," Ann. Math. Statist. 39, No. 2 (1968), 502-516.

[3] Blumenthal, S. and Cohen, A. "Estimation of the Larger of Two Normal Means," J. Amer. Statist. Assoc. 63, No. 32 (1968), 861-876.

[4] Chow, Y. S. and Robbins, H. "On the Asymptotic Theory of Fixed-Width Sequential Confidence Intervals for the Mean," Ann. Math. Statist. 36, No. 2 (1965), 457-462.

[5] Dudewicz, E. "Estimation of Ordered Parameters," Technical Report No. 60, Dept. of Operations Research, Cornell University, (1969).

[6] Dudewicz, E. and Tong, Y. L., "Optimal Confidence Intervals for the Largest Location Parameter," Statistical Decision Theory and Related Topics, Academic Press, New York (1971), 363-376.

[7] Saxena, K.M.L. and Tong, Y. L. "Interval Estimation of the Largest Mean of k Normal Populations with Known Variances," J. Amer. Statist. Assoc. 64, No. 325 (1969), 296-299.

[8] Starr, N. "On the Asymptotic Efficiency of a Sequential Procedure for Estimating the Mean," Ann. Math. Statist. 37, No. 5 (1966), 1173-1185.

[9] Tong, Y. L. "Multi-stage Interval Estimations of the Largest Mean of k Normal Populations," J. Roy. Statist. Soc. Ser. B 32, No. 2 (1970), 272-277.

[10] Tong, Y. L. "An Asymptotically Optimal Sequential Procedure for the Estimation of the Largest Mean," Unpublished Manuscript.

15

Proof of Theorem 2:   Since the proof given by Starr [8] of his Theorem 3 can be used with only trivial modification, our proof will be given in outline form only.

Define

$$\text{(A.1)} \qquad n^O = A^*\lambda, \quad V_n = k(n-1)S_n^2/\sigma^2, \quad \ell(n,\lambda) = kn(n-1)/A^*\lambda .$$

The stopping rule is:

$$\text{(A.2)} \qquad \text{stop for the smallest } n \geq m \text{ s.t. } V_n \leq \ell(n,\lambda),$$

where V_n has a chi square distribution with $k(n-1)$ degrees of freedom. Let $\epsilon > 0$ be given and define

$$\text{(A.3)} \qquad \alpha = (1-\epsilon)^{1/\omega}n^O, \quad \beta = (1+\epsilon)^{1/\omega}n^O$$

$$\Pi_1 = m^{-\omega}P(N = m)$$

$$\Pi_2 = \beta^{-\omega}P(m < N \leq \beta)$$

$$\Pi_3 = \sum_{n=m+1}^{\alpha} n^{-\omega}P(N = n)$$

$$\Pi_4 = \alpha^{-\omega}P(N \geq \alpha)$$

Noting that

$$P(V_m \leq \ell(m,\lambda)) = \left\{ \Gamma\left(\frac{k(m-1)}{2}\right)2^{k(m-1)/2} \right\}^{-1} \int_0^{\ell(m,\lambda)} x^{\frac{k(m-1)}{2}-1} e^{-x/2} dx$$

$$\geq \left\{ k(m-1)\Gamma\left(\frac{k(m-1)}{2}\right)2^{\frac{k(m-1)}{2}-1} \right\}^{-1} e^{-\ell(m,\lambda)/2}(\ell(m,\lambda))^{k(m-1)/2}$$

16

and using

$$EN^{-\omega} \geqq \Pi_1 + \Pi_2,$$

Starr's argument yields,

$$(A.4) \quad \liminf_{\lambda \to \infty}(n^o)^\omega EN^{-\omega} \geqq a(m,\omega)\lim_{\lambda \to \infty}\left\{(n^o)^\omega(\ell(m,\lambda))^{\frac{k(m-1)}{2}}\right\} + 1-\delta \quad (C<\delta = \delta(\varepsilon) < 1)$$

where

$$(A.5) \quad a(m,\omega) = \left\{m^\omega k(m-1)\Gamma(\frac{k(m-1)}{2})2^{\frac{k(m-1)}{2}-1}\right\}^{-1}.$$

The next step uses

$$(A.6) \quad EN^{-\omega} \leqq \Pi_1 + \Pi_3 + \Pi_4 .$$

First, it is easily checked that

$$(A.7) \quad \Pi_1 \leqq a(m,\omega)(\ell(m,\lambda))^{k(m-1)/2} .$$

Next,

$$(A.8) \quad \Pi_3 \leq \sum_{m+1}^{\alpha}\left[n^\omega\Gamma(\frac{k(n-1)}{2})2^{\frac{k(n-1)}{2}}\right]^{-1}\int_0^{\ell(n,\lambda)}\left[x^{\frac{k(n-1)+(\rho+1)}{2}-1}e^{-x/2}\right]x^{-(\rho+1)/2}dx .$$

Define

$$(A.9) \quad h(n,\lambda) = (\ell(n,\lambda)/k(n-1)) = n/A^*\lambda = n/n^o .$$

By the definition of α, (A.3),

$$(A.10) \quad h(n,\lambda) \leqq (1-\varepsilon)^{1/\omega} = 1-\xi \quad (0<\xi = \xi(\varepsilon) < 1) \quad \text{for all} \quad n \leqq \alpha,$$

so that,

(A.11) $\ell(n,\lambda) \leq k(n-1)(1-\xi) \leq k(n-1) + (\rho-1)$ for $n \geq 2$, $0 < 1-\xi < \rho < 1$.

From the fact that $\left[x^{\frac{k(n-1) + \rho-1}{2}} e^{-x/2} \right]$ achieves its maximum at $k(n-1) + \rho-1$, and (A.11), the integral in (A.8) is bounded above by

(A.12) $$(\ell(n,\lambda))^{\frac{k(n-1) + \rho-1}{2}} e^{-\ell(n,\lambda)/2} \int_0^{\ell(n,\lambda)} x^{-(\rho+1)/2} \, dx .$$

Furthermore,

(A.13) $$\Gamma(\frac{k(n-1)}{2}) \geq [\frac{k(n-1)}{2}]^{k(n-1)-2/2} e^{-k(n-1)/2}$$

so that the inequality (A.8) becomes

$$\Pi_3 \leq (1-\rho)^{-1} \sum_{m+1}^{\alpha} \{n^{-\omega}[k(n-1)]^{2-k(n-1)/2} e^{k(n-1)/2}\} e^{-\ell(n,\lambda)/2} (\ell(n,\lambda))^{k(n-1)/2}$$

(A.14)

$$= k(1-\rho)^{-1} \sum_{m+1}^{\alpha} \{n^{-\omega}(n-1)\} [e^{(1-h(n,\lambda))} h(n,\lambda)]^{k(n-m-1)/2} [e^{(1-h(n,\lambda))} h(n,\lambda)]^{km/2} .$$

Denote

(A.15) $$\ell(n,\lambda) = h(n,\lambda) e^{(1-h(n,\lambda))} .$$

From (A.9) and (A.10)

(A.16) $$(\Delta(n,\lambda))^{km/2} \leq e^{km/2} (n/A^*\lambda)^{km/2}$$

and

(A.17) $$\Delta(n,\lambda) \leq (1-\xi)e^{\xi} < 1 \qquad n \leq \alpha, \quad 0 < \xi < 1 .$$

Thus, using (A.16) and (A.17) in (A.14),

$$(A.18) \quad \Pi_3 \leq (A^*\lambda)^{-km/2} e^{km/2} k(1-\rho)^{-1} \sum_{m+1}^{\alpha} [n^{-\omega}(n-1)n^{km/2}](\Delta(n,\lambda))^{k(n-m-1)/2} .$$

From (A.17) and the ratio rule for series convergence, we obtain

$$(A.19) \quad \Pi_3 \leq G\lambda^{-km/2}$$

where G is a constant independent of λ. Combining (A.6), (A.7), and (A.19), gives

$$(A.20) \quad (n^o)^\omega EN^{-\omega} \leq a(m,\omega)(n^o)^\omega(\ell(m,\lambda))^{k(m-1)/2} + (n^o)^\omega \lambda^{-km/2}G + (1-\epsilon)^{-1}P(N \geq \alpha).$$

Definitions (A.1) imply that

$$(A.21) \quad \lim_{\lambda \to \infty}(n^o)^\omega(\ell(m,\sigma))^{k(m-1)/2} < \infty \implies \lim_{\lambda \to \infty}[(n^o)^\omega \lambda^{-km/2}G] = 0 .$$

Using (A.20), (A.21) and (2.11c) gives

$$(A.22) \quad \limsup_{\lambda \to \infty}(n^o)^\omega EN^{-\omega} \leq a(m,\omega)\lim_{\lambda \to \infty}[(n^o)^\omega(\ell(m,\lambda))^{k(m-1)/2}] + 1+\delta',$$
$$(0 < \delta' = \delta'(\epsilon) < 1) .$$

Combining (A.22) and (A.4) gives

$$(A.23) \quad \lim_{\lambda \to \infty}(n^o)^\omega EN^{-\omega} = 1 + a(m,\omega)\lim_{\lambda \to \infty}[(n^o)^\omega(\ell(m,\lambda))^{k(m-1)/2}] .$$

From (A.1), we have

$$(A.24) \quad (n^o)^\omega(\ell(m,\lambda))^{k(m-1)/2} = [km(m-1)]^{k(m-1)/2}(A^*\lambda)^{\omega-(k(m-1)/2)} .$$

Using (A.24) and (A.5) in (A.23) yields (2.20), completing the proof.