SEQUENTIAL ESTIMATION AND CLOSED SEQUENTIAL DECISION PROCEDURES[1]

By Edward Paulson

Queens College

1. Introduction and summary. In Chapter 10 of his classical work on sequential analysis [8], Wald started a program for dealing with multiple decision problems based on sequential estimation, and this work was to some extent further developed by Stein [7]. However, as far as the present writer is aware, this program was never completed.

In a recent paper [6] a sequential procedure for estimating the mean of a normal distribution was given and the results applied to the problem of deciding which of k non-overlapping intervals contains the mean. In the present paper, the main results of [6] are first derived in a different manner, using a slight modification of a procedure given by Wald (see Chapter 10 of [8]). This new derivation lends itself to dealing with other situations, and sequential confidence limits are worked out explicitly for two other cases, namely, for the variance and for the ratio of variances of normal distributions.

These results are then applied to get closed decision procedures for a number of decision problems, including (1) testing a hypothesis regarding the mean of a normal distribution against either one-sided or two-sided alternatives; (2) comparing the means of k experimental categories with a standard or control; (3) testing a hypothesis about the variance of a normal distribution; (4) deciding which of k non-overlapping intervals contains the variance; (5) testing a hypothesis about the ratio of variances; (6) comparing the variances of k experimental categories with a standard or control. In addition, a brief discussion of "mixed" problems, where we are concerned with finding a confidence interval for the parameter after a decision has been reached is given in Sections 3.1 and 3.3.

In some of these problems open sequential solutions (in which there is no upper bound to the number of measurements required to reach a decision) are already known which have some optimum properties. However, for administrative and other reasons it is often desirable to restrict consideration to closed sequential procedures, so as to have an upper bound to the time or cost of an experiment. This has recently been emphasized by Armitage in connection with medical applications [2]. All the sequential procedures of the present paper are closed, but the point of closure in Section 3 depends on s^2 if σ^2 is unknown.

2. Sequential confidence limits for the mean of a normal distribution. Let $X_1, X_2, \cdots$ be a sequence of independent random variables with a common

Received 14 October 1963; revised 24 February 1964.

[1] This research was supported by the National Science Foundation under Grant NSF-G23665.

normal distribution with mean m and variances σ^2. Let $\bar{x}_n = \sum_{r=1}^{n} X_r/n$, and let $f_n(m) = [\sigma(2\pi)^{\frac{1}{2}}]^{-n} \exp\{-\sum_{r=1}^{n} (X_r - m)^2/2\sigma^2\}$. It follows from Wald's work (see Chapter 10 of [8]) that $P[f_n(m + 2d)/f_n(m) > 1/\alpha$ for at least one n, $n = 1, 2, \cdots] \leqq \alpha$. Upon simplifying, this reduces to

$$P[m < \bar{x}_n - d - (\sigma^2 \log (1/\alpha))/2\,dn \quad \text{for at least one } n,\ n = 1, 2, \cdots] \leqq \alpha, \tag{2.1}$$

which is equivalent to

$$P[\bar{x}_n - d - (\sigma^2 \log (1/\alpha))/2\,dn \leqq m < \infty \quad \text{for every } n,\ n = 1, 2, \cdots] \geqq 1 - \alpha, \tag{2.2}$$

where all logarithms are to the base e, and d is a positive constant.

By starting with $f_n(m)/f_n(m - 2d)$, we obtain

$$P[-\infty < m \leqq \bar{x}_n + d + (\sigma^2 \log (1/\alpha))/2\,dn \quad \text{for every } n,\ n = 1, 2, \cdots] \geqq 1 - \alpha. \tag{2.3}$$

If we now let

$$u_n(\alpha) = \min_r\ (1 \leqq r \leqq n)[\bar{x}_r + d + (\sigma^2 \log (1/\alpha))/2\,dr]$$

and

$$v_n(\alpha) = \max_r\ (1 \leqq r \leqq n)[\bar{x}_r - d - (\sigma^2 \log (1/\alpha))/2\,dr],$$

then it follows from (2.2) and (2.3) that

$$P[v_n(\alpha/2) \leqq m \leqq u_n(\alpha/2) \text{ for every } n,\ n = 1, 2, \cdots] \geqq 1 - \alpha, \tag{2.4}$$

which agrees with Equation (5) of [6].

When σ^2 is unknown, suppose that an estimate s^2 of σ^2 is available such that fs^2/σ^2 has the χ^2 distribution with f degrees of freedom and that s^2 is independent of $\bar{x}_{n_0}$. If we let $a(\alpha) = [(1/\alpha)^{2/f} - 1](f/2)$, then using the results of [6], we have

$$P[\bar{x}_n - d - s^2a(\alpha)/2\,dn \leqq m < \infty \text{ for every } n,\ n \geqq n_0] \geqq 1 - \alpha \tag{2.5}$$

and

$$P[-\infty < m \leqq \bar{x}_n + d + s^2a(\alpha)/2\,dn \text{ for every } n,\ n \geqq n_0] \geqq 1 - \alpha. \tag{2.6}$$

If we let

$$u_n'(\alpha) = \min_r\ (n_0 \leqq r \leqq n)[\bar{x}_r + d + a(\alpha)s^2/2\,dr]$$

and

$$v_n'(\alpha) = \max_r\ (n_0 \leqq r \leqq n)[\bar{x}_r - d - a(\alpha)s^2/2\,dr],$$

then

$$P[v_n'(\alpha/2) \leqq m \leqq u_n'(\alpha/2) \text{ for every } n,\ n \geqq n_0] \geqq 1 - \alpha. \tag{2.7}$$

3. Applications to decision problems concerning the mean.

3.1 *Testing a hypothesis about m with one-sided alternatives.* Let H_0 denote the hypothesis that $m \leqq m_0$ and let H_1 denote the class of alternatives $m > m_0$. Suppose we ask for a closed sequential procedure for testing H_0 such that $P[\text{rejecting } H_0 \mid m \leqq m_0] \leqq \alpha$ and $P[\text{accepting } H_0 \mid m > m_0 + \Delta] \leqq \beta$. When σ is known, this can easily be solved as follows. We take measurements one at a time, and stop the experiment as soon as either (a) $\bar{x}_n - d - (\sigma^2 \log (1/\alpha))/2\,dn > m_0$ or (b) $\bar{x}_n + d + (\sigma^2 \log (1/\beta))/2\,dn < m_0 + \Delta$. If (a) occurs we reject H_0, and if (b) occurs we accept H_0. This rule is equivalent to stopping as soon as

$$\sum_{r=1}^{n} (X_r - m_0 - \Delta/2) > (\sigma^2 \log (1/\alpha))/2\,d - n(\Delta/2 - d)$$

or

$$\sum_{r=1}^{n} (X_r - m_0 - \Delta/2) < -(\sigma^2 \log (1/\beta))/2d + n(\Delta/2 - d).$$

For each d with $0 < d < \Delta/2$ the boundaries consist of two intersecting lines and we get a closed sequential procedure.

In order to get some information on what value to choose for d and on the efficiency of the resulting sequential scheme, a number of Monte Carlo experiments were carried out, and the results are summarized in Table I.

The data of Table I indicates that the sequential procedure with $d = 3\Delta/8$ is substantially more efficient than the procedure with $d = \Delta/4$, which had been recommended in [6]. At present we suggest using $d = 3\Delta/8$, not only for the present problem but in other applications of the results of Section 2.

TABLE I

*Average sample size for testing the hypothesis $m \leqq 0$ against $m \geqq .25$**

α	β	m	Sequential Procedure		
			$d = \Delta/4$	$d = 3\Delta/8$	SPRT
.05	.05	0, .25	119 (3)	92 (3)	85
		.125	184 (3)	146 (3)	139
.01	.01	0, .25	192 (4)	152 (5)	144
		.125	317 (5)	272 (6)	338

* The values $\sigma = 1$ and $\Delta = .25$ were used throughout. The estimated standard error of each empirically determined mean is given in parenthesis below the mean. The numbers in the SPRT column are the theoretical values for the sequential probability ratio test of Wald when the excess is neglected.

The data of Table I indicates that this procedure (with $d = 3\Delta/8$) seems to compare reasonably well with the sequential probability ratio test of Wald, and in addition to being a closed test has the further advantage that for small values of α and β it substantially reduces the maximun average sample size. In these respects it is somewhat similar to a modification of the Wald sequential test given by Anderson [1]. Although an exact comparison with the procedures of Anderson is not possible at present, the sequential test of this section with $d = 3\Delta/8$ seems to be somewhat less efficient than the Anderson procedures when $\alpha = \beta = .05$, but seems to be about as efficient when $\alpha = \beta = .01$.

When σ is known, closed sequential procedures were given by Armitage and Schneiderman [3] for this problem for nine different combinations of α and β. The present procedure has the advantage that it can be used for any combination of α and β.

A more important advantage of the present procedure is that it can easily be modified to solve the problem when σ is unknown. In the past the problem when σ is unknown has usually been dealt with by means of a sequential t test, but this amounts to changing the problem rather than solving it in the form specified. Recently an open sequential solution for σ unknown was given by Hall [5], who also gave a closed sequential solution for the special case $\alpha = \beta$, which is similar to but not identical with the solution given here. A closed sequential for any α and β can now be obtained as follows. We start by taking a sample of n_0 measurements, compute $s^2 = \sum_{r=1}^{n_0} (X_r - \bar{x}_{n_0})^2/(n_0 - 1)$, take $f = n_0 - 1$, and then take measurements one at a time, and stop the experiment and reject H_0 as soon as $\sum_{r=1}^{n} (X_r - m_0 - \Delta/2) > s^2a(\alpha)/2d - n(\Delta/2 - d)$, and stop and accept H_0 as soon as $\sum_{r=1}^{n} (X_r - m_0 - \Delta/2) < -s^2a(\beta)/2d + n(\Delta/2 - d)$. It follows from (2.5) and (2.6) that we still have $P[\text{rejecting } H_0 \mid m \leqq m_0] \leqq \alpha$ and $P[\text{accepting } H_0 \mid m > m_0 + \Delta] \leqq \beta$.

When a decision has been reached whether H_0 is accepted or rejected, that is whether $m \leqq m_0$ or $m > m_0$, it may often be useful to have a confidence interval for m. If n denotes the number of measurements required to reach a decision about H_0 and σ is known, it follows from (2.4) that $P[v_n(\gamma/2) \leqq m \leqq u_n(\gamma/2)] \geqq 1 - \gamma$. If the resulting interval $(v_n(\gamma/2), u_n(\gamma/2))$ is too large to be useful, additional measurements can be taken. For example let n' equal the number of additional measurements required after H_0 is rejected in order to have $u_{n+n'}(\gamma/2) - v_{n+n'}(\gamma/2) \leqq L$ (where $L > 2d$). Then if $n'' = n$ when H_0 is accepted and $n'' = n + n'$ when H_0 is rejected, we can still assert that $P[v_{n''}(\gamma/2) \leqq m \leqq u_{n''}(\gamma/2)] \geqq 1 - \gamma$. When σ is unknown, from (2.7) we have a similar result, namely $P[v'_{n''}(\gamma/2) \leqq m \leqq u'_{n''}(\gamma/2)] \geqq 1 - \gamma$. These ideas are used in Section 3.3.

3.2 *Testing a hypothesis about m with two-sided alternatives.* Let H_0 now denote the hypothesis that $m = m_0$ and H_1 the set of alternatives $m \neq m_0$. We briefly consider the problem of obtaining a closed sequential procedure for testing the hypothesis H_0 so that

$$P[\text{rejecting } H_0 \mid m = m_0] \leqq \alpha \text{ and } P[\text{accepting } H_0 \mid |m - m_0| \geqq \Delta] \leqq \beta.$$

This can be easily accomplished by combining two one-sided closed sequential tests of Section 3.1. That is, we test $m = m_0$ against $m \leqq m_0 - \Delta$ with Type I and II errors of $(\alpha/2, \beta)$ and test $m = m_0$ against $m \geqq m_0 + \Delta$ with Type I and II errors of $(\alpha/2, \beta)$. Then we reject H_0 if either of the two component tests reject m_0, otherwise we accept H_0. When this is explicitly worked out, it reduces to the following rule:

Accept H_0 as soon as both

$$\sum_{r=1}^{n} (X - m_0) > [\sigma^2 \log (1/\beta)]/2d - n(\Delta - d) \quad \text{and}$$

$$\sum_{r=1}^{n} (X_r - m_0) < -[\sigma^2 \log (1/\beta)]/2d + n(\Delta - d).$$

Reject H_0 as soon as either

$$\sum_{r=1}^{n} (X_r - m_0) > [\sigma^2 \log (2/\alpha)]/2d + nd \quad \text{or}$$

$$\sum_{r=1}^{n} (X_r - m_0) < -[\sigma^2 \log (2/\alpha)]/2d - nd.$$

We again recommend choosing $d = 3\Delta/8$.

If σ^2 is unknown, we start by taking n_0 measurements, calculate $s^2 = \sum_{r=1}^{n_0} (X_r - \bar{x}_{n_0})^2/(n_0 - 1)$, take $f = n_0 - 1$ and then taking measurements one at a time, we use the above procedure with σ^2 replaced by s^2, $\log (1/\beta)$ replaced by $a(\beta)$ and $\log (2/\alpha)$ replaced by $a(\alpha/2)$.

3.3 *Comparing the means of k experimental categories with an unknown standard.* We first introduce some additional notation. Let π_0 denote the standard (or control) category, let $\pi_1, \pi_2, \cdots, \pi_k$ denote the k experimental categories, and let X_{jr} denote the rth measurement with category π_j. We assume that for each r, X_{jr} is normally distributed with mean m_j and variance σ_j^2, and that for all r and j ($j = 0, 1, \cdots, k$ and $r = 1, 2, \cdots$) all measurements are independent. An experimental category π_ν ($\nu = 1, 2, \cdots, k$) is said to be superior to the standard or control if $m_\nu > m_0$.

We will first consider the problem of finding a closed sequential procedure for classifying each of the k experimental categories as superior or non-superior which will satisfy the requirement that the probability is $\geqq 1 - \alpha$ that *all* experimental categories with means $\leqq m_0$ or $\geqq m_0 + \Delta$ are classified correctly.

First, let $Z_{jr} = X_{jr} - X_{0r}$, let $\bar{z}_{jn} = \sum_{r=1}^{n} Z_{jr}/n$, let $\sigma_{zj}^2 = \sigma_j^2 + \sigma_0^2$. Let

$$u_j(d, n, \bar{z}_{jn}, \alpha) = \bar{z}_{jn} + d + [\sigma_{zj}^2 \log (1/\alpha)]/2\, dn,$$

$$v_j(d, n, \bar{z}_{jn}, \alpha) = \bar{z}_{jn} - d - [\sigma_{zj}^2 \log (1/\alpha)]/2\, dn,$$

$$u_{jn}(d, \alpha) = \min_r (1 \leqq r \leqq n)[u_j(d, r, \bar{z}_{jr}, \alpha)],$$

$$v_{jn}(d, \alpha) = \max_r (1 \leqq r \leqq n)[v_j(d, r, \bar{z}_{jr}, \alpha)].$$

When the variances are known the desired closed sequential procedure can be obtained as follows. At the first stage of the experiment we start with one measurement on all $k + 1$ categories. At the rth stage ($r = 2, 3 \cdots$) we take one measurement with the standard category and one measurement with each of the experimental categories not yet classified. We classify an experimental category π_ν as inferior as soon as $u_\nu(d, n, \bar{z}_{\nu n}, \alpha/k) < \Delta$ and classify π_ν as superior as soon as $v_\nu(d, n, \bar{z}_{\nu n}, \alpha/k) > 0$. The experiment is terminated as soon as all the categories are classified. As mentioned earlier, we recommend using $d = 3\Delta/8$ to attempt to minimize the number of observations required to classify all the experimental categories.

We now consider a mixed decision and estimation problem. Suppose in addition to classifying all k experimental categories subject to the requirement that all categories π_ν with m_ν either $\leqq m_0$ or $> m_0 + \Delta$ are classified correctly with probability $\geqq 1 - \alpha$, we also want to have a simultaneous confidence interval for $\{m_j - m_0\}$ ($j = 1, 2, \cdots, k$) when the experiment is terminated with joint confidence coefficient $1 - \gamma$ so that the width of the confidence interval for each category classified as superior shall not exceed $L (L > \Delta)$. This can be accomplished in the following manner. We stop as before taking measurements with any category when it is classified as inferior, but continue taking measurements with any category π_j after it has been classified as superior until $u_{jn}(d, \gamma/2k) - v_{jn}(d, \gamma/2k) \leqq L$. If n_j is the number of measurements taken with category π_j when the experiment is terminated, we easily see that

$$P[v_{jn_j}(d, \gamma/2k) \leqq m_j - m_0 \leqq u_{jn_j}(d, \gamma/2k) \text{ for every } j, j = 1, 2, \cdots, k] \geqq 1 - \gamma,$$

and obviously the width of the confidence interval for each category classified as superior cannot exceed L.

When the $k + 1$ categories have a common unknown variance, we start by taking n_0 measurements from each of the $(k + 1)$ populations, let

$$s^2 = \sum_{j=0}^{k} \sum_{r=1}^{n_0} (X_{jr} - \bar{x}_{jn_0})^2/(k + 1)(n_0 - 1),$$

take $f = (k + 1)(n_0 - 1)$, and taking measurements one at a time, use the procedure just described with σ_{zj}^2 replaced by $2s^2$, $\log(1/\alpha)$ replaced by $a(\alpha) = [(1/\alpha)^{2/f} - 1]f/2$, and $u_{jn}(d, \alpha)$ and $v_{jn}(d, \alpha)$ defined as the min and max with $(n_0 \leqq r \leqq n)$ instead of $(1 \leqq r \leqq n)$.

3.4 *The determination of* n_0. We conclude Section 3 with a brief discussion of how to choose n_0 when the variance or variances are unknown. A reasonably efficient choice of n_0 is important, since the efficiency of the sequential procedure is reduced if n_0 is taken either too small or too large. Although an optimum rule for selecting n_0 is unknown, it is hoped that the following somewhat tentative procedure will be useful.

First we consider the situation in which no knowledge of σ^2 is available. Let

$\delta = \min(\alpha, \beta)$ for the problem of Section 3.1 and let $\delta = \min(\alpha/2, \beta)$ for the problem of Section 3.2. Then for the problems of Sections 3.1 and 3.2 we suggest selecting n_0 so that $a(\delta)$ does not exceed its limiting value $\log(1/\delta)$ by more than (say) 25 per cent, that is, select the smallest value of n_0 so that $a(\delta) \leqq 1.25 \log(1/\delta)$, while for the problem of Section 3.3, when several populations have a common unknown variance, we suggest selecting n_0 so that $a(\alpha) \leqq 1.1 \log(1/\alpha)$.

When a moderately accurate estimate $\hat{\sigma}^2$ of σ^2 is available from past experience it may be desirable to modify the preceding procedure for selecting n_0. It seems difficult to give any precise rule, so we will attempt to illustrate what is involved by means of an example. Consider the problem of Section (3.1) and assume $\alpha = \beta = .02$. Then when no a priori knowledge of σ^2 is available, the preceding procedure would result in taking $n_0 = 20$. Suppose now that a moderately accurate estimate of $\hat{\sigma}^2$ is available for σ^2, let $N(\hat{\sigma})$ denote the number of measurements required by the corresponding fixed sample size procedure when $\sigma = \hat{\sigma}$, and consider the following three cases: (1) $N(\hat{\sigma}) = 120$, (2) $N(\hat{\sigma}) = 50$, (3) $N(\hat{\sigma}) = 25$. In Case (1) it would seem reasonable to increase n_0, say to $n_0 = 42$, so that $a(.02)$ is within ten per cent of its limiting value, in Case (2) n_0 would presumably be left unchanged, while in Case (3) it would seem reasonable to decrease n_0, say to $n_0 = 15$.

4. Sequential confidence limits for the variance. We again let $\{X_r\}$ ($r = 1, 2, \cdots$) denote a sequence of independent and normally distributed random variables with mean m and variance σ^2. Let $s_n^2 = \sum_{r=1}^{n} (X_r - \bar{x}_n)^2/(n-1)$ and let

$$g_n(s_n^2 \mid \sigma^2) = \frac{1}{\Gamma((n-1)/2)} \left(\frac{n-1}{2\sigma^2}\right)^{(n-1)/2} (s_n^2)^{(n-3)/2} \exp\left[-\frac{(n-1)s_n^2}{2\sigma^2}\right]$$

denote the probability density function of s_n^2, and let $n^*(\alpha)$ denote the largest integer contained in $1 + \log(1/\alpha)/\log \lambda$, where $\lambda > 1$. Then it follows from the work of Cox [4] that

$$(4.1) \quad P[g_n(s_n^2 \mid \lambda^2\sigma^2)/g_n(s_n^2 \mid \sigma^2) > 1/\alpha \text{ for at least one } n, n = 2, 3, \cdots] \leqq \alpha.$$

After simplification, this can be written in the equivalent form

$$(4.2) \quad \begin{aligned} &P[s_n^2(\lambda^2 - 1)/2\lambda^2\{\log \lambda + (\log(1/\alpha))/(n-1)\} \\ &\qquad \leqq \sigma^2 < \infty \text{ for all } n, n = 2, 3, \cdots] \geqq 1 - \alpha. \end{aligned}$$

In the same manner

$$P[g_n(s_n^2 \mid \sigma^2)/g_n(s_n^2 \mid \sigma^2/\lambda^2) < \alpha \text{ for at least one } n, n = 2, 3, \cdots] \leqq \alpha.$$

After simplification, this reduces to

$$(4.3) \quad \begin{aligned} &P[0 < \sigma^2 \leqq s_n^2(\lambda^2 - 1)/2\{\log \lambda - (\log(1/\alpha))/(n-1)\} \\ &\qquad \text{for all } n, n > n^*(\alpha)] \geqq 1 - \alpha. \end{aligned}$$

If we now let

$$u_{1n}(\alpha) = \infty \qquad \text{for } n \leqq n^*(\alpha)$$
$$= \min_r (n^*(\alpha) < r \leqq n)[s_r^2(\lambda^2 - 1)/2\{\log \lambda - (\log (1/\alpha))/(r-1)\}] \qquad \text{for } n > n^*(\alpha)$$
$$v_{1n}(\alpha) = \max_r (1 \leqq r \leqq n)[s_r^2(\lambda^2 - 1)/2\lambda^2\{\log \lambda + (\log (1/\alpha))/(r-1)\}],$$

it follows from (4.2) and (4.3) that

(4.4) $P[v_{1n}(\alpha/2) \leqq \sigma^2 \leqq u_{1n}(\alpha/2) \text{ for all } n, n = 2, 3, \cdots] \geqq 1 - \alpha.$

5. Applications to decision problems involving the variance.

5.1 *Testing a hypothesis about σ^2 with one sided alternatives.* Let H_0 denote the hypothesis that $\sigma^2 \geqq \sigma_0^2$ and let H_1 denote the set of alternatives $\sigma^2 < \sigma_0^2$. We now ask for a closed sequential procedure for testing H_0 satisfying the requirements that

$$P[\text{rejecting } H_0 \mid \sigma^2 \geqq \sigma_0^2] \leqq \alpha \text{ and } P[\text{accepting } H_0 \mid \sigma^2 < \sigma_0^2/w^2] \leqq \beta,$$

where $w > 1$. Using the results obtained in Section 4, we obtain a solution in the following manner. We take measurements one at a time and stop the experiment and (a) accept H_0 when

$$s_n^2(\lambda^2 - 1)/2\lambda^2\{\log (\lambda) + (\log (1/\beta))/(n-1)\} > \sigma_0^2/w^2,$$

(b) reject H_0 when

$$s_n^2(\lambda^2 - 1)/2\{\log \lambda - (\log (1/\alpha))/(n-1)\} < \sigma_0^2 \quad \text{and} \quad n > n^*(\alpha).$$

This is equivalent to accepting H_0 as soon as

$$\sum_{r=1}^{n} (X_r - \bar{x}_n)^2 > [2\sigma_0^2\lambda^2/w^2(\lambda^2 - 1)][(n-1) \log \lambda + \log (1/\beta)]$$

and rejecting H_0 as soon as

$$\sum_{r=1}^{n} (X_r - \bar{x}_n)^2 < [2\sigma_0^2/(\lambda^2 - 1)][(n-1) \log \lambda - \log (1/\alpha)].$$

For each λ with $1 < \lambda < w$ we get a closed sequential procedure. Although the optimum choice of λ is unknown, on the basis of some preliminary calculations we recommend taking $\lambda = 1 + .7(w - 1)$. A number of sampling experiments were carried out to obtain some information as to the efficiency of the sequential procedure when $\lambda = 1 + .7(w - 1)$, and the results are summarized in Table II.

5.2. *Deciding which of k non-overlapping intervals contains σ^2.* Let $I_1 = [0, c_1)$, $I_2 = [c_1, c_2)$, $I_3 = [c_2, c_3) \cdots$, and $I_k = [c_{k-1}, \infty)$ denote k non-overlapping intervals whose union is $[0, \infty)$. Let D_j denote the decision that σ^2 falls in I_j ($j = 1, 2, \cdots, k$). We assume that on the basis of practical considerations we can

TABLE II
*Average sample size for testing $\sigma^2 \geqq 1$ against $\sigma^2 \leqq \frac{2}{3}$**

α	β	σ^2	Sequential Procedure	
			$\lambda = 1 + .7(w - 1)$	SPRT
.05	.05	$\frac{2}{3}$	84 (2)	73
		.81	114 (2)	106
		1	56 (2)	56

* The estimated standard error of each empirically determined mean is given in parenthesis below the mean.

The numbers in the SPRT column are the theoretical values of the sequential probability ratio test of Wald, neglecting the excess.

find an indifference zone (a_j, b_j) about each end point c_j in which it does not matter whether decision D_j or D_{j+1} is taken. Let $W(D_j, \sigma^2)$ denote the error in taking decision D_j when σ^2 is the true value of the unknown variance. We let $a_0 = 0$ and $b_k = \infty$, and assume the following form for the error function: for each j $(j = 1, 2, \cdots, k)$

$$W(D_j, \sigma^2) = 0 \qquad \text{if } a_{j-1} < \sigma^2 < b_j$$
$$= 1 \qquad \text{otherwise.}$$

We now ask for a closed sequential procedure for choosing one of the k decisions $D_1, D_2, \cdots, D_k$ so that the probability of making an error shall be $\leqq \alpha$. Let $w^2 = \min (b_1/a_1, b_2/a_2, \cdots, b_{k-1}/a_{k-1})$, let I_j' denote the interval (a_{j-1}, b_j) for $j = 1, 2, \cdots, k$ and take $\lambda = 1 + .7(w - 1)$. Then we get a solution as follows. We take measurements one at a time, and stop the experiment and make a decision as soon as the interval $[v_{1n}(\alpha/2), u_{1n}(\alpha/2)]$ falls inside one at the intervals $I_1', I_2', \cdots, I_k'$. If $[v_{1n}(\alpha/2), u_{1n}(\alpha/2)]$ falls in interval I_j', we choose decision D_j, and if it falls in the intersection of I_j' and I_{j+1}', we can choose between D_j and D_{j+1} at random. It follows directly from (4.4) that the probability of making an error is $\leqq \alpha$.

6. Sequential confidence limits for the ratio of variances. Let $\{X_r\}$ and $\{Y_r\}$ $(r = 1, 2, \cdots)$ be two independent sequences of independent normally distributed random variables with means m_x and m_y and variances σ_x^2 and σ_y^2 respectively.

Let $\phi^2 = \sigma_x^2/\sigma_y^2$, and let $F_n = \sum_{r=1}^n (X_r - \bar{x}_n)^2 / \sum_{r=1}^n (Y_r - \bar{y}_n)^2$. We again let $n^*(\alpha)$ denote the largest integer contained in $1 + \log (1/\alpha)/\log \lambda$, and let

$$u(n, F_n, \alpha) = \frac{\lambda F_n\{\lambda - \alpha^{1/(n-1)}\}}{\lambda\alpha^{1/(n-1)} - 1}$$

$$v(n, F_n, \alpha) = \frac{F_n\{\lambda - (1/\alpha)^{1/(n-1)}\}}{\lambda\{\lambda(1/\alpha)^{1/(n-1)} - 1\}}$$

$$u_{2n}(\alpha) = \min_r (n^*(\alpha) < r \leqq n)[u(r, F_r, \alpha)],$$

$$v_{2n}(\alpha) = \max_r (n^*(\alpha) < r \leqq n)[v(r, F_r, \alpha)].$$

Let

$$h_n(F_n \mid \phi^2) = \frac{\Gamma(n-1)}{[\Gamma((n-1)/2)]^2} \frac{F^{(n-3)/2}}{\phi^{n-1}[1 + F_n/\phi^2]^{n-1}}$$

denote the probability density function of F_n. It follows at once from [4] that

$$P[h_n(F_n \mid \lambda^2\phi^2)/h_n(F_n \mid \phi^2) > 1/\alpha \text{ for at least one } n, n = 2, 3, \cdots] \leqq \alpha$$

where $\lambda > 1$. This reduces to

$$(6.1) \quad P[v(n, F_n, \alpha) \leqq \phi^2 < \infty \text{ for all } n, n > n^*(\alpha) \cdots] \geqq 1 - \alpha.$$

In the same manner

$$P[h_n(F_n \mid \phi^2)/h_n(F_n \mid \phi^2/\lambda^2) < \alpha \text{ for at least one } n, n = 2, 3, \cdots] \leqq \alpha$$

which reduces to

$$(6.2) \quad P[0 < \phi^2 \leqq u(n, F_n, \alpha) \text{ for all } n, n > n^*(\alpha)] \geqq 1 - \alpha.$$

Combining (6.1) and (6.2), we get

$$(6.3) \quad P[v_{2n}(\alpha/2) \leqq \phi^2 \leqq u_{2n}(\alpha/2) \text{ for all } n, n > n^*(\alpha/2)] \geqq 1 - \alpha.$$

7. Applications to problems involving the ratio of variances.

7.1 *A hypothesis involving the ratio of variances.* Let H_0 denote the hypothesis that $\sigma_x^2/\sigma_y^2 \leqq 1$, and let H_1 denote the set of alternatives $\sigma_x^2/\sigma_y^2 > 1$. We now look for a closed sequential procedure for testing this hypothesis subject to the restrictions $P[\text{rejecting } H_0 \mid \phi^2 \leqq 1] \leqq \alpha$ and $P[\text{accepting } H_0 \mid \phi^2 \geqq R^2] \leqq \beta$, where $R > 1$. This can be obtained as follows. Let $n^* = \min [n^*(\alpha), n^*(\beta)]$. We start by taking $n^* + 1$ pairs of measurements, and then take one pair of measurements at a time, and stop the experiment and accept H_0 as soon as $n > n^*(\beta)$ and $u(n, F_n, \beta) < R^2$, and stop the experiment and reject H_0 as soon as $n > n^*(\alpha)$ and $v(n, F_n, \alpha) > 1$. For each λ with $1 < \lambda < R$ we get a closed sequential procedure. By analogy with the problem of Section 6, we tentatively suggest taking $\lambda = 1 + .7(R - 1)$.

7.2 *Comparing the variances of k experimental categories with an unknown standard.* We have a standard category π_0 and k experimental categories $\pi_1, \pi_2, \cdots, \pi_k$ with corresponding variances $\sigma_0^2, \sigma_1^2, \cdots, \sigma_k^2$. An experimental category π_j is defined to be superior to the standard if $\sigma_j^2 < \sigma_0^2$. We will consider the problem of finding a closed sequential procedure for classifying the k experimental categories as superior or non-superior with respect to the standard so that the probability should be $\geqq 1 - \alpha$ of correctly classifying *all* experimental categories π_j such that $\sigma_j^2 \geqq \sigma_0^2$ or $\sigma_j^2 \leqq \sigma_0^2/R^2$, where $R > 1$.

First let $\phi_j^2 = \sigma_j^2/\sigma_0^2$, let

$$F_{jn} = \sum_{r=1}^{n} (X_{jr} - \bar{x}_{jn})^2 / \sum_{r=1}^{n} (X_{0r} - \bar{x}_{0n})^2.$$

We can obtain a closed sequential procedure for this problem as follows. We start by taking $n^*(\alpha/k) + 1$ measurements each of the $k + 1$ categories. At each subsequent stage we take one measurement with the standard and one with each of the experimental categories not yet classified. A category π_j ($j = 1, 2, \cdots, k$) is classified as superior as soon as $u(n, F_{jn}, \alpha/k)$ is <1, and is classified as not superior as soon as $v(n, F_{jn}, \alpha/k) > 1/R^2$. The experiment terminates as soon as all experimental categories are classified.

8. Concluding remarks. The methods of this paper can also be used to obtain sequential confidence limits for parameters of certain non-normal distributions such as the Poisson distribution, the exponential distribution, and the binomial distribution.

In this paper we only considered taking one measurement at a time from any category. In practice sequential sampling by groups may often be preferable to item by item sampling.

REFERENCES

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

[2] ARMITAGE, P. (1960). *Sequential Medical Trials*. Blackwell, Oxford.

[3] ARMITAGE, P. and SCHNEIDERMAN, M. A. (1962). A family of closed sequential procedures. *Biometrika* **49** 41–56.

[4] COX, D. R. (1952). Sequential tests for composite hypotheses. *Proc. Cambridge Philos. Soc.* **48** 290–299.

[5] HALL, W. J. (1962). Some sequential analogs of Stein's two-stage test. *Biometrika* **49** 367–378.

[6] PAULSON, E. (1963). A sequential procedure for choosing one of k hypothesis concerning the unknown mean of a normal distribution. *Ann. Math. Statist.* **34** 549–554.

[7] STEIN, C. (1948). The selection of the largest of a number of means (Abstract). *Ann. Math. Statist.* **19** 429.

[8] WALD, A. (1947). *Sequential Analysis*. Wiley, New York.