1. Introduction and Summary

The sequential probability ratio test is widely used to decide which of two hypotheses to accept or, more generally, to decide whether a parameter is large or small. A justification of its use is that the procedure possesses the optimum property of minimizing the expected sample size. However, as we shall show (Section 2), for the binomial distribution there exists a class of tests with a stronger optimum property.

This class of tests consists of curtailed single-sampling plans with acceptance number 0; in such plans, which have been widely used in sampling inspection, sample items are inspected one at a time until either an item is found to be defective, in which case the lot in question is rejected, or a designated number of items have been found to be nondefective, in which case the lot is accepted. Such a procedure never requires more observations for an individual sample sequence than any other test which gives the same probability of acceptance (between 0 and 1) at some value of \( p \), the true percentage defective (0 < \( p \) < 1), and which always leads to acceptance when \( p = 0 \); it requires fewer observations for some sample sequences (0 < \( p \) < 1). This optimal property is clearly a strong one.

Peculiarly enough, the curtailed single-sampling plan is not necessarily a sequential probability ratio test (Section 3). There seems to be a contradiction: two different tests each of which can be proved to be optimal. The resolution of the contradiction is that, when it arises, there is no sequential probability ratio test that gives as low a “power” as the curtailed single-sampling plan; the sequential probability ratio test that comes closest will give greater “power,” and it pays for this by requiring more observations on the average (Section 4).

The consideration of these two classes of tests reveals a limitation on the optimum property of the sequential probability ratio test as it is generally defined, at least for cases involving a discrete variate. That is, there can be more than one set of tests, not completely overlapping, for which similar optimum properties hold. We consider other limitations on the sequential probability ratio test (Section 5). The existence of such tests is related to the uniqueness of the tests; it is shown that only one sequential probability ratio test achieves a given pair of probabilities of error (Section 6). Finally, this study suggests that the usual way of defining these tests may not be the best way to satisfy the needs of the user of the tests (Section 7).

2. A Special Class of Sequential Procedures and a Strong Optimal Property

Consider sampling from a binomial distribution with parameter \( p \); that is, let
\[
(2.1) \quad \Pr\{X = 1\} = p, \\
\Pr\{X = 0\} = 1 - p = q,
\]
say. It may be convenient to think of sampling from a lot of items, each of which can be classified as a defective, \( X = 1 \), or a nondefective, \( X = 0 \). Then \( p \) is the proportion of defectives in the lot, which is assumed to be large enough so that it can be treated as an infinite population. If the proportion \( p \) is sufficiently small, we wish to accept the lot; if \( p \) is large, we wish to reject it.

A sequential sampling procedure consists in taking observations in sequence: \( x_1, x_2, \cdots \), and at some point in the sequence stopping observation and deciding to accept or reject the lot. More precisely, at least one observation \( x_1 \) is taken. After \( m \) observations are taken \((m = 1, 2, \cdots)\), there is a rule (based on \( x_1, \cdots, x_m \)) which indicates whether to stop sampling and accept the lot, to stop sampling and reject the lot, or to take one more observation and use the rule based on \( m + 1 \) observations. In this paper we consider only nonrandomized procedures.\(^2\) We also assume that for each \( p \) the probability is 1 of coming to a decision; that is, the probability of stopping sampling before more than \( n \) observations are taken approaches 1 as \( n \) approaches infinity. For any specific procedure we designate by \( L_p \) the probability of accepting the lot when the proportion defective is \( p \) \((0 \leq p \leq 1)\). (This is the operating characteristic.)

Let \( C \) be the curtailed single-sampling plan defined as follows: After taking \( m \) observations \((m = 1, \cdots, n - 1)\) stop sampling and reject the lot if \( x_m = 1 \) and continue sampling if \( x_m = 0 \); after taking \( n \) observations stop sampling and reject the lot if \( x_n = 1 \) and accept the lot if \( x_n = 0 \). In other words, if no defective appears among the first \( n \) sample items, accept the lot; if a defective appears, stop sampling and reject the lot.

The plan has one constant to be adjusted, namely, the integer \( n \). The probability of accepting the lot, \( L_p \), is the probability of \( n \) nondefectives in a sample of \( n \), namely
\[
(2.2) \quad L_p(C) = (1-p)^n = q^n.
\]
It is evident that \( L_0(C) = 1 \), \( L_p(C) \) decreases, and \( L_1(C) = 0 \). One way to fix \( n \) is to impose a requirement on \( L_p(C) \) at some particular value of \( p \), for example, the requirement that \( L_{p_1}(C) \leq \beta \) for some \( p_1 \) \((0 < p_1 < 1)\) and \( \beta \) \((0 < \beta < 1)\); that is, that the probability of accepting the lot should not exceed \( \beta \) if the proportion of defectives is as high as \( p_1 \). Then \( n \) is chosen as the smallest integer for which
\[
(2.3) \quad (1-p_1)^n \leq \beta.
\]

**Theorem 1.** For any procedure \( E \) satisfying
\[
(2.4) \quad L_0(E) = 1,
\]
\[
(2.5) \quad L_{p_1}(E) \leq \beta,
\]
where \( 0 < \beta < 1 \), the number of observations required to reach a decision is for each sample sequence at least as large as the number for the curtailed single-sampling plan \( C \). If the number is the same for every sample, then \( E \) is identical to \( C \).

**Proof.** If \( p = 0 \), every observation is \( x_1 = 0 \). Condition (2.4) implies that for any procedure \( E \), the sequence of observations \( 0, 0, \cdots \) leads to acceptance. Since the probability must be 1 (for \( p = 0 \)) of reaching a decision after a finite number of observations, there must be a number \( n^* \) such that the lot is accepted at the \( n^* \)th observation when \( x_1 = 0, \cdots, x_{n^*} = 0 \). Then \( L_{p_1}(E) \) is at least equal to the probability of this sequence, and Condition (2.5) implies...
(2.6) \[(1 - p_i) n^* \leq \beta,\]
which implies \[n^* \geq n.\] For any sequence with \(x_1 = 0, \cdots, x_n = 0,\) the number of observations required to reach a decision is, therefore, the same for \(E\) and for \(C\) if \[n^* = n\] and larger for \(E\) than for \(C\) if \[n^* > n.\]

The preceding discussion implies that a sequence \(x_1 = 0, \cdots, x_m = 0\) leads to continuation for \(E\) at the \(mth\) stage, \(m = 1, \cdots, n - 1\) \ref{eq:2.6}. Now consider a sequence with \(x_1 = 0, \cdots, x_{l-1} = 0, x_l = 1, l = 1, \cdots, n,\) which under \(C\) leads to termination at the \(lth\) observation. Unless \(E\) has the same decision rule at the \(lth\) observation as \(C,\) such a sequence requires more observations to terminate under \(E\) than under \(C.\)

This optimum property is a strong one. For \(every\) sample sequence, \(C\) is at least as good as any other procedure \(E\) in that it leads to a decision with no more observations. If \(E\) is different from \(C\) for some sequences of observations, \(C\) is better than \(E\) in this sense. This, of course, implies the weaker optimum property that the expected sample size under \(C\) is no greater than that under \(E\) for every value of \(p.\)

The result may be applied directly to problems in which it is desired solely to specify the risk of acceptance or rejection at a single point. For example, it might be desired solely to specify the risk of accepting a product with a stated percentage of defectives—the consumer's risk—and not to specify the producer's risk. For such a problem, a curtailed single-sampling plan \(C\) would be indicated.

The problem typically considered is that in which it is desired to set risks at two points, say \(p_0\) and \(p_1,\) and to provide that at \(p_1\) the risk of acceptance be less than or equal to \(\beta\) and at \(p_0\) the risk of rejection be less than or equal to \(\alpha.\) If \(\alpha\) and \(\beta\) happen to satisfy exactly the equations \(1 - \alpha = (1 - p_0)^n\) and \(\beta = (1 - p_1)^n\) for the same integral value of \(n,\) then it follows from Theorem 1 that the curtailed single-sampling plan \(C\) with maximum sample size \(n\) will never require a larger average number of observations than any other test for which \(L_{p_0} \leq \beta\) and \(L_{p_0} \geq 1 - \alpha\); hence it is a uniformly best plan for all values of \(p\) for such cases as well.

3. Is the Curtailed Single-sampling Procedure with Acceptance Number 0 a Sequential Probability Ratio Test?

The sequential probability ratio test has been shown by Wald and Wolfowitz [1] to have certain optimum properties. Since the curtailed single-sampling procedure \(C\) has an optimum property, it is of interest to compare the two.

Let \(f_0(x)\) and \(f_1(x)\) be two probability functions or two densities corresponding to hypotheses \(H_0\) and \(H_1,\) respectively. The sequential probability ratio test for testing \(H_0\) against \(H_1,\) is defined by two numbers \(3 A\) and \(B (0 < B < 1 < A < \infty)\) and the following rules: after \(m\) observations \((x_1, \cdots, x_m)\) have been taken \((m = 1, 2, \cdots),\) if

\[(3.1)\quad B < \prod_{i=1}^{m} \frac{f_i(x_i)}{f_0(x_i)} < A,
\]

take another observation; if
accept the hypothesis $H_0$ that sampling is from $f_0(x)$ [i.e., reject the hypothesis $H_1$ that sampling is from $f_1(x)$]; and if

$$A \leq \prod_{i=1}^{m} \frac{f_i(x_i)}{f_0(x_i)},$$

reject the hypothesis $^4H_0$ (that is, accept $H_1$). Denote such a procedure by $W$ (for Wald [2]). The optimum property is that for any other procedure $E$ for which $L_{j_0}(E) \geq L_{j_0}(W)$ and $L_{j_1}(E) \leq L_{j_1}(W)$, the expected sample size is at least as large as for procedure $W$ when sampling from $f_0(x)$ and $f_1(x)$, respectively.

The optimum property of the curtailed single-sampling plan is stronger than that of the sequential probability ratio test (since the former is in terms of every sample sequence), and the optimality is relative to a wider class of procedures (that is, procedures meeting only a requirement at $p_1$) than that of the sequential probability ratio test. These facts suggest that the curtailed single-sampling plan must always be a sequential probability ratio test. However, we must examine a little more closely what this means. There is one sense in which the curtailed single-sampling plan is always a special case of the Wald test, and another in which it may be a special case but need not be.

The essential point is that the definition of a sequential probability ratio test involves the specification of two parameter values, $p_0$ and $p_1$, as well as two probabilities of error (accepting $H_0$ when $H_1$ is true and accepting $H_1$ when $H_0$ is true). Construction of a curtailed single-sampling plan $C$, on the other hand, involves the specification of only one parameter value, $p_1$, and only one probability, the probability of accepting a lot when $p = p_1$. To put the curtailed single-sampling plan in terms comparable to the sequential probability ratio test we must associate some $H_0$ with our $H_1$, and there is no unique way to do so.

Under these circumstances what does it mean to ask whether a curtailed single-sampling plan is a special case of a sequential probability ratio test? One possible meaning is to ask whether there exists any $H_0$ which when combined with the $H_1$ of the curtailed single-sampling plan will yield a sequential probability ratio test identical with the curtailed single-sampling plan. The answer to this question is affirmative, as can readily be proved.

For our special binomial problem, $H_0$ is the hypothesis that $p = p_0$; $H_1$, that $p = p_1$; and we shall require that $p_0 < p_1$. In order for the curtailed single-sampling plan to be a sequential probability ratio test in the sense just defined, there must exist some $p_0 < p_1$ and some numbers $A$ and $B$, with $B < 1 < A$, such that
From The Collected Works of Milton Friedman, compiled and edited by Robert Leeson and Charles G. Palm.

(3.4) \[ B < \frac{(1 - p_1)^m}{(1 - p_0)^m} < A, \quad m = 1, \ldots, n - 1, \]

(3.5) \[ \frac{(1 - p_1)^n}{(1 - p_0)^n} \leq B, \]

(3.6) \[ A \leq \frac{(1 - p_1)^{m-1}}{(1 - p_0)^{m-1}} \cdot \frac{p_1}{p_0}, \quad m' = 1, \ldots, n. \]

These inequalities correspond to the inequalities (3.1), (3.2), and (3.3), for the binomial case where \( f_i(x) = p_i^i (1 - p_i)^{i-x}, \) \( i = 0, 1, \) under the tentative assumption that the curtailed single-sampling plan is a sequential probability ratio test. Inequality (3.4) indicates continuation at the \( m \)th step (all 0’s), (3.5) indicates acceptance of the lot and \( H_0 \) (n 0’s), and (3.6) indicates rejection of the lot and \( H_0 \) (one 1).

If there exists a constant \( B \) to define the curtailed single-sampling plan as a sequential probability ratio test, it must satisfy (3.4) and (3.5); that is,

(3.7) \[ \frac{(1 - p_1)^n}{(1 - p_0)^n} \leq B < \frac{(1 - p_1)^m}{(1 - p_0)^m}, \quad m = 1, \ldots, n - 1. \]

Since the right-hand side of (3.7) is smallest when \( m \) has its largest value, the limits (3.7) are tightest when \( m = n - 1 \). Clearly for any value of \( p_0 \) one can find a \( B \) to satisfy (3.7).

If there exists a constant \( A \) to define the curtailed single-sampling plan as a sequential probability ratio test, it must satisfy (3.4) and (3.6); that is,

(3.8) \[ \frac{(1 - p_1)^m}{(1 - p_0)^m} < A \leq \frac{(1 - p_1)^{m-1}}{(1 - p_0)^{m-1}} \cdot \frac{p_1}{p_0}, \quad m = 1, \ldots, n - 1, \]

Since \( (1 - p_1)^m/(1 - p_0)^m \) is a decreasing function of \( m \), the limits in (3.8) are tightest when \( m = 1 \) and \( m' = n \). However, we require \( 1 < A \). Hence, \( A \) must satisfy

(3.9) \[ 1 < A \leq \frac{(1 - p_1)^{m-1}}{(1 - p_0)^{m-1}} \cdot \frac{p_1}{p_0}. \]

The question then is to determine the set of \( p_0 \) for which the third expression in (3.9) is greater than the first expression. The function \( (1 - p_0)^{n-1} p_0 \) has the derivative \( (1 - p_0)^{n-2} (1 - np_0) \). Hence, it increases from a value of 0 at \( p_0 = 0 \) to a maximum at \( p_0 = 1/n; \) then it decreases to \( p_0 = 1 \). Thus if \( p_1 \leq 1/n \), the third expression in (3.9) is greater than 1 for every \( p_0 < p_1 \), and an \( A \) can be found that satisfies (3.9). If \( p_1 > 1/n \), the third expression in (3.9) is greater than 1 for sufficiently small \( p_0 \) but not for all \( p_0 < p_1 \). Thus there exist values of \( p_0 \), for which an \( A \) can be found, but there will be some values of \( p_0 \) for which \( A \) cannot be found. The conclusion is that, given \( p_1 \), there always exists a \( p_0 \) such that the curtailed single-sampling plan is a sequential probability ratio test for \( p_1 \) and this \( p_0 \).

However, this is clearly not a fully satisfactory answer to the question whether a curtailed single-sampling plan \( C \) is a special case of the sequential probability ratio test. The proof of the optimum properties of the curtailed single-sampling plan holds for every null hypothesis \( H_0 \). Pick any such hypothesis, and let the values of \( \alpha \) and \( \beta \) be those given by some curtailed single-sampling plan \( C \) (that is, some \( n \)). Then we have proved that this test requires fewer observations...
at both \( p_0 \) and \( p_1 \) than any other test yielding equal or greater power. For this proof to be a special case of the proof of the optimum properties of the sequential probability ratio test, the question whether a curtailed single-sampling plan is a special case of the sequential probability ratio test must be given a more demanding meaning, namely, whether for every \( H_0 \) that is relevant (namely, for which \( p_0 < p_1 \)), there exists a sequential probability ratio test that is identical with the curtailed single-sampling plan.

The answer to this question is clear from the preceding analysis. If \( p_1 \leq 1/n \), then both (3.7) and (3.9) can be satisfied for every \( p_0 < p_1 \). If \( p_1 > 1/n \), Inequality (3.7) can be satisfied for every \( p_0 < p_1 \), but there, are always some values of \( p_0 < p_1 \) for which Inequality (3.9) cannot be satisfied. Hence, the curtailed single-sampling plan is not necessarily a sequential probability ratio test in this more demanding sense.

4. Relationship Between the Optimum Properties of the Tests

We seem to have a discrepancy. The curtailed single-sampling plan has a strong optimum property; the sequential probability ratio test has an optimum property; but there are cases when the curtailed single-sampling plan is not a sequential probability ratio test. How does this happen? The answer is that the theorem about the sequential probability ratio test asserts that for any \( H_0 \) and \( H_1 \) and any values of \( \alpha \) and \( \beta \), which can be achieved by a sequential probability ratio test, there exists no other test yielding the same or lower values of \( \alpha \) and \( \beta \) for which the average number of observations is smaller at either \( p_0 \) or \( p_1 \). But not all values of \( \alpha \) and \( \beta \) can be achieved by a sequential probability ratio test defined for a particular \( H_0 \) and \( H_1 \). Whenever, for a given \( n \), \( p_1 \), \( \beta \), the value of \( p_0 \) is such that the sequential probability ratio test does not reduce to the curtailed single-sampling plan, then there is no sequential probability ratio test defined for these \( p_0 \) and \( p_1 \) that yields a probability of acceptance at \( p_0 \) as small as that yielded by the curtailed single-sampling plan.

THEOREM 2. Given a curtailed single-sampling plan \( C \) with a probability of acceptance of \( \beta \) at \( p_1 \) and a probability of rejection of \( \alpha \) at \( p_0 \) (\( < p_1 \)), if this procedure is not a sequential probability ratio test defined in terms of \( p_0 \) and \( p_1 \), then there is no sequential probability ratio test that achieves these probabilities of errors.

PROOF. Suppose a sequential probability ratio test is not the same as the curtailed single-sampling plan for a given \( n \), \( p_0 \) and \( p_1 \), but yields the same value of \( \beta = (1 - p_1)^n \). Then it must require continuation of sampling for a sample of \( n \) items having a value of zero. (Otherwise it would lead to acceptance for this and some other cases as well; hence the probability of acceptance when \( p_1 \) is true would be higher than \( \beta \).) But this means that

\[
B < \frac{(1 - p_1)^n}{(1 - p_0)^n},
\]

since the right-hand side is the probability ratio for such a sample.

A standard inequality for a sequential probability ratio test is given by Wald ([2], p. 41):

\[
L_{p_0}(W) \leq BL_{p_0}(W).
\]

Hence
\[ L_{p_1}(W) < \frac{(1 - p_1)^n}{(1 - p_0)^n} L_{p_0}(W). \]

Since
\[ L_{p_1}(W) = (1 - p_1)^n = \beta, \]
we have
\[ 1 - \alpha = (1 - p_0)^n < L_{p_0}(W). \]

The point of the theorem is that at some \( p_0 \) and \( p_1 \) a curtailed single-sampling plan can achieve an \( \alpha \) and \( \beta \) that cannot be achieved by a sequential probability ratio test. The sequential probability ratio test for the given \( p_0, p_1, \) and \( \beta \) must achieve a smaller \( \alpha \), and this improvement is paid for by greater expected sample sizes at \( p_0 \) and \( p_1 \).

5. Probabilities of Errors for the Sequential Probability Ratio Test

It is evident that, at least for binomial distributions, sequential probability ratio tests cannot achieve all of the \((\alpha, \beta)\) points that can be achieved by (nonrandomized) sequential procedures. Consequently, it is possible for classes of tests other than the sequential probability ratio test to have similar optimal properties, as is the case for the curtailed single-sampling plan \( C \). This raises questions about the set \((\alpha, \beta)\) of points that can be achieved by the sequential probability ratio test. We shall not make an exhaustive study of this subject, but simply note a few points.

For the binomial distribution, a sequential probability ratio test is defined for a given \( p_0 \) and \( p_1 \) by two constants \( B \) and \( A \). At the \( n \)th step we continue sampling if
\[ B < \frac{p_1^n q_0^{n-m}}{p_0^n q_0^{-m}} < A, \]
where \( m \) is the number of defectives among the \( n \) observations. The inequalities can be rewritten as
\[ c_0 + sn < m < c_1 + sn, \]
where
\[ s = \log \frac{q_0}{q_1} / \log \frac{p_1 q_0}{p_0 q_1}, \]
\[ c_0 = \log B / \log \frac{p_1 q_0}{p_0 q_1}, \]
\[ c_1 = \log A / \log \frac{p_1 q_0}{p_0 q_1}. \]

The conditions \( B < 1 < A \) are equivalent to \( c_0 < 0 < c_1 \), and \( s > 0 \). In the plane of \( n \) and \( m \), the region (5.2) lies between two lines, each with slope \( s \) and with intercepts \( c_0 \) and \( c_1 \), respectively. Each sample point can be plotted in this plane as a point \((n, m)\) with integer-valued coordinates. Two procedures, defined by pairs of intercepts \((c_0, c_1)\) and \((c_0^*, c_1^*)\), are equivalent if there is no
point \((n, m)\), \(n \geq m \geq 0\), between the lines \(y = c_0 + sx\) and \(y = c_0^* + sx\) or between \(y = c_1 + sx\) and \(y = c_1^* + sx\).

**Theorem 3.** If the slope \(s\) is rational, there is a denumerable number of sequential probability ratio tests. If the slope is irrational, there is a non-denumerable number.

**Proof.** Suppose \(s = M/N\) (where \(M\) and \(N\) are relatively prime integers). Then a point \((n, m)\) is on the line \(y = c + sx\) for a value of \(c = (Nm - Mn)/N\), that is, a rational value of \(c\). The only lines needed in defining sequential probability ratio tests in this case are those with intercepts of the form shown. There is a denumerable number of such lines, and hence a denumerable number of pairs of such lines.

Now suppose \(s\) is irrational. Suppose the lower line of a sequential probability ratio test is \(y = c_0 + sx\) for some \(c_0 < -1\), and consider upper lines \(y = c_1 + sx\) for \(0 < c_1 < 1\). We shall show that any two values of \(c_1\) define different tests and hence that the number of tests is nonenumerable. Consider the points \((n, m)\) between \(y = sx\) and \(y = sx + 1\). (Because \(s\) is irrational, no point can fall on either line.) Let \(a_n\) be the distance of \((n, m)\) above the line \(y = sx\). Then \(a_n = \lfloor sn \rfloor + 1 - sn\), and \(1 - a_n = sn - \lfloor sn \rfloor\) is the fractional part of \(sn\). (Read \([x]\) as the largest integer not greater than \(x\).) It is known ([3], p. 71) that this sequence is uniformly distributed; that is,

\[
\lim_{n \to \infty} \frac{\text{number of } (1 - a_j) \leq z, j = 1, \ldots, n = z}{n} = \frac{1}{2}, \quad j = 1, \ldots, n = z.
\]

This implies that between the lines \(y = c_i^* + sx\) and \(y = c_i^{**} + sx\) there are an infinite number of points \((n, m)\). Since \(0 < s < 1\), some of these points \((n, m)\) above \(c_i^* + sx\) can be reached from points \((n - 1, m - 1)\) below \(c_i^* + sx\) (continuation points). Hence, the two procedures are different. Thus the number of procedures is at least equal to the number of choices of \(c_1\), which is nonenumerable.

The interesting problems considered in this paper occur not only for the binomial distribution, but also for other distributions. They clearly do so for other discrete distributions. What is less obvious is that they arise even when the distributions are continuous, if the distributions of the probability ratio are continuous, if the distributions of the probability ratio are discrete. As an extreme example, consider rectangular distributions with unit range. Let

\[
f_i(x) = \begin{cases} 
1, & \text{if } \theta_i - \frac{1}{2} \leq x \leq \theta_i + \frac{1}{2}, \\
0, & \text{otherwise,}
\end{cases}
\]

for \(i = 0, 1\). Then if \(\theta_0 < \theta_1\) and \(\theta_1 - \theta_0 < 1\), the possible values of the probability ratio are

\[
f(x) = \begin{cases} 
0, & \theta_0 - \frac{1}{2} \leq x < \theta_1 - \frac{1}{2} \\
1, & \theta_1 - \frac{1}{2} \leq x \leq \theta_0 + \frac{1}{2} \\
\infty, & \theta_0 + \frac{1}{2} < x \leq \theta_1 + \frac{1}{2}
\end{cases}
\]

where \(\infty = 1/0\). Then if \(B < 1 < A\) there is only one possible sequential probability ratio test. The procedure is to continue sampling as long as the observations fall in the interval
\[ \theta = \frac{1}{2} \leq x \leq \theta_0 + \frac{1}{2}, \] to stop and accept \( H_0 \) as soon as an observation is less than \( \theta_0 - \frac{1}{2} \), and to stop and reject \( H_0 \) as soon as an observation is greater than \( \theta_0 + \frac{1}{2} \). The probability of each error is 0. The expected sample size under each hypothesis is \( 1/(\theta - \theta_0) \). Clearly there are other procedures with smaller expected sample size. For example, one can take a single observation and accept \( H_0 \) if it is less than \( c(\theta_0 - \frac{1}{2}) \leq c \leq \theta_0 + \frac{1}{2}, \) and reject \( H_0 \) if it is at least equal to \( c \).

Consider now distributions for which the probability ratio has a continuous distribution under each hypothesis. First consider the procedures based on the probability ratio \( f_1(x)/f_0(x) \) of one observation; that is, procedures which reject \( H_0 \) if

\[ \frac{f_1(x)}{f_0(x)} > k. \]

As \( k \) varies from 0 to \( \infty \), \( \alpha \) varies from 1 to 0, \( \beta \) varies from 0 to 1, and the set of \( (\alpha, \beta) \) points forms a closed connected curve. As is well known, this curve is convex.

We can show that no sequential probability ratio test with \( B < 1 < A \) can achieve an \( (\alpha, \beta) \) point that is above and to the right of the set of \( (\alpha, \beta) \) points achieved by use of one observation; that is, such that there is an \( (\alpha, \beta) \) point based on one observation for which each coordinate is less than or equal to that of the point achieved by the sequential probability ratio test. If the sequential probability ratio test is not effectively based on one observation, the expected sample number under one or both hypotheses is greater than one. Then the existence of a single-observation test with both probabilities of error at least as small as those of the sequential probability ratio test would contradict the optimum-property result.

The result can be extended to cases where the conditions \( B < 1 < A \) are not satisfied. Furthermore, it can be shown that each \( (\alpha, \beta) \) point below the curve achieved by one-observation procedures can be achieved by a sequential probability ratio test. The argument is based on some continuity properties. The idea is that given an arbitrary \( (\alpha, \beta) \) below the curve for one-observation procedures, there is a one-observation procedure with the same \( \beta \) and greater \( \alpha \); then the \( A \) can be increased and \( B \) can simultaneously be decreased while holding \( \beta \) fixed and decreasing \( \alpha \). Since Wijsman [4] has studied this situation, we shall not go into it further.

6. Uniqueness of a Sequential Probability Ratio Test

We have discussed some aspects of the existence of a sequential probability ratio test with specified probabilities of errors at the two hypotheses. A related question is the uniqueness of the sequential probability ratio test. That is, given two probability density functions or discrete probability functions, \( f_0(x) \) and \( f_1(x) \), is there more than one sequential probability ratio test that achieves a given \( \alpha \) (probability of rejecting \( H_0 \) when \( H_0 \) is true) and \( \beta \) (probability of accepting \( H_0 \) when \( H_0 \) is not true)? The answer is that there is not. Weiss [5] has shown this for the case in which \( f_1(x)/f_0(x) \) has a continuous distribution with positive probability on every interval in \((0, \infty)\). Here we give a proof that covers the general case and is simpler than the one given by Weiss. In particular, the case of the binomial distribution is included.
Theorem 4. There is at most one sequential probability ratio test for testing \( H_0: f(x) = f_0(x) \) against \( H_1 = f(x) = f_1(x) \), that achieves a given \( \alpha \) and \( \beta \).

Proof. Two tests are considered the same if they differ only on sample sequences that have probability 0 under each hypothesis. Suppose there is a pair of numbers \( A \) and \( B \), \( B < 1 < A \), such that the corresponding sequential probability ratio test achieves \( \alpha \) and \( \beta \). We shall show that any other pair \( A^* \) and \( B^* \) which gives the same probabilities of error defines the same test. First, suppose \( B^* \leq B \), \( A \leq A^* \). By the optimum property of the sequential probability ratio test [1] the expected sample sizes are the same for the two procedures, computed under each hypothesis. However, any sample sequence terminating on the \( n \)th step under the second procedure must terminate under the first procedure by at least the \( n \)th step. Hence

\[
(6.1) \quad \sum_{m=1}^{n} p_1(m) \geq \sum_{m=1}^{n} p_{1*}(m),
\]

where \( p_1(m) \) and \( p_{1*}(m) \) are the probabilities of terminating at exactly \( m \) steps under \( H_i \). Since the expected sample size is the same for the two procedures, it follows that \( p_1(m) = p_{1*}(m) \). Thus for every sample sequence (except for a set of probability 0), the number of observations to reach a decision is the same under the two procedures. If a sample sequence leads to termination at the \( n \)th step and acceptance of \( H_0 \) under the second procedure (a probability ratio not greater than \( B^* \)), it must do so under the first (not greater than \( B \)). Similarly, a sample sequence leading to termination at the \( n \)th step and rejection of \( H_0 \) under the second procedure must do so under the first. Thus the two procedures are the same. The case of \( B \leq B^* \), \( A^* \leq A \) is treated similarly.

Now consider \( B^* \leq B \), \( A^* \leq A \). A sequence leading to acceptance at step \( m \) under the second procedure must lead to acceptance at least by step \( m \) under the first procedure because the sequence of probability ratios has not reached \( A^* \) by the \( m \)th step and at the \( m \)th step is less than or equal to \( B^* (\leq B) \). Since the probabilities of acceptance for the two procedures are the same under each hypothesis, each sample sequence that leads to acceptance for one procedure must do so for the other. Similarly, any sequence leading to rejection under one procedure must do so under the other. It follows that the test with acceptance numbers \( B \) and \( A^* \) must yield the same decision for each sample sequence as would either \((B^*, A^*)\) or \((B, A)\). From the first case we see that all three of these procedures are the same.

The case \( B \leq B^* \), \( A \leq A^* \) is treated similarly.||

7. Specification of Procedures

A statistician designing a test and accustomed to employing sequential probability ratio test for this purpose will typically ask the user to specify two probabilities of error, \( \alpha \) and \( \beta \) at two parameter values termed \( H_0 \) and \( H_1 \). On this basis the statistician selects a sequential probability ratio test defined for this \( H_0 \) and \( H_1 \). However, as was shown above, he may not be able to find one that gives precisely the specified \( \alpha \) and \( \beta \); if not, he is likely to choose one that comes as close as possible in the direction of yielding greater “power.” In this event, he may be able to do better for the user. There may exist a sequential probability ratio test defined for a different pair of alternative hypotheses, \( H_0^* \) and \( H_1^* \), that nonetheless does better for the \( H_0 \) and \( H_1 \) specified
by the user than the one chosen by the statistician; that is, which comes closer to yielding the specified \( \alpha \) and \( \beta \) at \( H_0 \) and \( H_1 \) and has a smaller average number of observations at \( H_0 \) and \( H_1 \). Even if the statistician asks the user to specify values \( H_0, H_1, \alpha \) and \( \beta \), there is no reason why he has to restrict himself to the subclass of tests defined for this \( H_0 \) and \( H_1 \). Yet the method of formulating the sequential probability ratio test suggests that he do so, and the usual “cookbook” rules tell him to do so.

We can make this point in a different way, particularly for the binomial case. As we have seen, if the number of defectives or values equal to 1 are plotted on the ordinate and the number of observations on the abscissa, a sequential probability ratio test is defined by a pair of parallel straight lines. It takes three parameters to describe any specific test—the slope of the lines and the two intercepts. Now it happens that the values of \( p_1 \) and \( p_0 \) specify the slope of the lines independently of the values of \( \alpha \) and \( \beta \) [see (5.3)], but this limits more narrowly than is necessary the set of values of \( \alpha \) and \( \beta \) that can be attained.\(^8\) On the one hand, for any set of values of \( p_1, p_0, \alpha \), and \( \beta \), there may be more than one triplet of slope and intercepts that will yield them. If one of these is the sequential probability ratio test, it will require no larger an average sample size than any of the others, but if none is, we cannot now in general say which one will have this property. On the other hand, for any triplet of slope and intercepts there are an indefinite number of combinations of \( p_1, p_0, \alpha \), and \( \beta \) for which that triplet is the sequential probability ratio test.

One direction of investigation is to reformulate the sequential tests in terms of specifications other than \( p_1, p_0, \alpha \) and \( \beta \) that may establish a closer connection between the questions asked of the user and the appropriate test; then to adapt the proofs of optimum properties accordingly. One reformulation that comes readily to mind is to specify a \( p^* \), the probability of one decision at \( p^* \), and the desired slope of the operating characteristic curve at \( p^* \). However, this may not be the most suitable formulation.

More fundamentally, there is no reason why the definition of a class of tests needs to be in the same terms as the questions asked the user. The former should be chosen for mathematical and statistical convenience; the latter in light of the operating function to be served by the test. A fully satisfactory sampling-inspection theory would separate the two and require as information from the user only such items as the cost of inspection and the value of cost of rejecting lots of various characteristics. The rest would be the statistician's business.

**References**


Research carried out at the Center for Advanced Study in the Behavioral Sciences, Stanford, California. The theorem of Section 2 (which stimulated this paper) was derived by Milton Friedman in 1945 while a member of the Statistical Research Group at Columbia University, of which Harold Hotelling was Official Investigator.

Charles Stein (in unpublished work) has considered the effect of randomization on some of the problems treated in this paper.

The conditions $B < 1 < A$ are not necessary for the definition of the procedure but are necessary for the proof of the optimum property.

A method of randomization is to randomize between accepting $H_0$ and continuing sampling when equality holds in (3.2) and to randomize between rejecting $H_0$ and continuing sampling when equality holds in (3.3).

More generally one could divide $\left( \frac{1}{2}, \theta_0 + \frac{1}{2} \right)$ into $\left( \theta_1 - \frac{1}{2}, c \right)$, (c, d), and (d, $\theta_0 + \frac{1}{2}$). If an observation falls in the first interval, accept $H_0$; if it falls in the second interval, take another observation; and if it falls in the third, reject $H_0$. This type of procedure (which may apply in other problems) gives a kind of randomization.

The proof in this paper is simpler than that of Weiss, but it uses a more powerful result, namely, the optimum property of the sequential probability ratio test, which is more difficult to prove.

We assume that he can compute exactly the operating-characteristic curve of the test, though in practice he may only be able to approximate it.

This is another way of presenting the proof in Section 4. The curtailed single-sampling plan is defined by two parallel straight lines which have a slope less than $1/n$ and intercepts on the y axis, one positive, the other negative, that differ by less than unity. For some values of $p_0$ and $p_1$ the sequential probability ratio test specifies a slope greater than $1/n$, hence cannot be equivalent to the curtailed single-sampling plan.