STopping times of two rank order sequential probability ratio tests for symmetry based on lehmann alternatives

by harrison D. Weed, Jr., ralph a. bradley and z. govindaraju

FSU statistics report M148
ONR technical report No. 31

THE DEVELOPMENT OF STATISTICAL METHODS FOR QUALITY CONTROL AND SURVEILLANCE TESTING

july, 1969
department of statistics
florida state university
Tallahassee, Florida

research supported in part by the army, navy and air force under Office of Naval research contract No. NONR-988(08), Task Order NR 042-004, in part by the National institute of General medical sciences, training Grant 2-T01-GM0913-07 and in part by the National science foundation, Grant NSF-CP 7847. Reproduction in whole or in part is permitted for any purpose of the United States Government.
1. **Introduction and Summary.** Savage and Savage (1965) considered generalized sequential probability ratio tests (SPRT's) based on dependent and not necessarily identically distributed random variables, and obtained conditions for the stopping time to be finite with probability one and the expected stopping time to be finite. Their results were motivated by rank order problems. Wilcoxon, Rhodes and Bradley (1963) developed sequential grouped rank tests for the two-sample problem using Lehmann alternatives. Bradley, Merchant and Wilcoxon (1966) gave some Monte Carlo results on the modified two-sample procedure, modified in the sense that at each stage all observations available are reranked. Savage and Sethuraman (1966) showed that, under general conditions, the modified two-sample rank order SPRT based on Lehmann alternatives surely terminates and that the moment generating function of the sample size is finite.

In this paper, two, modified, one-sample, rank order SPRT's for symmetry based on Lehmann alternatives are given. The first model (Model I) is discussed by Weed and Bradley (1969) and Weed (1968); the second model (Model II) was proposed by Govindarajulu (1968) and is developed here. The two models depend on two different choices of Lehmann alternatives. The

---

1Research supported at the Florida State University in part by the Army, Navy and Air Force through ONR Contract NONR-988(08), Task Order NR 042-004 and in part by the National Institute of General Medical Science, Training Grant 2-T01-GM 00913-007.

2Research supported at the University of Kentucky by the National Science Foundation, Grant NSF-GP 7847.

Reproduction in whole or in part is permitted for any purpose of the United States Government.
results of this paper were developed independently but, since proofs are essentially the same, the authors have combined their work. It is shown, under very general conditions, that the two procedures terminate with probability one and the moments of stopping times are finite for all alternatives within the classes of alternatives defined by Models I and II. Further, procedures based on approximate probability ratios are also studied when the distribution function at zero under the alternative hypothesis is unknown but estimated from the sample observations. Parent (1965) also developed a sequential procedure based on ranks with a Lehmann alternative but it is not an SPRT.

2. **Notation and Formulation of the Problem.** Let \( Z_1, Z_2, \ldots \) be independent and identically distributed random variables observed sequentially and having a continuous cdf \( F \). We wish to test the hypothesis,

\[
H_0 : \quad F(z) + F(-z) = 1 \quad \text{for all } z,
\]

that is, \( F \) is symmetric about zero. Two models for \( F \) are considered leading to alternatives to (1). Let \( H(z) = P(Z < z | Z > 0) = \{F(z) - F(0)\}/\{1 - F(0)\} \) and \( G(z) = P(Z < z | Z < 0) = \{F(0) - F(-z)\}/F(0) \) for \( z > 0 \) and \( H(z), G(z) = 0, \) for \( z < 0 \). Thus \( F(z) = F(0)\{1 - G(-z)\} \) for \( z < 0 \) and \( F(z) = H(z) + F(0)\{1 - H(z)\} \) for \( z > 0 \). We can rewrite (1) as

\[
H_0 : \quad H(z) = G(z) \quad \text{for all } z \quad \text{and} \quad F(0) = 1/2
\]

and take

\[
H_1 : \quad H(z) \neq C(z) \quad \text{for some } z.
\]

However, since probabilities of desired rank orders cannot be derived explicitly under \( H_1 \), we postulate structure for two models under alternative hypotheses. Thus, we have
(2) **Model I**: \( H_{aI} : H(z) = 1 - (1-C(z))^A \) for all \( z, A > 0, A \neq 1, \)
A specified, \( F(0) = A/(1+A) \)
and

(3) **Model II**: \( H_{aII} : H(z) = c^A(z) \) for all \( z, A > 0, A \neq 1, A \) specified,
\( F(0) = \lambda_0, \lambda_0 \) specified.

Notice that, if \( X = -Z, Z < 0 \), and if \( Y = Z, Z \geq 0 \), and if \( X \) and \( Y \) have conditional cdf's \( (C \) and \( H \) respectively) satisfying (2), then the conditional cdf's of \(-X \) and \(-Y \) satisfy (3): the converse statement is also true. The two models are not the same and have been discussed by Weed and Bradley (1969). They have also given examples of cdf's \( F(z) \) and associated pdf's \( f(z) = F'(z) \) for Model I: it is easy to generate examples for Model II in the same way.

When the experiment has reached stage \( t, Z_1, Z_2, \ldots, Z_t \) have been observed. Let \( X_1, \ldots, X_m \) denote the absolute values of those \( Z \)'s that are negative and let \( Y_1, \ldots, Y_n \) denote the nonnegative \( Z \)'s, \( m + n = t \). Note that \( m \) is binomially distributed with parameters \( t \) and \( \lambda = F(0), 0 < \lambda < 1 \). Let the combined sample of \( X \)'s and \( Y \)'s be denoted by \( W_1, \ldots, W_t \) and the ordered combined sample by \( W_{t1}, \ldots, W_{tt} \). Let \( C_m \) and \( H_n \) respectively denote the empirical cdf's of \( X_1, \ldots, X_m \) and \( Y_1, \ldots, Y_n \). Further, following Savage's (1959) definition, let \( \delta = (\delta_1, \ldots, \delta_t) \) where \( \delta_i = 1 \) or 0 according as \( W_{ti} \) corresponds to a negative or nonnegative \( Z \) respectively. Also, let

\[
(4) \quad L_t(A, \delta) = P_t(\Delta=\delta|A)/P_t(\Delta=\delta|A=1) = 2^{tP_t(\Delta=\delta|A)}
\]

where \( \Delta \) represents the random rank order associated with the observed \( \delta \) and \( P_t(\Delta=\delta|A) \) denotes the probability of the rank order \( \delta \) when either (2) or
(3) holds. The SPRT for testing $H_0$ against $H_{aI}$ or $H_{aII}$ is given by:

(i) Take one more observation if $a < L_t < b$,

(ii) Accept $H_0$ if $L_t \leq a$,

(iii) Reject $H_0$ if $L_t \geq b$, $t = 1, 2, \ldots$

where $0 < a < 1 < b$ are suitable constants (independent of $t$).

The number of stages before termination $T$ is defined as follows:

(6) $T = r$ if $a < L_t < b$ for $t = 1, \ldots, r-1$ and $L_r \geq b$ or $L_r \leq a$,

$r = 1, 2, \ldots$.

We investigate properties of the distribution of $T$.

3. Preliminary Results. We obtain explicit expressions for $L_t(A, \delta)$.

**LEMMA 1.** With the preceding notation and Model II, we have

(7) $L_t(A, \delta) = 2^t t! \lambda_0^m (1-\lambda_0)^n A^t \prod_{i=1}^{n} \{mC_n(W_i) + An_n(W_i)\}^{-1}.$

**Proof.** Following Savage (1959) and using the definitions of $G$ and $H$, we have

$$P(\Delta=\delta|A) = \left( \begin{array}{c} t \\ m \end{array} \right) \lambda_0^m (1-\lambda_0)^n P(\Delta=\delta|A, m), \lambda_0 = F(0),$$

and

$$P(\Delta=\delta|A, m) = m! n! \int \ldots \int \prod_{j=1}^{t} \{dG(y_j)\}^{\delta_j} \{dH(y_j)\}^{1-\delta_j}$$

$$= m! n! A \int \ldots \int \prod_{j=1}^{t} \{\delta_j + A(1-\delta_j)\}$$

$$= m! n! A \prod_{j=1}^{t} \left( \frac{1}{\sum_{i=1}^{j} \delta_i} + A(1-\sum_{j=1}^{i} \delta_i) \right)^{-1},$$

$$= m! n! A.$$
$P(\Delta=\delta|A,m)$ being the conditional probability that $\Delta = \delta$ given $m$. Note that $\sum_{j=1}^{i} \delta_j$ is the number of negative $Z$'s that are less than or equal to $W_{ti}$ in absolute value and that $(i-\sum_{j=1}^{i} \delta_j)$ is the number of nonnegative $Z$'s that are less than or equal to $W_{ti}$. Hence $\sum_{j=1}^{i} \delta_j = m G_m(W_{ti})$ and $(i-\sum_{j=1}^{i} \delta_j) = n H_n(W_{ti})$, $i = 1, \ldots, t$. Then

$$P(\Delta=\delta|A) = t! \lambda_0^m (1-\lambda_0)^n A^n \prod_{i=1}^{t} \left[ m G_m(W_{ti}) + An H_n(W_{ti}) \right]^{-1}.$$

The desired result follows since $P(\Delta=\delta|A=1) = 2^{-t}$ and the $W_{ti}$ in (7) may be replaced by the $W_i$ because of the symmetry of the product.

Analogously, for Model I, one can easily obtain

$$\tilde{L}_t(A,\delta) = 2^{-t} t! \{A/(1+A)\}^t \prod_{i=1}^{t} \left[ m \tilde{G}_m(W_i) + An \tilde{H}_n(W_i) \right]^{-1},$$

as given by Weed and Bradley (1969) where $\tilde{G}_m$ and $\tilde{H}_n$ are cumulative relative frequencies accumulated from the right instead of from the left as in $G_m$ and $H_n$ and as is more conventional. Notice that (9) is obtained from (7) by replacement of $G_m$ and $H_n$ by $\tilde{G}_m$ and $\tilde{H}_n$ respectively and since $\lambda_0 = A/(1+A)$. Hence, it is sufficient to study in detail the termination of the SPRT based on Model II and an analogous method holds for the SPRT based on Model I.

Let

$$tS_t = \log \tilde{L}_t(A,\delta).$$

Then, from (7), we have

$$S_t = \log 2 - 1 + \log\{A(1-\lambda_0)\} - \lambda_t \log\{A(1-\lambda_0)/\lambda_0\} - B_t + O(\log t/t)$$

where
(12) \[ B_t = t^{-1} \sum_{i=1}^{t} \log(\lambda_t G(W_i) + A(1-\lambda_t)H(W_i)), \]

(13) \[ \lambda_t = m/t, \]

and we have noted that \( t^{-1} \ln t! = -\ln t = -1 + O(\log t/t). \) We establish certain preliminary lemmas.

**Lemma 2.** Chernoff (1952). Let \( V_1, \ldots, V_t \) be independent and identically distributed random variables with mean \( E(V_1) \) and finite moment generating function in a neighborhood of zero. Then, for each \( \epsilon > 0 \) and sufficiently large \( t, \)

(14) \[ P[|\{(V_1 + \cdots + V_t)/t\} - E(V_1)| \geq \epsilon] \leq \rho_1^t(\epsilon), \quad 0 < \rho_1(\epsilon) < 1. \]

**Proof.** See Theorem 1 of Chernoff (1952).

**Corollary 1.** Consider \( \lambda_t \) in (13). Then, for every \( \epsilon > 0 \) and sufficiently large \( t, \) there exists a \( \rho_1 \) such that

(i) \[ P[|\lambda_t^{t} - \lambda| \geq \epsilon] \leq \rho_1^t(\epsilon), \quad 0 < \rho_1(\epsilon) < 1, \]

(15) \[ (ii) \quad P[\log(\lambda_t^{t} / \lambda) \geq \epsilon] \leq \rho_1^t(\epsilon, \lambda), \quad 0 \leq \rho_1(\epsilon, \lambda) < 1, \quad 0 < \lambda < 1, \]

(iii) \[ P[\log((1-\lambda_t)^{1-\lambda} / (1-\lambda)^{1-\lambda}) \geq \epsilon] \leq \rho_1^t(\epsilon, \lambda), \quad 0 < \lambda < 1. \]

**Proof.** Part (i) trivially follows from Lemma 2. Part (iii) follows by symmetry from (ii). Hence it suffices to prove (ii). One can write

\[ \lambda_t^{a_t} = \lambda \left( 1 + \frac{a_t}{\lambda} \right)^{a_t}, \quad a_t = \lambda_t - \lambda. \]
Thus
\[
\Pr\{\log(\lambda_t^{\lambda_t^{-\lambda}}) > \varepsilon\} \leq \Pr\{|a_t| > \varepsilon/2|\log \lambda\|\} + \Pr\{\log(1 + \frac{|a_t|}{\lambda}) > \varepsilon/2\}
\]
\[
\leq \rho_t^t(\varepsilon/2|\log \lambda\|) + \rho_t^t(\varepsilon\lambda/2) \leq \rho_t^t(\varepsilon, \lambda)
\]
since \(\lambda\) is fixed and is bounded away from zero and unity. This completes the proof of Corollary 1.

**Lemma 3.** Let \(U_i, i = 1, \ldots, t\), be independent uniform random variables on \([0,1]\). Thus, for given \(\varepsilon > 0\), we can find \(\delta_0\) and \(0 < \rho_2(\varepsilon) < 1\) such that, for sufficiently large \(t\) and \(0 < \delta < \delta_0\),

\[
\Pr\{t^{-1} \sum_{i=1}^{t} \log(1+\delta/U_i) > \varepsilon\} < \rho_2^t(\varepsilon).
\]

**Proof.** This result was given by Savage and Sethuraman (1966) with \(\delta_0\) replacing \(\delta\) in (16). The monotonicity of \(\log(1+\delta/U_i)\) with respect to \(\delta\) gives (16) from the earlier work.

**Lemma 4.** Sethuraman (1964). Let \(K(u)\) be a continuous cdf for a random variable \(U\) and let \(K_t(u)\) be the associated empirical cdf based on \(t\) independent observations on \(U\). Then, for each \(\delta > 0\) and \(t\) sufficiently large, there exists \(\rho_3(\varepsilon)\), \(0 < \rho_3 < 1\), such that

\[
\Pr(\sup_u |K_t(u) - K(u)| > \varepsilon) \leq \rho_3^t(\varepsilon).
\]

**Proof.** See Theorem 1 of the reference.
LEMMA 5. Let $\Omega_1(m) = \sup_z |G_m(z) - G(z)|$, $\Omega_2(n) = \sup_z |H_n(z) - H(z)|$ and

$$\Omega(t) = (1+A)|\lambda_t - \lambda| + \lambda \Omega_1(m) + A(1-\lambda)\Omega_2(n).$$

Then, for each $\varepsilon > 0$, there exists a $\rho_4(\varepsilon)$, $0 \leq \rho_4(\varepsilon) < 1$, such that, for sufficiently large $t$,

$$(18) \quad P(\Omega(t) \geq \varepsilon) \leq \rho_4^t(\varepsilon).$$

**Proof.** Consider

$$P(\Omega_1(m) > \varepsilon) = \sum_{r=0}^{t} P(\Omega_1(m) > \varepsilon | m=r)P(m=r)$$

$$= \sum_{|r/t - \lambda| \leq \varepsilon} P(\Omega_1(m) > \varepsilon | m=r)P(m=r)$$

$$+ \sum_{|r/t - \lambda| > \varepsilon} P(\Omega_1(m) > \varepsilon | m=r)P(m=r)$$

$$\leq \sum_{|r/t - \lambda| \leq \varepsilon} \rho_3^t(\varepsilon)P(m=r) + \sum_{|r/t - \lambda| > \varepsilon} P(m=r)$$

$$\leq \rho_3^t(\lambda - \varepsilon)(\varepsilon) \sum_{r \geq t(\lambda - \varepsilon)} P(m=r) + \rho_1^t(\varepsilon)$$

$$\leq \rho_3^t(\lambda - \varepsilon)(\varepsilon) + \rho_1^t(\varepsilon) \leq \rho_5^t(\varepsilon)$$

for $0 \leq \rho_5(\varepsilon) < 1$. In reaching this result we have used Lemma 4 in the first summation and Part (i) of Corollary 1 in the second. A similar result holds for $P(\Omega_2(n) \geq \varepsilon)$. Now

$$P(\Omega(t) > \varepsilon) \leq P((1+A)|\lambda_t - \lambda| > \varepsilon/2) + P[\lambda \Omega_1(m) > \lambda \varepsilon/2(1+\lambda)]$$

$$+ P[A(1-\lambda)\Omega_2(n) > A(1-\lambda)\varepsilon/2(\lambda + (1-\lambda)A)]$$

$$\leq \rho_1^t(\varepsilon/2(1+A)) + 2\rho_5^t(\varepsilon/2(1+A)) \leq \rho_4^t(\varepsilon), \quad 0 \leq \rho_4(\varepsilon) < 1.$$
LEMMA 6. Let $\mathcal{B}_t^{(1)} = t^{-1} \sum_{i=1}^{\infty} \log(\lambda G(Y_i) + A(1-\lambda)H(Y_i))$ and

\begin{align*}
(19) \quad \mathcal{B}_t^{(1)}(A,G,H) &= \int_0^\infty \log(\lambda G(z) + A(1-\lambda)H(z)) d\{\lambda G(z) + (1-\lambda)H(z)\} \\
&= \lambda E_{G}[\log(\lambda G(X) + A(1-\lambda)H(X))] + (1-\lambda)E_{H}[\log(\lambda G(Y) + A(1-\lambda)H(Y))].
\end{align*}

Then, for every $\varepsilon > 0$, there is a $\rho_6(\varepsilon)$, $0 \leq \rho_6(\varepsilon) < 1$, such that for large $t$

\begin{equation}
(20) \quad P\{|B_t^{(1)} - \mathcal{B}_t^{(1)}(A,G,H)| \geq \varepsilon\} \leq \rho_6^t(\varepsilon).
\end{equation}

Proof. We rewrite $\mathcal{B}_t^{(1)}$ as

\[ B_t^{(1)} = \lambda t^{-1} \sum_{i=1}^{m} \log(\lambda G(X_i) + A(1-\lambda)H(X_i)) + (1-\lambda) t^{-1} \sum_{j=1}^{n} \log(\lambda G(Y_j) + A(1-\lambda)H(Y_j)). \]

Let $V_i = \log(\lambda G(X_i) + A(1-\lambda)H(X_i))$, $i = 1, \ldots, m$, and $V_j^* = \log(\lambda G(Y_j) + A(1-\lambda)H(Y_j))$, $j = 1, \ldots, n$. For fixed $m$, the $V_i$ are independent and identically distributed random variables having a finite moment generating function. Application of Lemma 2 and Part (i) of Corollary 1 in a manner similar to that of the proof of Lemma 5 yields, for every $\varepsilon > 0$, a $\rho_7(\varepsilon)$, $0 \leq \rho_7(\varepsilon) < 1$, such that

\begin{equation}
(21) \quad P\left(\left|m^{-1} \sum_{i=1}^{m} V_i - EV_1\right| \geq \varepsilon\right) \leq \rho_7^t(\varepsilon).
\end{equation}

An analogous result holds for the $V_j^*$. Hence
\[
P\left( |B_t^{(1)} - B_\lambda(A, G, H)| \geq \varepsilon \right) \leq P\left( |\lambda_t \sum_{i=1}^{m-1} v_{1,i} - E(V_1)| \geq \varepsilon / 2 \right) + P\left( |(1-\lambda_t) \sum_{j=1}^{n-1} v_{1,j}^* - E(V_1^*)| \geq \varepsilon / 2 \right) \\
+ P\left( |(\lambda_t - \lambda) (E(V_1) - E(V_1^*))| \geq \varepsilon / 2 \right) \\
\leq P\left( |m^{-1} \sum_{i=1}^{m} v_{1,i} - E(V_1)| \geq \varepsilon / 4 \right) + P\left( |n^{-1} \sum_{j=1}^{n} v_{1,j}^* - E(V_1^*)| \geq \varepsilon / 4 \right) \\
+ P\left( |\lambda_t - \lambda| \geq \varepsilon_1 \right) \leq \rho_6^t(\varepsilon), \quad 0 \leq \rho_6(\varepsilon) < 1,
\]

where \( \varepsilon_1 = \varepsilon / 4 |\log(1 + A)| \) since \(|E(V_1) - E(V_1^*)| \leq |E(V_1)| + |E(V_1^*)| \leq 2 |\log(1 + A)| \).

**Lemma 7.** Let \( B_t^{(2)} = t^{-1} \sum_{i=1}^{t} \log(\lambda_t G(W_i) + A(1-\lambda_t) H(W_i)) - B_t^{(1)} \). Then, for every \( \varepsilon > 0 \), there is a \( \rho_6(\varepsilon) \), \( 0 \leq \rho_6(\varepsilon) < 1 \), such that for large \( t \)

(22) \[
P\left( |B_t^{(2)}| \geq \varepsilon \right) \leq \rho_6^t(\varepsilon).
\]

**Proof.** One can easily write

\[
B_t^{(2)} \leq t^{-1} \sum_{i=1}^{t} \log \left\{ \frac{\lambda \Omega_1(m) + A(1-\lambda) \Omega_2(n) + (G_m - AH_n) (\lambda_t - \lambda)}{\lambda G(W_i) + A(1-\lambda) H(W_i)} \right\};
\]

hence

(23) \[
B_t^{(2)} \leq m^{-1} \sum_{i=1}^{m} \log \{1 + \Omega(t) / \lambda G(Y_i)\} + n^{-1} \sum_{j=1}^{n} \log \{1 + \Omega(n) / A(1-\lambda) H(Y_j)\}
\]

where \( \Omega(t), \Omega_1(m) \) and \( \Omega_2(n) \) are as defined in Lemma 5. Now choose \( \delta_0 \) in accordance with Lemma 3 so that...
\[ P\left[ m^{-1} \sum_{i=1}^{m} \log \left( 1 + \Omega(t) / \lambda G(X_i) \right) \geq \varepsilon / 2 \right] \leq P\left( |\lambda_t - \lambda| \geq \delta_0 \right) + \mathbb{P}\{ \Omega(t) \geq \delta_0 \lambda \} \]
\[ + \sum_{\Omega(t)/\lambda, |r/t - \lambda| \leq \delta_0} P\left[ m^{-1} \sum_{i=1}^{m} \log \left( 1 + \Omega(t) / \lambda G(X_i) \right) \geq \varepsilon / 2 |m=r\right] \leq \rho_9(\varepsilon), \quad 0 \leq \rho_9(\varepsilon) < 1, \]

after using Corollary 1, Lemma 3 and Lemma 5. An analogous result holds for \( n^{-1} \sum_{i=1}^{n} \log \left( 1 + \Omega(t) / A(1 - \lambda) H(Y_j) \right) \). Hence,

\[ (24) \quad P(B_t^{(2)} \geq \varepsilon) \leq \rho_{10}(\varepsilon), \quad 0 \leq \rho_{10}(\varepsilon) < 1. \]

Similarly one can bound \(-B_t^{(2)}\) and obtain for \( \Lambda \geq 1 \),

\[ -B_t^{(2)} \leq t^{-1} \sum_{i=1}^{m} \log \left( 1 + \Omega(t) / \lambda G(X_i) \right) + t^{-1} \sum_{j=1}^{n} \log \left( 1 + \Omega(t) / A(1 - \lambda) H(Y_j) \right) \]
\[ \leq t^{-1} \sum_{i=1}^{m} \log \left( 1 + t \Omega(t) / i \right) + t^{-1} \sum_{j=1}^{n} \log \left( 1 + t \Omega(t) / Aj \right) \]
\[ \leq t^{-1} \sum_{i=1}^{m} \log \left( 1 + t \Omega(t) / i \right) < t^{-1} \int_{0}^{t} \log \left( 1 + t \Omega(t) / x \right) dx \]
\[ = (1 + \Omega(t)) \log \left( 1 + \Omega(t) \right) - \Omega(t) \log \Omega(t). \]

Thus

\[ (25) \quad P(B_t^{(2)} \leq -\varepsilon) \leq P\{ (1 + \Omega(t)) \log \left( 1 + \Omega(t) \right) - \Omega(t) \log \Omega(t) \geq \varepsilon \}
\[ = P\{ \Omega(t) \geq \varepsilon' \} \leq \rho_5(\varepsilon') = \rho_{11}(\varepsilon)^t \]

after use of Lemma 5 and where \( \varepsilon' \) depends only on \( \varepsilon \) while \( 0 \leq \rho_{11}(\varepsilon) < 1 \).
An analogous argument can be employed to cover the case with $A < 1$. This completes the proof of Lemma 7 as $P( |B_t^{(2)}| \geq \varepsilon ) \leq \rho_{10}^t(\varepsilon) + \rho_{11}^t(\varepsilon) \leq \rho_8^t(\varepsilon)$, $0 \leq \rho_8^t(\varepsilon) < 1$.

**Lemma 8.** Let $B_t$ and $B_\lambda^{(A,C,H)}$ be given in (12) and (19) respectively. Then, for every $\varepsilon > 0$, there exists a $\rho_{12}(\varepsilon)$, $0 \leq \rho_{12}(\varepsilon) < 1$, such that for sufficiently large $t$ we have

\[(26) \quad P\{ |B_t - B_\lambda^{(A,C,H)}| \geq \varepsilon \} \leq \rho_{12}^t(\varepsilon),\]

**Proof.** One can write $B_t$ as

\[
B_t = t^{-1} \sum_{i=1}^{t} \log \{ \lambda G(W_i) + A(1-\lambda)H(W_i) \} + t^{-1} \sum_{i=1}^{t} \log \left\{ \frac{\lambda C_m(W_i) + A(1-\lambda)H(W_i)}{\lambda G(W_i) + A(1-\lambda)H(W_i)} \right\} = B_t^{(1)} + B_t^{(2)}
\]

where $B_t^{(1)}$ and $B_t^{(2)}$ are respectively as defined in Lemmas 6 and 7. One can easily verify that $E(B_t^{(1)}) = B_\lambda^{(A,C,H)}$ and

\[
P\{ |B_t - B_\lambda^{(A,C,H)}| \geq \varepsilon \} \leq P\{ |B_t^{(1)} - B_\lambda^{(A,C,H)}| \geq \varepsilon/2 \} + P\{ |B_t^{(2)}| \geq \varepsilon/2 \} \leq \rho_{12}^t(\varepsilon)
\]

after applying Lemmas 6 and 7.
Lemma 9. Let \( S_t \) be given by (11). Define

\[
S_\lambda(A, \lambda_0, G, H) = \log 2 - 1 + \log(A(1-\lambda_0)) - \lambda \log(A(1-\lambda_0)/\lambda_0) - B_\lambda(A, G, H)
\]

where \( B_\lambda(A, G, H) \) is given by (19). Then, for every \( \varepsilon > 0 \), there is a \( \rho_{13}(\varepsilon) \), \( 0 \leq \rho_{13}(\varepsilon) < 1 \), such that for sufficiently large \( t \) we have

\[
P\{ |S_t - S_\lambda(A, \lambda_0, G, H)| \geq \varepsilon \} \leq \rho_{13}^t(\varepsilon).
\]

Proof. Consider

\[
S_t - S_\lambda(A, \lambda_0, G, H) = (\lambda_t - \lambda) \log(A(A(1-\lambda_0)) - B_t + B_\lambda(A, G, H) + o(\log t/t).
\]

Now, the desired result follows from Corollary 1 and Lemma 8.

4. The Basic Results. We are ready to give the main theorems.

Theorem 1. Let \( S_\lambda(A, \lambda_0, G, H) \) defined by (27) be not equal to zero and let \( T \) denote the number of stages before termination of the SPRT under Model II. Then

(i) \( P(T > t) < \rho^t \) for sufficiently large \( t \) and some \( 0 \leq \rho < 1 \),

(ii) \( P(T < \infty) = 1 \),

(iii) \( E(e^{\theta T}) < \infty \) for \( \theta \) in some interval \((-\delta, \delta)\), \( \delta > 0 \).

Proof. Parts (ii) and (iii) immediately follow from (i). If \( S_\lambda(A, \lambda_0, G, H) \neq 0 \),

\[
P(T \leq t) \geq P(L_t \leq a \text{ or } L_t \geq b) = P\{ S_t \leq (\log a)/t \text{ or } S_t \geq (\log b)/t \}
\]

\[
\geq P\{ |S_t - S_\lambda(A, \lambda_0, G, H)| \leq \varepsilon \} \geq 1 - \rho^t(\varepsilon), \ 0 \leq \rho(\varepsilon) < 1
\]

for sufficiently large \( t \). Hence (i) follows from Lemma 9. This completes the proof of Theorem 1.
Remark 1. When $H_0$ or $H_{aI}$ is true, $S_{\lambda}(A,\lambda_0,G,H) \neq 0$ provided that $A \neq 1$ and/or $\lambda_0 \neq 1/2$ and the SPRT terminates with probability 1. For example, under $H_0$ we have $H = G$ and $\lambda = 1/2$ and

$$S_{1/2}(A,\lambda_0,G,G) = \frac{1}{2} \log \left\{ 16A \lambda_0(1-\lambda_0)/(1+A)^2 \right\}.$$  \hspace{1cm} (31)

Next, we shall present the theorem pertaining to the sure termination of the SPRT based on Model I:

Theorem 2. Suppose $\tilde{S}_{\lambda}(A,G,H) \neq 0$ where

$$\tilde{S}_{\lambda}(A,G,H) = \log 2 - 1 + \log \left\{ A/(1+A) \right\} - \tilde{B}_{\lambda}(A,G,H)$$ \hspace{1cm} (32)

and

$$\tilde{B}_{\lambda}(A,G,H) = \int_0^\infty \log \left\{ \lambda(1-G(z)) + A(1-\lambda)(1-H(z)) \right\} d\{\lambda G(z) + (1-\lambda)H(z)\}.$$  

Let $\tilde{T}$ denote the number of stages before termination of the SPRT under Model I. Then the conclusions of Theorem 1 hold for $\tilde{T}$.

Proof. Notice that $L_t(A,\delta)$ for Model I is obtained from the expression for $L_t(A,\delta)$ for Model II when $\lambda_0 = A/(1+A)$ and through replacement of $G_m$ and $H_n$ by $\tilde{G}_m$ and $\tilde{H}_n$ respectively. Thus $\tilde{S}_{\lambda}(A,G,H)$ is obtained from $S_{\lambda}(A,\lambda_0,G,H)$ when $\lambda_0 = A/(1+A)$ and with replacement of $G$ and $H$ by $1-G$ and $1-H$ respectively before integration in $B_{\lambda}(A,G,H)$.

Remark 2. When $H_0$ or $H_{aI}$ is true, we have again that $\tilde{S}_{\lambda}(A,G,H) \neq 0$ and termination of the SPRT with probability 1 when $A \neq 1$. Under $H_0$,

$$\tilde{S}_{1/2}(A,G,G) = \log \{ 4A/(1+A)^2 \}.$$  \hspace{1cm} (33)
Throughout we assumed that $0 < \lambda < 1$. Now, let us consider the case $\lambda = 0$. Then, $m = 0$ and $n = t$ and consequently for Model II, we obtain

\[(34) \quad L_t(A, \delta) = 2^t(1-\lambda_0)^t \quad \text{or} \quad S_t = \log(2(1-\lambda_0)).\]

Hence $P(S_t \leq \log a/t) = 1$ for $\lambda_0 > 1/2$ and sufficiently large $t$ and $P(S_t \geq \log b/t) = 1$ for $\lambda_0 < 1/2$ and sufficiently large $t$. For Model I by setting $\lambda_0 = A/(1+A)$ we observe that the sequential procedure terminates with probability one for all $A > 0$. Analogously we can cover the case $\lambda = 1$.

In the following we shall list some of the properties of $S_{\lambda}(A, \lambda_0, G, H)$ and $\tilde{S}_{\lambda}(A, G, H)$.

**Lemma 10.**

(i) $S_{\lambda}(1, \lambda_0, G, G) = \log 2 + (1-\lambda)\log(1-\lambda_0) + \lambda \log \lambda_0$.

(ii) $S_{\lambda}(1, 1/2, G, G) = 0$.

(iii) $S_{\lambda}(A, \lambda_0, G, \ell(G))$ is independent of $G$ where $\ell(\cdot)$ denotes a distribution function on $[0, 1]$.

(iv) $S_{1-\lambda}(1/A, 1-\lambda_0, H, G) = S_{\lambda}(A, \lambda_0, G, H)$.

(v) $S_{1/2}(A, \lambda_0, G, G) = \log[4(A\lambda_0(1-\lambda_0)^{1/2}/(1+A))]$ and is zero only when $\lambda_0 = 1/2$ and $A = 1$.

(vi) For each $A$ there exists a unique $C(A)$ lying between 1 and $A$ such that

\[S_{1/2}(A, 1/2, G, G^C) = 0. \quad \text{Further} \quad 1/C(1/A) = C(A).\]
Proof. Properties (i) - (v) trivially follow from the definition of $S_\lambda(A, \lambda_0, C, H)$. In (vi) we assume that the Lehmann alternatives as well as the sampling distribution have zero medians. Thus, when $\lambda = \lambda_0 = 1/2$, $S_\lambda(A, \lambda_0, C, H)$ is equal to one half the parameter studied by Savage and Sethuraman (1966). Hence $S_{1/2}(A, 1/2, C, H)$ enjoys all the properties given in Lemma 4 of Savage and Sethuraman. In particular, the pairs of values of $(A, C)$ for which $S_{1/2}(A, 1/2, C, C^C) = 0$ will be those values tabulated in the reference.

**Lemma 11.** Let $\tilde{S}_{C/(1+C)}(A, C, C^C) = \tilde{S}(A, C)$.

(i) $\tilde{S}_\lambda(1, C, H) = 0$.

(ii) $\tilde{S}(1/A, 1/C) = \tilde{S}(A, C)$.

(iii) $\tilde{S}(A, C)$ is strictly increasing in $C$ if $A > 1$.

(iv) $\tilde{S}(A, 1) < 0$ for $A \neq 1$.

(v) $\tilde{S}(A, A) > 0$ for $A \neq 1$.

(vi) For each $A$ there is a unique $C(A)$ lying between 1 and $A$ such that $\tilde{S}(A, C) = 0$.

**Proof.** Recall $\tilde{S}_\lambda(A, C, H)$ and $\tilde{B}_\lambda(A, C, H)$ in (32). Part (i) follows since $\tilde{B}_\lambda(1, C, H) = -1$.

Consider $\tilde{S}(A, C) = \log 2 - 1 + \log(A/(1+A)) - \tilde{B}(A, C)$ where $\tilde{B}(A, C) = \tilde{B}_{C/(1+C)}(A, C, C^C)$. That is,

$$\tilde{B}(A, C) = \int_0^1 \log \left( \frac{Ct + A t^C}{C^C + t} \right) dt. \left( \frac{t^C}{C^C + t} + \frac{C}{C^C + t} \right).$$

One can easily show that $\tilde{B}(1/A, 1/C) = \tilde{B}(A, C) - \log A$ and this implies (ii).
To prove (iii), one can write

\[ \tilde{B}(A,C) = A^{-1} \int_0^1 \log \left( \frac{Ct}{1+C} + \frac{At}{1+C} \right) \, d \left( \frac{Ct}{1+C} + \frac{At}{1+C} \right) \]

\[ + \frac{C}{1+C} (1-1/A) \int_0^1 \log \left( \frac{Ct}{1+C} + \frac{At}{1+C} \right) \, dt \]

\[ = \frac{1}{A(C+1)} \left[ (C+A) \left\{ \log \left( \frac{C+A}{C+1} \right) - 1 \right\} + C(A-1) \int_0^1 \log \left( \frac{Ct+At}{1+C} \right) \, dt \right]. \]

(36)

Now (iii) follows from (36) because \( \tilde{B}(A,C) \) is strictly decreasing in \( C \) for \( A > 1 \) and \( C > 0 \) as can be seen from an examination of the derivative of \( \tilde{B} \) with respect to \( C \). Also, from (36), we have \( \tilde{B}(A,1) = \log \{ (1+A)/2 \} - 1 \) and hence \( \tilde{S}(A,1) = \log \{ 4A(1+A)^{-2} \} \), \( \tilde{S}(A,1) < 0 \) for \( A \neq 1 \).

Direct substitution with use of (36) yields

\[ \tilde{S}(A,A) = \frac{A-1}{A+1} \int_0^1 \log \left( \frac{2}{1+t^{A-1}} \right) \, dt. \]

It follows immediately that \( \tilde{S}(A,A) > 0 \) for \( A \neq 1 \) and \( \tilde{S}(1,1) = 0 \).

The final part of the lemma follows directly from (iii), (iv) and (v) for \( A > 1 \). If \( A < 1 \), similar to (iii), \( \tilde{S}(A,C) \) is strictly decreasing in \( C \) and again (vi) follows for \( A < 1 \).

This completes the proof of Lemma 11.
5. The Case Where $\lambda_0$ Is Unknown. So far we have assumed that $F(0)$ is specified under $H_{aI}$ or $H_{aII}$. Indeed, Weed and Bradley (1969) assume throughout that Model I applies to the population sampled and take $H_0: A = 1$ and $H_a: A = A_1$ specified. Thus they always take $F(0) = A/(1+A)$. However, if $\lambda_0$ is unknown, a reasonable procedure would be to replace $\lambda_0$ by $\lambda_t$ in (8) and base the sequential procedure on $L_t^*(A, \delta)$ where

$$L_t^*(A, \delta) = 2^{t!} \lambda_t^m (1-\lambda_t)^n A^m \prod_{i=1}^{i=n} \{m \cdot \gamma_i + \lambda_i \cdot \gamma_i \}^{-1}.$$  

If $tS_t^* = \log L_t^*(A, \delta)$, then

$$S_t^* = \log 2 - 1 + \log\{A(1-\lambda_t)\} - \lambda_t \log\{A(1-\lambda_t)/\lambda_t\} - B_t + o(\log t/t),$$

where $B_t$ is defined in (12). Also define

$$S_\lambda^*(A, G, H) = \log 2 - 1 + \log\{A(1-\lambda)\} - \lambda \log\{A(1-\lambda)/\lambda\} - E_\lambda(A, G, H)$$

where $E_\lambda(A, G, H)$ is given by (19). Then

$$S_t^* - S_\lambda^*(A, G, H) = (\lambda - \lambda_t) \log A + \log(\lambda_t/\lambda)^{1-\lambda} + \log\{(1-\lambda_t)^{1-\lambda}\} + \log\{(1-\lambda)^{(1-\lambda)}\}. $$

$$- B_t - B_\lambda(A, G, H) + o(\log t/t).$$
Applying Corollary 1, Parts (ii) and (iii), and Lemma 8, we assert that for every \( \varepsilon > 0 \) there exists a \( \rho, 0 \leq \rho < 1 \), such that

\[
P\{ |S^*_t - S^*_\lambda(A,G,H)| \geq \varepsilon \} \leq \rho t \quad \text{for sufficiently large } t.
\]

Hence we have the following theorem:

**THEOREM 3.** Let \( 0 < \lambda < 1 \) and let \( S^*_\lambda(A,G,H) \) defined by (39) be not equal to zero. Also, let \( T^*_\lambda \) denote the number of stages before termination of the SPRT based on \( L^*_t(A,\delta) \). Then, the conclusions of Theorem 1 hold for \( T^*_\lambda \).

We could also suppose that \( \lambda_0 \) is not specified in Model I and consider a modified SPRT based on \( \tilde{L}^*_t(A,\delta) \) where

\[
\tilde{L}^*_t(A,\delta) = 2^t t \lambda_0 t^{m(1-\lambda)} n^{t} \prod_{i=1}^{\tilde{\sigma}^*} \left[ m \tilde{G}_m(W_i) + A n \tilde{H}_n(W_i) \right]^{-1}.
\]

Then

\[
\tilde{S}^*_t = \frac{1}{t} \log \tilde{L}^*_t(A,\delta) = \log 2 - 1 + \log(A(1-\lambda)) - \lambda t \log(A(1-\lambda_t)/\lambda_t) - \tilde{B}_t + O(\log t/t).
\]

Define

\[
\tilde{S}^*_t(A,G,H) = \log 2 - 1 + \log(A(1-\lambda)) - \lambda \log(A(1-\lambda)/\lambda) - \tilde{B}_\lambda(A,G,H),
\]

where \( \tilde{B}_\lambda(A,G,H) \) is given following (32). Then, we are led to the following theorem:
THEOREM 4. Let $0 < \lambda < 1$ and let $\tilde{S}_\lambda^*(A,G,H)$ defined by (44) be not equal to zero. Also, let $\tilde{T}^*$ denote the number of stages before termination of the SPRT based on $L_t^*(A,\delta)$. Then, the conclusions of Theorem 1 hold for $\tilde{T}^*$.

In Theorems 3 and 4, we assumed that $0 < \lambda < 1$. Now, let us consider the case $\lambda = 0$. Then, since $m \equiv 0$ and $n \equiv t$, we obtain

$$L_t^*(A,\delta) = L_t^*(A,\delta) = 2^t$$
and hence

$$S_t^* = S_t^* = \log 2 \geq \log b/t$$

for $t > t_0 = \log b/\log 2$. Consequently the SPRT's terminate finitely.

Analogously, when $\lambda = 1$, $m \equiv t$, $n \equiv 0$ and $L_t^*(A,\delta) = L_t^*(A,\delta) = 2^t$.

In the following we will list some of the properties of $S_\lambda^*(A,G,H)$.

**LEMMA 12.** (i) $S_\lambda^*(1,G,H) = \log 2 + (1-\lambda)\log(1-\lambda) + \lambda \log \lambda$.

(ii) $S_{1/2}^*(1,G,H) = 0$.

(iii) $S_\lambda^*(A,G,\ell(G))$ is independent of $G$ where $\ell(\cdot)$ denotes a distribution function on $[0,1]$.

(iv) $S_\lambda^*(A,G,H) = S_{1-\lambda}^*(1/A,H,G)$.

(v) $S_{1/2}^*(A,G,C) = \log(2 A^{1/2}/(1+A)) < 0$ if $A \neq 1$.

(vi) $S_\lambda^*(A,G,G^C)$ is strictly increasing (decreasing) in $C$ for $A > 1$ ($A < 1$).

(vii) $S_\lambda^*(A,G,G^C) = \log 2 + (1-\lambda)\log A + (1-\lambda)\log(1-\lambda) + \lambda \log \lambda - \log(\lambda + A(1-\lambda)) + \lambda(\lambda-1)(\lambda-C)/A + \frac{\lambda}{1-\lambda} t^{1-B} - 1 dt$.

(viii) If $S_\lambda^*(A,G,G^A) > 0$ for $A > 1$, then there exists a unique $C(A)$ lying between 1 and $A$ such that $S_\lambda^*(A,G,G^C) = 0$.

**Proof.** Properties (i) - (v) trivially follow from the definition of $S_\lambda^*(A,G,H)$. In order to prove (vi), consider
\[
B_\lambda(A,G,G^C) = \frac{1}{A} \int_0^1 \log(\lambda t + A(1-\lambda)t^C)d[\lambda t + A(1-\lambda)t^C]
\]

\[
+ \lambda(1 - 1/A) \int_0^1 \log(\lambda t + A(1-\lambda)t^C)dt
\]

\[
= A^{-1}[\lambda + (1-\lambda)A]\log(\lambda + (1-\lambda)A) = \{\lambda + (1-\lambda)A\}A^{-1}
\]

\[
+ \lambda(1 - 1/A) \int_0^1 \log(\lambda t + A(1-\lambda)t^C)dt
\]

which implies that \(B_\lambda(A,G,G^C)\) is strictly decreasing (increasing) in \(C\) for \(A > 1\) \((A < 1)\). Hence, \(S^*_\lambda(A,G,G^C)\) is strictly increasing in \(C\) when \(A > 1\). Next consider

\[
\int_0^1 \log(\lambda t + A(1-\lambda)t^C)dt = \int_0^1 \log t\ dt + \int_0^1 \log(\lambda + A(1-\lambda)t^{C-1})dt
\]

\[
= -1 + \log(\lambda + A(1-\lambda)) - A(C-1) \int_0^1 \frac{dt}{\frac{\lambda}{1-\lambda} t^{1-C}}.
\]

Hence, substituting this in \(B_\lambda(A,G,G^C)\) and the resultant expression in \(S^*_\lambda(A,G,G^C)\), we obtain (vii). Now (viii) follows from (v), (vi) and (vii).

Analogously one can study some of the properties of \(S^*_\lambda(A,G,H)\).

For the sake of brevity they are not presented here.

6. **Discussion and Concluding Remarks.** Theorems 1 and 2 establish the sure termination of the SPRT for Models I and II under very general conditions. Theorems 3 and 4 establish the sure termination of the "modified" SPRT for Models I and II with \(F(0)\) not specified. Further, in Theorems 1-4, the \(X's\) could be dependent upon the \(Y's\) provided \(\lambda^*_t\) has the large
deviation property given by Part (1) of (15). For instance m could be a hypergeometric variable. Although continuity of F is assumed in the derivation of \( P(\Delta = \delta | A) \), the expressions for \( L_t(A, \delta) \) given by (17) and (19) and the conclusions of Theorems 1-4 are valid even if F denotes an arbitrary distribution function because F can be made continuous by the known continuization process.
REFERENCES


**Abstract**

In this paper, two, modified, one-sample, rank order SPRT's for symmetry based on Lehmann alternatives are given. The first model (Model I) is discussed by Weed and Bradley (1969) and Weed (1968); the second model (Model II) was proposed by Govindarajulu (1968) and is developed here. The two models depend on two different choices of Lehmann alternatives. The results of this paper were developed independently but, since proofs are essentially the same, the authors have combined their work. It is shown, under very general conditions, that the two procedures terminate with probability one and the moments of stopping times are finite for all alternatives within the classes of alternatives defined by Models I and II. Further, procedures based on approximate probability ratios are also studied when the distribution function at zero under the alternative hypothesis is unknown but estimated from the sample observations.