STOPPING TIME OF A RANK-ORDER SEQUENTIAL PROBABILITY RATIO TEST BASED ON LEHMANN ALTERNATIVES*,[1]

By I. Richard Savage and J. Sethuraman

Florida State University, Indian Statistical Institute and Stanford University

1. The problem. We consider sequential probability ratio tests based on ranks for the two sample problem. The hypotheses used in computing the probability ratio are that the two sampled populations are identical and that one of the sampled populations has a distribution function which is a specified power of the other sampled distribution function. It is shown that this procedure terminates with probability 1 and the moments of the stopping time are finite. These results apply to whatever populations are actually sampled.

The problem to be discussed was presented by Robert Berk in a letter of July, 1964 to I. Richard Savage. It has also appeared in the work of E. A. Parent (1965) with which J. Sethuraman was familiar. Again, it appears in an appropriate theoretical context in Hall, Wijsman and Ghosh (1965, p. 594). Miss Sarla D. Merchant considered the problem in her unpublished Florida State University Masters' thesis of 1962.

To be specific we are concerned with the following situation: (X_1, Y_1), (X_2, Y_2), $\cdots$ are independently and identically distributed bivariate random variables with a joint distribution $H(\cdot, \cdot)$ which has continuous marginal distributions $F(\cdot)$ and $G(\cdot)$. We wish to test the null hypothesis $H_0 : X, Y$ *are independent, and* $G = F$ against the alternative hypothesis $H_1 : X, Y$ *are independent, and* $G = F^A$ where $A > 0$, $A \neq 1$ is a known constant. At the nth state of experimentation the available information is the ranks of $(Y_1, \cdots, Y_n)$ among $(X_1, \cdots, X_n, Y_1, \cdots, Y_n)$. We shall use a sequential probability-ratio test based on ranks (see Savage and Savage (1965)). If the distribution of (X, Y) is $H(\cdot, \cdot)$ with marginals $F(\cdot)$ and $G(\cdot)$ and $S(A, H) = S(A, F, G) \neq 0$ (for definition, see (12)) it is shown in Section 3 (Theorem 3) that this sequential test terminates with probability 1 and that the moment generating function of the required sample size is finite.

2. Notations, test procedure, preliminaries and lemmas. When the experiment has proceeded up to the nth stage we have observed (X_1, Y_1), (X_2, Y_2), $\cdots$, (X_n, Y_n). Let the combined sample be denoted by $Z_1, Z_2, \cdots, Z_{2n}$ and the ordered combined sample by $Z_{n1}, Z_{n2}, \cdots, Z_{n2n}$. Let $F_n(\cdot)$ and $G_n(\cdot)$ be the

Received 6 December 1965; revised 1 April 1966.

* Dedicated to the memory of Frank Wilcoxon

[1] This work was supported by the Office of Naval Research Contract No. Nonr 988(13) and by the National Science Foundation Research Grant No. GP-3836.

empirical distribution function of $X_1, X_2, \cdots, X_n$ and $Y_1, Y_2, \cdots, Y_n$, respectively.

The information that will be retained and used for sequential testing is the ranks $s_1 < s_2 < \cdots < s_n$ of the ordered $Y_1, Y_2, \cdots, Y_n$ among the combined sample. We note that the statistic $(s_1, s_2, \cdots, s_n)$ is equivalent to $(G_n(Z_{ni}), i = 1, \cdots, 2n)$ which in turn is equivalent to $(F_n(Z_{ni}), G_n(Z_{ni}), i = 1, \cdots, 2n)$.

The sequential probability-ratio test based on ranks for testing H_0 against H_1 depends on $L_n(A, F_n, G_n) = P_{H_1}(s_1, \cdots, s_n)/P_{H_0}(s_1, \cdots, s_n)$ and can be written in the form

$$\begin{aligned} &\text{Take one more observation if} \quad a \leqq L_n(A, F_n, G_n) \leqq b \\ &\text{accept } H_0 \text{ if} \quad L_n(A, F_n, G_n) < a \\ &\text{reject } H_0 \text{ if} \quad L_n(A, F_n, G_n) > b \\ &n = 1, 2, \cdots \end{aligned} \tag{1}$$

where $0 < a < 1 < b$ are suitable constants (independent of n) and

$$L_n(A, F_n, G_n) = A^n(2n)!/n^{2n} \textstyle\prod_1^{2n} (F_n(Z_{ni}) + AG_n(Z_{ni})). \tag{2}$$

Relation (2) can be easily deduced from Savage (1956, Corollary 7.a.1).

The number of stages before termination, N, is defined by

$$N = \begin{cases} k \text{ if} \quad a \leqq L_n \leqq b \quad \text{for} \quad n = 1, \cdots, k-1 \quad \text{and} \\ \qquad\qquad L_n > b \quad \text{or} \quad L_n < a \quad \text{for} \quad n = k \\ \infty \text{ if} \quad a \leqq L_n \leqq b \quad \text{for every } n. \end{cases} \tag{3}$$

Let

$$nS_n = \log L_n(A, F_n, G_n). \tag{4}$$

From (2) we have

$$S_n = \log 4A - 2 - T_n + O((\log n)/n) \tag{5}$$

where $T_n = n^{-1} \sum_1^{2n} \log (F_n(Z_{ni}) + AG_n(Z_{ni}))$.

From here assume that (X, Y) has a joint distribution $H(\cdot, \cdot)$ with continuous marginal distributions $F(\cdot)$ and $G(\cdot)$. Before formulating the basic Lemma 1, we find the auxiliary results (8), (9) and (10).

Let

$$\begin{aligned} &\Omega_1(n) = \sup_z |F_n(z) - F(z)|, \qquad \Omega_2(n) = \sup_z |G_n(z) - G(z)|, \\ &\Omega(n) = \Omega_1(n) + A\Omega_2(n). \end{aligned} \tag{6}$$

From Theorem 1 of Sethuraman (1964), for each $\delta > 0$ there is a $\rho_1(\delta) < 1$ such that

$$P\{\Omega_i(n) \geqq \delta\} \leqq \rho_1^{\,n}(\delta) \tag{7}$$

for sufficiently large n, $i = 1, 2$. Hence for large n,

$$(8) \quad \begin{aligned} P\{\Omega(n) \geqq \delta\} &\leqq P\{\Omega_1(n) \geqq \delta/(1 + A)\} + P\{A\Omega_2(n) \geqq A\delta/(1 + A)\} \\ &\leqq 2\rho_1^n(\delta/(1 + A)) \leqq \rho_2^n \end{aligned}$$

where $\rho_2 < 1$.

Now, let $W_1, W_2, \cdots, W_n$ be independent and identically distributed random variables with mean $E(W_1)$ and finite moment generating function in a neighborhood of 0. For each $\epsilon > 0$ and sufficiently large n,

$$(9) \qquad P\{|[(W_1 + \cdots + W_n)/n] - E(W_1)| \geqq \epsilon\} \leqq \rho_3^n(\epsilon)$$

where $\rho_3(\epsilon) < 1$ as can be seen from Theorem 1 of Chernoff (1952). In (9) let $W_i = \log(1 + (\delta/U_i))$ where U_i has the uniform distribution on [0, 1]. The moment generating function of W_1 is finite for $\delta > 0$ and $t < 1$ and $E(W_1) \to 0$ as $\delta \to 0$. Thus from (9), given $\epsilon > 0$ we can find δ_0 such that for sufficiently large n

$$(10) \qquad P\{n^{-1} \textstyle\sum_1^n \log(1 + (\delta_0/U_i)) \geqq \epsilon\} < \rho_4^n$$

where $\rho_4 < 1$.

In analogy to the random variables S_n and T_n (see Equation 5) we define the parameters

$$(11) \quad T(A, H) = T(A, F, G) = \int \log(F(x) + AG(x))(dF(x) + dG(x))$$

and

$$(12) \qquad S(A, H) = S(A, F, G) = \log 4A - 2 - T(A, F, G).$$

LEMMA 1. *Given $\epsilon > 0$, there exists $\rho < 1$ such that, for sufficiently large n*

$$(13) \qquad P\{|T_n - T(A, F, G)| \geqq 3\epsilon\} < \rho^n.$$

PROOF.

$$(14) \quad \begin{aligned} T_n &= n^{-1} \textstyle\sum_1^{2n} \log(F_n(Z_{ni}) + AG_n(Z_{ni})) \\ &= n^{-1} \textstyle\sum_1^{2n} \log(F_n(Z_i) + AG_n(Z_i)) \\ &= n^{-1} \textstyle\sum_1^{2n} \log(F(Z_i) + AG(Z_i)) \\ &\quad + n^{-1} \textstyle\sum_1^{2n} \log\{1 + [(F_n(Z_i) - F(Z_i)) + A(G_n(Z_i) - G(Z_i))]/ \\ &\qquad [F(Z_i) + AG(Z_i)]\} \\ &= T_n^{(1)} + T_n^{(2)} \qquad \text{(say)}. \end{aligned}$$

Let $W_i = \log(F(X_i) + AG(X_i)) + \log(F(Y_i) + AG(Y_i))$. Then W_1 has a finite moment generating function and $E(W_1) = T(A, F, G)$. Thus, from (9), for sufficiently large n

$$(15) \qquad P\{|T_n^{(1)} - T(A, F, G)| \geqq \epsilon\} \leqq \rho_5^n.$$

Now,

$$(16)\quad |T_n^{(2)}| \leqq n^{-1}\sum_1^{2n} \log \{1 + [\Omega(n)/(F(Z_i) + AG(Z_i))]\}$$
$$\leqq n^{-1}\sum_1^n \log [1 + (\Omega(n)/F(X_i))] + n^{-1}\sum_1^n \log [1 + (\Omega(n)/AG(Y_i))]$$

where $\Omega(n)$ is defined in (6). Choose δ_0 so that (10) holds. From (10) and (8)

$$P\{n^{-1}\sum_1^n \log [1 + (\Omega(n)/F(X_i))] \geqq \epsilon\}$$
$$\leqq P\{\Omega(n) \geqq \delta_0\}$$
$$(17)\quad + P\{n^{-1}\sum_1^n \log [1 + (\Omega(n)/F(X_i))] \geqq \epsilon, \Omega(n) < \delta_0\}$$
$$\leqq \rho_3^n + P\{n^{-1}\sum_1^n \log [1 + (\delta_0/F(X_i))] \geqq \epsilon\}$$
$$\leqq \rho_3^n + \rho_4^n \leqq \rho_5^n,$$

for sufficiently large n and $\rho_5 < 1$. Arguing in a similar fashion for the second term on the right hand side of (16) we have for sufficiently large n,

$$(18)\quad P\{|T_n^{(2)}| \geqq 2\epsilon\} \leqq \rho_6^n,$$

where $\rho_6 < 1$. Combining (15) and (18) we can find $\rho < 1$ such that (13) is true.

LEMMA 2. *Given $\epsilon > 0$, we can find $\rho < 1$ such that for sufficiently large n*

$$(19)\quad P\{|S_n - S(A, F, G)| \geqq 3\epsilon\} \leqq \rho^n.$$

PROOF. This follows immediately from (5) and Lemma 1.

3. Main theorem and discussion.

THEOREM 3. *Let $S(A, F, G) \neq 0$. Then*

(i) $P(N > n) < \rho^n$ *for sufficiently large n, and some $\rho < 1$.*

(ii) $P(N < \infty) = 1$.

(iii) $E(e^{tN}) < \infty$ *for t in some interval $(-\infty, \theta)$ where $\theta > 0$.*

PROOF. (ii) and (iii) are immediate from (i) which we shall prove. We first note that

$$(20)\quad P(N > n) \leqq P\{a < L_n < b\}.$$

Next, if $S(A, F, G) \neq 0$,

$$(21)\quad P(a < L_n < b) = P\{(\log a)/n < S_n < (\log b)/n\}$$
$$\leqq P\{|S_n - S(A, F, G)| > \epsilon\}$$

for sufficiently large n. Then (20), (21), and Lemma 2 establish (i).

Lemma 4 lists some properties of the parameter $S(A, F, G)$, defined in (12). In this lemma h will stand for a cumulative distribution function on [0, 1].

LEMMA 4. (i) $S(1, F, G) = 0$.

(ii) $S(A, F, h(F))$ *is independent of F.*

(iii) $S(A, F, G) = S(1/A, G, F)$.

(iv) $S(A, F, F) < 0$ *for* $A \neq 1$.
(v)[2] $S(A, F, F^A) > 0$ *for* $A \neq 1$.
(vi) $S(A, F, F^B)$ *is strictly increasing in* B *for* $A > 1$.
(vii) *For each* A, *there is a unique* $B(A)$, *lying between* 1 *and* A, *satisfying the equation*

$$S(A, F, F^B) = 0. \tag{22}$$

Further $B(A) = 1/B(1/A)$.

PROOF. (i), (ii), (iii) and (iv) follow directly from (12), the definition of $S(A, F, G)$. Now,

$$\begin{aligned} S(A, F, F^B) &= \log 4A - 2 - (1/A) \int_0^1 \log (t + At^B)\, d(t + At^B) \\ &\quad -(1 - (1/A)) \int_0^1 \log (t + At^B)\, dt \qquad (23) \\ &= \log 4A - 2 - (1/A)\{(1 + A) \log (1 + A) - (1 + A)\} \\ &\quad - (1 - (1/A)) \int_0^1 \log (t + At^B)\, dt. \end{aligned}$$

For $A > 1$, (23) shows that $S(A, F, F^B)$ is a strictly increasing function in B, establishing (vi). Equation (23) can be further simplified by writing $\log (t + At^B) = \log t + \log (1 + At^{B-1})$ and integrating by parts.

$$\begin{aligned} S(A, F, F^B) &= -2 \log \tfrac{1}{2}(A^{\frac{1}{2}} + A^{-\frac{1}{2}}) \\ &\quad + (A - 1)(B - 1) \int_0^1 (1/A + t^{1-B})\, dt, \end{aligned} \tag{24}$$

and

$$\begin{aligned} S(A, F, F^B) &= \int_0^1 \{[(A - 1)(B - 1)/A + t^{1-B}] \\ &\quad - (A - 1)^2/(4A + (A - 1)^2 t)\}\, dt. \end{aligned} \tag{25}$$

When $A = B < 1$, the integrand in the right hand side of (25) is > 0 for t in $(0, 1)$ and thus $S(A, F, F^A) > 0$ for $A < 1$. From (ii) and (iii), $S(A, F, F^A) = S(A, F^{1/A}, F) = S(1/A, F, F^{1/A})$. Hence $S(A, F, F^A) > 0$ for $A > 1$ also. This establishes (v). Now, (vii) follows from all the above.

4. Discussion. Theorem 3 establishes the sure termination of the sequential procedure (1) under very general conditions including H_0 and H_1. The restriction $S(A, F, G) \neq 0$ does not appear necessary to insure termination. It should be noted that Theorem 3 includes the possibility that the pair of random variables X_i and Y_i are dependent. It is strongly conjectured that sure termination will occur whenever X_i and Y_i are independent.

Although a hypothesis such as H_0 often occurs in nonparametric work, the hypothesis H_1 is special. At the present time we are not able to do other interesting special cases. The general case would be $H_{1'}: F(\cdot) \neq G(\cdot)$. The general case

---

[2] We are grateful to Dr. E. A. Parent for supplying us a proof of this property.

is perhaps no more difficult than the interesting special case of two normal populations with common variances and a known difference in means. The methods of this paper would certainly apply to situations with other sampling patterns, for example, not taking the same number of observations from each population at each stage or taking more than one observation from each population at some stages.

Lemma 2 appears to be the first "large deviation" result for a problem involving ranks. Since the random variable S_n has a more involved structure than those considered by Chernoff and Savage (1958), it would appear possible to prove large deviation results for their problem. Furthermore, it would be very useful to obtain the exact rates of exponential convergence to zero in such situations rather than mere upper bounds as in Lemma 2.

TABLE I

Pairs (A, B) satisfying (22)

A	B	A	B	A	B
0.0567	0.2500	0.44	0.6635	0.72	0.8485
0.0656	0.2667	0.45	0.6710	0.73	0.8544
0.0764	0.2857	0.46	0.6784	0.74	0.8602
0.0898	0.3077	0.47	0.6857	0.75	0.8660
0.1067	0.3333	0.48	0.6930	0.76	0.8718
0.1284	0.3636	0.49	0.7001	0.77	0.8775
0.1570	0.4000	0.50	0.7072	0.78	0.8832
0.1953	0.4444	0.51	0.7142	0.79	0.8888
		0.52	0.7212	0.80	0.8944
0.25	0.5016	0.53	0.7281	0.81	0.9000
0.26	0.5113	0.54	0.7349	0.82	0.9055
0.27	0.5209	0.55	0.7417	0.83	0.9111
0.28	0.5303	0.56	0.7484	0.84	0.9165
0.29	0.5396	0.57	0.7550	0.85	0.9220
0.30	0.5487	0.58	0.7616	0.86	0.9274
0.31	0.5577	0.59	0.7681	0.87	0.9327
0.32	0.5665	0.60	0.7746	0.88	0.9381
0.33	0.5752	0.61	0.7810	0.89	0.9434
0.34	0.5837	0.62	0.7874	0.90	0.9487
0.35	0.5922	0.63	0.7937	0.91	0.9540
0.36	0.6005	0.64	0.8000	0.92	0.9592
0.37	0.6088	0.65	0.8062	0.93	0.9644
0.38	0.6169	0.66	0.8124	0.94	0.9696
0.39	0.6249	0.67	0.8185	0.95	0.9747
0.40	0.6328	0.68	0.8246	0.96	0.9798
0.41	0.6406	0.69	0.8307	0.97	0.9849
0.42	0.6484	0.70	0.8367	0.98	0.9899
0.43	0.6560	0.71	0.8426	0.99	0.9950

NOTE: There is a possible error of .0001 in the value of A among the first eight values and in the value of B in the remaining. If (A, B) satisfies (22), $(1/A, 1/B)$ also satisfies (22).

We now give a table of pairs (A, B) satisfying (22) of (vii) of Lemma 4. It is for this pair that Theorem 3 does not establish sure termination. The table was computed by Mr. P. DeYoung of Florida State University and Mr. Dale Borglum of Stanford University.

REFERENCES

[1] CHERNOFF, H. and SAVAGE, I. R. (1958). Asymptotic normality and efficiency of certain non-parametric test statistics. *Ann. Math. Statist.* **29** 972–994.

[2] CHERNOFF, H. (1952). A measure of the asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Ann. Math. Statist.* **23** 493–507.

[3] HALL, W. J., WIJSMAN, R. A., GHOSH, J. K. (1964). The relationship between sufficiency and invariance with applications in sequential analysis. *Ann. Math. Statist.* **36** 575–614.

[4] PARENT, E. A. (1965). Sequential ranking procedures. Stanford University, Department of Statistics, *Tech. Rept. No. 80.*

[5] SAVAGE, I. R. (1956). Contributions to the theory of rank order statistics. *Ann. Math. Statist.* **27** 590–615.

[6] SAVAGE, I. R. and SAVAGE, L. J. (1965). Finite stopping time and finite expected stopping time. To appear in *J. Roy. Statist. Soc. Ser. B.* **27** 284–289.

[7] SETHURAMAN, J. (1964). On the probability of large deviations of families of sample means. *Ann. Math. Statist.* **35** 1304–1316.