SCHOOL OF OPERATIONS RESEARCH
AND INDUSTRIAL ENGINEERING
COLLEGE OF ENGINEERING
CORNELL UNIVERSITY
ITHACA, NEW YORK

TECHNICAL REPORT NO. 639

November 1984

ON THE RAMEY-ALAM SEQUENTIAL PROCEDURE FOR SELECTING THE MULTINOMIAL EVENT WHICH HAS THE LARGEST PROBABILITY

by

Robert E. Bechhofer
David M. Goldsman*

Research supported by

U.S. Army Research Office-Durham Contract DAAG-29-80-C-0036,

Office of Naval Research Contract N00014-81-K-0037

at Cornell University

Typed by Kathy King

*Presently at the School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA.

ON THE RAMEY-ALAM SEQUENTIAL PROCEDURE FOR SELECTING THE MULTINOMIAL EVENT WHICH HAS THE LARGEST PROBABILITY

Robert E. Bechhofer

School of Operations Research
and Industrial Engineering
Cornell University
Ithaca, New York 14853

David M. Goldsman

School of Industrial
and Systems Engineering
Georgia Institute of Technology
Atlanta, Georgia 30332

ABSTRACT

Ramey and Alam (1979) proposed a closed sequential procedure for selecting that one of $k \geq 2$ multinomial events which has the largest probability. They provided tables of constants necessary to implement their procedure and, based on these constants, they calculated certain performance characteristics of their procedure. In the course of devising a fundamentally different procedure and studying its performance, we had occasion to check their constants and associated performance characteristics. We found quite a few of them to be incorrect. In the present paper we provide extended sets of correct constants and associated performance characteristics (some of which were not provided by Ramey and Alam), and describe the methods that we used to calculate them.

1. INTRODUCTION AND SUMMARY

Ramey and Alam (1979) proposed a closed sequential procedure for selecting that one of $k \geq 2$ multinomial events which has the largest probability. They provided tables of constants ((r,N)-values) necessary to implement their procedure in order that it would guarantee a specified probability requirement, and at the same time possess a certain desirable property. These constants were then used to study the performance of their procedure relative to that of certain competing procedures. They concluded that their procedure was "uniformly better" than these procedures with respect to an important criterion of interest.

In the course of devising a fundamentally different competing procedure, Bechhofer and Goldsman (1985) (B-G), and studying its performance, we had occasion to check the accuracy of their constants and associated performance characteristics. We found quite a few of their calculated results to be incorrect. In the present paper we provide extended sets of correct (r,N)-values and associated performance characteristics (some of which were not provided by Ramey and Alam), and describe the methods that we used to calculate them. These corrected results do not change the basic conclusions of Ramey and Alam. Using their criterion of goodness, their procedure appears to be the best one proposed to that date for the problem under consideration. (Their procedure was constructed in such a way that it must perform at least as well as the sequential procedures to which it was being compared.)

With the corrected set of (r,N)-values and associated performance characteristics we were then able to make fairer comparisons between the performance of their procedure and that of our new procedure described in B-G (1985). In that paper we show that in several important respects our procedure is superior to theirs.

2. HISTORY OF THE PROBLEM

Let $\underset{\sim}{X}_j \equiv (X_{1j}, X_{2j}, \ldots, X_{kj})$ $(j = 1,2,\ldots)$ denote independent vector-observations from a single k-variate multinomial population (Π) having an unknown probability vector $\underset{\sim}{p} = (p_1, p_2, \ldots, p_k)$. Here p_i $(0 \leq p_i \leq 1,\ \sum_{i=1}^{k} p_i = 1)$ denotes the probability of the event E_i $(1 \leq i \leq k)$, the events $E_1, E_2, \ldots, E_k$ being mutually exclusive, and $X_{ij} = 1$ or 0 according as E_i does or does not occur on the jth vector-observation (hereinafter referred to simply as the jth observation) $(1 \leq i \leq k,\ j = 1,2,\ldots)$. Let $p_{[1]} \leq p_{[2]} \leq \cdots \leq p_{[k]}$ denote the ordered p_i-values. It is assumed that the experimenter has no prior knowledge concerning the values of the p_i or of the $p_{[j]}$; it is further assumed that the pairing of the $p_{[j]}$ with the E_i $(1 \leq i,j \leq k)$ is completely unknown. The category associated with $p_{[k]}$ will sometimes be referred to as the "best" category.

We are interested in statistical procedures which are devised to select the best category with appropriate control over the probability of a correct selection $(P\{CS\})$; we say that a <u>correct selection</u> has been made if the event associated with $p_{[k]}$ is the one selected. We limit consideration to procedures which guarantee the following probability requirement:

$$\text{Prob}\{CS\} \geq P^* \quad \text{whenever} \quad p_{[k]} \geq \theta^* p_{[k-1]}. \tag{2.1}$$

Here $\{\theta^*, P^*\}$ $(1 < \theta^* < \infty,\ 1/k < P^* < 1)$ are constants specified by the experimenter before the start of experimentation. We refer to (2.1) as an <u>indifference-zone</u> probability requirement.

<u>2.1 The Bechhofer-Elmaghraby-Morse Procedure</u>

The requirement (2.1) was first stated for this problem by Bechhofer, Elmaghraby and Morse (1959) (hereinafter referred to as B-E-M) who proposed a <u>single-stage</u> procedure to guarantee (2.1). For their procedure n independent observations are taken in a <u>single stage</u> of sampling from Π. Let $\underset{\sim}{y}_n = (y_{1n}, y_{2n}, \ldots, y_{kn})$ where $y_{in} = \sum_{j=1}^{n} x_{ij}$ $(1 \leq i \leq k)$ denotes the outcome of the experiment, and $y_{[1]n} \leq y_{[2]n} \leq \cdots \leq y_{[k]n}$ denote the ordered values of the y_{in}. The B-E-M procedure selects as "best" the category associated with $y_{[k]n}$, or selects one category at random from among all categories which have y_{in}-values tied with $y_{[k]n}$. Kesten and Morse (1959) proved that the $P\{CS\}$ is minimized with respect to $\underset{\sim}{p}$ subject to $p_{[k]} \geq \theta^* p_{[k-1]}$ when the p_i $(1 \leq i \leq k)$ are in the so-called <u>least-favorable</u> (LF-) configuration which is given for this procedure by

$$p_{[k]} = \theta^* p_{[k-1]}, \; p_{[k-1]} = p_{[1]}. \tag{2.2}$$

B-E-M choose n as the smallest integer that will guarantee (2.1) when (2.2) holds. They also provide tables of the $P\{CS\}$ under (2.2) for selected values of k, θ^* and n; a more comprehensive table (Table H.1) is given in Gibbons, Olkin and Sobel (1977).

2.2 The Cacoullos-Sobel Procedure

Cacoullos and Sobel (1966) (hereinafter referred to as C-S) proposed a <u>closed</u> inverse sampling sequential selection procedure to guarantee (2.1). For their procedure, observations are taken <u>one-at-a-time</u> until the count of any one category is equal to N, and at that point the category having this largest count is selected as best (no randomization ever being necessary). C-S proved that (2.2) is the LF-configuration for their procedure (as it was for the B-E-M procedure). They choose N as the smallest integer that will guarantee (2.1) when (2.2) holds; C-S also provide tables of N to implement their procedure. It is to be noted that the maximum possible number of observations that need be taken by the C-S procedure is $k(N-1)+1$.

2.3 The Alam Procedure

Alam (1971) proposed an <u>open</u> sequential selection procedure to guarantee (2.1). The procedure is said to be open since it is not possible, before the experiment starts, to state an upper bound (as can be given for the C-S procedure) on the number of observations required to terminate sampling. For Alam's procedure observations are taken <u>one-at-a-time</u> until $y_{[k]m} - y_{[k-1]m} = r$ where $y_{im} = \sum_{j=1}^{m} x_{ij}$, and $y_{[1]m} \leq \cdots \leq y_{[k]m}$ are the ordered y_{im}-values after m observations have been taken $(1 \leq i \leq k$, $m = 1,2,\ldots)$; r is prespecified as the smallest integer that will guarantee (2.1). (Actually, Alam's proof that (2.1) is guaranteed is given only for $k = 2$.) Values of r which are exact for $k = 2$ and are approximate (being based on very limited Monte Carlo (MC) sampling) for $k = 3,4,5$ are provided by Alam in his Table I; based on his r-values he also gives MC estimates of the expected number of vector-observations $(E\{n\})$ to terminate his procedure when (2.2) holds.

Alam's procedure has the same form as an open sequential selection procedure that had been proposed earlier by Bechhofer, Kiefer and Sobel (1968) (hereinafter referred to as B-K-S), and which they had proved will guarantee (2.1) for all $k \geq 2$ and $\{\theta^*, P^*\}$ when (2.2) holds. The procedure given in Section 5.3.3 of B-K-S can be simplified in a straightforward way using their equation (12.9.4) with $t = 1$ to obtain the Alam procedure. B-K-S set r equal to the smallest integer equal to or greater than $[\log\{(k-1)P^*/(1-P^*)\}]/\log \theta^*$, and this choice of r will guarantee (2.1) for all $k \geq 2$ and $\{\theta^*, P^*\}$ when (2.2) holds. However, this r is very conservative for $k > 2$, i.e., it may result in $P\{CS|LF\} >> P^*$, and a smaller r would reduce the overshoot $(P\{CS|LF\}-P^*)$ as well as the expected number of observations when (2.2) holds.

2.4 The Ramey-Alam Procedure

Ramey and Alam (1979) (hereinafter referred to as R-A) proposed a closed sequential selection procedure to guarantee (2.1). That procedure is a composite of the C-S and Alam procedures in that observations are taken one-at-a-time until either $y_{[k]m} = N$ or $y_{[k]m} - y_{[k-1]m} = r$ $(m = 1,2,...)$, whichever occurs first. The (r,N)-pair under consideration is incorporated in the stopping rule, and all pairs which guarantee (2.1) are eligible for consideration. R-A propose to choose that pair which minimizes $E\{n\}$ when the p_i $(1 \leq i \leq k)$ are in the configuration (2.2). (The (r,N)-pair chosen in this way will not in general have the same r-value (N-value) as the Alam (C-S) procedure when all guarantee (2.1); however, for simplicity of notation we use the same symbols.) For the R-A procedure we shall hereinafter denote $E\{n\}$ when the p_i $(1 \leq i \leq k)$ are in the configuration (2.2) by $E\{n|LF\}$ (although (2.2) has not yet been proved to be LF for their procedure); we shall denote $E\{n\}$ when the p_i $(1 \leq i \leq k)$ are in the configuration $p_{[1]} = p_{[k]}$, i.e., the equal-parameter (EP-) configuration, by $E\{n|EP\}$. Since the $P\{CS\}$ is an increasing function of r and N, the C-S and the Alam procedures can be regarded as special cases of the R-A procedure (the C-S (Alam) procedure being the R-A procedure with $r = \infty$ $(N = \infty)$). Thus it is clear that in terms of minimizing $E\{n|LF\}$, the R-A procedure can perform no worse than either the Alam procedure or the C-S procedure when each uses its appropriate value of (r,N), r or N to guarantee (2.1).

Remark 2.1: A perhaps more interesting criterion for choosing (r,N) for the R-A procedure might be to again limit consideration to all pairs which guarantee (2.1) but then to choose that pair which minimizes $E\{n|EP\}$, obtaining in this way a minimax procedure within the class being considered. This is analogous to the approach adopted by Tamhane and Bechhofer (1977,1979) for the normal means selection problem.

R-A in their Table 1 give for selected values of (k,P^*,θ^*), the values of (r,N) which they claim minimize $E\{n|LF\}$ subject to (2.1). Using these (r,N)-values they also provide in Table 1, associated "exact" or estimated values of $E\{n|LF\}$; they state that their results for $k = 2$ and 3 are based on exact calculations while their estimates for $k = 4$, 5 and 6 are based on Monte Carlo simulations, each employing 1000 replications. Using these results they then conclude, based on the entries in their Table 2, that the R-A procedure yields values of $E\{n|LF\}$ that are no larger (and sometimes smaller) than those of the B-E-M, C-S and Alam procedures when (r,N), n, N and r, respectively, are each chosen to guarantee (2.1). Of course, this had to be the case for the C-S and Alam procedures, but it was of interest to determine by how much $E\{n|LF\}$ was reduced.

In the course of devising a fundamentally different procedure than that of R-A, and which could be regarded as a competing one to theirs, we had occasion to check their (r,N)-values and associated $E\{n|LF\}$-values. (They did not provide $P\{CS|LF\}$- or $E\{n|EP\}$-values.) We found quite a few of their (r,N)-values to be incorrect for $k \geq 3$; also, even when their (r,N)-values were correct, the associated $E\{n|LF\}$-values (given in their Table 1) were often incorrect. In the next section we provide correct sets of (r,N)-values and associated $P\{CS|LF\}$-, $E\{n|LF\}$- and $E\{n|EP\}$-values for $k = 2(1)5,10$ and the same (P^*,θ^*)-combinations as in R-A's Table 1. R-A gave $E\{n|LF\}$-values in Table 1 to two decimal places; in many cases we were able to give $E\{n|LF\}$- and $E\{n|EP\}$-values correct to four decimal places. In Section 4 we describe the methods that we used to calculate these quantities.

3. CORRECT (OR IMPROVED) RAMEY-ALAM (r,N)-VALUES, AND ASSOCIATED P{CS|LF}-, E{n|LF}- AND E{n|EP}-VALUES

We give in Tables IA, IIA, IIIA, IVA and VA for $k = 2,3,4,5$ and 10, respectively, our calculated correct (or improved) values of (r,N) and the associated $P\{CS|LF\}$, $E\{n|LF\}$ and $E\{n|EP\}$ for the Ramey-Alam procedure.

In Tables IB, IIB, IIIB and IVB for $k = 2,3,4$ and 5, respectively, we give the original R-A (r,N)-values which are incorrect; the correct values of $P\{CS|LF\}$ and $E\{n|LF\}$ associated with each incorrect (r,N)-value are given in their respective columns. (R-A did not give $P\{CS|LF\}$.) It will be seen that the (r,N) errors are of two types: i) The original (r,N)-value yields a $P\{CS|LF\} < P^*$, or ii) the original (r,N)-value yields a $P\{CS|LF\} > P^*$ but a different (r,N)-value yields a $P\{CS|LF\} > P^*$ with a smaller $E\{n|LF\}$ than that yielded by the original one. Thus in Table IB for $(k = 2,\ P^* = 0.95,\ \theta^* = 2.4)$ the incorrect $(r,N) = (3,6)$ yields $P\{CS|LF\} = 0.91130 < P^* = 0.95$ while in Table IA the corresponding correct $(r,N) = (4,9)$ yields $P\{CS|LF\} = 0.95477 > P^* = 0.95$. In Table IIB for $(k = 3,\ P^* = 0.75,\ \theta^* = 2.0)$ the incorrect $(r,N) = (3,6)$ yields $P\{CS|LF\} = 0.77287 > 0.75$ with $E\{n|LF\} = 8.8250$ while the correct $(r,N) = (4,5)$ yields $P\{CS|LF\} = 0.75556 > 0.75$ with $E\{n|LF\} = 8.8087 < 8.8250$; the remaining three incorrect (r,N)-values in Table IIB are associated with $P\{CS|LF\}$-values $< P^*$. (Additionally, several of the $E\{n|LF\}$ entries in the R-A Table 1 are incorrect even though their (r,N)-value is correct. For example, for $k = 3$, $P^* = 0.90$, $\theta^* = 2.0$ and 1.6, R-A give $E\{n|LF\} = 15.01$ and 34.77 respectively, while our Table IIA gives $E\{n|LF\} = 16.5142$ and 37.8247, respectively.)

All of the entries associated with $P\{CS|LF\}$, $E\{n|LF\}$ and $E\{n|EP\}$ for $k = 2$ and 3 and certain entries for $k = 4$ and 5 were calculated using an iterative method which yields <u>exact</u> results (see Section 4 and the Appendix), and are <u>correct</u> to the number of decimal places given. It would have been prohibitively costly to calculate exact values when N is large, and hence all of the entries not calculated using the iterative method were computed using Monte Carlo (MC) sampling. Below each entry computed using MC sampling we give, in parentheses, the estimated standard error of the number above it, while to the left of each such entry we give, in parentheses, the number of independent replications (in thousands) on which the corresponding entries are based. Thus, e.g., in Table IIIA for $P^* = 0.90$, $\theta^* = 1.6$ we have $(r,N) = (7,33)$. The associated estimated values of $P\{CS|LF\}$, $E\{n|LF\}$ and $E\{n|EP\}$ are 0.9098, 63.33 and 89.82, respectively, with estimated standard errors of 0.0018, 0.20 and 0.29 based on 24,000, 24,000 and 12,000 replications, respectively. All $P\{CS|LF\}$ estimates obtained by MC sampling and recorded in Tables IIIA, IVA and VA are at least one standard error greater than the specified P^*. In some instances a slightly smaller N-value than the one recorded might also have yielded $P\{CS|LF\} > P^*$; however, in order to save computing costs we did not attempt to pin down N exactly. It is for this reason that we refer to our results as correct (or <u>improved</u>).

Although R-A did not give any results for $k = 10$, we have done so in Table VA. Our purpose in obtaining estimates for $k = 10$ (instead of for $k = 6$ as R-A had done) was to study how $E\{n|LF\}$ and $E\{n|EP\}$ changed with increasing k; the results for the R-A procedure were then compared with the corresponding Bechhofer and Goldsman (1985) results in the latter paper.

TABLE IA

Correct Ramey-Alam (r,N)-values for k = 2
with associated values of $P\{CS|LF\}$, $E\{n|LF\}$ and $E\{n|EP\}$

p^*	θ^*	(r,N)	$P\{CS\|LF\}$	$E\{n\|LF\}$	$E\{n\|EP\}$
0.75	3.0	(1,1)	0.75000	1.0000	1.0000
	2.4	(2,2)	0.79137	2.4152	2.5000
	2.0	(2,3)	0.77366	3.0864	3.2500
	1.6	(3,5)	0.75588	5.9560	6.2578
0.90	3.0	(2,12)	0.90000	3.2000	3.9985
	2.4	(3,6)	0.91130	5.7177	6.9434
	2.0	(4,8)	0.90325	8.8993	10.5862
	1.6	(5,21)	0.90057	17.0012	21.4820
0.95	3.0	(3,6)	0.95222	5.2514	6.9434
	2.4	(4,9)	0.95477	8.4651	11.3796
	2.0	(5,14)	0.95365	13.0906	17.8976

TABLE IB

Original incorrect Ramey-Alam (r,N)-values for k = 2
with associated values of $P\{CS|LF\}$ and $E\{n|LF\}$

p^*	θ^*	(r,N)	$P\{CS\|LF\}$	$E\{n\|LF\}$
0.95	2.4	(3,6)†	0.91130	5.7177

†Incorrect value of (r,N); the correct values of $P\{CS|LF\}$ and $E\{n|LF\}$ associated with this incorrect (r,N) are given in their respective columns.

TABLE IIA

Correct Ramey-Alam (r,N)-values for k = 3 with associated values of $P\{CS|LF\}$, $E\{n|LF\}$ and $E\{n|EP\}$

P*	θ*	(r,N)	$P\{CS\|LF\}$	$E\{n\|LF\}$	$E\{n\|EP\}$
0.75	3.0	(2,3)	0.79615	3.6844	4.2469
	2.4	(2,5)	0.76008	4.7031	5.5431
	2.0	(4,5)	0.75556	8.8087	9.6290
	1.6	(4,12)	0.75722	18.2404	20.9289
0.90	3.0	(3,5)	0.90037	6.7618	8.7405
	2.4	(4,8)	0.91043	11.7850	15.5062
	2.0	(4,15)	0.90149	16.5142	23.6043
	1.6	(6,30)	0.90207	37.8247	53.0739
0.95	3.0	(4,7)	0.95051	9.7676	13.7312
	2.4	(4,14)	0.95090	13.8371	22.8192
	2.0	(5,22)	0.95052	22.4013	37.0271

TABLE IIB

Original incorrect Ramey-Alam (r,N)-values for k = 3 with associated values of $P\{CS|LF\}$ and $E\{n|LF\}$

P*	θ*	(r,N)	$P\{CS\|LF\}$	$E\{n\|LF\}$
0.75	2.4	(2,4)†	0.74679	4.4339
	2.0	(3,6)†	0.77287	8.8250
	1.6	(4,11)†	0.74692	17.4847
0.90	2.4	(3,20)†	0.89896	9.7068

†Incorrect value of (r,N); the correct values of $P\{CS|LF\}$ and $E\{n|LF\}$ associated with this incorrect (r,N) are given in their respective columns.

TABLE IIIA

Correct Ramey-Alam (r,N)-values for k = 4
with associated values of $P\{CS|LF\}$, $E\{n|LF\}$ and $E\{n|EP\}$

P*	θ*	(r,N)	P{CS|LF}	E{n|LF}	E{n|EP}
0.75	3.0	(2,4)	0.77178	5.1483	6.3858
	2.4	(3,5)	0.77309	9.2210	10.9002
	2.0	(3,8)	0.75253	13.3713	16.1677
	1.6	(4,20)	0.75141	30.0511	37.2898
0.90	3.0	(3,7)	0.90912	9.8687	14.7088
	2.4	(4,10)	0.91069	16.9796	24.3486
	2.0	(5,15)	0.90602	27.9279	39.2176
	1.6	(7,33)	(24) 0.9098 (0.0018)	(24) 63.33 (0.20)	(12) 89.82 (0.29)
0.95	3.0	(4,9)	0.95774	13.7471	22.1928
	2.4	(5,13)	0.95310	22.2965	34.6525
	2.0	(6,21)	0.95340	36.2551	57.5616

TABLE IIIB

Original incorrect Ramey-Alam (r,N)-values for k = 4
with associated values of $P\{CS|LF\}$ and $E\{n|LF\}$

P*	θ*	(r,N)	P{CS|LF}	E{n|LF}
0.75	3.0	(2,5)†	0.78931	5.4949
	2.0	(3,9)†	0.76612	14.0139
	1.6	(4,19)†	0.74770	29.6979
0.90	3.0	(3,8)†	0.91932	10.2283
	2.4	(4,9)†	0.89696	16.1806
	2.0	(4,19)†	0.88985	23.6407
	1.6	(6,44)†	(8) 0.8932 (0.0035)	(8) 57.67 (0.40)
0.95	2.0	(5,23)†	0.94033	31.1514

†Incorrect value of (r,N); the correct values of $P\{CS|LF\}$ and $E\{n|LF\}$ associated with this incorrect value are given in their respective columns.

TABLE IVA

Correct Ramey-Alam (r,N)-values for k = 5
with associated values of $P\{CS|LF\}$, $E\{n|LF\}$ and $E\{n|EP\}$

P*	θ*	(r,N)	$P\{CS\|LF\}$	$E\{n\|LF\}$	$E\{n\|EP\}$
0.75	3.0	(2,5)	0.75436	6.6614	8.7608
	2.4	(3,6)	0.76827	12.6540	15.5788
	2.0	(3,11)	0.75041	18.8469	24.2930
	1.6	(5,19)	(24) 0.7533 (0.0028)	(24) 48.60 (0.14)	(12) 58.56 (0.20)
0.90	3.0	(3,8)	0.90462	12.5819	19.9499
	2.4	(4,11)	0.90458	21.8719	32.4554
	2.0	(5,18)	(24) 0.9093 (0.0019)	(24) 37.76 (0.12)	(12) 56.94 (0.18)
	1.6	(7,36)	(36) 0.9016 (0.0016)	(36) 84.33 (0.21)	(12) 122.00 (0.38)
0.95	3.0	(4,10)	0.95730	17.3659	29.8049
	2.4	(5,15)	(24) 0.9577 (0.0013)	(24) 29.05 (0.09)	(12) 48.16 (0.13)
	2.0	(6,23)	(24) 0.9518 (0.0014)	(24) 47.27 (0.15)	(12) 78.04 (0.21)

TABLE IVB

Original incorrect Ramey-Alam (r,N)-values for k = 5
with associated values of $P\{CS|LF\}$ and $E\{n|LF\}$

P*	θ*	(r,N)	$P\{CS\|LF\}$	$E\{n\|LF\}$
0.75	1.6	(4,23)†	(8) 0.7257 (0.0050)	(8) 40.07 (0.28)
0.90	2.0	(4,21)†	(8) 0.8776 (.0037)	(8) 30.61 (0.21)
	1.6	(7,45)†	(4) 0.9113 (0.0045)	(4) 88.30 (0.76)
0.95	2.4	(4,28)†	(8) 0.9446 (0.0026)	(8) 24.42 (0.17)
	2.0	(5,25)†	(8) 0.9409 (0.0026)	(8) 40.64 (0.26)

†Incorrect value of (r,N); the correct values of $P\{CS|LF\}$ and $E\{n|LF\}$ associated with this incorrect value are given in their respective columns.

TABLE VA

Correct Ramey-Alam (r,N)-values for k = 10
with associated values of $P\{CS|LF\}$, $E\{n|LF\}$ and $E\{n|EP\}$

P*	θ*	(r,N)	$P\{CS\|LF\}$	$E\{n\|LF\}$	$E\{n\|EP\}$
0.75	3.0	(3,6)	(4) 0.7803 (0.0065)	(4) 20.04 (0.13)	(4) 27.21 (0.13)
	2.4	(3,10)	(4) 0.7605 (0.0067)	(4) 32.22 (0.27)	(4) 44.66 (0.32)
	2.0	(4,14)	(8) 0.7626 (0.0048)	(8) 58.62 (0.29)	(4) 76.17 (0.42)
	1.6	(5,33)	(20) 0.7572 (0.0030)	(20) 129.56 (0.50)	(4) 174.73 (1.26)
0.90	3.0	(4,9)	(24) 0.9047 (0.0019)	(24) 31.33 (0.08)	(4) 48.23 (0.18)
	2.4	(4,16)	(12) 0.9060 (0.0027)	(12) 50.36 (0.24)	(4) 85.87 (0.52)
	2.0	(5,25)	(24) 0.9030 (0.0019)	(24) 87.61 (0.29)	(4) 145.87 (0.85)
	1.6	(7,58)	(16) 0.9069 (0.0023)	(16) 206.67 (0.88)	(4) 350.84 (2.24)
0.95	3.0	(4,13)	(12) 0.9566 (0.0019)	(12) 36.17 (0.17)	(4) 71.70 (0.36)
	2.4	(5,19)	(12) 0.9529 (0.0019)	(12) 63.26 (0.28)	(4) 116.16 (0.53)
	2.0	(6,31)	(16) 0.9538 (0.0017)	(16) 107.39 (0.43)	(4) 196.93 (1.00)

4. METHOD OF CALCULATION OF $P\{CS\}$- and $E\{n\}$-VALUES

As mentioned earlier, most of the $P\{CS|LF\}$-, $E\{n|LF\}$- and $E\{n|EP\}$-calculations for each (k,P^*,θ^*) and (r,N) pair were carried out using an iterative method, described in Appendix A, which yields exact results. For each (k,P^*,θ^*) combination an efficient search procedure was carried out on the (r,N)-grid in order to determine the eligible set of (r,N)-values which yielded $P\{CS|LF\} \geq P^*$; within the eligible set, that (r,N) pair was found which minimized $E\{n|LF\}$. This tedious and costly search was aided by the fact that $P\{CS|LF\}$ and $E\{n|LF\}$ are increasing functions of both r and N; thus for each fixed r, the smallest N which guaranteed $P\{CS|LF\} \geq P^*$ yielded an (r,N) in the eligible set. Because of cost considerations, all searches were restricted to $N \leq 50$ (except for $k = 10$ where we searched to $N \leq 60$).

We checked the exact results obtained using our iterative method against exact results obtained using complete enumeration of sample paths and found them to agree; however, the latter could be calculated only for moderate values of (k,r,N). Next we checked the results obtained using our simulation program against the exact results obtained using our iterative method, and found them to be in very close agreement--well within sampling error based on large numbers of replications. Since MC simulation was much less costly to carry out than exact calculations, especially for large N, the former was used in such cases. Because of the cross-checking that we have conducted, we have confidence in the results cited.

The MC simulations were carried out on CDC 6500 and CDC 6600 computers at Purdue University and on an IBM 3081 at Cornell University. The random variates were generated using the IMSL (1982) subroutines GGUBS or GGMTN. Our exact results were calculated using the machines on Purdue's VAX network.

5. CONCLUDING REMARKS

As noted in Section 3 and supported by the entries in the tables, quite a few of the R-A (r,N)-values are incorrect, several of them considerably so. We suspect that the incorrect entries are a consequence of errors in the methods of calculation and/or in the MC simulation program that was used (or perhaps the simulations were not based on a sufficient number of replications). However, the correct results support the basic Ramey-Alam conclusion--namely that their procedure yields $E\{n|LF\}$-values which are no larger than (and sometimes quite a bit smaller than) those obtained with the corresponding B-E-M, C-S or Alam procedures. Whether or not minimizing $E\{n|LF\}$ is the most meaningful criterion when comparing procedures is an open question. In B-G (1985) we focus not only on minimizing $E\{n|LF\}$ but also on minimizing $E\{n|EP\}$.

APPENDIX

ITERATIVE METHOD OF CALCULATING EXACT VALUES OF $P\{CS|LF\}$, $E\{n|LF\}$ AND $E\{n|EP\}$ FOR THE RAMEY-ALAM PROCEDURE

For given (k,P^*,θ^*,r,N), we can calculate the $P\{CS|LF\}$, $E\{n|LF\}$ and $E\{n|EP\}$ resulting from the implementation of the Ramey-Alam procedure; the conjectured LF-configuration is given by (2.2). We then search various values of r and N to find that (r,N) pair which minimizes $E\{n|LF\}$ while satisfying $P\{CS|LF\} \geq P^*$.

Consider the counts $y_{1t},y_{2t},\ldots,y_{kt}$ at stage t of sampling, $t = 1,2,\ldots$. The R-A procedure terminates sampling when either $y_{[k]t} - y_{[k-1]t} = r$ or $y_{[k]t} = N$. Define $\#(\ell_1,\ell_2,\ldots,\ell_k)$ to be the number of distinct paths of the sampling process $\{\underset{\sim}{y}_t, t = 1,2,\ldots\}$ which lead to procedure termination exactly when $(y_{1t},y_{2t},\ldots,y_{kt}) = (\ell_1,\ell_2,\ldots,\ell_k)$.

<u>Example</u>: Suppose $k = 2$, $r = 2$ and $N = 3$. Then $\#(2,0) = 1$ since only one path of the sampling process leads to termination exactly when $(y_{1t},y_{2t}) = (2,0)$, viz., $\underset{\sim}{y}_1 = (1,0)$, $\underset{\sim}{y}_2 = (2,0)$.

Obviously,

$$\#(\ell_1,\ldots,\ell_k) = \left[\begin{array}{l}\text{total number of}\\ \text{paths to}\\ (\ell_1,\ldots,\ell_k)\end{array}\right] - \left[\begin{array}{l}\text{number of paths to}\\ (\ell_1,\ldots,\ell_k)\ \text{for which}\\ \text{the R-A procedure}\\ \text{terminates } \underline{\text{en route}}\end{array}\right].$$

Example: Suppose $k = 2$, $r = 2$ and $N = 3$. We wish to calculate $\#(3,1)$. Noting that the R-A procedure terminates [en route to $\underset{\sim}{y}_4 = (3,1)$] if $\underset{\sim}{y}_2 = (2,0)$, we have:

$$\#(3,1) = \binom{4}{1} - \begin{bmatrix}\text{Number of paths} \\ \text{from } (2,0) \text{ to } (3,1)\end{bmatrix} \#(2,0) = 4 - \binom{2}{1}\cdot 1 = 2.$$

It is clear that

$$\#(\ell_1,\ell_2,\ldots,\ell_k) = \binom{\sum_{i=1}^k \ell_i}{\ell_1,\ell_2,\ldots,\ell_k}$$

$$- \sum_{j_1=0}^{\ell_1} \sum_{j_2=0}^{\ell_2} \cdots \sum_{j_k=0}^{\ell_k} \binom{\sum_{i=1}^k (\ell_i - j_i)}{\ell_1 - j_1, \ell_2 - j_2, \ldots, \ell_k - j_k} \#(j_1,\ldots,j_k) \qquad (1)$$

where $\binom{\sum_{i=1}^k n_i}{n_1,n_2,\ldots,n_k} = \dfrac{(\sum_{i=1}^k n_i)!}{n_1!n_2!\ldots n_k!}$ is the multinomial coefficient.

Remark: By symmetry, $\#(\ell_1,\ldots,\ell_k) = \#$(any permutation of $\ell_1,\ldots,\ell_k$). Hence, we need only explicitly calculate the following $\#(\ell_1,\ldots,\ell_k)$'s:

$$\#(j_2+r, j_2, j_3, \ldots, j_k), \quad 0 \le j_2 \le N-r-1,\ j_2 \ge j_3 \ge \cdots \ge j_k \qquad (2)$$

and

$$\#(N, j_2, j_3, \ldots, j_k), \quad N-r \le j_2 \le N-1,\ j_2 \ge j_3 \ge \cdots \ge j_k \qquad (3)$$

Any $\#(\ell_1,\ldots,\ell_k)$ which is not a permutation of (2) or (3) must equal zero; this is so because it is impossible for the R-A procedure to terminate at such $(\ell_1,\ldots,\ell_k)$ values.

Remark: The $\#(\ell_1,\ldots,\ell_k)$'s are to be computed in an iterative manner:

A. Initialize all $\#(\cdot)$'s to equal zero.

B. Using (1), left-lexicographically calculate only those $\#(\cdot)$'s of the forms (2) or (3). By the above Remarks, we obtain (with no further calculations) all $\#(\cdot)$'s which are permutations of the forms (2) or (3); all other $\#(\cdot)$'s equal zero:

$$\left.\begin{array}{l} \#(j+r,j,0,\dots,0), \\ \#(j+r,j,1,0,\dots,0), \\ \#(j+r,j,1,1,0,\dots,0), \\ \quad\vdots \\ \#(j+r,j,1,1,\dots,1) \\ \#(j+r,j,2,0,\dots,0) \\ \quad\vdots \\ \#(j+r+1,j+1,0,\dots,0) \\ \quad\vdots \\ \#(N-1,N-r-1,N-r-1,\dots,N-r-1) \end{array}\right\} \#(\cdot)\text{'s of the form (2)}$$

$$\left.\begin{array}{l} \#(N,N-r,0,\dots,0) \\ \#(N,N-r,1,0,\dots,0) \\ \quad\vdots \\ \#(N,N-1,N-1,\dots,N-1) \end{array}\right\} \#(\cdot)\text{'s of the form (3)}$$

The left-lexicographic order of calculation is necessary since the computation of $\#(\ell_1,\dots,\ell_k)$ involves all of the previous $\#(\cdot)$'s. "On the fly" storage of the values of these previous $\#(\cdot)$'s avoids recursive re-computation in (1).

Suppose that $\underset{\sim}{p}$ is in the LF-configuration. Without loss of generality, assume that $p_1 = \theta^* p_i$, $i = 2,\dots,k$; i.e., $p_1 = p_{[k]}$. Then,

$$P\{CS|LF\} = \sum_{j_2=0}^{N-r-1} \cdots \sum_{j_k=0}^{N-r-1} \#(M+r,j_2,j_3,\dots,j_k)p_1^{M+r}p_2^{j_2}p_3^{j_3}\cdots p_k^{j_k}$$
$$+ \sum_{j_2=0}^{N-1} \cdots \sum_{j_k=0}^{N-1} \#(N,j_2,j_3,\dots,j_k)p_1^{N}p_2^{j_2}p_3^{j_3}\cdots p_k^{j_k},$$

where $M \equiv M(\underset{\sim}{j}) \equiv \max(j_2,j_3,\dots,j_k)$.

Letting $p \equiv p_2 = \dots = p_k$, $\theta \equiv \theta^*$, and $S \equiv S(\underset{\sim}{j}) \equiv \sum_{i=2}^{k} j_i$, we have:

$$P\{CS|LF\} = \sum_{j_2=0}^{N-r-1} \cdots \sum_{j_k=0}^{N-r-1} \#(M+r,j_2,j_3,\dots,j_k)\theta^{M+r}p^{M+r+S}$$
$$+ \sum_{j_2=0}^{N-1} \cdots \sum_{j_k=0}^{N-1} \#(N,j_2,j_3,\dots,j_k)\theta^{N}p^{N+S}. \qquad (4)$$

We remind the reader that many of the $\#(\cdot)$'s in (4) equal zero.

Now,

$$E\{n|LF\}$$
$$= \sum_{i=1}^{k} \begin{bmatrix} \text{contribution to } E\{n|LF\} \\ \text{given that cell } i \text{ is chosen} \end{bmatrix}$$
$$= \begin{bmatrix} \text{contribution to } E\{n|LF\} \\ \text{given that cell 1 is chosen} \end{bmatrix} + (k-1)\begin{bmatrix} \text{contribution to } E\{n|LF\} \\ \text{given that cell 2 is chosen} \end{bmatrix}$$
$$= \sum_{j_2=0}^{N-r-1} \cdots \sum_{j_k=0}^{N-r-1} (M+r+S)\#(M+r,j_2,\dots,j_k)\theta^{M+r}p^{M+r+S}$$
$$+ \sum_{j_2=0}^{N-1} \cdots \sum_{j_k=0}^{N-1} (N+S)\#(N,j_2,\dots,j_k)\theta^{N}p^{N+S}.$$
$$+ (k-1) \sum_{j_2=0}^{N-r-1} \cdots \sum_{j_k=0}^{N-r-1} (M+r+S)\#(j_2,M+r,j_3,\dots,j_k)\theta^{j_2}p^{M+r+S}$$
$$+ (k-1) \sum_{j_2=0}^{N-1} \cdots \sum_{j_k=0}^{N-1} (N+S)\#(j_2,N,j_3,\dots,j_k)\theta^{j_2}p^{N+S}.$$
$$= \sum_{j_2=0}^{N-r-1} \cdots \sum_{j_k=0}^{N-r-1} (M+r+S)\#(M+r,j_2,\dots,j_k)p^{M+r+S}[\theta^{M+r}+(k-1)\theta^{j_2}]$$
$$+ \sum_{j_2=0}^{N-1} \cdots \sum_{j_k=0}^{N-1} (N+S)\#(N,j_2,\dots,j_k)p^{N+S}[\theta^{N}+(k-1)\theta^{j_2}]. \qquad (5)$$

$E\{n|EP\}$ is obtained by setting $\theta = 1$ and $p = 1/k$ in (5).

We note that the above expressions can be written somewhat more efficiently for computational purposes; however, they become more complex in appearance.

ACKNOWLEDGMENTS

We wish to thank Paul Sanchez for helpful discussions which led to an improvement in the accuracy of our computer programs. We are indebted to Dr. Radhika Kulkarni and Professor Ajit Tamhane for constructive criticisms of the manuscript. Some of our computations were carried out at Purdue University, and we wish to acknowledge their generosity in allowing us to use their computing facilities. This research was supported by U.S. Army Research Office-Durham Contract DAAG-29-81-K-0168 and Office of Naval Research Contract N00014-81-K-0037 at Cornell University.

BIBLIOGRAPHY

Alam, K. (1971). On selecting the most probable category. Technometrics 13, 843-850.

Bechhofer, R.E., Elmaghraby, S. and Morse, N. (1959). A single-sample multiple-decision procedure for selecting the multinomial event which has the largest probability. Ann. Math. Statist. 30, 102-119.

Bechhofer, R.E. and Goldsman, D.M. (1985). Truncation of the Bechhofer-Kiefer-Sobel sequential procedure for selecting the multinomial event which has the largest probability. Comm. Statist.--Simula. Computa., B14, this issue, following article.

Bechhofer, R.E., Kiefer, J. and Sobel, M. (1968). Sequential Identification and Ranking Procedures (with special reference to Koopman-Darmois populations), Chicago: The University of Chicago Press.

Cacoullos, T. and Sobel, M. (1966). An inverse-sampling procedure for selecting the most probable event in a multinomial distribution. Multivariate Analysis (ed. by P.R. Krishnaiah), New York: Academic Press, 423-455.

Gibbons, J.D., Olkin, I. and Sobel, M. (1977). Selecting and Ordering Populations, New York: John Wiley.

INTERNATIONAL MATHEMATICAL AND STATISTICAL LIBRARIES (1982), IMSL Manual (vol. II, 7th ed.), Houston, Texas: IMSL.

Kesten, H. and Morse, N. (1959). A property of the multinomial distribution. Ann. Math. Statist., 30, 120-127.

Ramey, J.T. and Alam, K. (1979). A sequential procedure for selecting the most probable multinomial event. Biometrika, 66, 171-173.

Tamhane, A.C. and Bechhofer, R.E. (1977). A two-stage minimax procedure with screening for selecting the largest normal mean. Commun. Statist.--Theor. Meth. A6, 1003-1033.

Tamhane, A.C. and Bechhofer, R.E. (1979). A two-stage minimax procedure with screening for selecting the largest normal mean (II): an improved PCS lower bound and associated tables. Commun. Statist.--Theor. Meth. A8, 337-358.