Adaptive Designs
IMS Lecture Notes - Monograph Series (1995) Volume 25

NEW PROCEDURES FOR GROUP-TESTING BASED ON THE HUFFMAN LOWER BOUND AND SHANNON ENTROPY CRITERIA

LIFANG HSU

Le Moyne College

Abstract

Our goal is to devise an efficient, possibly optimal method for identifying all defective units or determining that there aren't any among N given units. To do this, we use adaptive group-testing methodology assuming a binomial random sample of N independent and identically distributed units with known probability q of each unit being good. The optimality desired is to minimize the expected number of tests required. But this optimality may be infeasible. Two procedures (R_{HLB} and R_{1A}) for the group-testing problem are studied. Procedure R_{HLB} is based on the Huffman lower bound and Shannon-entropy criteria. All of the algorithms introduced have low design complexity and yet provide near-optimal results. Both procedures are adaptive in the sense that the present test can depend on the results of any or all previous tests. For $N = 5$, one of the procedures introduced can be shown to be optimal for selected values of q. It is conjectured that this procedure is optimal for all values of q. It is conjectured that this procedure is optimal for all values of q and $N > 5$.

Received March 1993; revised July 1994.

AMS 1991 subject classifications. Primary 62N10.

Key words and phrases. Adaptive group testing, binomial set, contaminate set, expected group-test number.

1. Introduction. In the problem of group testing, we are concerned with the classification of each one of a number N of given units into one of two distinct categories which we call satisfactory and unsatisfactory (or simply, good and defective). The characteristic feature of a group-test is that a simultaneous test of x units $(1 \leq x \leq N)$ is performed with only two possible results: (i) either all x are good or (ii) at least one of the x units is defective. In the second case, it is not known (unless $x = 1$) which one or how many of the x units are defective, and we call this set of size x a *contaminated set.* The model considered is that the N units are the realization of N independent and identically distributed Bernoulli random variables with common, known probability q of being good and $p = 1- q$ of being defective; any such set for which we have no further knowledge about the units will be called a *binomial set.* The ideal goal is to devise an adaptive strategy for identifying all the defective units or determining that there aren't any among the given N units with the minimum expected number of groups tested. An unqualified reference to optimality implies optimality in this sense.

The first published paper on the group-testing problem (in the present-day context) was by Dorfman (1943). Under Dorfman's procedure, for a prevalence rate of defectiveness equal to.01, there is a savings of close to 80% over testing one at a time. Sobel and Groll (1959) developed a nested procedure, called R_1. Under this procedure, for $p = .01$, there is a savings of 91.68% over testing one at a time. Sobel (1967) proposed a procedure R_∞ which allows mixing (i.e., testing a mixture of units from a binomial and a contaminated set) only when the size of contaminated set is two or three. However, R_∞ is not an optimal procedure. Friedman (1982) wrote a dynamic programming optimality algorithm for the group-testing problem. However, although Friedman's recursive algorithm is not complicated, the amount of computation required for its solution is enormous, even for small values of N. With advanced computers it was only possible to obtain results for $N \leq 5$ and for selected values of q. Thus, the optimal procedure remains unknown in general.

It is conjectured that the construction of an optimal group-testing procedure is a *NP-complete* problem and a heuristic approach to this problem is therefore desirable. *NP-completeness* is defined as the class of all problems that are in *NP* and in *NP-hard. NP* is defined as the class of all problems that can be solved by nondeterministic algorithms that run in polynomial-time. *NP-hard* is the class of problems for which a deterministic polynomial-time algorithm for its solution can be used to obtain a deterministic polynomial-time algorithm for every problem in *NP*. For examples of *NP-complete* problems, see Chapter 9 of Reingold, Nievergelt, and Deo (1977). Chen, Hsu and Sobel (1987) proposed a procedure R_1 based on the Shannon-entropy criteria that was based on such a heuristic approach. The Shannon-entropy criteria was described in Section 2. They showed that R_1 is optimal at every stage with respect to the Shannon-entropy criteria, but it is not necessarily optimal with respect to the expected total number of group tests required.

In this paper we apply some coding theory concepts to obtain an improved heuristic procedure. Starting with n binomial units, we can regard them as ordered. Since each unit is good or defective, there are 2^n possible states of nature, one of which is true. If we represent each test that succeeds by the digit 'zero' and each test that fails by the digit 'one', then a particular set of tests outcomes is identical with a particular 'word' of a binary code. Then the expected number of tests required is identical with the expected word length (i.e., the cost) of the code. Huffman (1952) gives a routine (i.e. an encoding scheme) for finding the code with the smallest cost which we call the *Huffman cost (HC)*. We describe a *Huffman lower bound* procedure, R_{HLB} in Section 2 that is based on choosing a sample size x for the next group test that minimizes the *average Huffman cost* over the possible actions for the next group test. It can be seen that R_{HLB} is not optimal with respect to the expected total number of group tests required. In Section 3, we introduce a procedure, R_{1A}, that is based on a combination of both the Huffman lower bound and the Shannon-entropy criteria. Explicit instructions for carrying out R_{HLB} and R_{1A} are given in Tables 1 and 3, respectively, for all q when $N = 1$ through 5.

2. The Huffman lower bound testing procedure. The Huffman lower bound testing procedure, R_{HLB}, is based on choosing a sample size x for the next group test that minimizes the average Huffman cost over the possible actions for the next group test. The following three lemmas play a fundamental role in the derivation of the R_{HLB} procedure which is described below. Let G be the current situation. By the phrase all states of nature, we mean the collection of possible sets of binomial and contaminated sets consistent with G. An *action* is the action of selecting a binomial set, a contaminated set, or a subset of these for testing.

LEMMA 1 [Sobel (1967)]. *Given a contaminated set C of size $m > 1$, and given also that a proper subset of C of size x with $1 \leq x \leq m$ contains at least one defective, then the conditional distribution associated with the $m - x$ remaining units is precisely that of $m - x$ independent Bernoulli random variables with common original probability $q$of being good.*

LEMMA 2 [Chen, Hsu and Sobel (1987)]. *The list of the sets currently (i.e., at stage G) known to be contaminated determines the distribution P_G of the states of nature conditional on being in stage G, and this list is a sufficient statistic for all information gathered up to stage G.*

LEMMA 3 [Chen, Hsu and Sobel (1987)]. *The list of the sets known to be contaminated at stage G determines the number t of all possible actions at stage G.*

Using the above three lemmas, we define the R_{HLB} procedure.

2.1 Derivation of the Huffman lower bound testing procedure. By Lemma 2, there exists a conditional distribution $P_G = (p_1, ..., p_k)$ for each stage G, where k is the number of states of nature consistent with G and p_i, $i = 1, ..., k$, is the probability of the ith state; of course

$$\sum_{i=1}^{k} p_i = 1.$$

Furthermore, Lemma 2 tells us that P_G contains all the information about the current stage G. By Lemma 3, there exists an action space $A_G = (A_1, ..., A_t)$ where $A_i, i = 1, ..., t$, are all the possible actions that may be taken in the current situation G. An action A_i is a success if all items tested under A_i are not defective.

Let $S_G(A_i)$ and $F_G(A_i)$ be the probabilities that A_i is a success or a failure, respectively, and let P_{A_i,s_G} and P_{A_i,F_G}, respectively, be the probability vectors associated with success and failure at stage G when we take action A_i given P_G. HC$\{P_{A_i,S_G}\}$ and HC$\{P_{A_i,F_G}\}$ are defined to be the Huffman cost after applying the Huffman (1952) encoding scheme to P_{A_i,S_G} and P_{A_i,F_G}, respectively. AHC$(P_G|A_i)$ denotes the average Huffman cost for taking action A_i when the initial probability vector is P_G:

(2.1)

$$AHC(P_G|A_i) = S_G(A_i)HC\{P_{A_i,S_G}\} + F_G(A_i)HC\{P_{A_i,F_G}\}.$$

The procedure R_{HLB} is defined by choosing $i*$ such that
(2.2)

$$\mathrm{AHC}(P_G|A_{i_*}) = \min_{1\leq i\leq t}\{AHC(P_G|A_i)\},$$

and then taking action A_{i*} at stage G. Note that, if T is the number of group-tests required for identifying all defective items in given G, then the expected value of T under R_{HLB} is
(2.3)

$$E(T|P_G, R_{HLB})$$

$$= 1 + S_G(A_{i*})E(T|P_{A_i S_G}, R_{HLB}) + F_G(A_{i*})E(T|P_{A_\imath, F_G}, R_{HLB}).$$

REMARK 1. Note that (2.3) does not require recursive calculation back to the first action in order to find A_{i*}. At each stage, optimization is based only on the current possible actions. Hence, R_{HLB} is substantially easier to implement than the procedure that is optimal with respect to the expected number of group-tests required.

REMARK 2. Although the R_{HLB} procedure is optimal with respect to the Huffman Lower Bound criteria given above, it is not necessary optimal with respect to the expected number of group-tests required.

REMARK 3. Clearly, if $E(T|P_G, R_{HLB})$ in (2.3) were equal to $\mathrm{HC}(P_G)$ [i.e., if $E(T|_G, R_{HLB})$ attains the Huffman cost by applying the Huffman (1952) encoding scheme to P_G], then R_{HLB} would be the optimal group-testing procedure. However, $E(T|P_G, R_{HLB}) \neq HC(P_G)$ does not mean that R_{HLB} is not optimal because the Huffman cost is not always attainable for a group-testing problem.

REMARK 4. Explicit instructions for carrying out the R_{HLB} procedure are given for $N = 1$ through 5 for all q in Table 1.

REMARK 5. The numerical results indicate that R_{HLB} is the optimal procedure for q close to 1, but this remains to be proved.

2.2 Illustration of the R_{HLB} procedure. Suppose we have $N = 5$ units and know that the probability a unit is good is $q = 0.98$. As indicated in Table 1, the first test-group is to be of size $x = 5$, i.e., testing all 5 units simultaneously will give the smallest average Huffman cost at the initial stage. If a success occurs, the experiment is over. If

Table 1.
Test size and polynomial coefficients* for the expected number of tests required to classify a binomial set of size N under R_{HLB} procedure

n	Test Size x	Range of q	1	q	q^2	q^3	q^4	q^5
2	1	0.0000 – 0.6180	2					
	2	0.6180 – 1.0000	3	−1	−1			
3	1	0.0000 – 0.5970	3					
	2	0.5970 – 0.6180	4	−1	−1			
	2	0.6180 – 0.7071	5	−3	−1	1		
	3	0.7071 – 0.8385	5	−2	−1	−1		
	3	0.8385 – 1.0000	7	−3	−6	3		
4	1	0.0000 – 0.5970	4					
	1	0.5970 – 0.6180	5	−1	−1			
	1	0.6180 – 0.6358	6	−3	−1	1		
	2	0.6358 – 0.7071	7	−5	0	1	−1	
	2	0.7071 – 0.7777	7	−4	−1	−1	1	
	4	0.7777 – 0.8385	7	−3	−1	−1	−1	
	4	0.8385 – 0.8532	9	−4	−6	3	−1	
	4	0.8532 – 1.0000	10	−5	−7	3		
5	1	0.0000 – 0.5893	5					
	2	0.5893 – 0.5970	6	−1	−1			
	2	0.5970 – 0.6180	7	−2	−2			
	2	0.6180 – 0.6358	8	−4	−2	1		
	2	0.6358 – 0.7071	9	−7	1	0	−1	1
	2	0.7071 – 0.7778	9	−6	0	−1	1	−1
	2	0.7778 – 0.8034	9	−5	−1	−1	−1	1
	5	0.8034 – 0.8385	10	−5	−2	−1	−1	
	5	0.8385 – 0.8827	12	−6	−7	3	−1	
	5	0.8827 – 1.0000	14	−8	−8	4	−3	2

* The integer shown is the coefficient of the power of q at the top of the column and the terms are then added to form the expected number of tests required for classifying a binomial set of size n. The entry x indicates that the next test is on x units taken from the only set available, i.e., the binomial set.

a failure occurs, then it follows from equation (2.2) that the next test-group will be ofsize 2, chosen at random from the 5. Label the sequence of groups tested at the initial stage by a, b, c, d, e and the sequence of groups tested at the second stage by a, b. If (a, b) yields a collective failure, then (a, b) contains a contaminated set and hence is a binomial set. Continuing in a similar manner yields the diagram in Figure 1. To derive the diagram, start with a binomial set of size 5 : $\{a, b, c, d, e\}$: the tree branches at each experiment with one path for success and another for failure; the sets that follow are distinguished with braces indicating a binomial set and the parenthesis indicating a contaminated set; the star indicates that the continuation of the procedure follows as given above or to the left for the same state of nature.

From Table 1, we obtain the expected number of group-tests required to identify defective items when $q = 0.98$ and $N = 5$:

(2.4)

$$E(T|R_{HLB}) = 14 - 8q - 8q^2 + 4q^3 - 3q^4 + 2q^5 = 1.2823.$$

This expectation for R_{HLB} is the same as that found by Sobel (1967) and later by Friedman (1982) for the optimal procedure. Hence, for $N = 5$ and $q = .98$ the R_{HLB} is optimal.

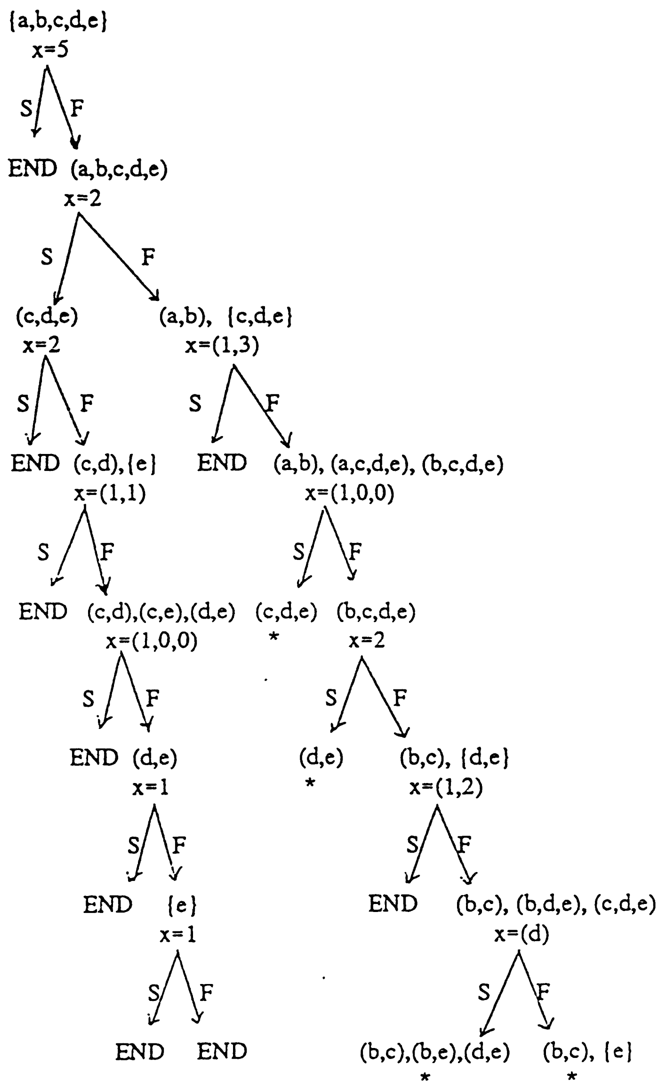
3. A combined Huffman lower bound and Shannon entropy based procedure. Chen, Hsu and Sobel (1987) proposed a procedure R_1 based on choosing that sample size x for the next group test which maximizes the Shannon-entropy reduction was defined there in terms of $P_G = (p_1, ..., p_k)$ and $A_G = (A_1, ..., A_1)$ by

(3.1)

$$I(P_G|A_i) = -S_G(A_i)\log_2(S_G(A_i)) - F_G(A_i)\log_2(F_G(A_i)),$$

where $S_G(A_i)$ and $F_G(A_i)$ are the probabilities that A_i leads to success and failure, respectively. The expected number of tests, $E(T|R_1)$, are given for $N = 1$ through 5 for all q in Table 2. The resulting value of $E(T|R_{HLB})$ and $E(T|R_1)$ is a function of q, and in particular, they are piecewise polynomials in q. At some point (for example, $q = .8827$ for $N = 5$), $E(T|R_{HLB})$ "jumps" from one polynomial to another. We refer to the polynomial expressions on each side of such a jump as *adjacent polynomials.* Not only may $E(T|R_{HLB})$ and $E(T|R_1)$ be discontinuous in q they need not be decreasing in q. To eliminate these negative

Figure 1.



{ } denotes binomial sets.
() denotes contaminated sets.
x=(1,y,z) indicates that a test of size 1 from the first set, size y from the second set and size z from the third set is to be performed.
* Continue as designated to the left for this set.

Table 2.
Test size and polynomial coefficients* for the expected number of tests required to classify a binomial set of size N under R_1 procedure

n	Test Size x	Range of q	1	q	q^2	q^3	q^4	q^5
2	1	0.0000 – 0.6180	2					
	2	0.6180 – 1.0000	3	−1	−1			
3	1	0.0000 – 0.6180	3					
	2	0.6180 – 0.7549	5	−3	−1	1		
	3	0.7549 – 1.0000	5	−2	−1	−1		
4	1	0.0000 – 0.6180	4					
	2	0.6180 – 0.7549	7	−5	0	1	−1	
	3	0.7549 – 0.8192	7	−3	−2	−2	2	
	4	0.8192 – 1.0000	8	−4	−2	−1		
5	1	0.0000 – 0.6180	5					
	2	0.6180 – 0.7549	9	−7	1	0	−1	1
	3	0.7549 – 0.7759	10	−6	−2	−4	7	−3
	3	0.7759 – 0.8192	10	−6	−3	−1	4	−2
	3	0.8192 – 0.8484	10	−5	−3	−2	2	
	5	0.8484 – 0.8567	11	−7	−2	−1	−1	2
	5	0.8567 – 1.0000	11	−7	−2	0	0	−1

* The integer shown is the coefficient of the power of q at the top of the column and the terms are then added to form the expected number of tests required for classifying a binomial set of size n. The entry x indicates that the next test is on x units taken from the only set available, i.e., the binomial set.

features, we introduce here a modified procedure R_{1A}. The cases in which q is and is not close to one are considered separately.

- CASE 1: q is not close to 1. The R_{1A} procedure is carried out in a manner similar to R_1 with but one modification. Suppose $I_1 = [a, b]$ and $I_2 = [b, c]$ are the intervals of q for two adjacent polynomials, and suppose these polynomials are not continuous at their common boundary point b. Then we extend their range so that they overlap and look for a new dividing point q_0 where two adjacent polynomials are equal. If there is one such q_0 in $[a, c]$, we use the I_1's polynomial on $[a, q_0]$ and the I_2's polynomial on $[q_0, c]$. If there is no such q_0 in $[a, c]$, we eliminate one of those two polynomials completely (whichever is uniformly larger) and repeat the process by looking at the resulting two adjacent polynomials.

- CASE 2: q is close to 1. For any given N, compare the polynomial from the R_{HLB} procedure for the highest range of q with the corresponding polynomial obtained in Case 1. Suppose the polynomials for the highest range of q covers the interval $[1, 1]$ for the R_{1A} procedure given for Case 1 and $[b, 1]$ for the R_{HLB} procedure and let $a < b$. We equate these two polynomials and proceed as in Case 1, looking for a new dividing point q_0. If there is no point q_0 in $[b, 1]$ for which the two polynomials are equal, we eliminate one of these two polynomials completely (whichever is uniformly larger) and repeat the process by looking at the polynomials for previously intervals, say, $[a, b]$. We then use the new dividing point to define the strategy for R_{1A} for q close to 1.

There are four main consequences of these modifications to the R_{1A} procedure:

1. $E(T|R_{1A})$ is continuous at all values of q;

2. $E(T|R_{1A})$ is decreasing in q;

3. For $q < .0618$, we test one unit at a time. This agrees with Ungar (1959), and the fact that one-at-a-time testing is known to be optimal for these q values;

4. $E(T|R_{1A}) \leq E(T|R_{HLB})$, and $E(T|R_{1A}) \leq E(T|R_1)$ for all values of q.

As an illustration, suppose we start with $N = 3$ units. From Chen, Hsu and Sobel (1987), polynomials for the expected tests required to classify a binomial set of size 1 through 5 using the R_1 procedure are given in Table 2. A portion of Table 2 is reproduced here to give the polynomials defining the expected tests required to identify the defective items in a binomial set of size 3 for all q :

x	Range of q	1	q	q^2	q^3
1	$0.0000 - 0.6180$	3			
2	$0.6180 - 0.7549$	5	-3	-1	1
3	$0.7549 - 1.0000$	5	-2	-1	-1

Since $E(T|R_1)$ jumps at $q = 0.7549$, proceed to create $E(T|R_{1A})$ by examining the polynomial for $[.6180, .7549]$and $[.7549, 1]$. Since they match at no point in $[.7549, 1]$, we examine the interval $[.6180, .7549]$ and find a match at .7071. Thus we obtain the new dividing point $q_0 = .7071$ where two adjacent polynomials are equal. To make the required modification for Case 2, compare the polynomial defining $E(T|R_{HLB})$ for the highest range of q with the corresponding one obtained in Case 1. Polynomials for the expected tests required for classifying a binomial set of size 1 through 5 using R_{HLB} procedure are given in Table 1. Since the polynomials for the highest range of q covers the interval $[.7071, 1]$ for the R_{1A} procedure given in Case 1 and $[.8385, 1]$ for the R_{HLB} procedure, we equate these two polynomials and find a new dividing point $q_0 = 8431$. We then use the new dividing point to define the strategy of R_{1A} for q close to 1. The explicit instructions for carrying out R_{1A} procedure are given in Table 3 for all q and for $N = 1$ through 5.

4. Concluding remarks. For $N \leq 5$, the R_{1A} procedure agreed with a result by Friedman (1982) for all the 9 values of q that he used for the calculation. Hence R_{1A} is optimal for these 9 selected values of q. It is believed that R_{1A} is the optimal procedure for all values of q and N, but this has not been fully confirmed.

Since Sobel's (1967) procedure R_∞ is the best known, we define the efficiency of the procedures R_{HLB} and R_{1A} by

$$\begin{aligned} Re(R_{HLB}|q) &= \{E(T|R_\infty)/E(T|R_{HLB})\} * 100\%; \\ Re(R_{1A}|q) &= \{E(T|R_\infty)/E(T|R_{1A})\} * 100\%. \end{aligned}$$

The values of $E(T|R_\infty)$, $E(T|R_{HLB})$, $E(T|R_{1A})$ and Huffman cost (HC) are given in Table 4 for $N = 3, 4, 5$, and for $q = 0.75$, 0.80,

Table 3.
Test size and polynomial coefficients* for the expected number of tests required to classify a binomial set of size N under R_{1A} procedure

n	Test Size x	Range of q	1	q	q^2	q^3	q^4	q^5
2	1	0.0000 – 0.6180	2					
	2	0.6180 – 1.0000	3	−1	−1			
3	1	0.0000 – 0.6180	3					
	2	0.6180 – 0.7071	5	−3	−1	1		
	3	0.7071 – 0.8431	5	−2	−1	−1		
	3	0.8431 – 1.0000	7	−3	−6	3		
4	1	0.0000 – 0.6180	4					
	2	0.6180 – 0.7071	7	−5	0	1	−1	
	2	0.7071 – 0.7862	7	−4	−1	−1	1	
	4	0.7862 – 0.8431	8	−4	−2	−1		
	4	0.8431 – 1.0000	10	−5	−7	3		
5	1	0.0000 – 0.6180	5					
	2	0.6180 – 0.7071	9	−7	1	0	−1	1
	2	0.7071 – 0.7549	9	−6	0	−1	1	−1
	3	0.7549 – 0.7834	9	−5	−1	−2	1	
	3	0.7834 – 0.8186	10	−5	−3	−5	7	−2
	5	0.8186 – 0.8431	11	−7	−2	−7	0	−1
	5	0.8431 – 0.8910	13	−8	−7	−8	0	−1
	5	0.8910 – 1.0000	14	−8	−8	−8	−3	2

* The integer shown is the coefficient of the power of q at the top of the column and the terms are then added to form the expected number of tests required for classifying a binomial set of size n. The entry x indicates that the next test is on x units taken from the only set available, i.e., the binomial set.

Table 4.
Expected number of group tests required*.

	$N=3$		$N=4$		$N=5$	
q	R_{00} R_{HLB} R_{1A}	HLB	R_{00} R_{HLB} R_{1A}	HLB	R_{00} R_{HLB} R_{1A}	HLB
.75	2.51562	2.46875	3.33203	3.27343	4.15723	4.08870
.80	2.24800	2.18400	3.00800[1]	2.96320	3.73568	3.68960
.85	1.95737	1.89325	2.53487[2]	2.47986	3.15529[3]	3.07193
.90	1.62700	1.59800	2.01700	1.97020	2.44868	2.40097
.95	1.30712	1.30700	1.50463	1.46900	1.71354	1.68100
.99	1.06030	1.06000	1.10020	1.09100	1.14059	1.13100

* $E(T)$ is the same for R_{00}, R_{HLB} and R_{1A} to 5 decimal places and efficiency with respect to HLB is 100.00%,except as noted:
[1] For R_{HLB}, $E(T) = 3.03840$ with 98.99% efficiency.
[2] For R_{HLB}, $E(T) = 2.58536$ with 98.05% efficiency.
[3] For R_{HLB}, $E(T) = 3.16287$ with 98.76% efficiency.

0.85, 0.90, 0.95 and 0.99. Note that, if $E(T|R_{1A})$ is not equal to the HC, it does not necessarily mean that R_{1A} is not optimal. The reason is simply that the Huffman cost is not always attainable for a group-testing. Since procedure R_{HLB} achieves 100% efficiency for $N = 3$, 4, and 5 and $q \geq 0.90$, it indicates that R_{HLB} is the optimal procedure for q close to 1, but this remains to be proved. For $N = 4$ and 5, and $q = 0.80$ and 0.85, the efficiency of R_{HLB} drops below 100%; hence R_{HLB} is not optimal for all q. The efficiency of R_{1A} is 100% to 5 decimal places. Therefore, R_{1A} is the best explicitly known procedure (to date) in the sense of minimizing the expected number of tests required to classify a binomial set of size ≤ 5.

Acknowledgments. The author would like to thank Professor Milton Sobel for reading an earlier version of this article and for helpful discussions. Thanks are also given to Professor Nancy Flournoy for her constructive comments.

References

CHEN, P., HSU, L., and SOBEL, M. (1987). Entropy based optimal group-testing procedures. *Probability in the Engineering and Informational Sciences* **1** 497-509.

DORFMAN, R. (1943). The detection of defective members of large population. *The Annals of Mathematical Statistics,* **14** 436-440.

FRIEDMAN, S. (1982). The optimal solution of the group testing problem, unpublished manuscript.

HUFFMAN, D.A. (1952). A method for the construction of minimum redundancy codes. *Proceedings of Institute of Radio Engineers* **40** 1098-1103.

REINGOLD, E.M., NIEVERGELT, J., and DEO, N. (1977). *Combinatorial Algorithms: Theory and Practice.* Englewood Cliffs, NJ: Prentice-Hall.

SOBEL, M. and GROLL, P.A. (1959). Group-testing to eliminate efficiently all defectives in a binomial sample. *Bell System Technology Journal* **38** 1179-1252.

SOBEL, M. (1967). Optimal group-testing. *Proceedings of the Colloquium on Information Theory,* Debrecen, Hungary, 411-488.

UNGAR, P. (1959). Cut-off point for group-testing. *Communications on Pure and Applied Mathematics* **13** 49-54.

DEPARTMENT OF MATHEMATICS
LE MOYNE COLLEGE
SYRACUSE, NY 13214-1399