ISSN-0019-5324

Indian Journal of Mathematics

Volume 46 Numbers 2-3 December 2004

THE ALLAHABAD MATHEMATICAL SOCIETY

10, C.S.P. Singh Marg, Allahabad, U.P., 211001 (India)

The Society is registered under the Societies Registration Act XXI of 1860 and has reciprocity arrangement with various Societies including the American Mathematical Society, the Canadian Mathematical Society and the London Mathematical Society.

INDIAN JOURNAL
OF MATHEMATICS
Vol. 46, Nos. 2-3, 2004, 303-321

THE PERFORMANCE OF ONE TYPE STEP-WISE GROUP SCREENING DESIGNS

M.M. MANENE AND R.O. SIMWA

(*Received* 1 August 2001; *Revised* 11 February 2005)

In this paper, we evaluate the performance of one type step-wise group screening designs in which group-factors contain equal number of factors in the initial step. We shall examine one type step-wise group screening designs without errors in observations, considering the possibility of cancellation of effects within group-factors. Expressions for the expected total number of runs and the expected number of active factors detected are obtained. The performance of one type step-wise group-screening designs is then compared with the performance of multistage group screening designs. Group screening designs have both biological and industrial applications.

1. Introduction

In scientific experimentation, situations arise in which a large number of potentially important factors must be examined. In such situations, there is often a need for an efficient method of factors screening, due to limited resources required for testing. One such method introduced by Watson [9] is the two stage group screening. This method has been generalized to more than two stages by Li [1] and Patel [6].

A basic assumption in group screening is that the direction of all suspected effects are known or can be correctly assumed a priori. Under this assumption factor levels can be assigned in such a way that there is no cancellation of effects within a group-factor. In practice, however, this assumption may be inadequate.

The notion of step-wise group screening was introduced by Patel and Manene [7] basing their method on the group testing procedure introduced by Sterrett [8]. Odhiambo and Manene [5] considered step-wise group screening design with errors in observations. Manene [2] has referred to step-wise group screening designs

Key words and phrases : one type step-wise designs, group-factors, initial step, type one search steps, cancellation of effects, screening efficiency.

considered by earlier authors as one type step-wise designs and extended these to two type step-wise design.

Mauro and Smith [3] considered the performance of two-stage group screening when the assumption of known effect direction is false. Odhiambo [4] generalized this approach and assessed the performance of multistage group screening, when there is a non-zero probability of cancellation of effects within a group factor. The purpose of this paper is to use the approach in Odhiambo [4] to assess the performance of the one type step-wise screening, when there is a non-zero probability of cancellation of effects within a group-factor. The performance of the one type step-wise group-screening designs shall be compared to the performance of multi-stage group screening designs.

2. Assumptions

Suppose that f factors are to be screened for their effect on the response. For detecting the factors having major effect, it is usually adequate to assume a first-order linear model

$$y_u = \beta_0 + \sum_{i=1}^{f} \beta_i x_{ui} + \varepsilon_u , \qquad (2.1)$$

where y_u is the u^{th} response, β_0 is a constant common to every response, β_i $(i \geq 1)$ is the linear effect of the i^{th} factor in the u^{th} run, and ε_u is the u^{th} error term.

In addition to (2.1), we shall assume that;

(i) All factors, have independently the same probability p_1 of having a positive effect and the same probability p_2 of having a negative effect. Thus the probability of a factor being active (defective) is $p = p_1 + p_2$.

(ii) All active factors have the same absolute effect, $\Delta \geq 0$, that is

$$|\beta_i| = \begin{cases} \Delta, & \text{if factor } i \text{ is active} \\ 0, & \text{if factor } i \text{ is inactive.} \end{cases}$$

(iii) The screening procedure is performed without experimental errors. This means that we take $\varepsilon_u = 0$ in (2.1).

These are basically the same assumptions made by Mauro and Smith (1982) and Odhiambo (1986).

Suppose that it is designed to classify f factors as active or inactive using the one type step-wise group screening procedure. The initial step of this procedure consists of dividing the f factors into g_1 first order group factors each of size k_1 ($f = k_1 g_1$). The first order group factors are then tested for their effect and those found to be effective are set aside. In step one of type one search steps, we start with any effective first order group-factor and test factors within it one by one till we find an active factor.

We set aside factors that are found to be inactive, keeping the active factor separate. In step two of type one search steps, the remaining factors are re-grouped in a group-factor which is then tested for its effect. The test procedure carried out in step one and step two of type one search steps are repeated successively in the subsequent type one search steps till the analysis terminate with a test on a non-effective group-factor or with a group-factor of size one. Type one search steps are performed on all the first order group-factors found to be effective in the initial step.

Since there is no experimental error, it is possible to use designs with the smallest number of runs in the initial step, i.e. the number of runs required to test g group factors is $g+1$, where the one extra run is the control run. This control may be used at every step of the one type step-wise design.

3. The Expected Number of Runs

Suppose that there are f factor divided into g group-factors in the initial step, such that each group-factor contains exactly k-factors. Let X denote the number of factors with positive effects and Y denote the number of factors with negative effects contained in a group factor of size k in the initial step. Let

p_1 = Prob. a factor chosen at random has a positive effect.

p_2 = Prob. a factor chosen at random has a negative effect.

Then

$$\text{Prob.}(X = x, Y = y) = \begin{cases} \dfrac{k!}{x!\,y!\,(k-x-y)!} p_1^x p_2^y (1-p_1-p_2)^{k-x-y}, & x+y \le k \\ 0, & \text{otherwise,} \end{cases}$$

which is the trinomial distribution with parameters k_1, p_1, p_2. This distribution will subsequently be denoted by $B(x; y; p_1, p_2, k)$.

At any step of type one step-wise design, a group-factor is active if it contains at least one active factor. However, an active group-factor will have a significant effect, i.e. will be effective if and only if the factor effects do not cancel completely within the group-factor. Let us define

p^* = Prob. (a group factor at the initial step is active) and $\theta =$ Prob. (a group factor at the initial step is effective). Then

$$p^* = 1 - \text{Prob.}(x=0, y=0) = 1 - (1 - p_1 - p_2)^k \tag{3.1}$$

and

$$\theta = 1 - \text{Prob.}(x = y)$$

$$= 1 - (1 - p_1 - p_2)^k - \sum_{x=1}^{k/2} B(x; x; p_1, p_2, k)$$

$$= p^* - \theta^* \quad (3.2)$$

where

$$\theta^* = p^* - \theta, \tag{3.3}$$

is the probability of cancellation within a group-factor at the initial step.

Let M denote the number of active group-factors and N denote the number of effective group-factors in the initial step. Then,

$$f(m) = \text{Prob.}(M = m) = \begin{cases} \binom{g}{m} p^* (1 - p^*)^{g-m}, & m = 0, 1, \cdots g \\ 0, & \text{otherwise} \end{cases} \tag{3.4}$$

and

$$f(n) = \text{Prob.}(N = n) = \begin{cases} \binom{g}{n} \theta^n (1 - \theta)^{g-n}, & n = 0, 1, \cdots g \\ 0, & \text{otherwise.} \end{cases} \tag{3.5}$$

Thus, we have

$$E(M) = gp^* \quad \text{and} \quad E(N) = g\theta. \tag{3.6}$$

Let $p_k^*(x, y)$ be the probability that a group-factor that has been identified as effective in the initial step contains exactly x factors with positive effects and y factors with negative effects for $x, y = 0, 1, 2, \cdots k,\ x \neq y$. This gives

$$p_k^*(x, y) = \frac{1}{\theta}\frac{k!}{x!\,y!} p_1^x p_2^y (1 - p_1 - p_2)^{k-x-y}\ (x + y \neq 0, x \neq y). \tag{3.7}$$

Denote by $E_k R_{(x,y)}$ $(x, y = 0, 1, 2, \cdots, k, x \neq y, 0 \leq x + y < k)$, the expected number of runs required to analyze a defective group-factor of size k that was found to be effective (either positive or negative) at the initial step if it contains x factors with positive effects and y factors with negative effects. If $x \neq 0$ and $y = 0$, then $E_k(R_{(x,y)})$ reduces to $E_k(R_{(x,0)})$ which is equal to $E_k(R_x)$ given by Patel and Manene [7]. Thus

$$E_k(R_{(x,0)}) = \frac{xk}{x+1} + x + \frac{x}{x+1} - \frac{2x}{k}, x = 1, 2, \cdots, k. \tag{3.8}$$

Similarly if $x = 0$ and $y \neq 0$,

$$E_k(R_{(0,y)}) = \frac{yk}{y+1} + y + \frac{y}{y+1} - \frac{2y}{k}, y = 1, 2, \cdots, k. \tag{3.9}$$

It follows that

$$E_k(R_{(0,x)}) = E_k(R_{(x,0)})$$

and

$$E_k(R_{(x,y)}) = E_k(R_{(y.x)}),\ x \neq y.$$

However, if $x = y$, i.e. the number of active factors with positive effects equals the number of active factors with negative effects, then the effects cancel completely and such a group-factor is dropped from further analysis in the initial step. Thus,

$$E_k(R_{(x,x)}) = 0,\ \ x = 0, 1, 2, \cdots, \frac{k}{2}. \tag{3.10}$$

We now prove two lemmas, which will be required to prove the theorem.

LEMMA 3.1. *The expected number of runs required to analyze a defective group-factor of size k containing one factor with positive effect and two factors with negative effects is given by*

$$E_k\,(R_{(1,2)}) = \frac{13}{18}k + \frac{29}{9} - \frac{7}{k} - \frac{7}{3k\,(k-1)}. \tag{3.11}$$

PROOF. There are two cases to consider, namely

(i) the first active factor detected has a positive effect;

(ii) the first active factor detected has a negative effect;

Considering case (i), the probability that the first factor rested has a positive effect is $\frac{1}{k}$, and the probability that the $(l+1)^{th}$ factor tested is the first active factor and has a positive effect is

$$\prod_{w=1}^{l} \frac{k-(w+1)}{k-(w-1)}\,\frac{1}{k-l}.$$

On the average we shall require $1+1+E_{k-1}\,(R_{(0,2)})$ tests to complete the test procedure if the first factor tested is active and has a positive effect, and $l+1+1+E_{k-(l+1)}\,(R_{(0,2)})$ tests to complete the test procedure if the $(l+1)^{th}$ factor tested is the first active factor and has a positive effect.

Considering case (ii), the probability that the first factor tested has a negative effect is $\frac{2}{k}$ and the probability that the $(l+1)^{th}$ factor tested is the first active factor and has a negative effect is

$$\prod_{w=l}^{l} \frac{k-(w+1)}{k-(w-1)}\,\frac{2}{k-l}.$$

On the average we shall require $1+1+E_{k-1}\,(R_{(1,1)})$ tests to complete the test procedure if the first factor tested is active and has a negative effect, and $l+1+1+E_{k-(l+1)}\,(R_{(1,1)})$ tests to complete the test procedure if the $(l+1)^{th}$ factor tested is the first active factor and has a negative effect. It follows that

$$E_k\,(R_{(1,2)}) = \frac{1}{k}\{1+1+E_{k-1}(R_{(0,2)})\}$$

$$+\sum_{l=1}^{k-3}\Pi_{w=1}^{l}\frac{k-(w+1)}{k-(w-1)}\frac{1}{k-l}(l+1+1+E_{k-(l+1)}(R_{(0,2)}))$$

$$+\frac{2}{k}\{1+1+E_{k-1}(R_{(1,1)})\}$$

$$+\sum_{l=1}^{k-3}\Pi_{w=1}^{l}\frac{k-(w+1)}{k-(w-1)}\frac{2}{k-l}\{l+1+1+E_{k-(l+1)}(R_{(1,1)})\}. \tag{3.12}$$

Rewriting (3.12) by using (3.9) and (3.10), we obtain

$$E_k(R_{(1,2)})=\frac{k-1}{k}\frac{1}{k-1}(2+\frac{2}{3}(k-1)+2+\frac{2}{3}-\frac{4}{k-1})$$

$$+\frac{k-2}{k}\frac{1}{k-1}(3+\frac{2}{3}(k-2)+2+\frac{2}{3}-\frac{4}{k-2})$$

$$+\frac{k-3}{k}\frac{1}{k-1}(4+\frac{2}{3}(k-3)+2+\frac{2}{3}-\frac{4}{k-3})$$

$$+\cdots$$

$$+\frac{2}{k}\frac{1}{k-1}(k-1+\frac{2}{3}(2)+2+\frac{2}{3}-\frac{4}{2})$$

$$+\frac{2}{k}\frac{k-1}{k-1}(2)+\frac{2}{k}\frac{k-2}{k-1}(3)$$

$$+\frac{2}{k}\frac{k-3}{k-1}(4)+\frac{2}{k}\frac{k-4}{k-1}(5)$$

$$+\cdots+\frac{2}{k}\frac{2}{k-1}(k-1).$$

That is,

$$E_k\,(R_{(1,2)}) = \frac{1}{k\,(k-1)} \sum_{i=1}^{k-2} (i+1)(k-i) + \frac{2}{3k\,(k-1)} \sum_{i=1}^{k-2} (k-i)^2$$

$$+ \frac{8}{3k\,(k-1)} \sum_{i=1}^{k-2} (i+1) - \frac{4\,(k-2)}{k\,(k-1)}$$

$$+ \frac{2}{k\,(k-1)} \sum_{i=1}^{k-2} (k-i)(i+1).$$

Simplifying we obtain equation (3.11). This completes the proof.

LEMMA 3.2. *The expected number of runs required to analyze an effective group-factor containing one factor with positive effects and three factors with negative effects are given by*

$$E_k\,(R_{(1,3)}) = \frac{47}{60}k + \frac{529}{120} - \frac{9}{k} - \frac{7}{2k\,(k-1)}. \tag{3.13}$$

PROOF. We again consider the two cases highlighted in the proof of Lemma 3.1 above. Considering case (i), the probability that the first factor tested has a positive effect is $\frac{1}{k}$ and the probability that the $(l+1)^{th}$ factor tested is the first active factor and has a positive effect is

$$\Pi_{w=1}^{l} \frac{k-(w+3)}{k-(w-1)} \frac{1}{k-l}.$$

On the average we shall need $1+1+E_{k-1}\,(R_{(0,3)})$ tests to complete the test procedure if the first factor tested is active and has a positive effect, and $l+1+1+E_{K-(l+1)}\,(R_{(0,3)})$ tests to complete the test procedure if the $(l+1)^{th}$ factor tested is the first active factor and has a positive effect.

For case (ii), the probability that the first factor tested has a negative effect is $\frac{3}{k}$ and the probability that the $(l+1)^{th}$ factor tested is the first active factor and has a negative effect is

$$\Pi_{w=1}^{l} \frac{k-(w+3)}{k-(w-1)} \frac{3}{k-l}.$$

On the average we shall need $1+1+E_{k-1}(\mathrm{R}_{(1,2)})$ tests to complete the test procedure if the first factor tested is active and has a negative effect; and $l+1+1+E_{k-(l+1)}(R_{(1,2)})$ tests to complete the test procedure if the $(l+1)^{th}$ factor tested is the first active factor and has a negative effect. Thus

$$E_k(R_{(1,3)}) = \frac{1}{k}(1+1+E_{k-1}(R_{(0,3)}))$$

$$+\sum_{l=1}^{k-4} \Pi_{w=1}^{l} \frac{k-(w+3)}{k-(w-1)} \frac{1}{k-l}(l+1+1+E_{k-(l+1)}(R_{(0,3)}))$$

$$+\frac{3}{k}(1+1+E_{k-1}(R_{(1,2)}))$$

$$+\sum_{l=1}^{k-4} \Pi_{w=1}^{l} \frac{k-(w+3)}{k-(w-1)} \frac{3}{k-l}(l+1+1++E_{k-(l+1)}(R_{(1,2)}).) \qquad (3.14)$$

Using (3.9) and Lemma 3.1 in (3.14), expanding the resulting expression and simplifying we obtain

$$E_k(R_{(1,3)}) = \frac{47}{60}k + \frac{529}{120} - \frac{9}{k} - \frac{7}{2k(k-1)}. \qquad (3.15)$$

This completes the proof of Lemma 3.2. Using the same argument as that used in the proof of Lemmas 3.1 and 3.2, we obtain

$$E_k(R_{(1,4)}) = \frac{37}{45}k + \frac{497}{90} - \frac{11}{k} - \frac{14}{3k(k-1)} \qquad (3.16)$$

$$E_k(R_{(2,3)}) = \frac{77}{180}k + \frac{259}{90} - \frac{9}{2k} - \frac{7}{3k(k-1)} \qquad (3.17)$$

$$E_k\,(R_{(1,5)}) = \frac{107}{126}k + \frac{1663}{252} - \frac{13}{k} - \frac{35}{6k\,(k-1)} \tag{3.18}$$

and

$$E_k\,(R_{(2,4)}) = \frac{28}{45}k + \frac{217}{45} - \frac{8}{k} - \frac{14}{3k\,(k-1).} \tag{3.19}$$

Let R_s^0 denote the number of tests required to analyze a group-factor of size k that is known to be effective. Then

$$E\,(R_s^0) = \sum_{x=0}^{k} \sum_{y=0}^{k} E_k\,(R_{(x,y)})\, p_k\,(x,y)$$

$$= \sum_{x=0}^{k} \sum_{y=0}^{k} E_k\,(R_{(x,y)}) \frac{B\,(x;y;p_1,p_2,k)}{\theta}\ (x+y \le k). \tag{3.20}$$

Denote by R_s the number of tests required to analyze all the factors in the N effective group-factors type one search steps. Then

$$R_s = NE\,(R_s^0). \tag{3.21}$$

Let R be the total number of runs required to analyze the f factors under investigation. Then

$$R = R_I + R_s, \tag{3.22}$$

where $R_I = 1 + \frac{f}{k}$ is the number of runs in the initial step.

Using (3.6), (3.20) and (3.21) in (3.22) and taking the expected value on both sides we obtain

$$E\,(R) = 1 + \frac{f}{k} + \frac{f}{k} \sum_{x=0}^{k} \sum_{y=0}^{k} E_k\,(R_{(x,y)})\, B\,(x;y;p_1,p_2,k).$$

$$(x+y \le k) \tag{3.23}$$

THEOREM 3.1. *Let R denote the total number of runs required to classify as effective or non-effective all the 'f' factors under investigation in a one type step-wise group-screening experiment. Then*

$$E(R) \cong 1 + \frac{f}{k} + \frac{f}{k}[\sum_{j=1}^{4} \frac{k(k-1)\cdots(k-j+1)}{j!}(1=p)^{k-j}(p_1^j + p_2^j)$$

$$\times(\frac{j}{j+1}k + j + \frac{j}{j+1} - \frac{2}{k}j)]$$

$$+\frac{k-2}{15120}\frac{f}{k}p_1 p_2(420(13k^3 + 45k^2 - 184k + 84)\times p(1-p)^{k-3}$$

$$+21(k-3)(94k^3 + 435k^2 - 1609k + 660)(p_1^2 + p_2^2)(1-p)^{k-4}),$$

where 'p_1' is the a-priori probability of a factor being effective with positive effect and 'p_2' is the a-priori probability of a factor being effective with negative effect ($p = p_1 + p_2$), and k is the size of the group-factor at the initial step.

PROOF. From (3.23), we have

$$E(R) = 1 + \frac{f}{k} + \frac{f}{k}\sum_{x=0}^{k}\sum_{y=0}^{k} E_k[R_{(x,y)}]B(x; y; p_1, p_2, k) \quad (x + y \le k).$$

Since p, the apriori probability of a factor being effective is usually small, we can assume that there are no more than four effective factors in an effective group-factor. This assumption is possible since the probability of getting more than four effective factors in an effective group-factor is very small. Using this assumption we have

$$E(R) \cong 1 + \frac{f}{k} + \frac{f}{k}[\sum_{j=1}^{4}(\frac{jk}{j+1} + j + \frac{j}{j+1} - \frac{2j}{k})[B(0; j; p_1, p_2, k)$$

$$+ B(j; 0; p_1, p_2, k)] + \sum_{x=1}^{4}\sum_{y=1}^{4} E_k[R_{(x,y)}]B(x; y; p_1, p_2, k)]$$

$$(x + y \le 4, x \ne y). \tag{3.24}$$

Now

$$\sum_{x=1}^{4}\sum_{y=1}^{4} E_k\,[R_{(x,y)}]\,B\,(x;y;p_1,p_2,k)$$

$$E_k\,[R_{(1,2)}]\,B\,(1;2;p_1,p_2,k)+E_k\,[R_{(2,1)}]\,B\,(2;1;p_1,p_2,k)$$

$$+E_k\,[R_{(1,3)}]\,B\,(1;3;p_1,p_2,k)+E_k\,[R_{(3,1)}]\,B\,(3;1;p_1,p_2,k)$$

$$=\frac{k-2}{15120}p_1p_2(420\,(13k^3+45k^2-184k+84)(p(1-p)^{k-3}$$

$$+21(k-3)\,(94k^3+435k^2-1609k+660)\,(p_1^2+p_2^2)\,(1-p)^{k-4}$$

$$(x+y\le 4,\quad x\ne y,\ x+y\le k). \tag{3.25}$$

Using (3.11), (3.15) and noting that $E_k\,(R_{(x,y)})=E_k\,(R_{(y,x)})$ and that $p=p_1+p_2$.

The proof of the theorem follows by using (3.25) in (3.24) and simplifying the resulting expression.

4. Expected Number of Active Factors Detected

Let A_I denote the number of active factors within active group-factors declared non-effective in the initial step. Then

$$A_I=(M-N)kp. \tag{4.1}$$

Denote by $E_k\,[I_{(x,y)}]$, the expected number of active factors declared inactive in type one search steps, from among the k factors within an initial step group-factor if the group-factor contains x factors with positive effects and y factors with negative effects. Obviously

$$E_k\,[I_{(x,0)}]=E_k\,[I_{(0,y)}]=E_k\,[I_{(x,x)}]=0 \tag{4.2}$$

and

$$E_k\,[I_{(x,y)}]=E_k\,[I_{(y,x)}],\quad x\ne y\,. \tag{4.3}$$

To obtain $E_k\ [I_{(x,y)}],\ x \neq y,\ \ x, y = 1, 2, k, x + y \leq k$; we consider the order in which the active factors occur for given x and y and the corresponding number of undetected active factors in each case. As an illustration, consider the case when x = 1 and y = 2. Then the possible orders in which active factors appear are as indicated in Table I, using '–' for negative effect and '+' for positive effect.

Table I

Determining $I_{(1,2)}$

Possibilities			$I_{(1,2)}$
–	–	+	2
–	+	–	2
+	–	–	0

Thus

$$E_k\ [I_{(1,2)}] = 2 \times \frac{2}{3} + 0 \times \frac{1}{3} = \frac{4}{3}. \tag{4.4}$$

Similarly

$$E_k\ [I_{(1,3)}] = 1 \tag{4.5}$$

$$E_k\ [I_{(1,4)}] = \frac{4}{5} \tag{4.6}$$

$$E_k\ [I_{(2,3)}] = \frac{14}{5} \tag{4.7}$$

$$E_k\ [I_{(1,5)}] = \frac{2}{3} \tag{4.8}$$

and

$$E_k\ [I_{(2,4)}] = \frac{32}{15}. \tag{4.9}$$

Let $\overline{A}_s^0$ denote the number of undetected active factors in a group-factor of size k which was found to be effective in the initial step. Then

$$E[\overline{A}_s^0] = \sum_{x=0}^{k} \sum_{y=0}^{k} E_k\,[I_{(x,y)}]\,P_k\,(x,y).$$

$$(x + y \leq k) \tag{4.10}$$

Denote by $\overline{A}_s$ the number of undetected active factors within the N group-factors of size k found to be effective in the initial step. Further, let $\overline{A}$ denote the total number of undetected active factors from among the f factors under investigation. Then

$$A_s = NE[\overline{A}_s^0] \tag{4.11}$$

and

$$\overline{A} = \overline{A}_I + \overline{A}_s. \tag{4.12}$$

Thus

$$E[\overline{A}] = E[\overline{A}_I] + E[\overline{A}_s] \tag{4.13}$$

Using (3.3), (3.6), (4.1), (4.10) and (4.11) in (4.13), we obtain

$$E[\overline{A}] = fp\theta^* + \frac{f}{k}\sum_{x=0}^{k} \sum_{y=0}^{k} E_k\,[I_{(x,y)}]\,B\,(x; y; p_1, p_2\ k) \quad (x + y \leq k). \tag{4.14}$$

THEOREM 4.1. *Let $\overline{A}$ denote the total number of undetected active factors in a one type step-wise group screening design in which an active factor has either a positive effect or a negative effect with respective probabilities p_1 and $p_2\ (p_1 + p_2 = p)$. Then*

$$E(\overline{A}) = fp\theta^* + \frac{(k-1)(k-2)}{6} fp_1 p_2\,[4p\,(1-p)^{k-3} + (k-3)(p_1^2 + p_2^2)(1-p)^{k-4}] \tag{4.15}$$

where f is the number of factors under investigation, k is the size of each group-factor at the initial step and θ^* is the probability of cancellation within a group-factor of size k.

PROOF. In equation (4.14), if we assume that an initial step group-factor contains no more than four active factors, then using equations (4.4) and (4.5), we have

$$\frac{f}{k}\sum_{x=0}^{k}\sum_{y=0}^{k} E_k\,[I_{(x,y)}]\,B\,(x;y;p_1,p_2,k) \cong \frac{f}{k}\sum_{x=0}^{4}\sum_{y=0}^{4} E_k\,[I_{(x,y)}]\,B\,(x;y;p_1,p_2,k)$$

$$= \frac{f}{k}\{\frac{4}{3}[B\,(1;2;p_1,p_2,k) + B\,(2;1;p_1,p_2,k) + B\,(1;3;p_1,p_2,k)$$

$$+B\,(3;1;p_1,p_2,k)\}. \qquad (4.16)$$

Simplifying (4.16) and substituting the resulting expression in (4.14), we obtain (4.15), which completes the proof.

Let A be the total number of active factors that are detected in a one type step-wise group screening design. Then

$$E\,(A) = fp - E\,(\overline{A})$$

$$\cong fp - \{fp\theta^* + \frac{(k-1)(k-2)}{6} fp_1 p_2$$

$$\times [4p\,(1-p)^{k-3} + (k-3)\,(p_1^2 + p_2^2)(1-p)^{k-4}]\}, \qquad (4.17)$$

where fp is the expected total number of active factors among the f factors under investigation.

5. Comparison of Relative Performance of One Type Step-Wise Group Screening and s-Stage Group Screening

To measure the efficiency of the one type step-wise group screening design, we need to obtain the efficiency of detecting active factors and the relative testing cost. Let us define

$$\phi_A = \frac{100E[\overline{A}]}{fp} \qquad (5.1)$$

as a percentage measure of the efficiency of a one type step-wise group screening strategy for detecting the active factors.

As another measure of the efficiency of a one type step-wise group screening procedure we define the relative testing cost,

$$E_R = \frac{100E(R)}{f+1} \tag{5.2}$$

as the ratio, expressed in percentage, of the expected number of runs required by a one type step-wise group screening design to the number of runs required to test the factors individually.

A large value of ϕ_A or a smaller value of E_R indicates better performance on the average, but both measures should be considered in assessing the performance of a group screening strategy since they trade off with one another. Only if one group screening design has both a high efficiency ϕ_A and smaller value of the relative testing cost E_R than another screening strategy, can the former be said to be definitely better than the latter.

Expressions for $E(R)$ and $E(\overline{A})$ for the one type step-wise group screening design are given in Theorem 3.1 and equation (4.17) respectively. Odhiambo [6] gave the corresponding expressions for an s-stage group screening design as

$$E(R_s) = 1 + g + \sum_{r=2}^{s-1} \pi_{j=1}^{r} g_j \theta_{r-1} + f\theta_{s-1} \tag{5.3}$$

and

$$E(\overline{A}_s) = fp\,(1 - \sum_{r=1}^{s-1} \theta_r^*) \tag{5.4}$$

respectively, where g_j $(j = 1, 2, \cdots, s-1)$ is the number of the j^{th} order group-factors, θ_r $(r = 1, 2, \cdots s-1)$ probability that an r^{th} order group-factor is effective and θ_r^* is the probability of cancellation in an r^{th} order group-factor $(r = 1, 2, \cdots, s-1, s \geq 2)$.

Table II shows the relative performance of one type step-wise group screening designs as compared with the two stage group screening designs. The table indicates that two-stage group screening designs are more efficient in detecting active factors than the one type step-wise group-screening designs. However, the one type step-wise group screening designs have lower relative testing cost.

From Table III we observe that one type step-wise group screening designs are more efficient in detecting active factors than the three stage group-screening designs. For $p \geq 0.030$, we observe that the one type step-wise group screening procedure has also a lower relative testing cost. Thus one type step-wise grouping screening is definitely better than the three stage group screening for $p \geq 0.030$.

Table II. Relative performance of type one step-wise and two ~~step~~-stage group screening designs for $f =$ 100 and specified values of p_1 and p_2.

	Type One step-wise group screening					Two stage group screening		
P	p_1	p_2	k	Min. E_R	ϕ_A	k	Min. E_R	ϕ_A
.001	.0006	.0004	46	5.5	99.92	32	7.18	99.97
.002	.0012	.0008	33	7.4	99.85	23	9.7	99.96
.003	.0018	.0012	27	8.87	99.77	19	11.62	99.93
.004	.0024	.0016	24	10.13	99.68	17	13.24	99.90
.005	.003	.002	21	11.38	99.64	15	14.64	99.88
.008	.0048	.0032	12	14.03	99.41	12	18.15	99.81
.009	.0054	.0036	11	14.85	99.34	11	19.17	99.80
.01	.006	.004	11	15.61	99.20	11	20.12	99.75
.02	.012	.008	8	21.88	98.36	8	27.67	99.52
.03	.018	.012	7	26.69	97.61	7	33.38	99.22
.04	.024	.016	6	30.72	96.75	6	38.03	99.02
.05	.03	.02	5	34.25	96.17	5	42.17	98.97
.06	.036	.024	5	37.36	94.78	5	45.71	98.56
.07	.042	.028	5	40.17	93.30	5	49.05	98.10
.08	.048	.032	5	42.7	91.72	5	52.17	97.60

Table III. Relative performance of type one step-wise and three-stage group screening designs for $f = 100$ and specified values of p_1 and p_2.

	Type One step-wise group screening					Three stage group screening			
P	p_1	p_2	k	Min. E_R	ϕ_A	k_1	k_2	Min. E_R	ϕ_A
.002	.0012	.0008	33	7.4	99.85	70	8	5.54	99.63
.003	.0018	.0012	27	8.87	99.77	54	7	6.92	99.52
.004	.0024	.0016	24	10.13	99.68	44	6	8.15	99.44
.005	.003	.002	21	11.38	99.64	40	6	9.25	99.23
.008	.0048	.0032	12	14.03	99.41	30	5	12.19	98.95
.009	.0054	.0036	11	14.85	99.34	28	5	13.06	98.84
.01	.006	.004	11	15.61	99.20	27	5	13.91	98.61
.02	.012	.008	8	21.88	98.36	18	4	21.08	97.75
.03	.018	.012	7	26.69	97.61	16	4	27.00	96.33
.04	.024	.016	6	30.72	96.75	15	4	32.23	94.82
.05	.03	.02	5	34.25	96.17	15	4	36.96	92.71
.06	.036	.024	5	37.36	94.78	15	4	41.27	90.70
.07	.042	.028	5	40.17	93.30	16	4	45.23	87.9
.08	.048	.032	5	42.7	91.72	17	4	48.86	85.31

6. Applications of Group Screening Designs

Group screening designs can be applied in blood screening for infectious diseases as HIV (Human Immunodeficiency Virus) during blood, donations for blood banks in Hospitals. Inspection of pooled samples of the blood may lead to detection and exclusion of the infected donated blood thus enhancing the safety of the blood recipients. In industry, group screening designs are used to identify the best catalyst for a chemical reaction from a large number of chemicals which act as catalysts. Group screening designs have also been used as a cost effective approach for screening chemical compounds as part of the drug discovery process in pharmaceutical companies. On the other hand, group screening designs with allowance of cancellation of effects, can be applied to cases where we anticipate that the factors that we wish to detect may interact with the possibility of neutralization, leading to non-detection of some of these factors.

REFERENCES

1. C.H. LI, A sequential method for screening experimental variables, *Jour. Amer. Statist. Assoc.* **57** (1962) 455-477.

2. M.M. MANENE, On two-type step-wise group screening designs, *Proceedings of theFifth Scientific Conference of the East Central and Southern Africa Network of International Biometric Society* (1997) 57-62.

3. C.A. MAURO AND D.E. SMITH, The performance of two-stage group screening in factor screening experiments, *Technom.* **24** (1982) 325-330.

4. J.W. ODHIAMBO, The performance of multistage group screening designs, *Commun. Statist. Theor. Meth.* **15** 8 (1986) 2467-2481.

5. J.W. ODHIAMBO AND M.M. MANENE, Step-wise group screening designs with errors in observations, *Commun. Statist. Theor. Meth.* **16** 10 (1987) 3095-3115.

6. M.S. PATEL, Group screening with more than two stages,*Technom.* **4** (1962) 209-217.

7. M.S. PATEL AND M.M. MANENE, Step-wise group screening with equal prior probabilities and no errors in observations, *Commun. Statist. Simula and Computa* **16** 3 (1987) 817-833.

8. A. STERRETT, On the detection of defective members of large populations, *Annals of Math. Statist.* **28** (1957) 1033-1036.

9. G.S. WATSON, A study of the group screening method, *Technom.* **3** (1961) 371-388.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF NAIROBI
P.O. BOX 30197
NAIROBI
KENYA

Indian Journal of Mathematics
Volume 46, Numbers 2-3, December 2004

CONTENTS

The Allahabad Mathematical Society Founded on the 25th of November, 1958, by Professor B.N. Prasad, an eminent mathematician and analyst, is a non-profit scientific organisation for promoting the cause of advanced study and research in various branches of mathematics including theoretical physics and mathematical statistics. The Society provides an outlet for original researches through its publications and organises lectures, symposia and conferences.

Publications of the Society

Indian Journal of Mathematics (IJM) : Three numbers per volume are published annually; it endeavors to publish significant research papers in different branches of mathematics and mathematical statistics.

Price : US $ 270 for each of the Volumes 45 (2003), 46 (2004) and 47 (2005) or equivalent thereof in Indian rupees. For institutions running research degree programmes in SAARC countries Indian Rs. 1200 for each of the Volumes 45 (2003), 46 (2004) and 47 (2005).

Bulletin of the Allahabad Mathematical Society (Bulletin, ams) : Bulletin ams, aims at publishing comparatively shorter research papers in all areas of mathematics, pure and applied, high quality research survey/expository articles on current topics, research announcements, book reviews, abstracts of papers presented at meetings of the Society and reports on the activities of the Society.

Price : US $ 100 for each of the Volumes 17 (2002), 18 (2003) and 19 (2004) or equivalent thereof in Indian rupees. For institutions running research degree programmes in SAARC countries Indian Rs. 600 for each of the Volumes 17 (2002), 18 (2003), and 19 (2004).

Lecture Note Series : Numbers in this series form texts/monographs on areas of current interest.

Contributors to Indian Journal of Mathematics should submit three copies of the manuscript to the Editor or Secretary for Publications, Allahabad Mathematical Society, 10 C.S.P. Singh Marg, Allahabad, U.P., 211001, India. The soft copy of the manuscript (pdf file) may also be submitted to ams10@sancharnet.in; pramila8@sancharnet.in for faster processing. Submission of a paper implies that the paper is original and is not submitted elsewhere. Authors are requested to leave margins of at least 5 cms. on both sides of the pages. Special instructions for the typesetter, where necessary, should be marked in light pencil on the left margin. Figures, if any, should be sent in camera ready form. The latest issue of the journal should be consulted for style. For details visit our website www.amsallahabad.org.

A galley proof will be sent to the first author unless otherwise indicated. Authors are requested to retain a copy of the paper as manuscript is not sent with the proof.

25 reprints are supplied gratis to the first author. Additional reprints can be obtained on payment by placing order at the time of proof correction.

Enquiries regarding subscriptions to the publications (current and back volumes) and membership of the Society should be addressed to :

The Treasurer, Allahabad Mathematical Society, 10, C.S.P. Singh Marg, Allahabad, U.P., 211001, India.

All correspondence of general nature related to the Society and its activities should be addressed to :

The Secretary, Allahabad Mathematical Society, 10, C.S.P. Singh Marg, Allahabad, U.P., 211001, India.