COMMUN. STATIST.—THEORY METH., 21(9), 2527-2539 (1992)

STEP-WISE SCREENING OF DEFECTIVE FACTORS

M.S. PATEL
Moi University, Eldoret, Kenya
M.M. MANENE
University of Nairobi, Kenya

Key Words and Phrases: Step-wise designs, group factors, initial step, subsequent steps, expected number of runs.

ABSTRACT

It is generally assumed that the defective factors of a population have the same a-priori probability of being defective. However, with some knowledge of the population, we can relax on this assumption and assume the population may consist of factors with unequal a-priori probabilities of being defective. Step-wise screening is developed to detect these factors with minimum expected number of runs assuming that there are no errors in the observations. Comparison is done with an equivalent two-stage group screening experiment.

1. INTRODUCTION

The problem of detecting defective factors of a population and eliminating them is of paramount importance not only in Biological Sciences but also has wide applications in Industry. Obviously, it is almost impossible to examine each factor of the population. It is essential to reduce the cost involved in the exercise. With this objective in mind, the concept of group

screening was first introduced by Dorfman (1943), which was later followed up extensively by Sobel and Groll (1949), Sterrett (1957), Watson (1961), Patel (1962), Patel and Ottieno (1984), Odhiambo and Patel (1986) Kleijnen (1989). This paper extends the approach made by Sterrett (1957) called Step-wise Screening by Patel and Manene (1987).

2. ASSUMPTIONS AND DESIGN STRUCTURE

With some prior knowledge of the population, it is possible to assume that all factors have different a-priori probabilities of being defective. The total number of factors, 'f' thus can be divided into a fixed number 'g' of group-factors in initial step such that $f = \sum_{i=1}^{g} k_i$, where k_i is the number of factors in i^{th} group-factor. All factors in the i^{th} group-factor have independently the same probability 'p_i' of being defective.

A defective factor within the i^{th} group-factor has a positive effect $\Delta_i (i=1,2,\ldots,g)$. There are no errors in the observations.

The step-wise group screening experiment is performed in steps as follows: In the initial step, the 'f' factors are divided into 'g' group-factors such that i^{th} group contains 'k_i' factors $(i=1,2,\ldots,g)$. These groups are called group-factors. Those that are indentified as non-defective are set aside. In step two, we start with any group-factor that is declared defective in the initial step and examine the factors within it one by one till we detect a defective factor. We set aside the factors which are identified as non-defective, keeping the factor declared defective separate. The remaining factors are then grouped into a group-factor which is tested in step three. Steps two and

minates with a group-factor declared non-defective or with a group-factor consisting of a single factor. Steps two onwards are called subsequent steps and are performed for all group-factors indentified as defective in the initial step.

3. EXPECTED NUMBER OF RUNS

Since the screening is done without errors in observations, we shall use designs with the smallest number of runs; i.e., the number of runs required to test m factors or group-factors is m + 1, where the extra one run is the control run. This control run will be used at every step of the step-wise experiment.

Let p_i be the probability that a factor in the i^{th} group-factor of size k_i in the initial step is defective (i=1,2,...,g). If p_i^* is the corresponding probability that the i^{th} group-factor is defective, then

$$p_i^* = \sum_{j=1}^{k_i} \binom{k_i}{j} p_i^{\,j} (1-p_i)^{k_i-j} = 1-(q_i)^{k_i} \qquad (3.1)$$

where

$$q_i = 1 - p_i$$

In the initial step, all the g group-factors are tested. Thus we require

$$R_I = g + 1$$

runs. (3.2)

Let $p_{k_i}(j)$ denote the probability that the group-factor identified as defective contains exactly j defective factors. Then

$$p_{k_i}(j) = (1-q_i^{k_i})^{-1}\binom{k_i}{j} p_i^{\,j} \quad (1-p_i)^{k_i-j} \qquad (3.3)$$

$$(j=1,2,\ldots\ldots,k_i).$$

Let $E_{k_i}(R_j)$ denote the average number of runs required to classify as defective or non-defective all the factors within i^{th} group-factor of size k_i in the subsequent steps if it contains exactly j defective factors. Then according to Patel and Manene (1987),

$$E_{k_i}(R_j) = \frac{jk_i}{j+1} + j + \frac{j}{j+1} - \frac{2j}{k_i}. \qquad (3.4)$$

Let R_{si} be the number of runs required to analyse the i^{th} group-factor in the initial step which is known to be defective. Then

$$E(R_{si}) = \sum_{j=1}^{k_i} E_{k_i}(R_j)p_{k_i}(j)$$

$$= (1-q_i^{k_i})^{-1}[(k_i + 1) + K_i p_i - 2p_i - \frac{1}{p_i}(1-q_i^{k_i+1})] \qquad (3.5)$$

Define a random variable U_i such that

$$U_i = \begin{cases} 1 \text{ with probability } p_i^* \\ 0 \text{ otherwise} \end{cases} \qquad (i = 1,2,\ldots\ldots,g)$$

Then the number $\underline{r}$ of defective group-factors in the initial step is given by

$$\underline{r} = \sum_{i=1}^{g} U_i \qquad (3.6)$$

and

$$E(\underline{r}) = \sum_{} P_i^* \qquad (3.7)$$

Let R_s denote the number of runs required to classify as defective or non-defective all the factors within the r group-factors found to be defective in the initial step. Then

$$R_s = \sum_{i=1}^{g} E(R_{si})U_i \tag{3.8}$$

Theorem 1

Let R be the total number of runs required to screen out the defective factors from among the f factors under investigation if the factors with probability p_i of being defective are grouped into a single group-factor of size k_i in the initial step ($i=1,2,\ldots,g$). Then

$$E(R) = 1+2g+f+\sum_{i=1}^{g}[k_ip_i - 2p_i - \frac{1}{p_i}(1-q_i^{k_i+1})].$$

Proof

The number of runs required in the initial step is

$$R_I = g+1.$$

In the subsequent steps we require

$$R_s = \sum_{i=1}^{g} E(R_{si})U_i \tag{3.9}$$

runs. The expected total number of runs is given by

$$E(R) = R_I + E(R_s)$$

$$= (g+1) + \sum_{i=1}^{g} E(R_{si})E(U_i)$$

$$= 1 + 2g + f + \sum_{i=1}^{g}[k_ip_i - 2p_i - \frac{1}{p_i}(1 - q_i^{k_i+1})]$$

This proves the theorem

Corollary 1

For small values of p_i's (i=1,2,....,g) the expected total number of runs is given by

$$E(R) = 1 + g - \sum_{i=1}^{g} 2p_i + \frac{3}{2}\sum_{i=1}^{g} k_i p_i + \frac{1}{2}\sum_{i=1}^{g} k_i^2 p_i$$

Proof

For small values of p_i's,

$$\frac{1}{p_i}(1 - q_i^{k_i+1}) = k_i + 1 - \frac{1}{2}(k_i^2 + k_i)p_i \qquad (3.11)$$

upto order p_i. The corollary follows immediately on substituting this value in (3.10).

4. OPTIMUM DESIGNS

In this section, we shall use corollary 1 to obtain estimates of the sizes of the group-factors in the initial step that minimize the expected number of runs. We shall also give an expression for the minimum expected number of runs. The expressions for the group sizes in the initial step and the minimum expected number of runs are approximate because of approximation in corollary 1.

Theorem 2

Assuming p_i, i.e., the a-priori probability of a factor in the i^{th} group-factor in the initial step to be defective to be small, the size k_i of the i^{th} group-factor which minimizes the expected total number of runs in a step-wise group screening

$$k_i = (f + \frac{3}{2}g)\ (p_i \sum_{i=1}^{g} 1/p_i)^{-1} - \frac{3}{2} \qquad (i=1,2,\ldots,g)$$

and the corresponding minimum value of E(R) is given by

$$\min E(R) = 1 + g - \frac{25}{8} \sum_{i=1}^{g} p_i + \frac{1}{8}(3g + 2f)^2\ (\sum_{i=1}^{g} 1/p_i)^{-1}$$

Proof

We have to minimize the expected total of number of runs in corollary 1 subject to the condition $f = \sum_{i=1}^{g} k_i$. Corollary 1 can now be re-written as

$$E(R) = F(k_1,k_2,\ldots,k_{g-1})$$

$$= 1 + g - 2(\sum_{i=1}^{g} p_i) + \frac{3}{2}[k_1p_1 + \ldots + k_{g-1}p_{g-1}$$

$$+ (f-k_1-k_2 - \ldots - k_{g-1})p_g]$$

$$+ \tfrac{1}{2}[k_1^2p_1 + \ldots + k^2_{g-1}p_{g-1} + (f - k_1 - k_2 -$$

$$\ldots\ldots - k_{g-1})^2p_g]. \qquad (4.1)$$

Assuming continuous variations in k_i's, critical values of k_i's are obtained by solving the equations $\frac{\partial F}{\partial k_i} = 0$ $(i=1.2\ldots,g-1)$, which imply $\frac{3}{2}(p_i - p_g) + k_ip_i - (f - k_1 - k_2 - \ldots - k_{g-1})p_g = 0$; i.e.,

$$\frac{k_1 + \frac{3}{2}}{1/p_1} = \frac{k_2 + \frac{3}{2}}{1/p_2} = \ldots\ldots = \frac{k_g + \frac{3}{2}}{1/p_g} = \frac{\frac{3g}{2} + f}{\sum_{i=1}^{g} 1/p_i} \qquad (4.2)$$

$$k_i \simeq (\frac{3}{2}g + f)(p_i \sum_{i=1}^{g} 1/p_i)^{-1} - \frac{3}{2} \qquad (i=1,2,\ldots,g). \quad (4.3)$$

The values of k_i's given in (4.3) can be easily shown to be the points of minimum for E(R). Substituting these values k_i in the expression for E(R) in corollary 1 we obtain

$$\min E(R) = 1 + g - \frac{25}{8} \sum_{i=1}^{g} p_i + \frac{1}{8}(3g + 2f)^2 \left(\sum_{i=1}^{g} 1/p_i\right)^{-1} \quad (4.4)$$

This completes the proof of the theorem.

5. COMPARISON OF STEP-WISE DESIGN WITH GROUP-FACTORS OF UNEQUAL SIZES WITH EQUIVALENT STEP-WISE DESIGN WITH GROUP-FACTORS OF EQUAL SIZES.

In this section, we shall compare the minimum expected of runs in step-wise group-screening when the group-factors are unequal in sizes with that one when the group-factors are of equal sizes.

When screening with group-factors of equal sizes 'k' and without errors in observations, Patel and Manene (1987) showed that the expected number of runs in a step-wise design is given by

$$E(R) = 1 + fp + \frac{2fq}{k} + f - \frac{f}{kp}(1 - q^{k+1})$$

$$= 1 + \frac{3fp}{2} + \frac{f}{k} - \frac{2fp}{k} + \tfrac{1}{2}fkp \qquad (5.1)$$

upto order p; where p is the prior probability of a factor to be defective ($q = 1 - p$). Assuming continuous variation, the value of k that minimizes E(R) is given by

$$k = \left(\frac{2-4p}{p}\right)^{\frac{1}{2}} \qquad (p < \tfrac{1}{2}). \qquad (5.2)$$

Substitute this value of k in (5.1) and we get

$$\min E(R) = 1 + \frac{3fp}{2} + f\left(\frac{p}{2-4p}\right)^{\frac{1}{2}} - 2fp\left(\frac{p}{2-4p}\right)^{\frac{1}{2}} + \frac{f}{2}(2-4p)^{\frac{1}{2}}p^{\frac{1}{2}}. \qquad (5.3)$$

Theorem 3

A step-wise group screening design with group-factors of unequal sizes where the i^{th} group factor contains factors with probability p_i of being defective has fewer runs than the corresponding step-wise design with the same number of group-factors but of equal sizes each containing factors with a-priori probability p of being defective provided $p_i \leq p$ (i=1,2,..,g).

Proof

We have to show that min E(R) given in theorem 2 is less than or equal to min E(R) given in (5.3), i.e.,

$$1 + g - \frac{25}{8}\sum_{i=1}^{g} p_i + \frac{1}{8}(3g + 2f)^2\left(\sum_{i=1}^{g} 1/p_i\right)^{-1} \leq 1 + \frac{3}{2}fp + \frac{fp^{\frac{1}{2}}}{(2-4p)^{\frac{1}{2}}} - \frac{3fp^{3/2}}{(2-4p)^{\frac{1}{2}}} + \frac{f}{2}(2-4p)^{\frac{1}{2}}p^{\frac{1}{2}}. \qquad (5.4)$$

Substituting $g = \frac{f}{k}$ where $k = \left(\frac{2-4p}{p}\right)^{\frac{1}{2}}$, inequality (5.4) becomes

$$-\frac{25}{8}\sum_{i=1}^{g} p_i + \frac{1}{8}(3g + 2f)^2\left(\sum_{i=1}^{g} 1/p_i\right)^{-1} \leq -\frac{25}{8}pg + \frac{1}{8}(3g + 2f)^2\frac{p}{g}\ ; \qquad (5.5)$$

$$\frac{25}{8}(pg - \sum_{i=1}^{g} p_i) + \frac{1}{8}(3g + 2f)^2 \quad [(\sum_{i=1}^{g} 1/p_i)^{-1} - \frac{p}{g}] \leq 0; \tag{5.6}$$

i.e.,

$$g^2 \leq \sum_{i=1}^{g} p_i \left(\sum_{i=1}^{g} 1/p_i \right) \tag{5.7}$$

which follows from Cauchy - Schwarz inequality. This completes the proof.

The minimum expected number of runs in a step-wise design when screening is done with unequal sizes could also be completed with the minimum expected number of runs in an equivalent two stage group screening design. This is indicated in tables I(a) and I(b).

When screening with unequal group sizes, Ottieno and Patel (1984) gave the expected number of runs in a two stage design as

$$E(R) = 1 + g + f - \sum_{i=1}^{g} k_i(1-p_i)^{k_i} \simeq 1 + g + \sum_{i=1}^{g} k^2{}_i p_i \tag{5.8}$$

for small values of p_i's. The values of k_i's which minimize E(R) in (5,8) are

$$k_i \simeq f(p_i \sum_{i=1}^{g} 1/p_i)^{-1} \qquad (i=1,2,\ldots,g). \tag{5.9}$$

Substituting (5.9) in (5.8) Ottieno and Patel gave min E(R) as

$$\min E(R) \simeq 1 + g + g^2(\sum_{i=1}^{g} 1/p_i)^{-1} \tag{5.10}$$

TABLE I

Optimum group sizes in the initial step and expected number of runs for selected probabilities for f = 100 when screening without errors.

(a) $p_i \leq p = 0.010$, $g = 7$

i	p_i	k_i
1	0.004	23.71
2	0.005	18.67
3	0.006	15.31
4	0.007	12.91
5	0.008	11.11
6	0.009	9.71
7	0.010	8.58
	Total	100.00

min E(R) = 13.42.

For equivalent two stage group screening design min E(R) = 17.13.

(continued)

6. GROUP SCREENING PLANS FOR STEP-WISE DESIGNS

In this section, we give group screening plans which minimize the expected number of runs as illustrations.

Tables I(a) and I(b) give the optimum group sizes in the initial step and the minimum expected number of runs for selected probabilities for f = 100, when screening is done without

TABLE I Continued

(b) $p_i \leq p = 0.035, \quad g = 13.$

i	p_i	k_i
1	0.008	17.10
2	0.009	15.03
3	0.010	13.38
4	0.013	9.94
5	0.015	8.42
6	0.017	7.25
7	0.020	5.94
8	0.022	5.26
9	0.025	4.45
10	0.027	4.01
11	0.030	3.46
12	0.033	3.01
13	0.035	2.75
	Total	100.00

min E(R) = 22.06.

For equivalent two stage screening design min E(R) = 22.45.

given by (4.3) and (4.4) respectively. The value of min E(R) for the equivalent two stage group screening is given in (5.10).

BIBLIOGRAPHY

Dorfman, R. (1943). The detection of defective members of large populations. Annals of Mathematical Statistics 14:436

Kleijnen, J.P.C. (1987). Review of Random and Group-Screening Designs. Special Issue of Communications in Statistics on "Experiments in Factor Screening." 16(10): 2885-2900.

Odhiambo, J.W. and Patel, M.S. (1986). On multistage Group-Screening Designs. Communications in Statistics, 15(5): 1627-1645.

Patel, M.S. (1962). Group screening with more than two stages. Technometrics, 4(2):209-217.

Patel, M.S. and Ottieno, J.A.M. (1984). Two-stage Group Screening Designs with equal prior probabilities and no errors in Decisions. Communications in Statistics, 13(9): 1147-1159.

Patel, M.S. and Manene, M.M. (1987). Step-wise Group-Screening with equal prior probability and no Errors in observations. Communications in Statistics, 16(3): S & C Series.

Sobel, M. and Groll, P.A. (1949). Group-testing to eliminate efficiently all defectives in a binomial sample. Bell System Techn. Journal 38: 1179-1252.

Sterrett, A. (1957). On the detection of defective members of large populations. Annals of Mathematical statistics, 28: 1033.

Watson, G.S. (1961). A study of group screening method, Technometrics 3: 371-388.

Received February 1991; Revised January 1992