Science and Technology, 8 (2003) 153-165

Step-Wise Group Screening Designs with Unequal A-Priori Probabilities and Errors in Observations

M.M. Manene

Department of Mathematics, University of Nairobi, P.O. Box 30197, Nairobi, Kenya, Email: mathuon@iconnect.co.ke.

تصاميم الغربلة التدريجية للمجموعات ذات الاحتمالات السابقة غير المتساوية والأخطاء في الملاحظات

م.م. مانين

خلاصـــة : يهـــتم هذا البحث بأداء طرق الغربلة التدريجية ذات الاحتمالات السابقة غير المتساوية بدلالة عدد التجارب المتوقعة وعـــدد الحـــد الأقصى المتوقع للقرارات الخاطئة ، وفيها يتم إيجاد الطريقة المثلى لتصاميم الغربلة التدريجية للاحتمالات السابقة غير المتساوية لحالة معرفة اتجاه العوامل المعتلة والملاحظات القابلة للخطأ ، وفيها أيضا يتم إيجاد دالة التكاليف المناسبة وإيجاد حجم المجموعة التي تجعل جملة التكاليف في قيمتها الصغرى.

ABSTRACT: The performance of step-wise group screening with unequal a-priori probabilities in terms of the expected number of runs and the expected maximum number of incorrect decisions is considered. A method of obtaining optimal step-wise designs with unequal a-priori probabilities is presented for the case in which the direction of each defective factor is assumed to be known a - priori and observations are subject to error. An appropriate cost function is introduced and the value of the group size which minimizes the expected total cost is obtained.

KEYWORDS: Step-Wise, Designs, Group-Factors, Initial Step, Subsequent Steps, Expected Number of Runs, Incorrect Decisions, Cost Function.

1. Introduction

There are investigations where a large number of factors needs to be examined. In such a situation we have to run an experiment to identify the influential factors. The group screening procedure aims at reducing the size of the experiment, thus conserving resources.

The method of group testing was first introduced by Dorfman (1943), who proposed that instead of testing each blood sample individually for the presence of a rare disease, blood samples be pooled and analysed together.

Watson (1961) considered two stage group screening designs with and without errors in observations and with equal prior probabilities. In the same paper, he laid down the device of using different group sizes when prior probabilities differ. Li (1962) and Patel (1962) generalized Watson's method to more than two stages. Both these authors considered multistage group-screening designs with equal prior probabilities and without errors in observations. Ottieno and Patel (1984) extended the idea of two stage group screening with unequal prior probabilities to include situations when no prior information is available so that no natural partitioning can be assumed. Odhiambo and Patel (1986) generalized this approach to multi-stage designs.

The group testing procedure first considered by Sterrett (1957) has been extended by Manene (1985), Patel and Manene (1987), Odhiambo and Manene (1987) and Manene (1997) in what they have called step-wise group-screening designs and they have approached the problem from the

3. Expected Number of Runs

Suppose that f factors are divided into a fixed number 'g' of group-factors in the initial step such that the i^{th} group-factor is of size $k_i \left(\sum_{i=1}^{g} k_i = f \right)$. The '$g$' group-factors are tested in

$$\begin{aligned} R(g) &= g + 4 - g \pmod 4 \\ &= g + h (h = 1,2,3,4) \end{aligned} \tag{3.1}$$

runs. In the subsequent steps, factors within the defective group-factors are tested as explained earlier.

Let $\hat{G}_i$ be the estimate of the main effect of the i^{th} group-factor in the initial step, with δ_i effective factors each with effect $\Delta_i > 0; (\delta_i = 1,2,....,k_i)$. Then $E(G_i) = \delta_i \Delta_i$ and Var $\left(\hat{G}_i\right) = \sigma^2 / g + h$.

Define a random variable W_i by

$$W_i = \hat{G}_i / \sigma \sqrt{g + h} \tag{3.2}$$

Then

$$E(W_i) = \delta_i u_i \text{ and Var } (W_i) = 1, \text{ where } u_i = \Delta_i / \sigma \sqrt{g + h} \tag{3.3}$$

Consider the hypothesis $H_0 : \delta_i u_i = 0$ alternative $u_i \delta_i = 0$. Let α_{Ii} be the level of significance for testing the i$^{\text{th}}$ group-factor in the initial step and denote by $\Pi_{\text{Ii}}(\delta_i u_i, \alpha_{\text{Ii}})$ the power function of the test. Then

$$\Pi_{\text{Ii}}(\delta_i u_i, \alpha_{\text{Ii}}) = 1 - \phi(Z(\alpha_{\text{Ii}}) - \delta_i u_i) \tag{3.4}$$

Where $\phi(.)$ denotes the standard normal distribution function and $Z(\alpha_{\text{Ii}})$ satisfies

$$\alpha_{\text{Ii}} = 1 - \phi(Z(\alpha_{\text{Ii}})) \tag{3.5}$$

Thus if $\delta_i = 0$ for $\mu_i = 0$, then $\Pi_{\text{Ii}}(0, \alpha_{\text{Ii}}) = \alpha_{\text{Ii}}$ and if $\delta_i \neq 0$ and Δ_1 / σ is large, then $\Pi_{\text{Ii}}(\delta_i u_i, \alpha_{\text{Ii}})$ tends to 1.

Let Π^*_{Ii} denote the probability that the i^{th} group factor is declared defective in the initial step. Then

$$\Pi^*_{\text{Ii}} = \sum_{\delta_i = 0}^{ki} \binom{k_i}{\delta_i} p_i^{\delta_i} (1 - p_i)^{k_i - \delta_i} \Pi_{\text{Ii}}(\delta_i, u_i, \alpha_{\text{Ii}}) \tag{3.6}$$

where p_i is the probability that a factor in the i^{th} group-factor in the initial step is defective. Define a random variable U_i such that

$$U_i = \begin{cases} 1 & \text{with probability } \Pi^*_{\text{Ii}} \\ 0 & \text{otherwise} \quad (i = 1,2,...g) \end{cases} \tag{3.7}$$

Then

$$E(U_i) = \Pi^*_{\text{Ii}} \tag{3.8}$$

$$E_{k_i}^{*}\left(R_j\right)=\begin{cases} k_i \quad \text{for} \quad j=0 \\ \dfrac{jk_i}{j+1}+j+\dfrac{j}{j+1}-\dfrac{j\left(k_i+j-2\right)}{k_i\left(k_i-1\right)}+\alpha_i^{*}\left\{\dfrac{k_i}{j+1}-\dfrac{j}{j+1}-\dfrac{j}{k_i-1}+\dfrac{j^2}{k_i\left(k_i-1\right)}\right\} \\ -\dfrac{j\left(1-\xi_i\right)\left(k_i-j\right)}{k_i\left(k_i-1\right)} \qquad \text{for} \quad j=1,2,.....,k_i \end{cases} \tag{3.14}$$

where α_i^{*} is as already defined and $\xi_i=0$ if $\alpha_i^{*}=0$ and 1 otherwise.

Let R_{si} be the number of runs required to analyse the i^{th} group-factor once it has been declared defective in the initial step. Then

$$\begin{aligned} E\left(R_{si}\right)&=\sum_{j=1}^{k_i}E_{k_i}^{*}\left(R_j\right)p_{k_i}^{*}\left(j\right) \\ &=k_i-\frac{k_i}{\Pi_{Ii}}\left[1-\left(1-\overline{\beta}_i^{*}\right)^{k_i}\right]-\frac{1}{\Pi_{Ii}^{*}}\left[\left(2-\xi_i\right)\overline{\beta}_i^{*}+\xi_i\overline{\beta}_i^{*2}\right] \\ &+\frac{1}{\Pi_{Ii}}\left[k_i+1+k_i\overline{\beta}_i^{*}-\frac{1}{\overline{\beta}_i^{*}}\left\{1-\left(1-\overline{\beta}_i^{*}\right)^{k_i+1}\right\}\right] \\ &+\frac{\alpha_i^{*}}{\Pi_{Ii}^{*}}\left[\frac{1}{\overline{\beta}_i^{*}}\left\{1-\left(1-\overline{\beta}_i^{*}\right)^{k_i+1}-k_i\overline{\beta}_i^{*}\left(1-\overline{\beta}_i^{*}\right)^{k_i}-1-\overline{\beta}_i^{*}+\overline{\beta}_i^{*2}\right\}\right] \end{aligned} \tag{3.15}$$

Using (3.13) and (3.14). If R_s is the number of runs required to analyse all the group factors declared defective in the initial step, then

$$R_s=\sum_{i=1}^{g}U_iE\left(R_{si}\right) \tag{3.16}$$

Theorem 3.1: The expected total number of runs in a step-wise group screening design with g (fixed) group-factors in the initial step such that the i^{th} group-factor is of size $k_i\left(i=1,2,....,g\right)$ is given by

$$\begin{aligned} E\left(R\right)&=h+f+2g-\sum_{i=1}^{g}\frac{1}{\overline{\beta}_i^{*}}\left(1-\alpha_i^{*}\right)\left\{1-\left(1-\overline{\beta}_i^{*}\right)^{k_i+1}\right\} \\ &+\sum_{i=1}^{g}k_i\overline{\beta}_i^{*}\left\{1-\frac{2-\xi_i}{k_i}-\frac{\alpha_i^{*}}{k_i}\right\}+\sum_{i=1}^{g}k_i\Pi_{Ii}^{*} \\ &-\sum_{i=1}^{g}k_i\left[1-\left(1-\overline{\beta}_i^{*}\right)^{k_i}\right]-\sum_{i=1}^{g}\alpha_i^{*}+\sum_{i=1}^{g}\left(\alpha_i^{*}-\xi_i\right)\overline{\beta}_i^{*2} \\ &-\sum_{i=1}^{g}k_i\alpha_i^{*}\left(1-\overline{\beta}_i^{*}\right)^{k_i} \end{aligned}$$

Where $\overline{\beta}_i^{*},\alpha_i^{*}$ and ξ_i are as defined earlier.

Proof : In the initial step we require $R_1=g+h$ runs $\left(h=1,2,3,4\right)$. The number of runs

$$\underline{d} = \sum_{i=1}^{g} k_i \overline{\beta}_i^{+} U_i \tag{4.2}$$

where U_i is as defined in (3.7).

Let $p_i^{(0)}$ be the probability that a factor chosen at random from the i^{th} initial step group- factor declared non-defective, is defective. Then

$$p_i^{(0)} = p_i \left(1-\Pi_{Ii}^{+}\right)/\left(1-\Pi_{Ii}^{*}\right) \tag{4.3}$$

Let p_i^{+} be the probability that a factor from the i^{th} group factor is non-defective given that it is declared defective. Then

$$p_i^{+} = \alpha_{si}\left(1-p_i'\right)/\overline{\beta}_i^{+} \tag{4.4}$$

Theorem 4.1 : Let M_R be the number of defective factors declared defective in a step-wise group screening design with g initial group-factors, the factors in the i^{th} group-factor of size k_i being defective with a priori probability $p_i\ (i=1,2,.....,g)$. Then

$$E\left(M_R\right) = \sum_{i=1}^{g} k_i p_i \Pi_{Ii}^{+} \gamma_{si}$$

Proof : The expected total number of factors declared defective in the subsequent steps is given by

$$E\left(\underline{d}\right) = \sum_{i=1}^{g} k_i \overline{\beta}_i^{+} E\left(U_i\right) = \sum_{i=1}^{g} k_i \overline{\beta}_i^{+} \Pi_{Ii}^{*} \tag{4.5}$$

The probability that a factor which is declared defective from the i^{th} group-factor, is defective is given by $1-p_i^{+}$.

Therefore

$$\begin{aligned} E\left(M_R\right) &= \sum_{i=1}^{g} k_i \overline{\beta}_i^{+}\left(1-p_i^{+}\right) E\left(U_i\right) \\ &= \sum_{i=1}^{g} k_i \left[\overline{\beta}_i^{+} - \alpha_{si}\left(1-P'\right)\right] \Pi_{Ii}^{*} \\ &= \sum_{i=1}^{g} k_i p_i \Pi_{Ii}^{*} \gamma_{si} \end{aligned} \tag{4.6}$$

as required.

Let I_I denote the expected number of defective factors declared non-defective in the initial step. Then

$$\begin{aligned} I_I &= E\left[\sum_{i=1}^{g}\left(1-U_i\right) k_i p_i^{(0)}\right] \\ &= \sum_{i=1}^{g} k_i p_i \left(1-\Pi_{Ii}^{+}\right) \end{aligned} \tag{4.7}$$

Theorem 4.2 : In a step-wise group screening design with errors in observations and unequal a-priori probabilities, the expected number of defective factors declared non-defective in the subsequent steps is given by

$$I_s = \sum_{i=1}^{g} k_i p_i \Pi_{Ii}\left(1-\gamma_{si}\right).$$

$$Max\ \ I=\sum_{i=1}^{g}k_i p_i-\sum_{i=1}^{g}k_i p_i\gamma_{si}\Pi_{Ii}\left(\phi_{Ii},\alpha_{Ii}\right)$$
$$+\sum_{i=1}^{g}k_i\alpha_{si}\left\{\left(\alpha_{Ii}\,q_i^{k_i}+\left(1-q_i^{k_i}\right)\Pi_{Ii}\left(k_i\phi_{Ii},\alpha_{Ii}\right)\right)-p_i\Pi_{Ii}\left(\phi_{Ii},\alpha_{Ii}\right)\right\}$$

Proof:

$$\begin{aligned}I&=\sum_{i=1}^{g}k_i p_i-\sum_{i=1}^{g}k_i p_i\Pi_{Ii}^{+}\gamma_{si}+\sum_{i=1}^{g}k_i\alpha_{si}\left(\Pi_{Ii}^{*}-p_i\Pi_{Ii}^{+}\right)\\&=\sum_{i=1}^{g}k_i p_i-E\left(M_R\right)+E\left(M_u\right)\end{aligned}\qquad(4.11)$$

Hence I will take its maximum value when $E\left(M_R\right)$ is minimum and $E\left(M_u\right)$ is maximum. But $E\left(M_R\right)$ takes its maximum value when Π_{Iii}^{+} is replaced by $\Pi_{Ii}\left(\phi_{Ii},\alpha_{Ii}\right)$. That is

$$Min\ \ E\left(M_R\right)=\sum_{i=1}^{g}k_i\gamma_{si}p_i\Pi_{Ii}\left(\phi_{Ii},\alpha_{Ii}\right)\qquad(4.12)$$

$E\left(M_u\right)$ will take a minimum value when Π_{Iii}^{*} is replaced by its maximum value and Π_{Ii}^{+} is replaced by its minimum value. That is when Π_{Ii}^{*} is replaced by $\left\{\alpha_{Ii}\,q_i^{k_i}+\left(1-q_i^{k_i}\right)\Pi_{Ii}\left(k_i\phi_{Ii},\alpha_{Ii}\right)\right\}$ and Π_{Ii}^{+} is replaced by $\Pi_{Ii}\left(\phi_{Ii},\alpha_{Ii}\right)$.
Thus

$$Max\ \ E\left(M_u\right)=\sum_{i=1}^{g}k_i\alpha_{si}\left\{\alpha_{Ii}\,q_i^{k_i}+\left(1-q_i^{k_i}\right)\Pi_{Ii}\left(k_i\phi_{Ii},\alpha_{Ii}\right)-p_i\Pi_{Ii}\left(\phi_{Ii},\alpha_{Ii}\right)\right\}\qquad(4.13)$$

The result follows on using (4.12) and (4.13) in (4.11).

Corollary 4.2 : For large $\frac{\Delta_i}{\sigma}$ and small $p_i's$

$$Max\ \ I\approx\sum_{i=1}^{g}k_i\alpha_{si}\left[\alpha_{Ii}-p_i+\left(1-\alpha_{Ii}\right)p_i k_i\right]$$

Proof: For large $\frac{\Delta_i}{\sigma}$ and small $p_i's$, $\Pi_{Ii}\left(s_i\phi_{Ii},\alpha_{Ii}\right)\approx 1$, $\gamma_{si}\approx 1$ and $q_i^{k_i}\approx 1-k_i p_i$.

The result follows immediately on replacing these values by their approximations in the expression for *Max* I given in corollary 4.1.

5. Optimum sizes of initial group-factors in relation to total cost

We define the expected total cost(C) as a linear function of the expected number of runs and the expected number of incorrect decisions and obtain the sizes of the group-factors so that the expected total cost is minimum.
Let c_1 be the cost of inspection per run and c_2 be the loss for each incorrect decision made. Then the expected total cost is given by

$$C=c_1E\left(R\right)+c_2\ \ \text{Max}\ I$$

$$F\left(k_1,k_2,.....,k_g,\lambda\right)=c_1E\left(R\right)+c_2\ \text{Max}\ I+\lambda\left(\sum_{i=1}^{g}k_i-f\right)$$

where λ is the Lagrange multiplier. Assuming continuous variation in k_i, the critical value of k_i is obtained from the equations

$$\partial F/\partial k_i=0\quad;\quad i=1,2,...,g,\ \text{and}\ \partial F/\partial\lambda=0 \tag{5.3}$$

The theorem follows immediately on solving equations (5.3).

6. Examples of screening plans

The screening efficiency of step-wise group screening design with unequal group sizes can be measured in terms of the minimum expected total cost. A small value of (C) indicates better performance on the average. Examples of group screening plans which minimize the expected total cost (C) are given in Table 1 below. The corresponding values of $E(R)$ and $Max\ I$ are also given.

Table 1 : Optimum group-sizes obtained by minimizing expected total cost (C), when $\alpha_{Ii}=\alpha_I,\alpha_i^*=\alpha^*$ and $\alpha_{si}=\alpha_s$, for $f=100$ and for selected unequal apriori probabilities. The minimum (C) given is a relative figure using c_1 (the cost of observing a run) as the unit.

(a) $h=3,g=13,\alpha_I=\alpha^*=\alpha_s=0.05$, $c_2:c_1=3:5$, $p_i\le p=0.035$.

i	p_I	k_I
1	0.008	17.088
2	0.009	15.024
3	0.010	13.372
4	0.013	9.942
5	0.015	8.418
6	0.017	7.252
7	0.020	5.940
8	0.022	5.265
9	0.025	4.454
10	0.027	4.014
11	0.030	3.463
12	0.033	3.013
13	0.035	2.755
Total		**100.000**

$E(R)=28.577$, $Max\ I=0.770$, $Min\ C=29.039$. The corresponding value of min $E(R)=26.739$ when incorrect observations are not considered.

(b) $h=4,g=20,\alpha_I=\alpha^*=\alpha_s=0.05$, $c_2:c_1=3:5$, $p_i\le p=0.100$.

PATEL, M.S. and MANENE, M.M. 1987. Step-Wise Group Screening with Equal Prior Probabilities and no errors in Observations, *Comm. In Statistics Simula_and Comguta,* **16(3):** 817-833.

PLACKET, R.L. and BURMAN, J.P 1946. The Design of Optimum Multifactor Experiments. *Biometrika,* **33:** 305-325.

STERRET, A. 1957. On the Detection of Defective Members of Large Populations, *Annals of Mathematical Statistics*. **28:** 1033-1036.

WATSON, G.S. 1961. A study of the Group Screening Method, *Technometrics*, **3:** 371-388.

Received 30 December 2002
Accepted 20 November 2003