This thesis is my original work and has not been presented for a degree in any other University.

Signature ....

M.M. MANENE

This thesis has been submitted for examination with my approval as the University Supervisor.

Signature ....

M.S. PATEL

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

FURTHER INVESTIGATIONS OF GROUP SCREENING DESIGNS:- STEP-WISE DESIGNS

By

MOSES MWANGI MANENE

This thesis is submitted in fulfilment for the Degree of Doctor of Philosophy in Mathematical Statistics in the Department of Mathematics

UNIVERSITY OF NAIROBI

DECEMBER, 1985

CONTENTS

APPENDICES

Summary of Contents

The inspection of individual members of a large population is an expensive and tedious process. Often when testing the results of manufacture, the work can be reduced greatly by examining only a sample of the population and rejecting the whole if the proportion of defectives in the sample is unduly large. In many inspections however, the objective is to eliminate all the defective members of the population. This situation arises in manufacturing processes where the defect being tested for can result in disastrous failures. It also arises in certain inspections of human populations with say infectious diseases.

Where the objective is to weed out individual defective units, a sample inspection will not suffice. In this case, we need designs which will classify all the items in the population as defective or non-defective. Such designs are known as screening designs. Earlier work in this area was done by Dorfman [3] and Sterret [26]. Connor [1], Watson [28] and Patel [13], [14] have approached the problem from the point of view of designs of experiments and called these designs "Group screening designs". This thesis is along the lines of Sterret's paper [26]. The problem has been approached from the point of view of design of experiments.

Chapter I defines the concept of group screening designs and describes briefly the work done in this and related areas by several authors in the past. The chapter also lays down the assumptions which are used in this thesis.

In chapter II, step-wise group screening designs have been introduced and are studied assuming that all factors have the same a-priori probability of being defective. Optimum group sizes in the initial step have been determined considering only the expected total number of runs. A comparison of two-stage group screening design with step-wise group screening design is presented.

Chapter III extends the results of chapter II to the case where factors are defective with unequal a-priori probabilities. It is shown that under certain conditions, the minimum expected number of runs when screening is done under the assumption that factors are defective with unequal a-priori probabilities is smaller than the minimum expected number of runs when screening is done under the assumption that all factors are defective with the same a-priori probability.

In chapter IV, the optimum sizes of the group-factors for both the cases when we screen with equal and with unequal a-priori probabilities have been determined taking into consideration both the expected number of incorrect decisions and the expected number of runs. To balance the apparently opposite trends of the expected number of runs and the expected number of incorrect decisions, a cost function has been defined and optimum sizes of the group-factors determined by minimizing the cost function.

At the end, are given a series of tables which show some group screening plans resulting from the work that has been done in chapters II through IV. This appears in appendices I, II and III.

Throughout this thesis, it is assumed that the value of 'p', i.e., the a-priori probability of a factor to be defective is known heuristically. Thus no attempt is made to estimate 'p' in this thesis. The work has been extended to the case with more than one value of 'p'. For example in a manufacturing plant turning out hundreds of items every day, the probability of the plant producing defective items will vary from time to time due to assignable causes of variation which affect the production. Thus in such a case, it is reasonable to assume that items will be defective with unequal a-priori probabilities. Again we shall assume that the values of these a-priori probabilities are known heuristically. However, the optimum sizes of the group-factors will depend on the expected number of runs and the expected number of incorrect decisions.

Familiar calculus methods have been used to solve most of the problems in this thesis. The methods used include Newton - Raphson iterative method, the method of Lagrange's multipliers and ordinary differentiation.

ACKNOWLEDGEMENT

I desire to express my indebtedness to those who have assisted me at different stages in the preparation of this thesis.

My thanks are due to my supervisor, Professor M.S. Patel, for the constant encouragement, instructive discussions and valuable suggestions he has given me during the preparation of this thesis.

I am grateful to my mentor, Dr. J.W. Odhiambo, for his valuable suggestions, and for the interest and encouragement he accorded me during the course of this research.

To the members of staff Mathematics Department (University of Nairobi), I wish to express my gratitude for their encouragement.

Also my thanks go to my wife Ng'endo and daughter Wangari for their patience, understanding and constant encouragement during the preparation of this thesis.

I greatly appreciate the efficient manner in which Mrs. M. Okello typed this thesis.

Finally, I wish to express my intellectual indebtedness to the authors whose works are cited in the brief literature review given in section two of chapter one of this thesis. All have contributed much to my understanding of group screening designs.

CHAPTER I

INTRODUCTION

1.1 BASIC IDEAS IN GROUP-SCREENING DESIGNS

The problem of detecting defective factors in a large population consisting of defective and non-defective factors has been tackled in various ways. Designs used in this kind of investigation have been called screening designs. One such class of designs is the group-screening designs.

In group-screening designs, the factors or members of the population are divided into groups called group-factors. The group-factors are then tested for significance and classified as either defective or non-defective. If a group-factor is classified as non-defective, then it is dropped from further investigation since it is assumed that all the factors within that group-factor are non-defective. If, however a group-factor is classified as defective, individual factors from that group-factor are investigated further.

Group screening experiments can be carried out in several stages. In a two stage group screening design, the group-factors formed are tested in the first stage and factors from defective group-factors only are tested in the second stage. In a three stage group screening design, the first stage consists of dividing the factors into group-factors, known as first order group-factors, which are then tested and classified as defective or non-defective. In the second stage of the experiment, each first order group-factor classified as defective in the first stage is further divided into smaller group-factors called second order group-factors which are then tested and classified as defective or non-defective. Finally, in the third stage all the factors belonging to the second order group-factors found to be defective in the second stage are tested individually and classified as defective or non-defective. The three stage group-screening design can be extended to s-stage group-screening design ($s \geq 3$).

In a step-wise group-screening design, the analysis is carried out as follows:- In the initial step, the factors are divided into group-factors. The group-factors are then tested for their significance. Those that are found to be non-defective are set aside. In step two, we start with any defective group-factor and test the factors within it one by one till we find a defective factor. We set aside the factors which are found to be non-defective, keeping the defective factor separate. The remaining factors are then grouped into a group-factor. In step three, we test the group-factor obtained after step two is performed. If the group-factor is non-defective, we terminate the test procedure. If the group-factor is defective, we continue with step four. In step four, factors within a group-factor found to be defective in step three are tested one by one till a defective factor is found. Factors which are found to be non-defective are again set aside keeping the defective factor separate. The remaining factors are grouped into a group-factor. In step five, the group-factor obtained in step four is tested. The test procedure is repeated until the analysis terminates with a test on a non-defective group-factor. Steps two onwards are carried out for all the group-factors found to be defective in step one. In brief, the test procedure consists of testing the group-factors and the factors within the group-factors found to be defective, one by one till a defective factor is detected by several steps alternately.

The main objective of group-screening is to reduce the number of tests or observations by eliminating a large number of non-defective factors in a bunch thus reducing the cost of the experiment.

1.2 BRIEF REVIEW OF LITERATURE ON GROUP SCREENING DESIGNS

The concept of testing items in groups and testing individual items only if the group test is positive was first introduced by Dorfman [3] in 1943 as an economical method of testing blood samples of army inductees in order to detect the presence of infection. Dorfman proposed that rather than test each blood sample individually, portions of each of the samples could be pooled together and the pooled sample tested first. If the pooled sample was free of infection, all the inductees in the sample could be passed with no further tests. Otherwise the remaining portion of each of the blood samples would be tested individually. If the prevalence of infection were low, the expected total number of tests and thus the expected total cost of inspection, would be reduced.

The work of Dorfman was carried further by Sterret [26] in 1957. In Sterret's screening plan, individual items from a defective pooled sample were tested one at a time until a defective item was found. The remaining items from the defective pooled sample were again tested in a pool. If the result was negative, then the work was complete for that pooled sample. Otherwise testing items individually was continued until another defective item was found. The remaining items were again tested in a pool. The process was continued until all the defective items in the defective pooled - sample were weeded out. The basic argument behind Sterret's plan was that since Dorfman's plan worked well for low prevalence rate of defective items, this low prevalence of defective items makes the chance of exactly one defective in a defective pooled sample high enough to warrant a pooled test once a defective item has been found. Sterret's plan reduced the number of runs obtained using Dorfman's plan by as much as eight per-cent for a prevalence rate of five per-cent.

Graff L.F. and Roeloffs, R. [6] in 1972 extended the work done by Dorfman [3] to the case when a test error was present. They defined the cost as a linear function of the number of runs, the number of defective factors classified as non-defective and the number of non-defective factors classified as defective.

Sobel and Groll [24] in 1959 devised a sequential sampling scheme which minimized the expected number of tests required to classify all the factors in a population as defective or non-defective. They discussed group-testing procedures which could be used efficiently for smaller populations. They assumed that factors represent the items in a sequence of indipendent Bernoulli trials with probability q and p = 1 - q of being non-defective and defective respectively . In another paper [25], they extended the Binomial sequential group testing to the case when the prevalence rate of defectives 'p' is unknown. To estimate p, they used the maximum likelihood estimation procedure and the Baye's procedure .

Connor [1] was the first person to approach the group-testing problem from the point of view of designs of experiments. This was later followed by Watson [28] in 1961. Watson studied two-stage group screening designs with and without errors in observations using equal size groups. For the case where there were errors in observations, he obtained expressions for the power of the tests in the two stages. Assuming continuous variations in group-sizes, he obtained the optimum group-sizes by minimizing the total expected number of runs (tests) with respect to the group-sizes using ordinary calculus techniques. He also worked out expressions for the expected number of defective factors declared non-defective and for the expected number of non-defective factors declared defective.

In a technical report submitted to the Research triangle institute in 1962, Patel [13] maximized the expected number of correct decisions by a proper choice of the sizes of the critical regions in a two stage group screening design and compared the value with that obtained by maximizing the same in a single stage design. In a note on Watson's paper, Patel [15] proved that the expected number of defective factors declared defective is a non-decreasing function of the level of significance in the first stage; thus removing the doubt which Watson had. Patel [14] in another paper, extended the two-stage group screening procedure to multistage group screening procedure when responses are observed with negligible error. He restricted his work to the case when all the factors were defective with the same a-priori probability. In the same paper, he discussed the question of the choice of the number of stages which should be used. He showed that the optimum number of stages depended on the prevalence rate of defective factors. In yet another paper, Patel [16] has shown that the factors to be included in an experiment before it is carried out should have a-priori probability of being defective different from $\frac{1}{2}$. Patel [18], has worked out the condition under which the expected number of correct decisions in a two-stage experiment is at least equal to that in a single stage experiment while at the same time keeping the expected number of runs in the former experiment fewer than the expected number of runs in the latter experiment. In paper [17], Patel gives caution on when a two stage group screening method should not be used.

Li [8] in 1962, developed multi-stage designs for screening experimental variables and obtained results similar to those obtained by Patel [14]. Considering the likely case where at each stage, a defective group-factor contains only one defective factor, he showed that for maximum screening efficiency, each stage in a multi-stage experiment must have the same number of tests.

In 1962, Thompson [27] used the group screening method and the method of maximum likelihood to estimate the proportion 'p' of vectors capable of transmitting auster - yellow virus in a natural population of macrosteles fascifrons (Stal) - the six spotted leafhopper. William, G. Hunter and Reiji Mezaki [29] in 1964 used a group screening method to select the best catalyst from a list of possible catalysts for the oxidation of methane. They stated that by arranging possible catalysts for a reaction in logical groups and testing each group in a single run, the less active catalysts can be weeded out and the total number of runs reduced.

Finucan [4] in 1964 considered a multi-stage group screening design without errors in observations and in which all factors are defective with the same a-priori probability. He suggested the method of finite differences in solving for the optimum group-sizes in a two-stage group-screening design.

Curnow [2] in a note on G.S. Watson's paper [28], points out an error in derivation of some probabilities by G.S. Watson. Kleijnen [7] has compared group screening designs with other types of factor screening designs. He investigates the assumptions made by Watson [28] and derives some new results on two-stage group-screening by allowing the possibility of two factor interactions.

Samuels [23] states that the expected number of runs in a two-stage group-screening design is not a unimodal function of group size. He however confirms Dorfman's results.

In 1974, Garey and Hwang [5] obtained the optimal group-testing procedure for isolating a single defective in a finite set of n items containing at least one defective. They considered the case when the probability of each item to be defective is known and used a binary testing tree.

Patel and Ottieno [20] in 1984 approached two stage group-screening designs with equal prior probabilities of factors to be defective and with no errors in observations from the point of view of discontinuous variation in the sizes of group-factors. They used the method of finite differences to obtain optimum group sizes and compared their results with Watson's results obtained by assuming continuous variation in the sizes of group-factors. In another paper, Patel and Ottieno [21] have extended Watson's paper [28] to the case when items have unequal a-priori probabilities of being defective. They have considered the case where there are no errors in observations. They have shown that in the case of group screening from a population with unequal a-priori probabilities, the number of observations needed on the average is considerably smaller than that required in the case of a population with factors having the same a-priori probability of being defective. In another paper [19], they obtained optimum two-stage group screening designs with errors in observations by considering both the expected total number of runs and the expected total number of incorrect decisions. Optimum group sizes were obtained by minimizing the expected total number of runs for a fixed value of the expected total number of incorrect decisions and vice versa. As an alternative method of obtaining optimum two stage group screening designs, they defined the expected total cost of screening as a linear function of the expected total number of incorrect decisions and the expected total number of runs and obtained the group size that minimizes the expected total cost.

Odhiambo [12] in 1981 studied group screening designs with three stages. He assumed that different factors were defective with (i) the same a-priori probabilities and (ii) unequal a-priori probabilities. For each of these cases, he considered screening with and without errors in observations. For the case when there are errors in observations, he used orthogonal fractional factorial designs of the type obtained by Plackett and Burman [22] to derive theoretical results. He also studied multi-stage group-screening designs without errors in observations and assuming that the factors have unequal a-priori probabilities of being defective.

Mauro and Smith [9] in 1982 have examined the performance of two stage group screening designs when the assumption that the direction of possible effects are known or are correctly assumed a-priori is relaxed. The case of zero error variance is considered. They assumed that for all defective factors, the magnitude of the effect is the same but the direction of the effects could be different. To gauge the effect of cancellation, they define a percentage measure of efficiency of a screening stratefy for detecting the active factors. They also define the relative testing cost as another measure of screening efficiency. Mauro [10] in 1984 extended this work to the case when there are errors in observations. In their paper [11], Mauro and Burns have compared random balance screening strategy with two stage group screening designs. A screening model in which the effects of defective factors are additive is assumed. They found that the optimal group-screening strategy is generally better than the optimal random balance strategy at low type I error rates but begins to lose its advantage at higher type I error rates.

1.3. ASSUMPTIONS

The assumptions made in this thesis are essentially those made by Watson [28], later modified by Patel and Ottieno [19], [20] and [21]. When screening with equal a-priori probabilities, the assumptions are as follows:-

(1) All factors have, indipendently, the same a-priori probability 'p' of being defective.

(2) Defective factors have the same positive effect Δ.

(3) None of the factors interact.

(4) The required designs exist.

(5) The directions of possible effects are known.

(6) The errors of all observations are independently normal with a constant known variance, σ^2.

(7) The total number of factors is '$f = kg$', where 'g' is the number of group-factors in the initial step and 'k' is the number of factors in each of the group-factors.

When screening with unequal a-priori probabilities, the assumptions have been modified as follows:-

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

(ii) $p_i>0$, $i=1,2,..,g$, is taken as the probability that a factor in the i^{th} group-factor in the initial step is defective.

(iii) $\Delta_i > 0$, $i=1,2,\ldots,g$, is the effect of a factor within the i^{th} group-factor in the initial step.

(iv) None of the factors interact.

(v) The directions of possible effects are known.

(vi) The required designs exist.

(vii) The errors of all observations are independently normal with a constant known variance σ^2.

(viii) α_{Ii} is the level of significance for testing the i^{th} group-factor in the initial step and α_{si} is the level of significance for testing the factors within the i^{th} group-factor which is declared defective in the initial step ($i=1,2,\ldots,g$).

(ix) α_i^* is the probability that a group-factor consisting of factors from the i^{th} step group-factor is declared defective but on testing the individual factors, no factor is declared defective due to errors in observations.

Non orthogonal fractional factorial designs are used in this thesis when there are no errors in observations, whereas orthogonal fractional factorial designs of the type given by Plackett and Burman [22] are used when errors in observations are allowed. In the case of screening with errors in observations, the expression for the power of the test in the initial step has been obtained. Optimum group-sizes have been obtained using calculus methods.

CHAPTER II

STEP-WISE GROUP SCREENING DESIGNS WITH EQUAL A-PRIORI PROBABILITIES

2.1. SCREENING WITHOUT ERRORS

Let there be 'f' factors under investigation. The problem is to isolate defective factors with minimum number of observations (also called runs). With this objective in view, we first divide the 'f' factors into 'g' group-factors in step one. If each group-factor has k factors, then

$$f = kg \qquad (2.1.1).$$

The group-factors are then tested for their significance by an experiment consisting of (g + 1) runs. Those that are found to be non-defective are set aside. In step two, we start with any defective group-factor and test the factors within it one by one till we find a defective factor. We set aside the factors which are found to be non defective, keeping the defective factor separate. The remaining factors are then grouped into a group-factor. In step three, we test the group-factor obtained after step two is performed. If the group-factor is non-defective, we terminate the test procedure. If the group-factor is defective, we continue with step four. In step four, factors within a group-factor found to be defective in step three are tested one by one till a defective factor is found. Factors which are found to be non-defective are again set aside keeping the defective factor separate. The remaining factors are grouped into a group-factor. In step five, the group-factor obtained in step four is tested.

The test procedure is repeated until the analysis terminates with a test on a negative (non-defective) group-factor . The procedure will certainly terminate in a finite number of steps. If the probability of a factor to be defective is small, the probability of exactly one defective factor of a positive (defective) group-factor is high enough to warrant a group analysis once a defective factor is found. Steps two onwards are carried out for all the group-factors found to be defective in step one. This procedure differs from the procedure first introduced by Sterret [26] in that in the first step, the g group-factors are tested in a factorial experiment with (g + 1) runs.

Alternatively, if we use the control run used in step one in the subsequent steps, then steps two onwards could be performed in a series of experiments as follows:- In step two, we take one factor from each group-factor found to be defective in step one. The factors are then tested for their significance by an experiment. If no defective factor is observed, we take another set of factors one from each group-factor and test their significance . We repeat this procedure till at least one defective factor is observed. The non-defective factors are set aside, keeping the defective factor(s) separate. The remaining factors from a group-factor that contained a defective factor are set aside and grouped into a group-factor. This process is repeated until one defective factor from each group-factor found to be defective in the initial step has been isolated.

In the third step, the group-factors set aside in step two are tested in an experiment using the control test used in step one . Again the group-factors found to be non-defective in step three are set aside . In the fourth step, we proceed with a series of experiments as in step two until we isolate one defective factor from each group-factor found to be defective in step three. The remaining factors from each of the group-factors found to be defective in step three are grouped into a group-factor after step four is performed. Again the group-factors set aside in step four are tested in an experiment in step five. This procedure is repeated until the analysis terminates with all negative (non-defective) group-factors when all the defective factors have been isolated. Both these test procedures are equivalent; but when errors in the observations are allowed, it is convenient to use the alternative procedure to derive theoretical results. In brief, the test procedure consists of testing the group-factors and the factors within the group-factors found to be defective, one by one till a defective factor is detected by several steps alternately.

2.1.1 The expected number of runs

Let 'p' be the a-priori probability that a factor is defective. A group-factor is defective if it contains at least one defective factor. Let p^* be the probability that a group-factor in step one is defective. If j is the number of defective factors in such a group-factor, then

$$p^* = \sum_{j=1}^{k} \binom{k}{j} p^j q^{k-j}$$

$$= 1 - q^k \quad (2.1.2)$$

where

$$q = 1 - p \quad (2.1.3).$$

In the initial step, all the g group-factors are tested for significance. Thus the number of tests (runs) required in the initial step is given by

$$R_I = g + 1 \quad (2.1.4)$$

where the one extra test is the control test. This control test is used as a control test for the subsequent steps. Let r be the number of defective group-factors in the first step. Then the probability distribution of r is given by

$$f(r) = \begin{cases} \binom{g}{r} (p^*)^r (1 - p^*)^{g-r} & r = 0,1,2,..,g \\ 0 & \text{Otherwise} \end{cases} \quad (2.1.5).$$

Thus

$$E(r) = gp^*$$

$$= \frac{f}{k} (1 - q^k) \quad (2.1.6).$$

In the subsequent steps, the analysis of the r group-factors found to be defective in the initial step is continued as described in the earlier part of section 2.1. Let $P_k(j)$ denote the probability that a group-factor of size k contains exactly j defective factors if it is known to contain at least one defective factor. Then

$$P_k(j) = (1 - q^k)^{-1} \binom{k}{j} p^j (1 - p)^{k-j}$$

$$j=1,2,..,k \qquad (2.1.7).$$

Let $E_k(R_j)$ be the expected number of tests (runs) required to analyse a group-factor i.e., classify as defective or non-defective all the factors within a group-factor of size k which is known to be defective if it contains exactly j defective factors. To obtain an expression for $E_k(R_j)$, we start by considering a sequence of lemmas.

Lemma 2.1.1

$$E_k(R_1) = \frac{k}{2} + 1 + \frac{1}{2} - \frac{2}{k}$$

Proof

It is equally likely that the defective factor is found at any trial. Consequently the probability that it is found on any one trial is $\frac{1}{k}$. If the defective factor is found at the ℓ^{th} trial; $\ell=1,2,..,k-1$, then ℓ tests are needed to find it. The other test we need is the group test on a group-factor consisting of $(k-\ell)$ factors. This group-factor is non-defective if $j=1$. If the first $(k-1)$ factors tested are non-defective, then the k^{th} factor is the defective one. We need not test this factor since the initial group-factor of size k is known to contain at least one defective factor. Thus

$$E_k(R_1) = \frac{1}{k} \sum_{\ell=1}^{k-1} (\ell+1) + \frac{1}{k}(k-1) \qquad (2.1.8).$$

Simplifying (2.1.8) we obtain

$$E_k(R_1) = \frac{1}{k}\,\frac{(k-1)(k+2)}{2} + \frac{k-1}{k}$$

$$= \frac{k}{2} + 1 + \frac{1}{2} - \frac{2}{k} \qquad (2.1.9).$$

This completes the proof of the lemma.

Lemma 2.1.2

$$E_k(R_2) = \frac{2k}{3} + 2 + \frac{2}{3} - \frac{4}{k}$$

Proof

In this case, the approach is to find the first defective factor and thus reduce the problem to the one in which the group-factor has only one defective factor. This problem of a group-factor having only one defective factor was considered in lemma 2.1.1. The probability that the first factor tested is defective is $\frac{2}{k}$. If the first factor tested is defective, then on the average we require $\{1+1+E_{k-1}(R_1)\}$ tests to complete the test procedure. For $\ell=1,2,..,k-2$, the probability that the $(\ell+1)^{st}$ factor tested is the first defective factor to be found is $\prod_{w=1}^{\ell}\left(\frac{k-(w+1)}{k-(w-1)}\right)\frac{2}{k-\ell}$ and on the average, the number of runs required to complete the test procedure in this case is $\{(\ell+1)+1+E_{k-(\ell+1)}(R_1)\}$.

Hence

$$E_k(R_2) = \frac{2}{k}\{1+1+E_{k-1}(R_1)\}$$

$$+ \sum_{\ell=1}^{k-2}\left[\prod_{w=1}^{\ell}\left(\frac{k-(w+1)}{k-(w-1)}\right)\frac{2}{k-\ell}\{(\ell+1)+1+E_{k-(\ell+1)}(R_1)\}\right]$$

$$(2.1.10).$$

Rewriting (2.1.10) and using (2.1.9), it follows that

$$E_k(R_2) = \frac{2}{k}\ \frac{k-1}{k-1}\ \{2 + \frac{k-1}{2} + \frac{3}{2} - \frac{2}{k-1}\}$$

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

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

$$+ \ .\ .\ .\ .\ .\ .\ .\ .$$

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

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

$$= \frac{2}{k(k-1)} \sum_{m=1}^{k-1} (m+1)(k-m) + \frac{1}{k(k-1)} \sum_{m=1}^{k-1} (k-m)^2$$

$$+ \frac{3}{k(k-1)} \sum_{m=1}^{k-1} (k-m) - \frac{4(k-1)}{k(k-1)}$$

$$= \frac{2}{k(k-1)} \left[\frac{(k+1)k(k-1)}{6} + \frac{k(k-1)}{2}\right] + \frac{1}{k(k-1)} \left[\frac{k(k-1)(2k-1)}{6}\right]$$

$$+ \frac{3}{k(k-1)}\left[\frac{k(k-1)}{2}\right]. - \frac{4}{k}$$

i.e.,

$$E_k(R_2) = \frac{2k}{3} + 2 + \frac{2}{3} - \frac{4}{k} \qquad (2.1.11).$$

This proves the lemma.

<u>Lemma 2.1.3</u>

$$E_k(R_3) = \frac{3k}{4} + 3 + \frac{3}{4} - \frac{6}{4}$$

<u>Proof</u>

After finding one defective factor, the problem reduces to that considered in lemma 2.1.2. The probability that the first factor tested is defective is $\frac{3}{k}$ and the probability that for $\ell=1,2,\ldots,k-3$ the $(\ell+1)^{st}$ factor tested is the first defective is $\prod_{w=1}^{\ell}\left(\frac{k-(w+2)}{k-(w-1)}\right)\frac{3}{k-\ell}$. If the first factor tested is defective, then on the average we need $\{1+1+E_{k-1}(R_2)\}$ tests to complete the test procedure.

However, if for $\ell=1,2,3,\ldots,k-3$, the $(\ell+1)^{st}$ factor tested is the first defective, then on the average we shall need $\{(\ell+1)+1+E_{k-(\ell+1)}(R_2)\}$ runs to complete the test procedure.

Thus

$$E_k(R_3) = \frac{3}{k}\{1+1+E_{k-1}(R_2)\} + \sum_{\ell=1}^{k-3}\left[\prod_{w=1}^{\ell}\left(\frac{k-(w+2)}{k-(w-1)}\right)\frac{3}{k-\ell}\{(\ell+1)+1+E_{k-(\ell+1)}(R_2)\}\right] \qquad (2.1.12).$$

Using (2.1.11) we get

$$E_k(R_3) = \frac{3}{k}\,\frac{k-1}{k-1}\,\frac{k-2}{k-2}\{2+\frac{2(k-1)}{3}+\frac{8}{3}-\frac{4}{k-1}\}$$

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

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

$$+\cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot$$

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

$$+\frac{3}{k}\,\frac{1}{k-1}\,\frac{2}{k-2}\{(k-1)+\frac{2(2)}{3}+\frac{8}{3}-\frac{4}{2}\}.$$

$$= \frac{3}{k(k-1)(k-2)}\sum_{m=1}^{k-2}(m+1)(k-m)(k-m-1)$$

$$+ \frac{2}{k(k-1)(k-2)} \sum_{m=1}^{k-2} (k-m)^2 (k-m-1)$$

$$+ \frac{8}{k(k-1)(k-2)} \sum_{m=1}^{k-2} (k-m)(k-m-1)$$

$$- \frac{12}{k(k-1)(k-2)} \sum_{m=1}^{k-2} (k-m-1)$$

(2.1.13).

The summations

$$\sum_{m=1}^{k-2} m(k-m)(k-m-1) = \frac{(k+1)k(k-1)(k-2)}{12},$$

$$\sum_{m=1}^{k-2} (k-m)^2 (k-m-1) = \frac{k(k-1)(k-2)(3k-1)}{12}$$

(2.1.14)

and

$$\sum_{m=1}^{k-2} (k-m)(k-m-1) = \frac{k(k-1)(k-2)}{3}$$

have been obtained by Sterret [26].

Using these summations in (2.1.13) above, we get

$$E_k(R_3) = \frac{k+1}{4} + 1 + \frac{3k-1}{6} + \frac{8}{3} - \frac{6}{k}$$

$$= \frac{3k}{4} + 3 + \frac{3}{4} - \frac{6}{k} \qquad (2.1.15)$$

This completes the proof of the lemma.

We are now in a position to state and prove a more general result.

Theorem 2.1.1

The average number of tests required to analyse a defective group-factor of size k assuming that it contains exactly j defective factors is given by

$$E_k(R_j) = \frac{jk}{j+1} + j + \frac{j}{j+1} - \frac{2j}{k} \qquad (j=1,2,\ldots,k).$$

Proof

The proof follows by mathematical induction. The validity of the Theorem has been shown for j=1. We assume that the Theorem is true for j=n-1, $(1\le n-1\le k)$ that is

$$E_k(R_{n-1}) = \frac{(n-1)k}{n} + (n-1) + \frac{n-1}{n} - \frac{2(n-1)}{k}, \qquad (2.1.16).$$

We shall show that the Theorem is true for j=n. Now for j=n ,

$$E_k(R_n) = \frac{n}{k}\{1+1+E_{k-1}(R_{n-1})\} + \sum_{\ell=1}^{k-n}\left[\prod_{w=1}^{\ell}\left(\frac{k-(w+n-1)}{k-(w-1)}\right)\frac{n}{k-\ell}\{(\ell+1)+1 + E_{k-(\ell+1)}(R_{n-1})\}\right] \qquad (2.1.17).$$

The factor $\frac{n}{k}$ in the first term is the probability that the first factor tested is defective and $\{1 + 1 + E_{k-1}(R_{n-1})\}$ is the average number of runs required to perform the analysis if the first factor is defective. The value $\prod_{w=1}^{\ell}\left(\frac{k-(w+n-1)}{k-(w-1)}\right)$ is the probability that the first ℓ factors tested are non-defective; $\frac{n}{k-\ell}$ is the probability that the $(\ell+1)^{st}$ factor tested is defective. The term $\{(\ell+1)+1+E_{k-(\ell+1)}(R_{n-1})\}$ consists of the number of tests required to find the first defective factor on the $(\ell+1)^{st}$ trial, the group test on $k-(\ell+1)$ factors and the average number of tests required to complete the analysis with (n-1) defective factors in $k-(\ell+1)$ factors.

Substituting in (2.1.17) the values given in (2.1.16) we obtain

$$E_k(R_n) = \frac{n}{k}\,\frac{k-1}{k-1} \cdots \frac{k-(n-1)}{k-(n-1)}\left[2+\frac{n-1}{n}(k-1)+(n-1)+\frac{n-1}{n} - \frac{2(n-1)}{k-1}\right]$$

$$+ \frac{k-n}{k}\,\frac{n}{k-1}\,\frac{k-2}{k-2}\cdots\frac{k-(n-1)}{k-(n-1)}\left[3+\frac{n-1}{n}(k-2)+(n-1)+\frac{n-1}{n} - \frac{2(n-1)}{k-2}\right]$$

$$+ \frac{k-n}{k}\,\frac{k-n-1}{k-1}\,\frac{n}{k-2}\,\frac{k-3}{k-3}\cdots\frac{k-(n-1)}{k-(n-1)}\left[4+\frac{n-1}{n}(k-3)+(n-1) + \frac{n-1}{n} - \frac{2(n-1)}{k-3}\right]$$

$$+ \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot$$

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

$$\times\left[(k-n+1) + \frac{n-1}{n}(n) + (n-1) + \frac{n-1}{n} - \frac{2(n-1)}{n}\right]$$

$$+ \frac{k-n}{k}\,\frac{k-(n-1)}{k-1}\cdots\frac{k-(k-1)}{k-(k-n-1)}\,\frac{n}{k-(k-n)}\,\frac{k-(k-n+1)}{k-(k-n+1)}\cdots\frac{k-(n-1)}{k-(n-1)}$$

$$\times\left[(k-n+2) + \frac{n-1}{n}(n-1) + (n-1) + \frac{n-1}{n} - \frac{2(n-1)}{n-1}\right] \tag{2.1.18}$$

By rearranging and taking appropriate summations (2.1.18) becomes

$$E_k(R_n) = \frac{n}{k^P n}\sum_{m=1}^{k-n+1}(m+1)(k-m)(k-m-1)\cdots(k-m-n+2)$$

$$+ \frac{n-1}{k^P n}\sum_{m=1}^{k-n+1}(k-m)^2(k-m-1)\cdots(k-m-n+2)$$

$$+\frac{n^2-1}{{}_kP_n}\sum_{m=1}^{k-n+1}(k-m)(k-m-1)\cdots(k-m-n+2)$$

$$-\frac{2n(n-1)}{{}_kP_n}\sum_{m=1}^{k-n+1}(k-m-1)(k-m-2)\cdots(k-m-n+2) \qquad (2.1.19)$$

where

$${}_kP_n = k(k-1)(k-2)\cdots(k-n+1) \qquad (2.1.20).$$

The summations

$$\sum_{m=1}^{k-n+1} m(k-m)(k-m-1)\cdots(k-m-n+2) = \frac{(k+1)^P(n+1)}{n(n+1)},$$

$$\sum_{m=1}^{k-n+1}(k-m)^2(k-m-1)(k-m-2)\cdots(k-m-n+2) = \frac{(k+1)^P(n+1)}{(n+1)} - \frac{{}_kP_n}{n}$$

$$\sum_{m=1}^{k-n+1}(k-m)(k-m-1)\cdots(k-m-n+2) = \frac{{}_kP_n}{n} \qquad (2.1.21)$$

and

$$\sum_{m=1}^{k-n+1}(k-m-1)(k-m-2)\cdots(k-m-n+2) = \frac{(k-1)^P(n-1)}{n-1}$$

have been determined by Sterret [26].

Using these summations in (2.1.19) above we obtain

$$E_k(R_n) = \frac{k+1}{n+1} + 1 + \frac{(n-1)(k+1)}{n+1} - \frac{n-1}{n} + \frac{n^2-1}{n} - \frac{2n(n-1)}{(n-1)k}$$

$$= \frac{nk}{n+1} + n + \frac{n}{n+1} - \frac{2n}{k} \qquad (2.1.22)$$

This is exactly the value of $E_k(R_j)$ for $j=n$. Thus if the Theorem is true for $j=n-1$ $0\leq n-1\leq k)$ it is also true for $j=n$. But the Theorem is true for $j=1$ (c.f. lemma 2.1.1). Hence it is true for $j=2$ and in general for any j $(j=1,2,..,k)$. This completes the proof of the Theorem.

Let R_S^o denote the number of tests required to analyse a group-factor i.e., classify as defective or non-defective all the factors within a group-factor of size k that is known to be defective. Then

$$E(R_S^o) = \sum_{j=1}^{k} E_k(R_j)\, P_k(j) \qquad (2.1.23)$$

where $P_k(j)$ is as defined in (2.1.7).

Using (2.1.7) and Theorem 2.1.1 in (2.1.23) we get

$$E(R_S^o) = \sum_{j=1}^{k} \left\{ \frac{j(k+1)}{j+1} + j - \frac{2j}{k} \right\} \frac{1}{1-q^k} \binom{k}{j} p^j q^{k-j}$$

$$= \frac{1}{1-q^k} \{(k+1)(1-q^k)+kp-2p\} - \frac{k+1}{1-q^k} \sum_{j=1}^{k} \frac{1}{j+1} \binom{k}{j} p^j q^{k-j}$$

(2.1.24)

Next

$$(k+1) \sum_{j=1}^{k} \frac{1}{j+1} \binom{k}{j} p^j q^{k-j}$$

$$= (k+1)\left\{\frac{k}{2}\, pq^{k-1} + \frac{k(k-1)}{3\times 2}\, p^2 q^{k-2} + \cdots + \frac{1}{k+1}\, p^k\right\}$$

$$= \frac{1}{p}\left\{\frac{(k+1)k}{2!}\, p^2 q^{k-1} + \frac{(k+1)k(k-1)}{3!}\, p^3 q^{k-2} + \cdots + p^{k+1}\right\}$$

$$= \frac{1}{p}\{1-q^{k+1} - (k+1)pq^k\} \qquad (2.1.25).$$

Using (2.1.25) in (2.1.24) we obtain

$$E(R_S^o) = \frac{1}{1-q^k}\left[(k+1)(1-q^k)+kp-2p-\frac{1}{p}\{1-q^{k+1}-(k+1)pq^k\}\right]$$

$$= \frac{1}{1-q^k}\left[(k+1) + kp - 2p - \frac{1}{p}\{1-q^{k+1}\}\right] \qquad (2.1.26)$$

Let R_S denote the number of tests (runs) required to analyse all the factors in the r group-factors found to be defective in the first step. Then,

$$R_S = r\,E(R_S^o)$$
$$= \frac{r}{1-q^k}\left[(k+1) + kp - 2p - \frac{1}{p}\{1-q^{k+1}\}\right] \qquad (2.1.27).$$

Further let R be the total number of runs required to investigate the f factors. Then

$$R = R_I + R_S \qquad (2.1.28)$$

Theorem 2.1.2

Let R denote the total number of runs required to screen out the defective factors from among the 'f' factors under investigation in a step-wise group screening experiment where p is a-prior probability of any factor being defective and k is the size of the group-factor in the initial step, then

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

where

$$q = 1 - p.$$

Proof

In the first step, we have $g = \frac{f}{k}$ group-factors to test. Therefore the number of runs required in step one is

$$R_I = g + 1 \qquad (2.1.29)$$

the one extra test being the control test. The number of runs in the subsequent steps is

$$R_S = \frac{r}{1-q^k}\left[(k+1) + kp - 2p - \frac{1}{p}\{1-q^{k+1}\}\right] \qquad \text{(c.f. 2.1.27)},$$

where r is the number of group-factors found to be defective in step one. Then

$$E(R_S) = \left[(k+1) + kp - 2p - \frac{1}{p}\{1-q^{k+1}\}\right]\frac{E(r)}{1-q^k}$$

$$= \left[(k+1) + kp - 2p - \frac{1}{p}\{1-q^{k+1}\}\right]\frac{f}{k} \qquad (2.1.30)$$

using (2.1.6) •

The expected total number of runs is given by

$$E(R) = R_I + E(R_S) \qquad (2.1.31)$$

Using (2.1.29) and (2.1.30) we obtain

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

$$= 1 + fp + \frac{2fq}{k} + f - \frac{f}{kp}\{1 - q^{k+1}\} \qquad (2.1.32)$$

This completes the proof of Theorem 2.1.2.

<u>Corollary 2.1.1</u>

For small values of p, the expected total number of runs is given by

$$E(R) \simeq 1 + \frac{3fp}{2} + \frac{f}{k} - \frac{2fp}{k} + \frac{fkp}{2}$$

upto order p.

<u>Proof</u>

For small values of p,

$$\frac{f}{kp}\left[1 - (1-p)^{k+1}\right] \simeq \frac{f(k+1)}{k} - \frac{f(k+1)p}{2}, \quad \text{upto order } p \qquad (2.1.33)$$

Substituting this expression in (2.1.32), we get

$$E(R) \simeq 1 + fp + \frac{2fq}{k} - \frac{f}{k} + \frac{fkp}{2} + \frac{fp}{2}$$

$$= 1 + \frac{3fp}{2} + \frac{f}{k} - \frac{2fp}{k} + \frac{fkp}{2} \quad \text{upto order } p \qquad (2.1.34).$$

This completes the proof of corollary 2.1.1.

2.1.2. The Optimum size of the group-factor in the initial step

Theorem 2.1.3

Assuming p i.e. a-priori probability of a factor to be defective to be small, the size 'k' of the group-factor which minimizes the expected total number of runs in a step-wise group screening design is given by

$$k \simeq \left(\frac{2 - 4p}{p}\right)^{\frac{1}{2}}$$

provided $p<\frac{1}{2}$. The corresponding minimum expected total number of runs is given by

$$\text{Min } E(R) \simeq 1 + \frac{3fp}{2} + f(2p)^{\frac{1}{2}}(1 - 2p)^{\frac{1}{2}}.$$

Proof

Assuming continuous variation in k, the optimum group size is obtained by solving the equation

$$\frac{d}{dk} E(R) = 0$$

where E(R) is as given in corollary 2.1.1.

This implies

$$\frac{2p}{k^2} - \frac{1}{k^2} + \frac{p}{2} = 0,$$

i.e.,

$$4p - 2 + k^2 p = 0,$$

which gives

$$k^2 = \frac{2 - 4p}{p}$$

or

$$k = \left(\frac{2 - 4p}{p}\right)^{\frac{1}{2}} \qquad (2.1.35),$$

provided $p<\frac{1}{2}$.

The value of k given in (2.1.35) will be in the neighbourhood of the point of minimum of E(R) if

$$\frac{d^2}{dk^2} E(R) > 0 .$$

i.e. if

$$\frac{-4fp}{k^3} + \frac{2f}{k^3} > 0 .$$

which is true if $p<\frac{1}{2}$. Thus the value of k given in (2.1.35) is in the neighbourhood of the point of minimum of E(R). Substituting this value of k in the expression for E(R) given in corollary 2.1.1, we obtain

$$\min E(R) \simeq 1 + \frac{3fp}{2} + f\left(\frac{p}{2-4p}\right)^{\frac{1}{2}} - 2fp\left(\frac{p}{2-4p}\right)^{\frac{1}{2}} + \frac{f}{2}(2-4p)^{\frac{1}{2}}p^{\frac{1}{2}} \qquad (2.1.36a)$$

$$= 1 + \frac{3fp}{2} + f(2p)^{\frac{1}{2}}(1-2p)^{\frac{1}{2}} \qquad (2.1.36b).$$

This completes the proof of Theorem 2.1.3.

Next we wish to obtain the value of k that minimizes E(R) for arbitrary values of p. For arbitrary values of p,

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

c.f.(2.1.32).

The value of k that minimizes E(R) in (2.1.32) is a solution of the equation

$$\frac{d}{dk} E(R) = 0$$

i.e.,

$$\frac{-2q}{k^2} + \frac{1}{pk^2}\{1 - q^{k+1}\} + \frac{1}{kp} q^{k+1} \ln q = 0$$

which implies

$$1 - q^{k+1} - 2pq + kq^{k+1} \ln q = 0 \qquad (2.1.37)$$

Equation (2.1.37) is non linear in k and can be solved by Newton - Raphson iterative method. Let the initial approximation be the value of k obtained in (2.1.35). That is

$$k^o = \left(\frac{2 - 4p}{p}\right)^{\frac{1}{2}} \qquad (2.1.38).$$

Let us denote the left hand side of equation (2.1.37) by $\psi(k)$. Then the next better approximation of optimum k is given by

$$k = k^o - \frac{\psi(k^o)}{\psi'(k^o)} \qquad (2.1.39)$$

where

$$\psi'(k) = kq^{k+1}(\ln q)^2 \qquad (2.1.40)$$

The iterations may be continued until the desired level of accuracy is attained.

In the next Theorem, we give a sufficient condition for a step-wise design to be more efficient than a single stage design.

Definition We shall say that one design is more efficient than another if the expected number of runs in one is less than or equal to that in the other for all p $(0<p<1)$ with strict inequality holding true for at least one value of 'p' i.e. the probability of a factor to be defective.

Theorem 2.1.4

A step-wise group screening design with 'f' factors and 'g' group-factors in the initial step, each group-factor of size $k = \left(\frac{2 - 4p}{p}\right)^{\frac{1}{2}}$ where p is the prior probability of a factor being defective, assumed to be small, is more efficient than the corresponding single stage design.

Proof

The Theorem is true if

$$1 + \frac{3fp}{2} + f(2p)^{\frac{1}{2}}(1 - 2p)^{\frac{1}{2}} \leq f + 1 \qquad (2.1.41)$$

where the left hand side is the minimum expected number of runs in a step-wise group screening design as given in (2.1.36b).

Inequality (2.1.41) is true if

$$\left(1 - \frac{3p}{2}\right)^2 \geq 2p(1 - 2p)$$

i.e. if

$$(2 - 5p)^2 \geq 0 \qquad (2.1.42)$$

The inequality in (2.1.42) is strict for all values of p, $0<p<1$ with equality holding for $p=0.4$.

This proves theorem 2.1.4.

Since p is assumed to be small and the left hand side of inequality (2.1.41) holds for $p<\frac{1}{2}$, we consider only values of p for which

$$p<0.4 \qquad (2.1.43)$$

Using the fact that optimum value of k decreases as p increases and that the expected number of runs increases as p increases, one is tempted to argue that the maximum value of p for which a step-wise group-screening design is better than a corresponding single stage design can be obtained by putting k=2 and solving for p the inequality

$$fp + \frac{2fq}{k} + f - \frac{f}{kp}\left[1 - q^{k+1}\right] + 1 \leq f + 1 \qquad (2.1.44)$$

where the left hand side of (2.1.44) represents the expected total number of runs in a step-wise group screening design and the right hand side represents the number of runs in a single stage design. The inequality (2.1.44) is true if

$$- p + 3p^2 - p^3 \leq 0$$

i.e. if

$$p(p^2 - 3p + 1) \geq 0.$$

Solving the equation

$$p^2 - 3p + 1 = 0,$$

we obtain

$$p = \frac{3}{2} \pm \frac{\sqrt{5}}{2}$$

i.e.,

$$p = 1.5 - 1.118$$

$$= 0.382 \qquad \text{(since } p<1\text{)}.$$

Thus inequality (2.1.44) implies that

$$p<0.382. \qquad (2.1.45)$$

<u>Comment:</u>

Although the result obtained above agrees with that in (2.1.43), the argument is not generally correct.

2.1.3 A comparison of two stage group screening design with the step-wise group screening design

Let there be 'f' factors to be tested. The f factors are divided into 'g' group-factors of k factors each. Let 'p' be the a-priori probability that a factor is defective. In the two stage group screening procedure, each of the group-factors is tested for significance in the first stage. In the second stage, all the factors within the defective group-factors are tested. The probability that a group-factor is defective p* is given by

$$p^* = 1 - q^k \qquad \text{where } q = 1 - p \qquad (2.1.46).$$

The expected total number of runs required to test the f factors using this procedure is given by

$$E(R) = 1 + \frac{f}{k} + f(1 - q^k) \qquad \text{c.f. Watson } [28] \qquad (2.1.47).$$

Patel and Ottieno [20] have given the value of k that minimizes E(R) in (2.1.47) as

$$k \simeq \frac{1}{\sqrt{p}} + \frac{\sqrt{p}}{4} + \frac{p}{4} \qquad (2.1.48),$$

upto order p.

They gave the corresponding minimum value of E(R) as

$$\text{Min } E(R) \simeq 1 + 2fp^{\frac{1}{2}} - \frac{fp}{2} + \frac{2}{3}fp^{3/2} - \frac{19}{24}fp^2 \qquad (2.1.49),$$

upto order p^2.

The size 'k' of the group-factor in the initial step which minimizes the expected total number of runs in a step-wise group screening design is approximated as

$$k \approx \left(\frac{2-4p}{p}\right)^{\frac{1}{2}} \qquad \text{(c.f. (2.1.35))}$$

$$\approx \left(\frac{2}{p}\right)^{\frac{1}{2}} (1-p) \qquad \text{upto order } p \qquad (2.1.50).$$

For a step-wise group screening design,

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

(c.f. (2.1.32)).

$$\approx 1 + fp + \frac{2f(1-p)}{k} - \frac{f}{k} + \frac{f(k+1)}{2}p$$

$$- \frac{f(k^2-1)}{6}p^2 + f\frac{(k^2-1)(k-2)p^3}{24} \qquad (2.1.51),$$

upto order p^3.

Substituting the value of k in (2.1.50) in (2.1.51), we obtain

$$\text{Min } E(R) \approx 1 + fp + 2f\left(\frac{p}{2}\right)^{\frac{1}{2}} - f\left(\frac{p}{2}\right)^{\frac{1}{2}} (1-p)^{-1}$$

$$+ \frac{f}{2}\left[\left(\frac{2}{p}\right)^{\frac{1}{2}} (1-p) + 1\right]p - \frac{f\left[\frac{2}{p}(1-p)^2 - 1\right]}{6}$$

$$+ \frac{f}{24}\left[\frac{2}{p}(1-p)^2 - 1\right]\left[\left(\frac{2}{p}\right)^{\frac{1}{2}}(1-p) - 2\right]p^3$$

$$\approx 1 + 2f\left(\frac{p}{2}\right)^{\frac{1}{2}} + \frac{7}{6}fp - \frac{11\sqrt{2}}{12}fp^{3/2} + \frac{1}{2}fp^2 \qquad (2.1.52),$$

upto order p^2.

<u>Theorem 2.1.5</u>

The step-wise group screening design is more efficient than the corresponding two stage group screening design assuming p, i.e. the a-priori probability of a factor to be defective to be small if

$p \leq 0.26$.

Proof:

We have to show that if $p \leq 0.26$, then min E(R) obtained using the step-wise group screening procedure is less than or equal to min E(R) obtained using the two stage group screening procedure. That is we show that if $p \leq 0.26$,

$$1+2f\left(\frac{p}{2}\right)^{\frac{1}{2}} + \frac{7}{6}fp - \frac{11\sqrt{2}}{12}fp^{3/2} + \frac{1}{2}fp^{2}$$

$$\leq 1 + 2fp^{\frac{1}{2}} - \frac{fp}{2} + \frac{2}{3}fp^{3/2} - \frac{19}{24}fp^{2}$$

i.e. we show that if $p \leq 0.26$

$$2^{\frac{1}{2}} + \frac{7}{6}p^{\frac{1}{2}} - \frac{11\sqrt{2}}{12}p \leq 2 - \frac{1}{2}p^{\frac{1}{2}} + \frac{2}{3}p - \frac{31}{24}p^{3/2} \qquad (2.1.53).$$

But for $p \leq 0.26$,

$$\frac{2}{3}p - \frac{31}{24}p^{3/2} > 0 \qquad (2.1.54a)$$

Thus (2.1.53) is true if

$$2^{\frac{1}{2}} + \frac{7}{6}p^{\frac{1}{2}} - \frac{11\sqrt{2}}{12}p \leq 2 - \frac{1}{2}p^{\frac{1}{2}}$$

i.e. if

$$(2^{\frac{1}{2}} - 2) + \frac{5}{3}p^{\frac{1}{2}} - \frac{11\sqrt{2}}{12}p \leq 0 \qquad (2.1.54b).$$

The equation

$$(2^{\frac{1}{2}} - 2) + \frac{5}{3}p^{\frac{1}{2}} - \frac{11\sqrt{2}}{12}p = 0 \qquad (2.1.55),$$

has no real root. The left hand side of (2.1.55) is less than zero when $p=1$ and when $p=0$. Therefore inequality

(2.1.54b) is a strict inequality for all values of p $(0<p<1)$.

Thus inequality (2.1.53) holds if

$$p \leq 0.26.$$

This completes the proof of the Theorem.

2.2 SCREENING WITH ERRORS

In section 2.1, we did not allow errors in observations, i.e., a factor was correctly identified as either defective or non-defective. In this section, we shall allow errors in observations and work out corresponding results given in section 2.1.

2.2.1 The expected number of runs

Let there be f factors to be tested for significance. In step one, the f factors are divided into g group-factors of k factors each. The group-factors are then tested for their significance by an experiment. Those that are declared non-defective are set aside. In step two, one factor is taken from each group-factor declared defective in step one. The factors are then tested for their significance by an experiment. If no factor is declared defective, we take another set of factors one from each group-factor and test their significance. We repeat this procedure till at least one factor is declared defective. The factors declared non-defective are set aside, keeping the factor(s) declared defective separate. The remaining factors from a group-factor that contained a factor that is declared defective are set aside and grouped into a group-factor. This process is repeated until one factor is declared defective from each group-factor declared defective in step one. In the third step, the group-factors set aside in step two are tested in an experiment. Again the group-factors declared non-defective in step three are set aside.

In the fourth step, we proceed with a series of experiments as in step two until one factor is declared defective from each group-factor declared to be defective in step three. The remaining factors from each of the group-factors declared defective in step three are grouped into a group-factor after step four is performed. Again the group-factors set aside in step four are tested for their significance in an experiment in step five. This procedure is repeated until the analysis terminates with all group-factors declared non defective. Certainly the analysis will terminate in a finite number of steps. We allow the possibility that defective group-factors and factors may not be detected. Also non-defective group-factors and factors may be declared defective. Our objective is to determine the group size 'k' in the initial step which minimizes the expected number of tests (runs).

Let α_I be the level of significance of tests in step one. Thus α_I is the probability of declaring a non-defective group-factor defective in step one, i.e. the first kind of error.

Consider the hypothesis

H_0: a group-factor in step one is non-defective.

Alternative

H_1: a group-factor in step one is defective

(2.2.1).

In testing the significance of factors and group-factors, we shall use orthogonal fractional factorial plans of the type given by Plackett and Burman [22]. These are specially constructed two-level orthogonal designs for studying upto (4m-1) factors in 4m runs. In general the number of runs required by the orthogonal design to study m factors (or group-factors) is given by

$R(m) = 4\left[\frac{m}{4}\right]$ where $\left[\frac{m}{4}\right]$ is the smallest integer greater than $\frac{m}{4}$ except that $\left[\frac{m}{4}\right] = 0$ when m=0.

According to Patel and Ottieno [19],

$$4\left[\frac{m}{4}\right] = m + h \quad \text{where } h = 1,2,3,4 \qquad (2.2.2).$$

There are g group-factors to be tested in step one. Each group-factor has two levels, the lower level denoted by '0' and the upper level denoted by '1'. Thus for tests of significance we require an orthogonal plan for a 2^g factorial experiment. Now let $\hat{A}$ be the estimate of the main effect of any group-factor in step one with s defective factors each with effect $\Delta>0$ for s=1,2,..,k. Then

$$E(\hat{A}) = s\Delta \qquad (2.2.3)$$

and

$$\text{Var}(\hat{A}) = \frac{\sigma^2}{g+h}\,; \quad h = 1,2,3,4 \qquad (2.2.4)$$

where σ^2 is the error in observation.

Next define

$$z = \frac{\hat{A} - s\Delta}{\sqrt{\sigma^2/(g+h)}}$$

$$= y - s\phi_I \qquad (2.2.5)$$

where

$$y = \frac{\hat{A}}{\sqrt{\sigma^2/(g+h)}} \qquad (2.2.6)$$

and

$$\phi_I = \frac{\Delta}{\sqrt{\sigma^2/(g+h)}} \qquad (2.2.7).$$

Assuming that the observations are normally distributed, z is a standardized normal variate. We shall say that a step one group-factor is non-defective if s=0, which implies that $s\phi_1=0$. On the other hand, a first step group-factor will be defective if $s\phi_1 \neq 0$. Therefore the hypothesis (2.2.1) may be expressed as

$$H_o: s\phi_I = 0$$

against (2.2.8)

$$H_1: s\phi_I \neq 0$$

In testing the hypothesis (2.2.8) we shall use the normal deviate test if σ^2 is known otherwise we shall use the t test if σ^2 is estimated from the experiment.

Let $\Pi_I(s\phi_I,\alpha_I)$ denote the power of the test in step one. Then

$$\Pi_I(s\phi_I,\alpha_I) = \int_{z(\alpha_I)-s\phi_I}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz \qquad (2.2.9)$$

where $z(\alpha_I)$ is given by

$$\alpha_I = \int_{z(\alpha_I)}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz \qquad (2.2.10).$$

When s=0 or $\frac{\Delta}{\sigma} = 0$, we have

$$\Pi_I(0,\alpha_I) = \alpha_I \qquad (2.2.11).$$

When $s \neq 0$ and $\frac{\Delta}{\sigma}$ is large, then we have

$$\Pi_I(s\phi_I,\alpha_I) \simeq 1 \qquad (2.2.12).$$

Let p be the a-priori probability that a factor is defective. Then the probability that a group-factor in step one with s defective factors is declared defective is given by

$$\Pi_I^* = \sum_{s=0}^{k} \binom{k}{s} p^s (1-p)^{k-s} \Pi_I(s\phi_I,\alpha_I)$$

$$= (1-p)^k \alpha_I + \sum_{s=1}^{k} \binom{k}{s} p^s (1-p)^{k-s} \Pi_I(s\phi_I,\alpha_I) \qquad (2.2.13).$$

These results are the same as those given by Patel and Ottieno [19].

Let r be the number of group-factors declared defective in step one. Then the probability distribution of r is given by

$$f(r) = \binom{g}{r} \Pi_I^{*r} (1-\Pi_I^*)^{g-r} \qquad (2.2.14).$$

Thus

$$E(r) = g\Pi_I^*$$

$$= \frac{f}{k} \Pi_I^* \qquad (2.2.15)$$

In the subsequent steps, the analysis is continued as described in the earlier part of section (2.2.1) for the r group-factors declared defective in step one. In the first experiment of step two, we test r factors 1 factor from each of the r group-factors declared defective in step one. Each factor has two levels, the lower level denoted by '0' and the upper level denoted by '1'. Thus using the main effect plans of the type given by Plackett and Burman [22], we require

$$4\left[\frac{r}{4}\right] \qquad (2.2.16)$$

runs to test the significance of the r factors. The h runs (observations) used in step one can be repeatedly used in all the subsequent steps to make the experiments orthogonal. Let p' be the probability that a factor chosen at random from a group-factor in step one containing s defective factors that has been declared defective is defective. Then

$$p' = \frac{p}{\Pi_I^*} \sum_{s=1}^{k} \binom{k-1}{s-1} p^{s-1} (1-p)^{k-s} \Pi_I(s\phi_I, \alpha_I)$$

$$= \frac{p\Pi_I^*}{\Pi_I^*} \qquad \text{(c.f. Curnow [2])} \qquad (2.2.17),$$

where

$$\Pi_I^* = \sum_{s=1}^{k} \binom{k-1}{s-1} p^{s-1} q^{k-s}\, \Pi_I(s\varphi_I, \alpha_I)$$

is the probability that a group-factor containing at least one defective factor is declared defective in the initial step. Define a random variable δ as follows:-

$\delta = 0$ if a factor chosen at random from a group-factor that is declared defective in step one is non-defective

$= 1$ otherwise.

Then

$$\delta = 0 \quad \text{with probability } 1-p'$$

$$= 1 \quad \text{with probability } p'$$

Let α_s be the probability of declaring a non-defective factor defective and γ_s be the probability of declaring a defective factor defective in the subsequent steps.

Let

$$\beta^{+-} = (1-\delta)\alpha_s + \delta\,\gamma_s \qquad (2.2.18).$$

Then

$$\beta^{+} = \alpha_s \quad \text{with probability } 1-p'$$

$$= \gamma_s \quad \text{with probability } p'$$

Hence the average value of β^{+-} is given by

$$\bar{\beta}^{+} = \gamma_s p' + \alpha_s (1-p')$$

$$= \frac{1}{\Pi_I^*}\left[p(\gamma_s - \alpha_s)\Pi_I^{+} + \Pi_I^{*}\,\alpha_s\right]$$

$$= \frac{\bar{\beta}^*}{\Pi_I^*} \qquad (2.2.19)$$

where

$$\bar{\beta}^* = \left[p(\gamma_s - \alpha_s)\Pi_I^{+-} + \Pi_I^{*}\,\alpha_s\right] \qquad (2.2.20)$$

is the probability that a factor chosen at random is declared defective in the subsequent steps. Thus $\bar{\beta}^{+}$ may be interpretted as the probability that a factor chosen at random from a group-factor that is declared defective in step one is declared defective in the subsequent steps.

Out of the group-factors declared defective at any step, it is possible that due to errors in observations, we may find some group-factors from which no factor is declared defective on individual tests. Let α_s^* be the proportion of such group-factors. Obviously α_s^* will be different at every step. However for simplicity in algebra, we shall assume α_s^* to be of uniform value, say α^*.

Let us denote by $P_k^*(j)$ the probability that exactly j factors from a group-factor that is declared defective in step one are declared defective in the subsequent steps.

Then

$$P_k^*(j) = \begin{cases} 1 - \frac{1}{\Pi_I^*}\left[1 - (1 - \bar{\beta}^*)^k\right] & j=0 \\ \frac{1}{\Pi_I^*}\binom{k}{j}\bar{\beta}^{*j}(1 - \bar{\beta}^*)^{k-j} & j=1,2,\ldots,k \end{cases} \tag{2.2.21}$$

Let $E_k^*(R_j)$ be the expected number of tests (runs) required to declare exactly j factors defective from a group-factor of size k which has been declared defective in step one. To obtain an expression for $E_k^*(R_j)$ we start by considering a sequence of lemmas.

Lemma 2.2.1

$$E_k^*(R_0) = k$$

Proof

The proof is trivial since to declare all the k factors in the group-factor as non-defective we need to test all of them.

Lemma 2.2.2

$$E_k^*(R_1) = \frac{k}{2} + \frac{3}{2} - \frac{1}{k} + \frac{\alpha^* k}{2} - \frac{\alpha^*}{2} - \frac{\alpha^*}{k} - \frac{(1-\xi)}{k}$$

where

$$\xi = 0 \quad \text{if} \quad \alpha^* = 0$$
$$\quad = 1 \quad \text{otherwise}$$

Proof

It is equally likely that the one factor declared defective be found at any trial. Consequently the probability that it is found on any one trial is $\frac{1}{k}$. If the one factor declared defective is found on the ℓ^{th} trial, $\ell=1,2,\ldots,k-2$, then ℓ tests are needed to find it. The next test we need is the group test on a group-factor consisting of $(k-\ell)$ factors. If this group-factor is declared non-defective, we shall stop the test procedure otherwise we continue testing individual factors until all the $(k-\ell)$ factors are declared non-defective. If the $(k-1)^{st}$ factor is the one declared defective, then we have to test the k^{th} factor as well. However if the first k-1 factors tested are declared non-defective, we shall need to test the k^{th} factor to declare it defective only if $\alpha^* \neq 0$; otherwise we would declare it defective with probability 1 (this corresponds to the case when we have no errors in observations).

Thus

$$E_k^*(R_1) = \frac{1}{k}\left[\sum_{\ell=1}^{k-2}\{(\ell+1) + \alpha^* E_{k-\ell}^*(R_o)\} + k + (k-1) + \xi\right] \tag{2.2.22}$$

Using lemma 2.2.1, we get

$$E_k^*(R_1) = \frac{1}{k}\left[\sum_{\ell=1}^{k-2}(\ell+1) + \alpha^* \sum_{\ell=1}^{k-2}(k-\ell)\right] + 2 - \frac{(1-\xi)}{k}$$

i.e.

$$E^*_k(R_1) = \frac{k}{2} + \frac{3}{2} - \frac{1}{k} + \frac{\alpha^* k}{2} - \frac{\alpha^*}{2} - \frac{\alpha^*}{k} - \frac{(1-\xi)}{k} \qquad (2.2.23)$$

This proves the lemma.

Lemma 2.2.3

$$E^*_k(R_2) = \frac{2k}{3} + 2 + \frac{2}{3} - \frac{2}{k-1} + \alpha^*\{\frac{k}{3} - \frac{2}{3} - \frac{2}{k-1} + \frac{4}{k(k-1)}\} - \frac{2(1-\xi)(k-2)}{k(k-1)}$$

Proof

Here the approach is to find the first factor to be declared defective and thus reduce the problem to the one in which the group-factor has only one factor to be declared defective. This problem of a group-factor having only one factor to be declared defective was considered in lemma 2.2.2. The probability that the first factor tested is declared defective is $\frac{2}{k}$. If the first factor tested is declared defective, then on the average we require $\{1+1+E^*_{k-1}(R_1)\}$ runs to complete the test procedure. For $\ell=1,2,\dots,k-3$, the probability that the $(\ell+1)^{st}$ factor tested is the first to be declared defective is

$$\prod_{w=1}^{\ell}\left(\frac{k-(w+1)}{k-(w-1)}\right)\frac{2}{k-\ell}$$

and on the average the number of tests (runs) required to complete the test procedure in this case is $\{(\ell+1) + 1 + E^*_{k-(\ell+1)}(R_1)\}$. If the first k-2 factors tested are declared non-defective, then we need to test the other two factors to declare them defective i.e., we need k tests.

Hence

$$E^*_k(R_2) = \frac{2}{k}\{1 + 1 + E^*_{k-1}(R_1)\}$$

$$+ \sum_{\ell=1}^{k-3} \prod_{w=1}^{\ell}\left(\frac{k-(w+1)}{k-(w-1)}\right)\frac{2}{k-\ell}\{(\ell+1) + 1 + E^*_{k-(\ell+1)}(R_1)\}$$

$$+ \frac{2k}{k(k-1)} \qquad (2.2.24).$$

Substituting in (2.2.24) the values given by (2.2.23) we obtain

$$E^*_k(R_2) = \frac{2}{k}\frac{k-1}{k-1}\left[2+ \frac{k-1}{2} + \frac{3}{2} - \frac{1}{k-1} + \alpha^*\{\frac{k-1}{2} - \frac{1}{2} - \frac{1}{k-1}\} - \frac{(1-\xi)}{k-1}\right]$$

$$+ \frac{2}{k}\frac{k-2}{k-1}\left[3+ \frac{k-2}{2} + \frac{3}{2} - \frac{1}{k-2} + \alpha^*\{\frac{k-2}{2} - \frac{1}{2} - \frac{1}{k-2}\} - \frac{(1-\xi)}{k-2}\right]$$

$$+ \frac{2}{k}\frac{k-3}{k-1}\left[4+ \frac{k-3}{2} + \frac{3}{2} - \frac{1}{k-3} + \alpha^*\{\frac{k-3}{2} - \frac{1}{2} - \frac{1}{k-3}\} - \frac{(1-\xi)}{k-3}\right]$$

$$+ \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot$$

$$+ \frac{2}{k}\frac{3}{k-1}\left[(k-2) + \frac{3}{2} + \frac{3}{2} - \frac{1}{3} + \alpha^*\{\frac{3}{2} - \frac{1}{2} - \frac{1}{3}\} - \frac{(1-\xi)}{3}\right]$$

$$+ \frac{2}{k}\frac{2}{k-1}\left[(k-1) + \frac{2}{2} + \frac{3}{2} - \frac{1}{2} + \alpha^*\{\frac{2}{2} - \frac{1}{2} - \frac{1}{2}\} - \frac{(1-\xi)}{2}\right]$$

$$+ \frac{2}{k}\frac{1}{k-1}(k) \qquad (2.2.25)$$

By taking appropriate summations, (2.2.25) can be written as

$$E^*_k(R_2) = \frac{2}{k(k-1)}\sum_{m=1}^{k-1}(m+1)(k-m) + \frac{1}{k(k-1)}\sum_{m=1}^{k-2}(k-m)^2$$

$$+ \frac{3}{k(k-1)}\sum_{m=1}^{k-2}(k-m) - \frac{2(k-2)}{k(k-1)}$$

$$+ \frac{\alpha^*}{k(k-1)}\sum_{m=1}^{k-2}(k-m)^2 - \frac{\alpha^*}{k(k-1)}\sum_{m=1}^{k-2}(k-m)$$

$$- \frac{2\alpha^*(k-2)}{k(k-1)} - \frac{2(1-\xi)(k-2)}{k(k-1)}$$

$$= \frac{2}{k(k-1)}\left[\frac{(k+1)k(k-1)}{6} + \frac{k(k-1)}{2}\right]$$

$$+ \frac{1}{k(k-1)}\left[\frac{k(k-1)(2k-1)}{6} - 1\right] + \frac{3}{k(k-1)}\left[\frac{(k+1)(k-2)}{2}\right]$$

$$- \frac{2(k-2)}{k(k-1)} + \frac{\alpha^*}{k(k-1)}\left[\frac{k(k-1)(2k-1)}{6} - 1\right]$$

$$- \frac{\alpha^*}{k(k-1)}\left[\frac{(k+1)(k-2)}{2}\right] - \frac{2\alpha^*(k-2)}{k(k-1)}$$

$$- \frac{2(1-\xi)(k-2)}{k(k-1)} \qquad (2.2.26).$$

This gives the required result on further simplification.

Lemma 2.2.4

$$E_k^*(R_3) = \frac{3k}{4} + 3 + \frac{3}{4} - \frac{3(k-2+3)}{k(k-1)}$$

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

$$- \frac{3(1-\xi)(k-3)}{k(k-1)}$$

Proof

After one factor has been declared defective, the problem reduces to that considered in lemma 2.2.3. The probability that the first factor tested is declared defective is $\frac{3}{k}$ and the probability that for $\ell=1,2,..,k-3$, the $(\ell+1)^{st}$ factor tested is the first to be declared defective is $\prod_{w=1}^{\ell}\left(\frac{k-(w+2)}{k-(w-1)}\right)\frac{3}{k-\ell}$. If the first factor tested is declared defective, then on the average we need $\{1 + 1 + E_k^*(R_2)\}$ tests to complete the test procedure. However if for $\ell=1,2,..,k-3$

the $(\ell+1)^{st}$ factor tested is the first to be declared defective, then on the average we shall need $\{(\ell+1) + 1 + E^*_{k-(\ell+1)}(R_2)\}$ tests to complete the test procedure. Thus

$$E^*(R_3) = \frac{3}{k}\{1 + 1 + E^*_{k-1}(R_2)\}$$
$$+ \sum_{\ell=1}^{k-3}\left[\prod_{w=1}^{\ell}\left(\frac{k-(w+2)}{k-(w-1)}\right)\frac{3}{k-j}\ \{(\ell+1)+1+E^*_{k-(\ell+1)}(R_2)\}\right] \tag{2.2.27}$$

Using (2.2.26) we get

$$E^*_k(R_3) = \frac{3}{k}\,\frac{k-1}{k-1}\,\frac{k-2}{k-2}\left\{2 + \frac{2(k-1)}{3} + 2 + \frac{2}{3} - \frac{2}{k-2}\right.$$

$$\left. + \alpha^*\left(\frac{k-1}{3} - \frac{2}{3} - \frac{2}{k-2} + \frac{4}{(k-1)(k-2)}\right) - \frac{2(1-\xi)(k-3)}{(k-1)(k-2)}\right\}$$

$$+ \frac{3}{k}\,\frac{k-3}{k-1}\,\frac{k-2}{k-2}\left\{3 + \frac{2(k-2)}{3} + 2 + \frac{2}{3} - \frac{2}{k-3}\right.$$

$$\left. + \alpha^*\left(\frac{k-2}{3} - \frac{2}{3} - \frac{2}{k-3} + \frac{4}{(k-2)(k-3)}\right) - \frac{2(1-\xi)(k-4)}{(k-2)(k-3)}\right\}$$

$$+ \frac{3}{k}\,\frac{k-3}{k-1}\,\frac{k-4}{k-2}\left\{4 + \frac{2(k-3)}{3} + 2 + \frac{2}{3} - \frac{2}{k-4}\right.$$

$$\left. + \alpha^*\left(\frac{k-3}{3} - \frac{2}{3} - \frac{2}{k-4} + \frac{4}{(k-3)(k-4)}\right) - \frac{2(1-\xi)(k-5)}{(k-3)(k-4)}\right\}$$

$$+ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot$$

$$+ \frac{3}{k}\,\frac{2}{k-1}\,\frac{3}{k-2}\left\{(k-2) + \frac{2(3)}{3} + 2 + \frac{2}{3} - \frac{2}{2}\right.$$

$$\left. + \alpha^*\left(\frac{3}{3} - \frac{2}{3} - \frac{2}{2} + \frac{4}{3\times 2}\right) - \frac{2(1-\xi)}{3\times 2}\right\}$$

$$+ \frac{3}{k}\,\frac{1}{k-1}\,\frac{2}{k-2}\left\{(k-1) + \frac{2(2)}{3} + 2 + \frac{3}{2} - \frac{2}{1}\right.$$

$$\left. + \alpha^*\left(\frac{2}{3} - \frac{2}{3} - \frac{2}{1} + \frac{4}{2\times 1}\right)\right\}$$

$$= \frac{3}{{}_kP_3}\sum_{m=1}^{k-2}(m+1)(k-m)(k-m-1) + \frac{2}{{}_kP_3}\sum_{m=1}^{k-2}(k-m)^2(k-m-1)$$

$$+ \frac{8}{{}_kP_3}\sum_{m=1}^{k-2}(k-m)(k-m-1) - \frac{6}{{}_kP_3}\sum_{m=1}^{k-2}(k-m)$$

$$+ \frac{\alpha^*}{{}_kP_3}\sum_{m=1}^{k-2}(k-m)^2(k-m-1) - \frac{2\alpha^*}{{}_kP_3}\sum_{m=1}^{k-2}(k-m)(k-m-1)$$

$$- \frac{6\alpha^*}{{}_kP_3}\sum_{m=1}^{k-2}(k-m) + \frac{12\alpha^*(k-2)}{{}_kP_3} - \frac{6(1-\xi)}{{}_kP_3}\sum_{m=1}^{k-3} m$$

where ${}_kP_3 = k(k-1)(k-2)$.

Using the sums given in (2.1.21) in the equation above, we obtain

$$E_k^*(R_3) = \frac{3}{{}_kP_3}\left[\frac{(k+1)P_4}{12} + \frac{{}_kP_3}{3}\right] + \frac{2}{{}_kP_3}\left[\frac{{}_kP_3}{12}(3k-1)\right]$$

$$+ \frac{8}{{}_kP_3}\left[\frac{{}_kP_3}{3}\right] - \frac{6}{{}_kP_3}\left[\frac{(k-2)(k+1)}{2}\right]$$

$$+ \frac{\alpha^*}{{}_kP_3}\left[\frac{{}_kP_3}{12}(3k-1)\right] - \frac{2\alpha^*}{{}_kP_3}\left[\frac{{}_kP_3}{3}\right]$$

$$- \frac{6\alpha^*}{{}_kP_3}\left[\frac{(k-2)(k+1)}{2}\right] + \frac{12\alpha^*(k-2)}{{}_kP_3}$$

$$- \frac{6(1-\xi)}{{}_kP_3}\left[\frac{(k-2)(k-3)}{2}\right]$$

$$= \frac{k+1}{4} + 1 + \frac{(3k-1)}{6} + \frac{8}{3} - \frac{3(k+1)}{k(k-1)}$$

$$+ \alpha^* \{\frac{3k-1}{12} - \frac{2}{3} - \frac{3(k+1)}{k(k-1)} + \frac{12}{k(k-1)} \}$$

$$- \frac{3(1-\xi)(k-3)}{k(k-1)} \qquad (2.2.28)$$

This gives the required result on further simplification.

We are now ready to state and prove a more general result.

Theorem 2.2.1

In a step-wise group screening design in which all the group-factors are of the same size 'k' in the initial step, the average number of tests required to analyse a group-factor which is declared defective in step one, from which exactly j factors are declared defective in the subsequent steps is given by

$$E_k^*(R_j) = \frac{jk}{j+1} + j + \frac{j}{j+1} - \frac{j(k+j-2)}{k(k-1)}$$

$$+ \alpha^*\{\frac{k}{j+1} - \frac{j}{j+1} - \frac{j}{k-1} + \frac{j^2}{k(k-1)}\}$$

$$- \frac{j(1-\xi)(k-j)}{k(k-1)} \qquad \text{for } j=1,2,..,k,$$

where α^* is the proportion of group-factors declared defective at any step but due to errors in observations no factor from each such group-factor is declared defective on individual tests and $\xi=0$ if $\alpha^*=0$ and 1 otherwise.

Proof

The proof follows by mathematical induction. The validity of the Theorem has been shown for j=1. We assume that the Theorem is true for j=n-1 ($1 \leq n-1 \leq k$), that is

$$E_k^*(R_{n-1}) = \frac{(n-1)k}{n} + (n-1) + \frac{(n-1)}{n} - \frac{(n-1)\{k+(n-1)-2\}}{k(k-1)}$$

$$+ \alpha^*\{\frac{k}{n} - \frac{(n-1)}{n} - \frac{(n-1)}{k-1} + \frac{(n-1)^2}{k(k-1)}\}$$

$$- \frac{(n-1)(1-\xi)(k+1-n)}{k(k-1)} \qquad (2.2.29)$$

We shall show that the Theorem is true for j=n. Now for j=n,

$$E_k^*(R_n) = \frac{n}{k}\{1 + 1 + E_{k-1}^*(R_{n-1})\}$$

$$+ \sum_{\ell=1}^{k-n} \prod_{w=1}^{\ell} \left(\frac{k-(w+n-1)}{k-(w-1)}\right) \frac{n}{k-\ell} \{(\ell+1) + 1 + E_{k-(\ell+1)}^*(R_{n-1})\}$$

(2.2.30).

In the first term, $\frac{n}{k}$ is the probability that the first factor tested is declared defective and $\{1 + 1 + E_{k-1}^*(R_{n-1})\}$ is the average number of tests required to perform the analysis if the first factor tested is declared defective. The term

$$\prod_{w=1}^{\ell} \left(\frac{k-(w+n-1)}{k-(w-1)}\right) \frac{n}{k-\ell} \{(\ell+1) + 1 + E_{k-(\ell+1)}^*(R_{n-1})\}$$

in the summation is the product of the probability that for $\ell=1,2,\ldots,k-n$, the $(\ell+1)^{st}$ factor is the first factor to be declared defective and the average number of tests (runs) required to perform the analysis in that case. Substituting in (2.2.30) the values given by (2.2.29), we obtain

$$E_k^*(R_n) = \frac{n}{k}\frac{k-1}{k-1}\cdots\frac{k-(n-1)}{k-(n-1)}\left\{2 + \frac{(n-1)(k-1)}{n} + \frac{n^2-1}{n} - \frac{(n-1)(k+n-4)}{(k-1)(k-2)}\right.$$

$$\left. + \alpha^*\left[\frac{k-1}{n} - \frac{n-1}{n} - \frac{n-1}{k-2} + \frac{(n-1)^2}{(k-1)(k-2)}\right] - \frac{(n-1)(1-\xi)(k-n)}{(k-1)(k-2)}\right\}$$

$$+ \frac{n}{k}\frac{k-n}{k-1}\frac{k-2}{k-2}\cdots\frac{k-(n-1)}{k-(n-1)}\left\{3 + \frac{(n-1)(k-2)}{n} + \frac{n^2-1}{n}\right.$$

$$- \frac{(n-1)(k+n-5)}{(k-2)(k-3)} + \alpha^*\left[\frac{k-2}{n} - \frac{n-1}{n} - \frac{n-1}{k-3} + \frac{(n-1)^2}{(k-2)(k-3)}\right]$$

$$\left. - \frac{(n-1)(1-\xi)(k-n-1)}{(k-2)(k-3)}\right\}$$

$$+ \frac{k-n}{k}\frac{k-n-1}{k-1}\frac{n}{k-2}\frac{k-3}{k-3}\cdots\frac{k-(n-1)}{k-(n-1)}\left\{4 + \frac{(n-1)(k-3)}{n} + \frac{n^2-1}{n}\right.$$

$$- \frac{(n-1)(k+n-6)}{(k-3)(k-4)} + \alpha^*\left[\frac{k-3}{n} - \frac{n-1}{n} - \frac{n-1}{k-4} + \frac{(n-1)^2}{(k-3)(k-4)}\right]$$

$$\left. - \frac{(n-1)(1-\xi)(k-n-2)}{(k-3)(k-4)}\right\}$$

$$+ \; . \; . \; . \; . \; . \; . \; . \; . \; . \; . \; .$$

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

$$\times \frac{k-(n-1)}{k-(n-1)}\left\{(k-n+1) + \frac{n(n-1)}{n} + \frac{n^2-1}{n} - \frac{(n-1)(2n-3)}{n(n-1)}\right.$$

$$\left. + \alpha^*\left[\frac{n}{n} - \frac{n-1}{n} - \frac{n-1}{n-1} + \frac{(n-1)^2}{n(n-1)} + \frac{(n-1)(1-\xi)}{n(n-1)}(1)\right]\right\}$$

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

$$\times \frac{k-(n-1)}{k-(n-1)}\left\{(k-n+2) + \frac{(n-1)(n-1)}{n} + \frac{n^2-1}{n} - \frac{(n-1)(2n-5)}{(n-1)(n-2)}\right.$$

$$+\ \alpha^*\left[\frac{n-1}{n}-\frac{n-1}{n}-\frac{n-1}{n-2}+\frac{(n-1)^2}{(n-1)(n-2)}\right]\Bigg\} \qquad (2.2.31)$$

By rearranging and taking appropriate summations, (2.2.31) becomes

$$E^*_k(R_n) = \frac{n}{{}_kP_n}\sum_{m=1}^{k-n+1}\{(m+1)(k-m)(k-m-1)\ldots(k-m-n+2)$$

$$+\frac{n-1}{{}_kP_n}\sum_{m=1}^{k-n+1}\{(k-m)^2(k-m-1)\ \ldots\ (k-m-n+2)\}$$

$$+\frac{n^2-1}{{}_kP_n}\sum_{m=1}^{k-n+1}\{(k-m)(k-m-1)\ \ldots(k-m-n+2)\}$$

$$-\frac{(n-1)n}{{}_kP_n}\sum_{m=1}^{k-n+1}\{(k-m-1)(k-m-2)\ldots(k-m-n+2)\}$$

$$-\frac{(n-1)(n-2)n}{{}_kP_n}\sum_{m=1}^{k-n+1}\{(k-m-2)(k-m-3)\ldots(k-m-n+2)$$

$$+\frac{\alpha^*}{{}_kP_n}\sum_{m=1}^{k-n+1}\{(k-m)^2(k-m-1)(k-m-2)\ldots(k-m-n+2)\}$$

$$-\frac{(n-1)\alpha^*}{{}_kP_n}\sum_{m=1}^{k-n+1}\{(k-m)(k-m-1)(k-m-2)\ldots(k-m-n+2)\}$$

$$+\frac{\alpha^*(n-1)^2n}{{}_kP_n}\sum_{m=1}^{k-n+1}\{(k-m-2)(k-m-3)\ldots(k-m-n+2)\}$$

$$-\frac{\alpha^*(n-1)n}{{}_kP_n}\sum_{m=1}^{k-n+1}\{(k-m)(k-m-2)\ldots(k-m-n+2)\}$$

$$-\frac{n(n-1)}{{}_kP_n}(1-\xi)\sum_{m=1}^{k-n}\{(k-m-2)(k-m-3)\ldots(k-m-n+1)\} \qquad (2.2.32)$$

where ${}_kP_n = \frac{k!}{(k-n)!}$

Using (2.1.21) in (2.2.32) we obtain

$$E_k^*(R_n) = \frac{n}{{}_kP_n}\left[\frac{{}_{(k+1)}P_{(n+1)}}{n(n+1)} + \frac{{}_kP_n}{n}\right] + \frac{n-1}{{}_kP_n}\left[\frac{{}_{(k+1)}P_{(n+1)}}{(n+1)} - \frac{{}_kP_n}{n}\right]$$

$$+ \frac{n^2-1}{{}_kP_n}\left[\frac{{}_kP_n}{n}\right] - \frac{(n-1)n}{{}_kP_n}\left[\frac{{}_{(k-1)}P_{(n-1)}}{n-1}\right]$$

$$- \frac{(n-1)(n-2)n}{{}_kP_n}\left[\frac{{}_{(k-2)}P_{(n-2)}}{n-2}\right] + \frac{\alpha^*}{{}_kP_n}\left[\frac{{}_{(k+1)}P_{(n+1)}}{(n+1)} - \frac{{}_kP_n}{n}\right]$$

$$- \frac{(n-1)\alpha^*}{{}_kP_n}\left[\frac{{}_kP_n}{n}\right] + \frac{\alpha^*(n-1)^2 n}{{}_kP_n}\left[\frac{{}_{(k-2)}P_{(n-2)}}{n-2}\right]$$

$$- \frac{(n-1)n\alpha^*}{{}_kP_n}\left[\frac{{}_{(k-1)}P_{(n-1)}}{n-1} + \frac{{}_{(k-2)}P_{(n-2)}}{n-2}\right]$$

$$- \frac{n(n-1)(1-\xi)}{{}_kP_n}\left[\frac{{}_{(k-2)}P_{(n-1)}}{n-1}\right]$$

$$= \frac{k+1}{n+1} + 1 + \frac{n-1}{(n+1)}(k+1) - \frac{n-1}{n} - \frac{n^2-1}{n} - \frac{n}{k}$$

$$- \frac{n(n-1)}{k(k-1)} + \alpha^*\left\{\frac{k+1}{(n+1)} - \frac{1}{n} - 1 + \frac{1}{n} + \frac{(n-1)^2 n}{k(k-1)(n-2)}\right.$$

$$\left. - \frac{n}{k} - \frac{(n-1)n}{(n-2)k(k-1)}\right\}$$

$$- \frac{n(n-1)(k-n)}{(n-1)k(k-1)}(1-\xi)$$

i.e.,

$$E_k^*(R_n) = \frac{nk}{n+1} + n + \frac{n}{n+1} - \frac{n(k+n-2)}{k(k-1)}$$

$$+ \alpha^* \left\{ \frac{k}{n+1} - \frac{n}{n+1} - \frac{n}{k-1} + \frac{n^2}{k(k-1)} \right\}$$

$$- \frac{n(1-\xi)(k-n)}{k(k-1)} \qquad (2.2.33)$$

This is exactly the value of $E_k^*(R_j)$ for j=n. Thus if the Theorem is true for j=n-1 ($0 \leq n-1 \leq k$), it is also true for j=n. But the Theorem is true for j=1 (c.f. lemma 2.2.2). Hence it is true for j=2 and in general for any j, j=1,2,...,k. This completes the proof.

The Theorem does not apply to the case j=0 which is trivial and was considered in lemma 2.2.1. In special case, when $\alpha^*=0$, then $\xi=0$. This is the case when we have no errors in observations and the formula for $E_k^*(R_j)$ given in Theorem 2.2.1 coincides with that for $E_k(R_j)$ given in Theorem 2.1.1.

Let R_S^* denote the number of tests (runs) required to analyse a group-factor i.e., declare as defective or non-defective the factors within a group-factor of size k which has been declared defective in step one. Then we have the following corollary.

<u>Corollary 2.2.1</u>

In a step-wise design, the expected number of runs required to analyse a group-factor of size k which has been declared defective in step one is given by

$$E(R_S^*) = k - \frac{k}{\Pi_I^*} \left[1 - (1-\bar{\beta}^*)^k \right]$$

$$+ \frac{1}{\Pi_I^*}\left[k+1 + k\bar{\beta}^* - \frac{1}{\bar{\beta}^*}\{1 - (1-\bar{\beta}^*)^{k+1}\}\right]$$

$$- \frac{1}{\Pi_I^*}\left[(2-\xi)\bar{\beta}^* + \xi\,\bar{\beta}^{*2}\right]$$

$$+ \frac{\alpha^*}{\Pi_I^*}\left[\frac{1}{\bar{\beta}^*}\{1 - (1-\bar{\beta}^*)^{k+1} - k\bar{\beta}^*(1-\bar{\beta}^*)^k\}\right.$$

$$\left. - 1 + \bar{\beta}^{*2} - \bar{\beta}^*\right] .$$

Proof

$$E(R_S^*) = \sum_{j=0}^{k} E_k^*(R_j)\,P_k^*(j) \qquad (2.2.34).$$

Using (2.2.21), lemma 2.2.1 and Theorem 2.2.1 we get

$$E(R_S^*) = k - \frac{k}{\Pi_I^*}\left[1 - (1-\bar{\beta}^*)^k\right]$$

$$+ \frac{1}{\Pi_I^*}\sum_{j=1}^{k}\left\{\frac{jk}{j+1} + j + \frac{j}{j+1} - \frac{j(k+j-2)}{k(k-1)}\right\}\binom{k}{j}\bar{\beta}^{*j}(1-\bar{\beta}^*)^{k-j}$$

$$+ \frac{\alpha^*}{\Pi_I^*}\sum_{j=1}^{k}\left\{\frac{k}{j+1} - \frac{j}{j+1} - \frac{j}{k-1} + \frac{j^2}{k(k-1)}\right\}\binom{k}{j}\bar{\beta}^{*j}(1-\bar{\beta}^*)^{k-j}$$

$$- \frac{1-\xi}{k(k-1)}\frac{1}{\Pi_I^*}\sum_{j=1}^{k}(jk-j^2)\binom{k}{j}\bar{\beta}^{*j}(1-\bar{\beta}^*)^{k-j}$$

i.e.,

$$E(R_S^*) = k - \frac{k}{\Pi_I^*}\left[1 - (1-\bar{\beta}^*)^k\right]$$

$$+ \frac{1}{\Pi_I^*}\left[(k+1)\{1 - (1-\bar{\beta}^*)^k\} + k\bar{\beta}^* - \frac{1}{\bar{\beta}^*}\{1 - (1-\bar{\beta}^*)^{k+1}\right.$$

$$- (k+1)\,\bar{\beta}^*(1-\bar{\beta}^*)^k\} - \frac{1}{k(k-1)}\{k^2\bar{\beta}^*$$

$$+ k\bar{\beta}^*(1-\bar{\beta}^*) - 2k\bar{\beta}^* + k^2\bar{\beta}^{*2}\}\Big]$$

$$+ \frac{\alpha^*}{\Pi_I^*}\Big[- \{1 - (1-\bar{\beta}^*)^k\} + \frac{1}{\bar{\beta}^*}\{1 - (1 - \bar{\beta}^*)^{k+1}$$

$$- (k+1)\bar{\beta}^*(1-\bar{\beta}^*)^k\} - \frac{k\bar{\beta}^*}{k-1} + \frac{k\bar{\beta}^*(1-\bar{\beta}^*)+k^2\bar{\beta}^{*2}}{k(k-1)}\Big]$$

$$- \frac{1-\xi}{k(k-1)}\frac{1}{\Pi_I^*}\Big[k^2\bar{\beta}^* - k\bar{\beta}^*(1-\bar{\beta}^*) - k^2\bar{\beta}^{*2}\Big]$$

i.e.,

$$E(R_S^*) = k - \frac{k}{\Pi_I^*}\Big[1 - (1 - \bar{\beta}^*)^k\Big]$$

$$+ \frac{1}{\Pi_I^*}\Big[k + 1 + k\bar{\beta}^* - \frac{1}{\bar{\beta}^*}\{1 - (1 - \bar{\beta}^*)^{k+1}\}\Big]$$

$$- \frac{1}{\Pi_I^*}\Big[(2-\xi)\bar{\beta}^* + \xi\bar{\beta}^{*2}\Big]$$

$$+ \frac{\alpha^*}{\Pi_I^*}\Big[\frac{1}{\bar{\beta}^*}\{1 - (1 - \bar{\beta}^*)^{k+1} - k\bar{\beta}^*(1 - \bar{\beta}^*)^k\}$$

$$- 1 + \bar{\beta}^{*2} - \bar{\beta}^*\Big] \qquad (2.2.35).$$

This proves the corollary.

Let R_S denote the number of tests required to analyse the r group-factors declared defective in step one. Then

$$R_S = r\,E(R_S^*) \qquad (2.2.36)$$

We may now state and proof the following Theorem.

Theorem 2.2.2

The expected total number of runs in a step-wise group screening design with errors in observations, in which k is the size of each of the group-factor in step one and $\overline{\beta}^*$ is the probability of declaring a factor defective in the subsequent steps is given by

$$E(R) = h + \frac{2f}{k} + f - \frac{f}{k\overline{\beta}^*}\{1 - (1 - \overline{\beta}^*)^{k+1}\}(1-\alpha^*)$$

$$+ f\overline{\beta}^* \{1 - \frac{2-\xi}{k} - \frac{\alpha^*}{k}\} + f\Pi_I^*$$

$$- f\left[1 - (1 - \overline{\beta}^*)^k\right] - \frac{f}{k}\alpha^*$$

$$- \frac{f}{k}\overline{\beta}^{*2}\{\xi - \alpha^*\} - f\alpha^*\overline{\beta}^*(1 - \overline{\beta}^*)^k$$

where α^* is the proportion of group-factors declared defective at any step but due to errors in observations no factor from each such group-factor is declared defective on individual tests,

$$\xi = 0 \text{ if } \alpha^* = 0$$
$$= 1 \quad \text{otherwise.}$$

Proof

The number of runs in step one is

$$R_I = h + g$$
$$= h + \frac{f}{k} \qquad (h=1,2,3,4) \qquad (2.2.37).$$

The number of runs required in the subsequent steps is

$R_S = rE(R_S^*)$ from (2.2.36).

Using corollary 2.2.1, it follows that

$$E(R_S) = f\Pi_I^* - f\left[1 - (1 - \bar{\beta}^*)^k\right] + f + \frac{f}{k} + f\bar{\beta}^*$$

$$- \frac{f}{k\bar{\beta}^*}\left\{1 - (1 - \bar{\beta}^*)^{k+1}\right\} - \frac{f}{k}\left\{(2 - \xi)\bar{\beta}^* + \xi\bar{\beta}^{*2}\right\}$$

$$+ \frac{f\alpha^*}{k\bar{\beta}^*}\left\{1 - (1 - \bar{\beta}^*)^{k+1}\right\} - f\alpha^*(1 - \bar{\beta}^*)^k\bar{\beta}^*$$

$$- \frac{f\alpha^*}{k} + \frac{f\alpha^*\bar{\beta}^{*2}}{k} - \frac{f\alpha^*\bar{\beta}^*}{k} \qquad (2.2.38),$$

after replacing 'r' by $E(r) = g\Pi_I^*$ given in (2.2.15).

The expected total number of runs is now given by

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

$$= h + \frac{2f}{k} + f - \frac{f}{k\bar{\beta}^*}\{1 - (1 - \bar{\beta}^*)^{k+1}\}(1 - \alpha^*)$$

$$+ f\bar{\beta}^*\left\{1 - \frac{2-\xi}{k} - \frac{\alpha^*}{k}\right\} + f\Pi_I^*$$

$$- f\left[1 - (1 - \bar{\beta}^*)^k\right] - \frac{f}{k}\alpha^*$$

$$- \frac{f}{k}\bar{\beta}^{*2}\{\xi - \alpha^*\} - f\alpha^*\bar{\beta}^*(1 - \bar{\beta}^*)^k$$

(2.2.39)

using (2.2.37) and (2.2.38) putting the like terms together. This completes the proof of Theorem 2.2.2.

<u>Corollary 2.2.2</u>

For large values of $\frac{\Delta}{\sigma}$ and arbitrary values of p, the expected total number of runs in a step-wise group screening design is approximately equal to

$$h + \frac{2f}{k} + f - \frac{f(1-\alpha^*)}{k}\left[\frac{1-\left\{1-(1-\alpha_s)p-\alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}^{k+1}}{(1-\alpha_s)p + \alpha_s\left(1-(1-\alpha_I)q^k\right)}\right]$$

$$+ f\left[(1-\alpha_s)p + \alpha_s\{1-(1-\alpha_I)q^k\}\right]\left[1-\frac{2-\xi}{k}-\frac{\alpha^*}{k}\right]$$

$$+ f\left[1-(1-\alpha_I)q^k\right] - f\left[1-\left\{1-(1-\alpha_s)p - \alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}^k\right]$$

$$- \frac{f\alpha^*}{k} - \frac{f}{k}\left\{(1-\alpha_s)p + \alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}^2 \{\xi-\alpha^*\}$$

$$- f\alpha^*\left\{(1-\alpha_s)p + \alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}\left[1-(1-\alpha_s)p-\alpha_s\left(1-(1-\alpha_I)q^k\right)\right]^k .$$

Proof:

If $\frac{\Delta}{\sigma}$ is large, we have the following approximations

$$\Pi^*_I \simeq 1-(1-\alpha_I)q^k \qquad (2.2.40)$$

$$\Pi^+_I \simeq 1 \qquad (2.2.41)$$

$$\gamma_s \simeq 1 \qquad (2.2.42)$$

and

$$\bar{\beta}^* \simeq p(1-\alpha_s) + \alpha_s\{1-(1-\alpha_I)q^k\} \qquad (2.2.43)$$

The corollary follows immediately on substituting these approximations in (2.2.40). This completes the proof.

Corollary 2.2.3

If $\alpha_I = \alpha_s = \alpha^* = 0$, which is the case when we have no errors in observations, then

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

Proof

The proof follows on substituting $\alpha_I = \alpha_s = \alpha^* = 0$ in the expression for E(R) given in corollary 2.2.2 noting that $\xi = 0$ and using non-orthogonal designs. The value of E(R) given in corollary 2.2.3 coincides with that given in Theorem 2.1.2.

Corollary 2.2.4

For large values of $\frac{\Delta}{\sigma}$ and small values of p, the expected total number of runs in a step-wise group screening design is approximately equal to

$$h + \frac{f}{k} + f\alpha^* + f(1 - \alpha_s)p\left\{1 - \frac{(2-\xi)}{k} - k + \frac{1}{2}(k + 1)(1 - \alpha^*)\right\}$$

$$+ f(1 - \alpha_I)kp + f\alpha_I.$$

Proof

If $\frac{\Delta}{\sigma}$ is large, then α_I, α_s and α^* are relatively small. Thus if p is small, we have

$$1 - (1 - \alpha_I)q^k \approx (1 - \alpha_I)kp + \alpha_I \text{ upto order p} \qquad (2.2.44).$$

The corollary follows immediately on substituting the approximate value given above in the expression for E(R) in corollary 2.2.2 and approximating the resulting expression to terms of order p.

This completes the proof of the corollary.

2.2.2 The optimum size of the group-factor in the initial step

Theorem 2.2.3

Assuming p i.e., the a-priori probability of a factor to be defective to be small, and $\frac{\Delta}{\sigma}$ large, the size 'k' of the group-factor which minimizes the expected total number of runs in a step-wise group screening design with errors in observations is given by

$$k \approx \left[\frac{2 - 2(1-\alpha_s)(2-\xi)p}{2(1-\alpha_I)p - (1-\alpha_s)(1+\alpha^*)p}\right]^{\frac{1}{2}}$$

provided k is real, and the corresponding minimum value of E(R) is given by

$$\min\ E(R) \approx h+2f\left[(1-\alpha_I)p - \frac{1}{2}(1-\alpha_s)(1+\alpha^*)p\right]^{\frac{1}{2}}\left[1-(1-\alpha_s)(2-\xi)p\right]^{\frac{1}{2}}$$

$$+ f\left[\alpha^*+\alpha_I + \frac{1}{2}(1-\alpha_s)(3-\alpha^*)p\right]$$

where α_I is the probability of declaring a non-defective group-factor defective in the initial step, α_s is the probability of declaring a non-defective factor defective in the subsequent steps and α^* is the proportion of group-factors declared defective at any step but due to errors in observations no factor is declared defective on individual tests. The variable ξ takes the value 0 if $\alpha^*=0$ and the value 1 otherwise.

Proof

Assuming continuous variation in k, the optimum size of the group-factor is obtained by solving the equation

$$\frac{d}{dk} E(R) = 0,$$

where E(R) is as given in corollary 2.2.4.

This implies

$$\frac{1}{k^2}\left[1 - (1 - \alpha_s)(2 - \xi)p\right] + \frac{1}{2}(1 - \alpha_s)(1 + \alpha^*)p$$

$$- (1 - \alpha_I)p = 0$$

i.e.

$$k \simeq \left[\frac{2 - 2(1 - \alpha_s)(2 - \xi)p}{2(1-\alpha_I)p - (1-\alpha_s)(1+\alpha^*)p}\right]^{\frac{1}{2}} \qquad (2.2.45)$$

This value of k is real if

$$2(1-\alpha_s)(1-\frac{\xi}{2}) < 1$$

i.e. if

$$p < \frac{1}{2(1-\alpha_s)(1-\frac{\xi}{2})} \qquad (2.2.46)$$

The minimum value of the right hand side in inequality (2.2.46) is $\frac{1}{2}$. This implies that inequality (2.2.46) is true if

$$p < \frac{1}{2} \qquad (2.2.47).$$

Next we show that the value of k given in (2.2.45) is in the neighbourhood of the point of minimum of E(R) given in corollary 2.2.4. This is so if

$$\frac{d^2}{dk^2} E(R) > 0$$

i.e.,

$$\frac{1}{k^3}\left[1-(1-\alpha_s)(2-\xi)p\right] > 0$$

i.e.,

$$p < \frac{1}{2(1-\alpha_s)\left(1-\frac{\xi}{2}\right)}$$

which is condition (2.2.46).

Therefore the value of k given in (2.2.45) is in the neighbourhood of the point of minimum of E(R) given in corollary 2.2.4.

Substituting this value of k in the expression for E(R) in corollary 2.2.4, we obtain

$$\min E(R) \simeq h+2f\left[(1-\alpha_I)p - \frac{1}{2}(1-\alpha_s)(1+\alpha^*)p\right]^{\frac{1}{2}}\left[1-(1-\alpha_s)(2-\xi)p\right]^{\frac{1}{2}}$$

$$+ f\left[\alpha^* + \alpha_I + \frac{1}{2}(1-\alpha_s)(3-\alpha^*)p\right] \qquad (2.2.48).$$

This completes the proof of the Theorem.

Corollary 2.2.5

If $\alpha_I = \alpha_s = \alpha^* = 0$, which is the case when we have no errors in observations, the optimum size of the group-factor is given by

$$k \simeq \left[\frac{2-4p}{p}\right]^{\frac{1}{2}} \qquad \text{provided } p<\tfrac{1}{2},$$

and the corresponding minimum E(R) is given by

$$\min E(R) \simeq 1 + \frac{3fp}{2} + f(2p)^{\frac{1}{2}}(1-2p)^{\frac{1}{2}}.$$

Proof

The proof is obvious on substituting $\alpha_I = \alpha_s = \alpha^* = 0$ in (2.2.45) and in (2.2.48), noting that in this case h = 1 and $\xi = 0$. The values of k and min E(R) in corollary 2.2.5

coincides with those given in Theorem 2.1.3.

Next we wish to obtain the value of k that minimizes E(R) for arbitrary values of p and large $\frac{\Delta}{\sigma}$. For arbitrary values of p and large $\frac{\Delta}{\sigma}$,

$$E(R) \simeq h + \frac{2f}{k} + f - \frac{f(1-\alpha^*)}{k}\left[\frac{1-\left\{1-(1-\alpha_s)p-\alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}^{k+1}}{(1-\alpha_s)p+\alpha_s\left(1-(1-\alpha_I)q^k\right)}\right]$$

$$+ f\left[(1-\alpha_s)p+\alpha_s\{1-(1-\alpha_I)q^k\}\right]\left[1-\frac{2-\xi}{k}-\frac{\alpha^*}{k}\right]$$

$$+ f\left[1-(1-\alpha_I)q^k\right] - f\left[1-\left\{1-(1-\alpha_s)p-\alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}^k\right]$$

$$-\frac{f\alpha^*}{k} - \frac{f}{k}\left\{(1-\alpha_s)p+\alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}^2(\xi-\alpha^*)$$

$$- f\alpha^*\left\{(1-\alpha_s)p+\alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}$$

$$\times\left[1-(1-\alpha_s)p-\alpha_s\left(1-(1-\alpha_I)q^k\right)\right]^k$$

(c.f.corollary 2.2.2).

The value of k that minimizes E(R) given above is a solution of the equation

$$\frac{d}{dk}E(R) = 0$$

i.e.,

$$\frac{1}{k^2}\left[-2+\alpha^*+\left\{(1-\alpha_s)p+\alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}(2-\xi+\alpha^*)+\left\{(1-\alpha_s)p\right.\right.$$

$$\left.\left.+\alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}^2(\xi-\alpha^*)\right]$$

$$- (1-\alpha_I)q^k \ln q \left[-\alpha_s + \frac{(2-\xi)\alpha_s}{k} + \frac{\alpha_s\alpha^*}{k} + 1\right]$$

$$+ k\alpha_s(1-\alpha_I)q^k \ln q \{1-(1-\alpha_s)p - \alpha_s(1-(1-\alpha_I)q^k)\}^{k-1}$$

$$\times \left[1-\alpha^*\{(1-\alpha_s)p + \alpha_s(1-(1-\alpha_I)q^k)\}\right]$$

$$+\alpha_s(1-\alpha_I)q^k \ln q \left[\frac{2(\xi-\alpha_s^*)}{k} \{(1-\alpha_s)p + \alpha_s(1-(1-\alpha_I)q^k)\}\right.$$

$$\left. - \alpha^*\{1-(1-\alpha_s)p - \alpha_s(1-(1-\alpha_I)q^k)\}^k\right]$$

$$+ \frac{(1-\alpha^*)}{k^2}\left[\frac{1-\{1-(1-\alpha_s)p-\alpha_s(1-(1-\alpha_I)q^k)\}^{k+1}}{(1-\alpha_s)p+\alpha_s(1-(1-\alpha_I)q^k)}\right]$$

$$+ \frac{\alpha_s(1-\alpha^*)}{k}\left[\frac{(k+1)\{1-(1-\alpha_s)p-\alpha_s(1-(1-\alpha_I)q^k)\}^k}{(1-\alpha_s)p+\alpha_s(1-(1-\alpha_I)q^k)}\right.$$

$$\left.\times (1-\alpha_I)q^k \ln q\right]$$

$$- \frac{1-\alpha^*}{k}\left[\frac{\alpha_s(1-\alpha_I)q^k \ln q\left[1-\{1-(1-\alpha_s)p-\alpha_s(1-(1-\alpha_I)q^k)\}^{k+1}\right]}{\{(1-\alpha_s)p+\alpha_s(1-(1-\alpha_I)q^k)\}^2}\right]$$

$$= 0 \qquad (2.2.49)$$

The value of k that minimizes E(R) given in corollary 2.2.2 which is a solution to equation (2.2.49) is close to the value of k given in (2.2.45) and can be obtained using Newton - Raphson's iterative method on equation (2.2.49).

CHAPTER III

STEP-WISE GROUP SCREENING DESIGNS WITH UNEQUAL A-PRIORI PROBABILITIES

3.1. SCREENING WITHOUT ERRORS

In chapter 2, we assumed that every factor is defective with the same a-priori probability and thus divided the factors under investigation in the initial step into group-factors of equal sizes. It is quite possible however, that all factors may not be defective with the same a-priori probability. In such a case it is possible using certain criteria, to divide the factors under investigation in the initial step into group-factors of unequal sizes. For example in a manufacturing plant turning out hundreds of items everyday, the probability of the plant producing defective items will vary from time to time due to assignable causes of variation which affect the production. Thus it is reasonable to assume that all items are not defective with the same a-priori probability. Let $p_1, p_2, \ldots, p_g$ ($p_i \leq p$ $i=1,2,\ldots,g$) be a sequence of variables selected in some way from the unit interval (0,1). The p_i's can be selected either by using a systematic procedure or by some random process such as a table of random numbers. For the purpose of dividing the factors into group-factors, we shall identify p_i as the probability that a factor selected at random from the i^{th} group-factor is defective. Thus we have a situation where the factors to be tested in the i^{th} group-factor have a variable probability p_i of being defective. It is expected that this method of grouping the factors such that the factors to be tested in the i^{th} group-factor have probability p_i of being defective could reduce the expected number of runs needed to isolate defective items from the population.

<u>3.1.1</u> <u>The expected number of runs</u>

Let there be 'f' factors divided into 'g' group-factors in the initial step, where 'f' and 'g' are fixed. Let k_i be the number of factors in the i^{th} group-factor in the initial step $(i=1,2,..,g)$.

Then

$$f = \sum_{i=1}^{g} k_i \qquad (3.1.1).$$

Let p_i be the a-priori probability that a factor in the i^{th} group-factor in the initial step is defective $(i=1,2,..,g)$. It is possible to re-order p_i's so that $p_1 \leq p_2 \leq \ldots \leq p_i \leq \ldots \leq p_g \leq p < 1$. The value p could be the probability of factors being defective under the assumption that all factors are defective with the same a-priori probability. If p_i^* is the probability that the i^{th} group-factor of size k_i is defective, and j is the number of defective factors in it, then

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

$$= 1-q_i^{k_i} \qquad (3.1.2)$$

where

$$q_i = 1-p_i \qquad (3.1.3).$$

In the initial step, all the g group-factors are tested for significance. Thus the number of runs required in the initial step is given by .

$$R_I = g + 1 \qquad (3.1.4)$$

where the one extra run is the control run.

Define a random variable U_i such that

$U_i = 1$ with probability p_i^* if the i^{th} group-factor is defective,

$= 0$ otherwise

$(i=1,2,3,\ldots,g)$.

Then

$$E(U_i) = p_i^* = 1-q_i^{k_i} \qquad (3.1.5)$$

Let $P_{k_i}(j)$ denote the probability that the i^{th} defective group-factor contains exactly j defective factors.

Then

$$P_{k_i}(j) = \frac{1}{1-q_i^{k_i}} \binom{k_i}{j} p_i^j (1-p_i)^{k_i-j} \qquad (j=1,2,3,\ldots,k_i) \qquad (3.1.6)$$

Let $E_{k_i}(R_j)$ be the average number of tests (runs) required to analyse the i^{th} group-factor i.e. classify as defective or non-defective all the factors within the i^{th} group-factor of size k_i in the subsequent steps if it contains exactly j defective factors. Then using Theorem 2.1.1, we get

$$E_{k_i}(R_j) = \frac{jk_i}{j+1} + j + \frac{j}{j+1} - \frac{2j}{k_i} \qquad \begin{matrix}(j=1,2,\ldots,k_i)\\(i=1,2,\ldots,g)\end{matrix} \qquad (3.1.7)$$

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

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

$$= \frac{1}{1-q_i^{k_i}} \sum_{j=1}^{k_i} \left\{ \frac{jk_i}{j+1} + j + \frac{j}{j+1} - \frac{2j}{k_i} \right\} \binom{k_i}{j} p_i^j q_i^{k_i-j}$$

$$= \frac{1}{1-q_i^{k_i}} \left[(k_i+1)+k_i p_i - 2p_i - \frac{1}{p_i}\{1-q_i^{k_i+1}\} \right]$$

(3.1.8) [c.f. (2.1.26)].

Let R_S denote the number of tests required to analyse all the group-factors found to be defective in the initial step. Then

$$R_S = \sum_{i=1}^{g} E(R_{Si})\, U_i$$

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

(3.1.9).

Theorem 3.1.1

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

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

Proof

The number of runs required in the initial step is

$$R_I = 1+g \tag{3.1.10}$$

In the subsequent steps, we require

$$R_S = \sum_{i=1}^{g} \frac{1}{1-q_i^{k_i}} \left[(k_i+1)+k_ip_i-2p_i-\frac{1}{p_i}\{1-q_i^{k_i+1}\}\right]U_i \tag{3.1.11}$$

runs.

This implies that

$$E(R_S) = \sum_{i=1}^{g} \frac{1}{1-q_i^{k_i}} \left[(k_i+1)+k_ip_i-2p_i-\frac{1}{p_i}\{1-q_i^{k_i+1}\}\right]E(U_i)$$

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

using (3.1.5). Hence,

$$E(R_S) = g+f+\sum_{i=1}^{g} \left[k_ip_i-2p_i-\frac{1}{p_i}\{1-q_i^{k_i+1}\}\right] \tag{3.1.12}$$

The expected total number of runs is given by

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

$$= 1+2g+f+\sum_{i=1}^{g} \left[k_ip_i-2p_i-\frac{1}{p_i}\{1-q_i^{k_i+1}\}\right] \tag{3.1.13}$$

This proves the theorem.

Corollary 3.1.1

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

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

Proof

For small p_i i.e., the a-priori probability of a factor in the i^{th} group-factor to be defective,

$$\frac{1}{p_i}\left[1-q_i^{k_i+1}\right] \simeq \frac{1}{p_i}\left[(k_i+1)p_i - \frac{(k_i+1)k_i}{2}p_i^2 + \frac{(k_i+1)k_i(k_i-1)}{2\times 3}p_i^3 - \cdots\right]$$

$$\simeq k_i+1-\frac{k_i^2+k_i}{2}p_i \text{ upto order } p_i \tag{3.1.14}$$

Using (3.1.14) in (3.1.13) we get

$$E(R) \simeq 1+2g+f+\sum_{i=1}^{g}\left[k_i p_i - 2p_i - k_i - 1 + \frac{k_i^2 p_i}{2} + \frac{k_i p_i}{2}\right]$$

$$= 1+g-2\sum_{i=1}^{g} p_i + \frac{3}{2}\sum_{i=1}^{g} k_i p_i + \frac{1}{2}\sum_{i=1}^{g} k_i^2 p_i \tag{3.1.15}$$

This completes the proof of the corollary.

3.1.2 The optimum sizes of the group-factors in the initial step

Theorem 3.1.2

Assuming p_i i.e., the a-priori probability of a factor in the i^{th} group-factor 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 design is given by

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

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

$$\text{Min } E(R) \simeq 1+g - \frac{25}{8}\sum_{i=1}^{g} p_i + \frac{1}{8}(3g+2f)^2 \frac{1}{\sum_{i=1}^{g} \frac{1}{p_i}} .$$

Proof

The problem is to obtain k_i's which minimize the expected total number of runs given in corollary 3.1.1 subject to the condition

$$f = \sum_{i=1}^{g} k_i$$

The condition above, implies that

$$k_g = f - k_1 - k_2 - \cdots - k_{g-1} .$$

Substituting for k_g in (3.1.15), we get

$$E(R) = F(k_1, k_2, k_3, \ldots, k_{g-1})$$

$$\simeq 1+g-2\sum_{i=1}^{g} p_i + \frac{3}{2}\Big[k_1p_1+k_2p_2+\cdots+k_{g-1}p_{g-1} + (f-k_1-k_2-\cdots-k_{g-1})p_g\Big]$$
$$+ \frac{1}{2}\Big[k_1^2p_1+k_2^2p_2+\cdots+k_{g-1}^2p_{g-1} + (f-k_1-k_2-\cdots-k_{g-1})^2p_g\Big] . \qquad (3.1.16).$$

Assuming continuous variations in k_i's, critical values of k_i's are obtained by solving the equations

$$\frac{\partial F}{\partial k_i} = 0 \qquad (i=1,2,\ldots,g-1) \qquad (3.1.17),$$

which imply

$$\frac{3}{2}(p_i-p_g) + k_ip_i-(f-k_1-k_2-\dots-k_{g-1})p_g = 0$$

$$i=1,2,\dots,g-1$$

i.e.,

$$k_ip_i-(f-k_1-k_2-\dots k_{g-1})p_g=\frac{3}{2}(p_g-p_i)$$

which imply

$$(k_i+\frac{3}{2})p_i = (k_g+\frac{3}{2})p_g$$

$$i=1,2,\dots,g-1 \qquad (3.1.18).$$

Equations (3.1.18) imply

$$\frac{(k_i+\frac{3}{2})}{1/p_i} = \frac{(k_g+\frac{3}{2})p_g}{1/p_g}$$

$$i=1,2,\dots,g-1$$

$$(3.1.19).$$

i.e.,

$$\frac{(k_1+\frac{3}{2})}{1/p_1} = \frac{(k_2+\frac{3}{2})}{1/p_2} = \dots = \frac{(k_{g-1}+\frac{3}{2})}{1/p_{g-1}} = \frac{(k_g+\frac{3}{2})}{1/p_g}$$

$$= \frac{\frac{3g}{2} + \sum_{i=1}^{g} k_i}{\sum_{i=1}^{g} \frac{1}{p_i}} = \frac{\frac{3g}{2} + f}{\sum_{i=1}^{g} \frac{1}{p_i}} \qquad (3.1.20)$$

which gives

$$k_i \approx (\frac{3}{2}g + f)\frac{1}{p_i \sum_{i=1}^{g} \frac{1}{p_i}} - \frac{3}{2} \quad (i=1,2,\dots,g)$$

$$(3,1.21).$$

We now wish to show that the values of k_i's given in (3.1.21) are in the neighbourhood of points of minimum of E(R). This will be so if the second order derivative matrix

$$D = \left(\left(\frac{\partial^2 F}{\partial k_i \partial k_j}\right)\right) \quad \text{of dimension}$$

(g-1) x (g-1) is positive definite for values of k_i's given in (3.1.21) where

$$\frac{\partial^2 F}{\partial k_i^2} = p_i + p_g \qquad (i=1,2,\ldots,g-1)$$

and

$$\frac{\partial^2 F}{\partial k_i \partial k_j} = p_g \qquad (i \neq j = 1,2,\ldots,g-1).$$

Hence

$$D = \mathrm{Diag}(p_1, p_2, \ldots, p_{g-1}) + p_g J_{g-1}$$

where J_{g-1} is a (g-1) x (g-1) matrix of ones.
The matrix D is positive definite since all the elements along the leading diagonal are of the form

$p_i + p_g > 0$, $i=1,2,\ldots,g-1$; furthermore

$$\begin{vmatrix} p_1+p_g & p_g \\ p_g & p_2+p_g \end{vmatrix} = p_1p_2+p_1p_g+p_2p_g > 0,$$

$$\begin{vmatrix} p_1+p_g & p_g & p_g \\ p_g & p_2+p_g & p_g \\ p_g & p_g & p_3+p_g \end{vmatrix} = p_1p_2p_3+p_1p_2p_g+p_1p_3p_g+p_2p_3p_g > 0,$$

and in general

$$|D| = p_1p_2p_3\cdots p_{g-1}+p_2p_3\cdots p_{g-1}p_g+p_1p_3p_4\cdots p_{g-1}p_g+\cdots$$
$$+p_1p_2p_3\cdots p_{i-1}p_{i+1}p_{i+2}\cdots p_{g-1}p_g+\cdots$$
$$+p_1p_2p_3\cdots p_{g-2}p_g > 0.$$

Hence the value of k_i's given in (3.1.21) are in the neighbourhood of points of minimum for E(R) in corollary 3.1.1. Substituting these values of k_i's in the formula for E(R) given in corollary 3.1.1, we obtain

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

$$= 1+g-\frac{25}{8}\sum_{i=1}^{g} p_i+\frac{1}{8}(3g+2f)^2\sum_{i=1}^{g}\frac{1}{p_i\left(\sum_{i=1}^{g}\frac{1}{p_i}\right)^2}$$

i.e.,

$$\min E(R) \simeq 1+g-\frac{25}{8}\sum_{i=1}^{g} p_i+\frac{1}{8}(3g+2f)\frac{1}{\sum_{i=1}^{g}\frac{1}{p_i}} \qquad (3.1.22).$$

This completes the proof of Theorem 3.2.1.

In the theorem that follows, we shall show that min E(R) given in (3.1.22) is less than or equal to min E(R) given in (2.1.36a) under the assumption $p_1=p_2=\cdots=p_g=p$.

Theorem 3.1.3

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

Proof

The problem is to show that min E(R) given in Theorem 3.1.2 is less than or equal to min E(R) given in (2.1.36a). That is we show that

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

$$\leq 1+\frac{3}{2}fp+\frac{fp^{\frac{1}{2}}}{(2-4p)^{\frac{1}{2}}}-\frac{2fp^{3/2}}{(2-4p)^{\frac{1}{2}}}+\frac{f}{2}(2-4p)^{\frac{1}{2}}p^{\frac{1}{2}}$$

(3.1.23).

Substituting $g = \frac{f}{k}$ where $k = \left(\frac{2-4p}{p}\right)^{\frac{1}{2}}$ as given in (2.1.35), inequality (3.1.23) becomes

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

$$\leq 1+ \frac{3}{2}fp + g - 2gp + \frac{f^2 p}{2g}$$

i.e.,

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

$$\leq -\frac{25}{8}pg + \frac{1}{8}(3g + 2f)^2 \frac{p}{g} \qquad (3.1.24).$$

i.e.,

$$\frac{25}{8}(gp - \sum_{i=1}^{g} p_i) + \frac{1}{8}\left(3g + 2f\right)^2 \left(\frac{1}{\sum_{i=1}^{g} \frac{1}{p_i}} - \frac{p}{g}\right) \leq 0$$

(3.1.25),

which is true if

$$\frac{25}{8}\left(gp - \sum_{i=1}^{g} p_i\right) + \frac{25}{8}g^2\left(\frac{1}{\sum_{i=1}^{g} \frac{1}{p_i}} - \frac{p}{g}\right) \leq 0$$

i.e.,

$$-\sum_{i=1}^{g} p_i + \frac{g^2}{\sum_{i=1}^{g} \frac{1}{p_i}} \leq 0$$

i.e.

$$g^2 \leq \sum_{i=1}^{g} p_i \sum_{i=1}^{g} \frac{1}{p_i} \qquad (3.1.26),$$

which follows from Cauchy - Schwarz inequality

$$\left[\sum_{i=1}^{g} (p_i)^{\frac{1}{2}}\left(\frac{1}{p_i}\right)^{\frac{1}{2}}\right]^2 \leq \sum_{i=1}^{g} p_i \sum_{i=1}^{g} \frac{1}{p_i}$$

This proves the Theorem.

3.2. SCREENING WITH ERRORS

The problem of step-wise group screening with unequal a-priori probabilities of factors to be defective has been considered in section 3.1. While developing the theory, we assumed that there were no errors in observations. In this section, we shall allow errors in observations and work out corresponding results given in section 3.1. As in section 3.1, factors with the same a-priori probability of being defective will be put together in the same group-factor in the initial step, thus resulting in group-factors of unequal sizes in the initial step. For the purpose of experimentation, we shall follow the method described in section 2.2.

3.2.1. The expected number of runs

Let there be 'f' factors under investigation. In the initial step, the 'f' factors are divided into 'g' group-factors such that all the factors with the same a-priori probability of being defective are put in the same group-factor. Suppose the i^{th} group-factor has k_i factors, then we have

$$\sum_{i=1}^{g} k_i = f \qquad (i=1,2,..,g) \qquad (3.2.1),$$

where f and g are fixed.

Let $\hat{A}_i$ be the estimate of the main effect of the i^{th} group-factor containing 's_i' defective factors. If the effect of a defective factor within the i^{th} group-factor is $\Delta_i > 0$ $(i=1,2,\ldots,g)$, then

$$E(\hat{A}_i) = s_i \Delta_i \tag{3.2.2a}$$

and

$$\text{Var}\,(\hat{A}_i) = \frac{\sigma^2}{g+h} \qquad (h=1,2,3,4) \tag{3.2.2b}$$

Next define

$$z_i = \frac{\hat{A}_i - s_i \Delta_i}{\sqrt{\sigma^2/(g+h)}}$$

$$= y_i - s_i \phi_{Ii} \tag{3.2.3}$$

where

$$y_i = \frac{\hat{A}_i}{\sqrt{\sigma^2/(g+h)}} \tag{3.2.4}$$

and

$$\phi_{Ii} = \frac{\Delta_i}{\sqrt{\sigma^2/g+h}} \tag{3.2.5}$$

Assuming observations to be normal, z_i is a standard normal variate. We shall say that the i^{th} group-factor is non-defective if $s_i=0$, which implies that $s_i\phi_{Ii} = 0$. It is defective if $s_i\phi_{Ii} \neq 0$. Thus we wish to test the hypothesis

$$H_0: s_i\phi_{Ii} = 0$$

alternative (3.2.6)

$$H_1: s_i\phi_{Ii} \neq 0$$

Assuming σ is known, we shall use the normal deviate test, otherwise we would use a corresponding t-test.

The power of the test for the i^{th} group-factor is

$$\Pi_{Ii}(s_i\phi_{Ii},\ \alpha_{Ii}) = \int_{z(\alpha_{Ii})-s_i\phi_{Ii}}^{\infty} \frac{1}{\sqrt{2\Pi}} \exp\left(-\frac{z^2}{2}\right) dz \qquad (3.2.7)$$

where $z(\alpha_{Ii})$ is given by

$$\alpha_{Ii} = \int_{z(\alpha_{Ii})}^{\infty} \frac{1}{\sqrt{2\Pi}} \exp\left(-\frac{z^2}{2}\right) dz \qquad (3.2.8)$$

which is the size of the critical region for testing the significance of the i^{th} group-factor. When $s_i=0$ or $\frac{\Delta_i}{\sigma} = 0$, we have

$$\Pi_{Ii}(0,\alpha_{Ii}) = \alpha_{Ii} \qquad (3.2.9).$$

When $s_i \neq 0$ and $\frac{\Delta_i}{\sigma}$ is large, then we have

$$\Pi_{Ii}(s_i\phi_{Ii},\alpha_{Ii}) \approx 1 \qquad (3.2.10).$$

Let p_i $(i=1,2,...,g)$, be the a-priori probability that a factor within the i^{th} group-factor is defective. Then the probability that the i^{th} group-factor containing s_i defective factors is declared defective is given by

$$\Pi^{*}_{Ii} = \sum_{s_i=0}^{k_i} \binom{k_i}{s_i} p_i^{s_i} \left(1-p_i\right)^{k_i-s_i} \Pi_{Ii}(s_i\phi_{Ii},\alpha_{Ii}) \qquad (3.2.11).$$

Define a random variable U_i such that

$U_i=1$ with probability π^*_{Ii} if the i^{th} initial group-factor is declared defective,

$= 0$ otherwise

$i=1,2,\ldots,g.$ (3.2.12a).

Then

$$E(U_i) = \pi^*_{Ii} \qquad (3.2.12b).$$

In the subsequent steps, the analysis of the group-factors declared defective in step one is continued as described in section 2.2.

Let p_i' be the probability that a factor chosen at random from the i^{th} group-factor containing s_i defective factors that has been declared defective in step one is defective. Then

$$p_i' = \frac{p_i}{\pi^*_{Ii}} \sum_{s_i=1}^{k_i} \binom{k_i-1}{s_i-1} p_i^{s_i-1} q_i^{k_i-s_i} \pi_{Ii}(s_i\phi_{Ii},\alpha_{Ii})$$

$$= \frac{p_i \pi^+_{Ii}}{\pi^*_{Ii}} \qquad (3.2.13)$$

where

$$\pi^+_{Ii} = \sum_{s_i=1}^{k_i} \binom{k_i-1}{s_i-1} p_i^{s_i-1} q_i^{k_i-s_i} \pi_{Ii}(s_i\phi_{Ii},\alpha_{Ii})$$

(3.2.14)

following Curnow [2].

Define a random variable δ_i as follows;-

$\delta_i=0$ if a factor chosen at random from a group-factor that is declared defective in step one is non-defective

= 1 otherwise.

Then

$$\delta_i = 0 \text{ with probability } 1-p_i'$$
$$= 1 \text{ with probability } p_i' .$$

Let α_{si} be the probability of declaring a non-defective factor from the i^{th} group-factor defective and γ_{si} be the probability of declaring a defective factor from the i^{th} group factor defective in the subsequent steps.

Let

$$\beta_i^+ = (1-\delta_i)\alpha_{si} + \delta_i\gamma_{si} .$$

Then

$$\beta_i^+ = \alpha_{si} \quad \text{with probability } 1-p_i'$$
$$= \gamma_{si} \quad \text{with probability } p_i' \qquad (3.2.15).$$

Thus the average value of β_i^+ is given by

$$\bar{\beta}_i^+ = \gamma_{si}p_i' + \alpha_{si}(1-p_i')$$

$$= \frac{1}{\Pi_{Ii}^*}\left[p_i(\gamma_{si} - \alpha_{si})\Pi_{Ii}^+ + \Pi_{Ii}^*\alpha_{si}\right]$$

$$= \frac{\bar{\beta}_i^*}{\Pi_{Ii}^*} \qquad (3.2.16a)$$

where

$$\bar{\beta}_i^* = \left[p_i(\gamma_{si} - \alpha_{si})\Pi_{Ii}^+ + \Pi_{Ii}^*\alpha_{si}\right] \qquad (3.2.16b)$$

is the probability that a factor from the i^{th} initial step group-factor is declared defective in the subsequent steps. Thus $\bar{\beta}_i^+$ may be interpretted as the conditional probability that a factor chosen at random from the i^{th} group-factor that is declared defective in the initial step is declared defective.

Let α^*_{si} be the probability that a group-factor consisting of factors from the i^{th} initial group-factor is declared defective at any step but on testing individual factors within it, no factor is declared defective due to errors in observations. Obviously α^*_{si} will take different values at different steps. However for simplicity in algebra, we shall assume α^*_{si} to be of uniform value, say α^*_i .

Denote by $P^*_{k_i}(j)$ the probability that exactly j factors from the i^{th} group-factor that is declared defective in step one are declared defective in the subsequent steps. Then

$$P^*_{k_i}(j) = \begin{cases} 1 - \dfrac{1}{\Pi^*_{Ii}}\left[1-(1-\bar{\beta}^*_i)^{k_i}\right] & j=0 \\ \dfrac{1}{\Pi^*_{Ii}}\dbinom{k_i}{j}\bar{\beta}^{*j}(1-\bar{\beta}^*_i)^{k_i-j} & j=1,2,\ldots,k_i \end{cases} \qquad (3.2.17)$$

Let $E^*_{k_i}(R_j)$ be the expected number of runs required to analyse the i^{th} group-factor, i.e. declare as defective or non-defective all the factors within the i^{th} group-factor of size k_i which has been declared defective, if exactly j factors from it are declared defective in the subsequent steps. Then using lemma 2.2.1 and Theorem 2.2.1 we have

$$E^*_{k_i}(R_j) = k_i \qquad \text{for } j=0 \qquad (3.2.18a)$$

and

$$E^*_{k_i}(R_j) = \frac{jk_i}{j+1} + j + \frac{j}{j+1} - \frac{j(k_i+j-2)}{k_i(k_i-1)}$$
$$+ \alpha^*_i \left\{ \frac{k_i}{j+1} - \frac{j}{j+1} - \frac{j}{k_i-1} + \frac{j^2}{k_i(k_i-1)} \right\}$$
$$- \frac{j(1-\xi_i)(k_i-j)}{k_i(k_i-1)} \qquad \text{for } j=1,2,\ldots,k \qquad (3.2.18b)$$

where $\xi_i=0$ if $\alpha^*_i=0$ and 1 otherwise.

Denote by R_{Si} the number of tests (runs) required to analyse the i^{th} group-factor once it has been declared defective in the initial step. Then

$$E(R_{Si}) = \sum_{j=0}^{k_i} E^*_{k_i}(R_j)P^*_{k_i}(j)$$

$$= k_i - \frac{k_i}{\Pi^*_{Ii}}\left[1-(1-\bar{\beta}^*_i)^{k_i}\right]$$

$$+ \frac{1}{\Pi^*_{Ii}} \sum_{j=1}^{k_i} \left[\left\{ \frac{jk_i}{j+1} + j + \frac{j}{j+1} - \frac{j(k_i+j-2)}{k_i(k_i-1)} \right\} \times \binom{k_i}{j} (\bar{\beta}^*_i)^j (1-\bar{\beta}^*_i)^{k_i-j} \right]$$

$$+ \frac{\alpha^*_i}{\Pi^*_{Ii}} \sum_{j=1}^{k_i} \left[\left\{ \frac{k_i}{j+1} - \frac{j}{j+1} - \frac{j}{k_i-1} + \frac{j^2}{k_i(k_i-1)} \right\} \right.$$

$$\times \binom{k_i}{j} (\bar{\beta}_i^*)^j (1-\bar{\beta}_i^*)^{k_i-j} \Bigg]$$

$$- \frac{(1-\xi_i)}{\pi_{Ii}^*} \sum_{j=1}^{k_i} \frac{j(k_i-j)}{k_i(k_i-1)} \binom{k_i}{j} (\bar{\beta}_i^*)^j (1-\bar{\beta}_i^*)^{k_i-j}$$

i.e.

$$E(R_{Si}) = k_i - \frac{k_i}{\pi_{Ii}^*} \left[1-(1-\bar{\beta}_i^*)^{k_i}\right]$$

$$+ \frac{1}{\pi_{Ii}^*}\left[k_i+1+k_i\bar{\beta}_i^* - \frac{1}{\bar{\beta}_i^*}\{1-(1-\bar{\beta}_i^*)^{k_i+1}\}\right]$$

$$- \frac{1}{\pi_{Ii}^*}\left[(2-\xi_i)\bar{\beta}_i^* + \xi_i \bar{\beta}_i^{*2}\right]$$

$$+ \frac{\alpha_i^*}{\pi_{Ii}^*}\left[\frac{1}{\bar{\beta}_i^*}\{1-(1-\bar{\beta}_i^*)^{k_i+1} - k_i\bar{\beta}_i^*(1-\bar{\beta}_i^*)^{k_i}\} -1+\bar{\beta}_i^{*2}-\bar{\beta}_i^*\right] \quad (3.2.19),$$

using (2.2.35).

Let R_S denote the number of tests required to analyse all the group-factors declared defective in step one. Then

$$R_S = \sum_{i=1}^{g} U_i E(R_{Si}) \quad (3.2.20)$$

where U_i is as already defined in (3.2.12a).

<u>Theorem 3.2.1</u>

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 $(i=1,2,...,g)$ and $\bar{\beta}_i^*$

is the probability of declaring a factor within the i^{th} group-factor defective in the subsequent steps is given by

$$E(R) = h+f+2g - \sum_{i=1}^{g} \frac{1}{\bar{\beta}_i^*}(1-\alpha_i^*)\left\{1-(1-\bar{\beta}_i^*)^{k_i+1}\right\}$$

$$+ \sum_{i=1}^{g} k_i \bar{\beta}_i^* \left\{1- \frac{2-\xi_i}{k_i} - \frac{\alpha_i^*}{k_i}\right\} + \sum_{i=1}^{g} k_i \pi_{Ii}^*$$

$$- \sum_{i=1}^{g} k_i \left[1-(1-\bar{\beta}_i^*)^{k_i}\right] - \sum_{i=1}^{g} \alpha_i^* + \sum_{i=1}^{g} (\alpha_i^* - \xi_i)\bar{\beta}_i^{*2}$$

$$- \sum_{i=1}^{g} k_i \; \alpha_i^* \; \bar{\beta}_i^*(1-\bar{\beta}_i^*)^{k_i} .$$

where α_i^* is the probability that a group-factor consisting of factors from the i^{th} initial group-factor is declared defective but on testing individual factors within it, no factor is declared defective due to errors in observations and $\xi_i=0$ if $\alpha_i^*=0$ and 1 otherwise.

<u>Proof</u>

In step one, we require

$$R_I = h+g \text{ runs } (h=1,2,3,4). \qquad (3.2.21)$$

The number of runs required in the subsequent steps is

$$R_S = \sum_{i=1}^{g} U_i E(R_{Si}) \qquad \text{as given in } (3.2.20).$$

Using (3.2.19), it follows that

$$E(R_S) = \sum_{i=1}^{g} k_i \Pi_{Ii}^* - \sum_{i=1}^{g} k_i \left[1-(1-\bar{\beta}_i^*)^{k_i}\right]$$

$$+ \sum_{i=1}^{g} \left[k_i + 1 + k_i \bar{\beta}_i^* - \frac{1}{\bar{\beta}_i^*}\{1-(1-\bar{\beta}_i^*)^{k_i+1}\}\right]$$

$$- \sum_{i=1}^{g} \left[(2-\xi_i)\bar{\beta}_i^* + \xi_i \bar{\beta}_i^{*2}\right]$$

$$+ \sum_{i=1}^{g} \alpha_i^* \left[\frac{1}{\bar{\beta}_i^*}\{1-(1-\bar{\beta}_i^*)^{k_i+1} - k_i\bar{\beta}_i^*(1-\bar{\beta}_i^*)^{k_i}\}\right.$$

$$\left. - 1 + \bar{\beta}_i^{*2} - \bar{\beta}_i^*\right] \quad (3.2.22),$$

after replacing U_i by $E(U_i) = \Pi_{Ii}^*$ given in (3.2.12b).

The expected total number of runs is now given by

$$E(R) = R_I + E(R_S) \qquad (3.2.23).$$

Using (3.2.21) and (3.2.22) in (3.2.23), putting the like terms together, we obtain the expression for E(R) given in the Theorem. This completes the proof.

Corollary 3.2.1

For large values of $\frac{\Delta_i\text{'s}}{\sigma}$ and arbitrary values of p_i's, the expected total number of runs in a step-wise group screening design with the i^{th} group-factor of size k_i $(i=1,2,\ldots,g)$ is approximately equal to

$$h+f+2g-\sum_{i=1}^{g} \frac{(1-\alpha_i^*)\left[1-\left\{1-(1-\alpha_{si})p_i-\alpha_{si}\left(1-(1-\alpha_{Ii})q_i^{k_i}\right)\right\}^{k_i+1}\right]}{p_i(1-\alpha_{si}) + \alpha_{si}\{1-(1-\alpha_{Ii})q_i^{k_i}\}}$$

$$+ \sum_{i=1}^{g} k_i \left\{ p_i(1-\alpha_{si}) + \alpha_{si}\left(1-(1-\alpha_{Ii})q_i^{k_i}\right)\right\}\left\{1 - \frac{2-\xi_i}{k_i} - \frac{\alpha_i^*}{k_i}\right\}$$

$$+ \sum_{i=1}^{g} k_i \left\{1-(1-\alpha_{Ii})q_i^{k_i}\right\}$$

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

$$- \sum_{i=1}^{g} \alpha_i^* + \sum_{i=1}^{g} (\alpha_i^* - \xi_i)\left\{p_i(1-\alpha_{si}) - \alpha_{si}\left(1-(1-\alpha_{Ii})q_i^{k_i}\right)\right\}^2$$

$$- \sum_{i=1}^{g} \left[k_i \alpha_i^* \left\{(1-\alpha_{si})p_i - \alpha_{si}\left(1-(1-\alpha_{Ii})q_i^{k_i}\right)\right\}\right.$$

$$\left. \times \left\{1-(1-\alpha_{si})p_i - \alpha_{si}\left(1-(1-\alpha_{Ii})q_i^{k_i}\right)\right\}^{k_i}\right]$$

Proof

If $\frac{\Delta_i\text{'s}}{\sigma}$ are large, we have the following approximations

$$\Pi_{Ii}^* \simeq 1 - (1-\alpha_{Ii})q_i^{k_i} \qquad (3.2.24)$$

$$\Pi_{Ii}^+ \simeq 1 \qquad (3.2.25)$$

$$\gamma_{si} \simeq 1 \qquad (3.2.26)$$

and

$$\bar{\beta}_i^* \simeq (1-\alpha_{si})p_i + \alpha_{si}\left[1-(1-\alpha_{Ii})q_i^{k_i}\right] \qquad (3.2.27).$$

The corollary follows immediately on using these approximations in the expression for E(R) given in Theorem 3.2.1.

This completes the proof.

Corollary 3.2.2

For large values of $\frac{\Delta_i\text{'s}}{\sigma}$ and small values of p_i's, the expected total number of runs in a step-wise group screening design with g group-factors in the initial step, the i^{th} group-factor being of size $k_i (i=1,2,...,g)$ is approximately equal to

$$h+g+\sum_{i=1}^{g} k_i\alpha_i^* + \sum_{i=1}^{g}(1-\alpha_{si})k_i p_i\left\{1-\frac{2-\xi_i}{k_i}-k_i+\frac{1}{2}(1-\alpha_i^*)(k_i+1)\right\}$$

$$+\sum_{i=1}^{g}\alpha_{Ii}k_i+\sum_{i=1}^{g}k_i^2(1-\alpha_{Ii})p_i$$

Proof

If $\frac{\Delta_i\text{'s}}{\sigma}$ are large, then α_{Ii}'s , α_{si}'s and α_i^*'s are relatively small. Hence if p_i's are small, we have

$$1-(1-\alpha_{Ii})q_i^{k_i} \approx (1-\alpha_{Ii})k_i p_i + \alpha_{Ii}$$

upto order p_i (3.2.28).

The corollary follows immediately on substituting the approximate value given in (3.2.28) in corollary 3.2.1, approximating the resulting expression to terms of order p_i and rearranging similar terms. This completes the proof.

3.2.2 The optimum sizes of the group-factors in the initial step

Theorem 3.2.2

For large values of $\frac{\Delta_i\text{'s}}{\sigma}$ and small values of

p_i's, where p_i is the a-priori probability of a factor within the i^{th} group-factor to be defective, the size 'k_i' of the i^{th} group-factor that minimizes the expected number of runs is given by

$$k_i \approx \left(f + \sum_{i=1}^{g} \frac{\left[\alpha_i^* + \alpha_{Ii} + \frac{1}{2}(1-\alpha_{si})(3-\alpha_i^*)p_i\right]}{\left[2(1-\alpha_{Ii})-(1-\alpha_{si})(1+\alpha_i^*)\right]p_i} \right)$$

$$\times \frac{1}{\left[2(1-\alpha_{Ii})-(1-\alpha_{si})(1+\alpha_i^*)\right]p_i \sum_{i=1}^{g} \frac{1}{\left[2(1-\alpha_{Ii})-(1-\alpha_{si})(1+\alpha_i^*)\right]p_i}}$$

$$- \frac{\left[\alpha_i^* + \alpha_{Ii} + \frac{1}{2}(1-\alpha_{si})(3-\alpha_i^*)p_i\right]}{\left[2(1-\alpha_{Ii})-(1-\alpha_{si})(1+\alpha_i^*)\right]p_i} ,$$

where α_{Ii} is the probability of declaring the i^{th} non-defective group-factor defective, α_i^* is the probability that a group-factor consisting of factors from the i^{th} initial step group-factor is declared defective at any step but on testing individual factors within it, no factor is declared defective due to errors in observations and α_{si} is the probability of declaring a non-defective factor from the i^{th} group-factor defective.

Proof

The problem is to minimize E(R) given in corollary 3.2.2, subject to the condition

$$\sum_{i=1}^{g} k_i = f.$$

By the method of Lagrange's multiplier, let

$$F(k_1,k_2,\ldots,k_g,\lambda)$$

$$= h+g+\sum_{i=1}^{g} k_i\alpha_i^* + \sum_{i=1}^{g} k_i\alpha_{Ii}$$

$$+ \sum_{i=1}^{g} (1-\alpha_{si})k_ip_i\left[1-\frac{(2-\xi_i)}{k_i} - k_i + \frac{1}{2}(1-\alpha_i^*)(k_i+1)\right]$$

$$+ \sum_{i=1}^{g} k_i^2(1-\alpha_{Ii})p_i + \lambda\left(f-\sum_{i=1}^{g} k_i\right)$$

where λ is the Lagrange's multiplier.

Assuming continuous variation in k_i, the critical values of k_i are obtained by solving the equations

$$\frac{\partial F}{\partial k_i} = 0 \qquad (i=1,2,..,g)$$

and (3.2.29).

$$\frac{\partial F}{\partial \lambda} = 0$$

Conditions (3.2.29) imply

$$\left[2(1-\alpha_{Ii})-2(1-\alpha_{si})+(1-\alpha_{si})(1-\alpha_i^*)\right]k_ip_i$$

$$+\left[\alpha_i^*+\alpha_{Ii}+(1-\alpha_{si})p_i+\frac{1}{2}(1-\alpha_{si})(1-\alpha_i^*)p_i\right]$$

$$-\lambda = 0 \qquad (3.2.30)$$

and

$$\sum_{i=1}^{g} k_i = f \qquad (3.2.31)$$

From (3.2.30) we get

$$k_i = \frac{\lambda - \left[\alpha_i^* + \alpha_{Ii} + \frac{1}{2}(1-\alpha_{si})(3-\alpha_i^*)p_i\right]}{\left[2(1-\alpha_{Ii}) - (1-\alpha_{si})(1+\alpha_i^*)\right]p_i} \tag{3.2.32}$$

Summing (3.2.32) over i and solving for λ we get

$$\lambda = \left(f + \sum_{i=1}^{g} \frac{\left[\alpha_i^* + \alpha_{Ii} + \frac{1}{2}(1-\alpha_{si})(3-\alpha_i^*)p_i\right]}{\left[2(1-\alpha_{Ii}) - (1-\alpha_{si})(1+\alpha_i^*)\right]p_i} \right)$$

$$\times \frac{1}{\sum_{i=1}^{g} \frac{1}{\left[2(1-\alpha_{Ii}) - (1-\alpha_{si})(1+\alpha_i^*)\right]p_i}} \tag{3.2.33}$$

The Theorem follows immediately on substituting this value of λ in (3.2.32).

This completes the proof.

<u>Corollary 3.2.3</u>

If $\alpha_{Ii} = \alpha_{si} = \alpha_i^* = 0$, which is the case when we have no errors in observations,

$$k_i = \left(f + \frac{3g}{2}\right) \frac{1}{p_i \sum_{i=1}^{g} \frac{1}{p_i}} - \frac{3}{2}$$

<u>Proof</u>

The proof is obvious on substituting $\alpha_{si} = \alpha_i^* = \alpha_{Ii} = 0$ in the expression for k_i given in Theorem 3.2.2.

The value of k_i given in corollary 3.2.3 coincides with that given in Theorem 3.1.2.

CHAPTER IV

INCORRECT DECISIONS IN STEP-WISE GROUP SCREENING DESIGNS

4.1. SCREENING WITH EQUAL A-PRIORI PROBABILITIES

In section 2.2, we worked out the value of the optimum size of the initial group-factor taking into consideration only the expected total number of runs. In this section, we shall work out the value of the optimum size of the group-factor in the initial step, taking into consideration both the expected total number of runs and the expected total number of incorrect decisions.

4.1.1 Calculation of the expected total number of incorrect decisions

We shall consider the following cases of incorrect decisions;

(i) declaring defective factors as non-defective in the initial step,

(ii) declaring defective factors as non-defective in the subsequent steps and

(iii) declaring non-defective factors as defective in the subsequent steps.

Let $P_k^*(j)$ be the probability that exactly j factors are declared defective in the subsequent steps from a group-factor of size k that is declared defective in the initial step.

Then

$$P_k^*(j) = \begin{cases} 1 - \frac{1}{\pi_I^*} \{1-(1-\bar{\beta}^*)^k\} & j=0 \\ \frac{1}{\pi_I^*}\binom{k}{j} \bar{\beta}^{*\,j} (1-\bar{\beta}^*)^{k-j} & j=1,2,\ldots,k \end{cases}$$

as explained in (2.2.21).

Let $E_k(j)$ denote the expected number of factors declared defective from a group-factor that was declared defective in the initial step.

Then

$$E_k(j) = \frac{1}{\pi_I^*} k \bar{\beta}^*$$

$$= k \bar{\beta}^+ \qquad (4.1.1),$$

where $\bar{\beta}^+$ is the conditional probability that a factor chosen at random from a group-factor of size 'k' that is declared defective in the initial step is declared defective.

Let $p^{(o)}$ be the probability that a factor chosen at random from a group-factor that is declared non-defective in initial step is defective.

Then

$$p^{(o)} = \frac{p(1-\pi_I^+)}{1-\pi_I^*} \qquad \text{(c.f. Patel [18])} \qquad (4.1.2),$$

where π_I^+ and π_I^* are as defined in section 2.2.

Further let p^+ be the conditional probability that a factor is non-defective given that it is declared defective.

Lemma 4.1.1

$$p^+ = \frac{\alpha_s(1-p')}{\bar{\beta}^+}$$

where α_s and p' are as defined in section 2.2.

Proof

Let E_1 be the event that a non-defective factor from a group-factor that is declared defective in the initial step is declared defective in the subsequent steps and let E_2 be the event that a factor from a group-factor that is declared defective in the initial step is declared defective. Then

$$\text{Prob.}(E_1) = (1-p')\alpha_s \qquad (4.1.3)$$

and

$$\text{Prob.}(E_2) = p'\gamma_s + (1-p')\alpha_s \qquad (4.1.4)$$

where γ_s is the probability of declaring a defective factor defective in subsequent steps.

Hence,

$$p^+ = \text{Prob.}(E_1/E_2)$$

$$= \frac{\alpha_s(1-p')}{p'\gamma_s+(1-p')\alpha_s}$$

$$= \frac{\alpha_s(1-p')}{\bar{\beta}^+} \qquad \left[\text{c.f. } (2.2.19)\right]$$

(4.1.5).

This completes the proof of the lemma.

Theorem 4.1.1

Let M_R be the number of defective factors declared defective in a step-wise group screening design with f factors, the a-priori probability of a factor to be defective being 'p', then

$$E(M_R) = fp\, \Pi_I^+ \gamma_s$$

where γ_s is the probability of declaring a defective factor defective in the subsequent steps and Π_I^+ is as given in (2.1.18).

Proof

The total number of factors that are declared defective from the r ($r \leq g$) group-factors declared defective in the initial step is equal to

$$r\,k\,\bar{\beta}^+ \qquad \text{using (4.1.1).}$$

The probability that a factor which is declared defective is defective is given by

$$1-p^+ \qquad (4.1.6)$$

where p^+ is as given in lemma 4.1.1.

Therefore

$$M_R = rk\bar{\beta}^+(1-p^+) \qquad (4.1.7).$$

Replacing r by $E(r) = g\Pi_I^*$ we get

$$E(M_R) = f\Pi_I^*\bar{\beta}^+(1-p^+) \qquad (4.1.8).$$

Substituting the value of p^+ given in lemma 4.1.1 and noting that

$$\bar{\beta}^+ = \gamma_s p' + \alpha_s(1-p')$$

and

$$p' = \frac{p\Pi_I^+}{\Pi_I^*}$$

in (4.1.8) we obtain

$$E(M_R) = fp\Pi_I^+\gamma_s \qquad (4.1,9).$$

This completes the proof of the Theorem.

In the next Theorem, we shall obtain an expression for the expected number of defective factors declared non-defective in the subsequent steps.

Theorem 4.1.2

In a step-wise group-screening design with f factors and with errors in observations, each factor being defective with a-priori probability 'p', the expected number of defective factors declared non-defective in the subsequent steps is given by

$$I_S = fp\Pi_I^+(1-\gamma_s)$$

where γ_s is the probability of declaring a defective factor defective in the subsequent steps and Π_I^+ is as given in (2.2.18).

Proof

The expected total number of defective factors in the g group-factors is equal to

$$fp.$$

The number of defective factors declared non-defective in the initial step is equal to

$$(g-r)kp^{(o)} \qquad (4.1.10)$$

where $p^{(o)}$ is as given in (4.1.2).

The number of defective factors declared defective in the subsequent steps is M_R. If I_S denotes the expected number of defective factors declared non-defective in the subsequent steps, then

$$I_S = E\left[fp-(g-r)p^{(o)} - M_R\right]$$

$$= fp\Pi_I^{+}(1-\gamma_s) \qquad (4.1.11),$$

using (4.1.9) and replacing r by $E(r) = g\Pi_I^{*}$.
This proves the Theorem.

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

<u>Lemma 4.1.2</u>

$$I_I = fp(1-\Pi_I^{+})$$

<u>Proof</u>

$$I_I = E(g-r)kp^{(o)} \qquad \left(\text{c.f. } (4.1.10)\right)$$

$$= fp(1-\Pi_I^{+}) \qquad (4.1.12),$$

substituting for $E(r) = g\Pi_I^{*}$.

Hence the lemma.

In the Theorem that follows, we shall obtain an expression for the expected number of non-defective factors declared defective in the subsequent steps,

<u>Theorem 4.1.3</u>

Let M_u be the number of non-defective factors declared defective in the subsequent steps.

Then

$$E(M_u) = f\alpha_s(\Pi_I^* - p\Pi_I^+),$$

where α_s is the probability of declaring a non-defective factor defective in the subsequent steps, Π_I^* is the probability of declaring a group-factor defective in the initial step and p, Π_I^+ and f are as defined earlier.

Proof

The total number of factors that are declared defective from the r group-factors declared defective in step-one is

$$rk\bar{\beta}^+ \qquad \text{using} \qquad (4.1.1)$$

Thus

$$M_u = rk\bar{\beta}^+ p^+$$

where p^+ is the probability that a factor that is declared defective in the subsequent steps is non-defective. Therefore

$$E(M_u) = E(rk\bar{\beta}^+ p^+)$$

$$= f\alpha_s(\Pi_I^* - p\Pi_I^+) \qquad (4.1.13)$$

using (4.1.5) and noting $p' = \dfrac{p\Pi_I^+}{\Pi_I^*}$.

This completes the proof of the Theorem.

Let I denote the expected total number of incorrect decisions in a step-wise group screening design with errors in observations. Then we have the following Theorem.

Theorem 4.1.4

The expected total number of incorrect decisions in a step-wise group screening design with f factors each factor being defective with a-priori probability p is given by

$$I = fp - fp\Pi_I^+\gamma_s + f\alpha_s(\Pi_I^* - p\Pi_I^+)$$

where p, Π_I^+, Π_I^* and γ_s are as defined earlier.

Proof

The expected total number of incorrect decisions is obtained by adding I_I, I_S and $E(M_u)$, i.e.,

$$I = I_I + I_S + E(M_u)$$

$$= fp(1-\Pi_I^+) + fp\Pi_I^+(1-\gamma_s) + f\alpha_s(\Pi_I^* - p\Pi_I^+) \qquad (4.1.14),$$

using (4.1.12) and Theorems (4.1.2) and (4.1.3).

Simplifying (4.1.14) we get

$$I = fp - fp\Pi_I^+\gamma_s + f\alpha_s(\Pi_I^* - p\Pi_I^+) \qquad (4.1.15).$$

This completes the proof of the Theorem.

Corollary 4.1.1

For large $\frac{\Delta}{\sigma}$ and arbitrary p, the expected total number of incorrect decisions in a step-wise group screening design with errors in observations is approximately equal to

$$f\alpha_s\left[q - (1-\alpha_I)q^k\right] .$$

Proof

When $\frac{\Delta}{\sigma}$ is large, we have the following approximations

$$\Pi_I^* \approx \alpha_I q^k + (1-q^k)$$

$$= 1-(1-\alpha_I)q^k$$

$$\Pi_I^+ \approx 1$$

and

$$\gamma_s \approx 1.$$

The corollary follows immediately on using these approximations in (4.1.15).

Corollary 4.1.2

For large $\frac{\Delta}{\sigma}$ and small p, the expected total number of incorrect decisions in a step-wise group screening design with errors in observations is approximately equal to

$$f\alpha_s\left[(\alpha_I - p) + (1-\alpha_I)pk\right].$$

Proof

For small p,

$q^k \approx 1-kp$, upto order p.

The result follows on using this approximation in corollary 4.1.1.

4.1.2 Optimum size of the group-factor in the initial step considering the expected total number of runs and the expected total number of incorrect decisions

Since we cannot minimize both I and E(R) at the same time, we will try to minimize one of them while fixing the value of the other, for the following cases:-

(i) large $\frac{\Delta}{\sigma}$ and small p and

(ii) large $\frac{\Delta}{\sigma}$ and arbitrary p.

4.1.2.1 Optimum size of the group-factor in the initial step for large $\frac{\Delta}{\sigma}$ and small p

Theorem 4.1.5

For large $\frac{\Delta}{\sigma}$ and small p, i.e., a-priori probability of a factor to be defective, the size k of the group-factor in the initial step which minimizes the expected total number of runs for a fixed value of the expected total number of incorrect decisions ω say, in a step-wise group screening design with errors in observations is given by

$$k \approx \frac{\omega - f\alpha_s(\alpha_I - p)}{f\alpha_s(1-\alpha_I)p},$$

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

$$\min E(R) \approx h + f\left[\alpha^* + \alpha_I + \frac{1}{2}(1-\alpha_s)(3-\alpha^*)p\right]$$

$$+ \frac{f^2\alpha_s(1-\alpha_I)p}{\omega - f\alpha_s(1-\alpha_I)p}\left[1-(1-\alpha_I)(2-\xi)p\right]$$

$$+ \frac{\omega - f\alpha_s(\alpha_I - p)}{\alpha_s(1-\alpha_I)}\left[(1-\alpha_I) - \frac{1}{2}(1-\alpha_s)(1+\alpha^*)\right]$$

where α_I, α_s and α^* are as defined in section 2.2 and the variable ξ takes the value 0 if $\alpha^*=0$ and the value 1 otherwise.

Proof

For large $\frac{\Delta}{\sigma}$ and small p, we have

$$E(R) \approx h + \frac{f}{k} + f\alpha^* + f(1-\alpha_s)p\left\{1 - \frac{(2-\xi)}{k} - k + \frac{1}{2}(k+1)(1-\alpha^*)\right\}$$

$$+ f(1-\alpha_I)kp + f\alpha_I ,$$

and

$$I \approx f\alpha_s\left[(\alpha_I - p) + (1-\alpha_I)pk\right]$$

using corollaries 2.2.4 and 4.1.2 respectively. The problem is to minimize E(R) given above subject to the condition

$$f\alpha_s\left[(\alpha_I - p) + (1-\alpha_I)pk\right] = \omega \qquad \text{(fixed)}.$$

This is equivalent to solving this constraint. Thus the required value of k is

$$k \approx \frac{\omega - f\alpha_s(\alpha_I - p)}{f\alpha_s(1-\alpha_I)p} \qquad (4.1.16a).$$

Since I is an increasing function of k, ω should be chosen so that the value of k in (4.1.16a) does not exceed the value of k that minimizes E(R) in corollary (2.2.4). That is we choose the value of ω which satisfies the condition

$$\frac{\omega - f\alpha_s(\alpha_I - p)}{f\alpha_s(1-\alpha_I)} \leq \left[\frac{2-2(1-\alpha_s)(2-\xi)p}{2(1-\alpha_I)p-(1-\alpha_s)(1+\alpha^*)p}\right]^{\frac{1}{2}} \qquad (4.1.16b).$$

The expression on the right hand side is the value of k which minimizes E(R).

Inequality (4.1.16b) gives

$$\omega \leq f\alpha_s(1-\alpha_I)\left[\frac{2-2(1-\alpha_s)(2-\xi)p}{2(1-\alpha_I)p-(1-\alpha_s)(1+\alpha^*)p}\right]^{\frac{1}{2}}$$

$$+ f\alpha_s(\alpha_I - p) \qquad (4.1.16c).$$

The inequality (4.1.16c) is valid if

$$1-2(1-\alpha_s)(1-\tfrac{\xi}{2})p > 0$$

which is true if $p<\frac{1}{2}$ (c.f. (2.2.47)).

Substituting the value of k given in (4.1.16a) in the formula for E(R) given above we obtain

$$\min E(R) = h+f\left[\alpha^*+\alpha_I+\tfrac{1}{2}(1-\alpha_s)(3-\alpha^*)p\right]$$

$$+ \frac{f^2\alpha_s(1-\alpha_I)p}{\omega - f\alpha_s(1-\alpha_I)p}\left[1-(1-\alpha_I)(2-\xi)p\right]$$

$$+ \frac{\omega - f\alpha_s(\alpha_I - p)}{\alpha_s(1-\alpha_I)}\left[(1-\alpha_I) - \tfrac{1}{2}(1-\alpha_s)(1+\alpha^*)\right]$$

$$(4.1.17).$$

This completes the proof of the Theorem.

Interchanging the roles of E(R) and I in Theorem 4.1.5, we obtain the following Theorem.

<u>Theorem 4.1.6</u>

Assuming p to be small and $\frac{\Delta}{\sigma}$ large, the size 'k' of the group-factor which minimizes the expected total number of incorrect decisions subject to a fixed value of the expected total number of runs, say ν, is given by

$$\text{(i)}\quad k \approx \frac{\nu-h-f\{\alpha^*+\alpha_I+\frac{1}{2}(1-\alpha_s)(3-\alpha^*)p\}}{2fp\left[(1-\alpha_I)-\frac{1}{2}(1-\alpha_s)(1+\alpha^*)\right]} - \left[\left\{\frac{\nu-h-f\{\alpha^*+\alpha_I+\frac{1}{2}(1-\alpha_s)(3-\alpha^*)p\}}{2fp\{(1-\alpha_I)-\frac{1}{2}(1-\alpha_s)(1+\alpha^*)\}}\right\}^2 - \frac{1-(1-\alpha_s)(2-\xi)p}{2p\{(1-\alpha_I)-\frac{1}{2}(1-\alpha_s)(1+\alpha^*)\}}\right]^{\frac{1}{2}}$$

when

$$\nu > h+2f\left[(1-\alpha_I)p-\frac{1}{2}(1-\alpha_s)(1+\alpha^*)p\right]^{\frac{1}{2}}\left[1-(1-\alpha_s)(2-\xi)p\right]^{\frac{1}{2}} + f\left[\alpha^*+\alpha_I+\frac{1}{2}(1-\alpha_s)(3-\alpha^*)p\right]$$

$$\text{(ii)}\quad k = \left[\frac{2-2(1-\alpha_s)(2-\xi)p}{2(1-\alpha_I)p-(1-\alpha_s)(1+\alpha^*)}\right]^{\frac{1}{2}}$$

when

$$\nu = h+2f\left[(1-\alpha_I)p-\frac{1}{2}(1-\alpha_s)(1+\alpha^*)p\right]^{\frac{1}{2}}\left[1-(1-\alpha_s)(2-\xi)p\right]^{\frac{1}{2}}$$

$$+ f\left[\alpha^* + \alpha_I + \frac{1}{2}(1-\alpha_s)(3-\alpha^*)p\right]$$

<u>Proof</u>

The problem is to minimize

$$I = f\alpha_s\left[(\alpha_I - p) + (1-\alpha_I)pk\right],$$

subject to the condition

$$E(R) = h + \frac{f}{k} + f\alpha^* + f(1-\alpha_s)p\left\{1 - \frac{2-\xi}{k} - k + \frac{1}{2}(k+1)(1-\alpha^*)\right\}$$

$$+ f(1-\alpha_I)kp + f\alpha_I$$

$$= \nu \text{ (fixed)}.$$

This is equivalent to solving for k in the equation

$$fk^2p\left[(1-\alpha_I) - \frac{1}{2}(1-\alpha_s)(1+\alpha^*)\right]$$

$$+ \left[h - \nu + f\{\alpha^* + \alpha_I + \frac{1}{2}(1-\alpha_s)(3-\alpha^*)p\}\right]k$$

$$+ f\left[1-(1-\alpha_s)(2-\xi)p\right] = 0 \qquad (4.1.18).$$

Equation (4.1.18) implies

$$k = \frac{\nu - h - f\{\alpha^* + \alpha_I + \frac{1}{2}(1-\alpha_s)(3-\alpha^*)p\}}{2fp\left[(1-\alpha_I) - \frac{1}{2}(1-\alpha_s)(1+\alpha^*)\right]}$$

$$\pm \left[\left\{\frac{\nu - h - f\{\alpha^* + \alpha_I + \frac{1}{2}(1-\alpha_s)(3-\alpha^*)p\}}{2fp\left[(1-\alpha_I) - \frac{1}{2}(1-\alpha_s)(1+\alpha^*)\right]}\right\}^2 - \frac{1-(1-\alpha_s)(2-\xi)p}{2p\{(1-\alpha_I) - \frac{1}{2}(1-\alpha_s)(1+\alpha^*)\}}\right]^{\frac{1}{2}}$$

(4.1.19).

The smaller value of k gives the point of minimum for I. The values of k in (4.1.19) will be real if we choose $\nu \geq \min E(R)$ given in Theorem 2.2.3. i.e.,

$$\nu \geq h+2f\left[(1-\alpha_I)p - \frac{1}{2}(1-\alpha_s)(1+\alpha^*)p\right]^{\frac{1}{2}}$$

$$\times \left[1-(1-\alpha_s)(2-\xi)p\right]^{\frac{1}{2}}$$

$$+ f\left[\alpha^*+\alpha_I + \frac{1}{2}(1-\alpha_s)(3-\alpha^*)p\right] \qquad (4.1.20)$$

If ν is such that we have equality in (4.1.20) instead of inequality, then the value of k will be given by

$$k = \left[\frac{2-2(1-\alpha_s)(2-\xi)p}{2(1-\alpha_I)p-(1-\alpha_s)(1+\alpha^*)p}\right]^{\frac{1}{2}} \qquad (4.1.21),$$

which is exactly the value of k which minimizes E(R) in corollary 2.2.4.

This completes the proof.

4.1.2.2 Optimum size of the group-factor in the initial step for large $\frac{\Delta}{\sigma}$ and arbitrary p

Theorem 4.1.7

For large $\frac{\Delta}{\sigma}$ and arbitrary p, the group size 'k' which minimizes the expected total number of runs for a fixed value of the expected total number of incorrect decisions ω say, in a step-wise group screening design with errors in observations is given by

$$k \approx \frac{\log(f\alpha_s q - \omega) - \log f\alpha_s(1-\alpha_I)}{\log q}$$

where $q=1-p$, α_I and α_s are as defined earlier.

Proof

The problem is to minimize E(R) given in corollary 2.2.2 subject to the condition

$$I = \omega \qquad \text{(fixed)}$$

where I is as given in corollary 4.1.1.

i.e.,

minimize

$$E(R) \approx h + \frac{2f}{k} + f - \frac{f(1-\alpha^*)}{k}\left[\frac{1-\left\{1-(1-\alpha_s)p-\alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}^{k+1}}{(1-\alpha_s)p+\alpha_s\left(1-(1-\alpha_I)q^k\right)}\right]$$

$$+ f\left[(1-\alpha_s)p+\alpha_s\{1-(1-\alpha_I)q^k\}\right]\left[1-\frac{2-\xi}{k}-\frac{\alpha^*}{k}\right]$$

$$+ f\left[1-(1-\alpha_I)q^k\right] - f\left[1-\left\{1-(1-\alpha_s)p-\alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}^k\right]$$

$$- \frac{f\alpha^*}{k} - \frac{f}{k}\left\{(1-\alpha_s)p+\alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}^2\{\xi-\alpha^*\}$$

$$-f\alpha^*\left\{(1-\alpha_s)p+\alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}\left[1-\left(1-\alpha_s\right)p-\alpha_s\left(1-(1-\alpha_I)q^k\right)\right]^k$$

(4.1.22).

subject to the condition

$$I = f\alpha_s\left(q-(1-\alpha_I)q^k\right) = \omega \qquad \text{(fixed)}.$$

The optimum value of k is obtained by solving the constraint given above. This gives

$$k \approx \frac{\log(f\alpha_s q-\omega) - \log f\alpha_s(1-\alpha_I)}{\log q} \qquad (4.1.23).$$

Since I is an increasing function of k, ω should be chosen so that the value of k in (4.1.23) should not exceed the value of k that minimizes E(R) in (4.1.22) obtained using Newton - Raphson's iterative method on equation (2.2.49). The corresponding min E(R) is obtained by substituting the value of k in (4.1.23) in the expression for E(R) given in (4.1.22).

This completes the proof.

Next we would like to choose k such that I is minimum for fixed value of E(R), say ν. The problem is equivalent to minimizing

$$I = f\alpha_s\left[q-(1-\alpha_I)q^k\right],$$

subject to the condition

$$E(R)=h+\frac{2f}{k}+f-\frac{f(1-\alpha^*)}{k}\left[\frac{1-\left\{1-(1-\alpha_s)p-\alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}^{k+1}}{(1-\alpha_s)p+\alpha_s\left(1-(1-\alpha_I)q^k\right)}\right]$$

$$+ f\left[(1-\alpha_s)p+\alpha_s\{1-(1-\alpha_I)q^k\}\right]\left[1-\frac{2-\xi}{k}-\frac{\alpha^*}{k}\right]$$

$$+ f\left[1-(1-\alpha_I)q^k\right] - f\left[1-\left\{1-(1-\alpha_s)p-\alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}^k\right].$$

$$- \frac{f\alpha^*}{k} - \frac{f}{k}\left\{(1-\alpha_s)p+\alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}^2\{\xi-\alpha^*\}$$

$$- f\alpha^*\left\{(1-\alpha_s)p+\alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}\left[1-(1-\alpha_s)p-\alpha_s\left(1-(1-\alpha_I)q^k\right)\right]^k$$

$$= \nu \quad \text{(fixed)} \qquad (4.1.24).$$

The problem is equivalent to solving the constraint. Equation (4.1.24) can be solved for k using Newton - Raphson iterative method, taking the value of k given in Theorem 4.1.6 as the initial approximation. The required minimum value of I is obtained by substituting the value of k obtained as stated above in the expression for I in corollary 4.1.1.

4.1.3 Optimum size of the group-factor in the initial step in relation to the total cost

Let c_1 be the cost of inspection per run and c_2 be the loss incurred per incorrect decision. Then the expected total cost is

$$C = c_1\, E(R) + c_2 I \qquad (4.1.25)$$

i.e.,

$$C = c_1\left[h + \frac{2f}{k} + f - \frac{f}{k\bar{\beta}^*}\{1-(1-\bar{\beta}^*)^{k+1}\}(1-\alpha^*)\right.$$

$$+ f\bar{\beta}^*\{1-\frac{2-\xi}{k} - \frac{\alpha^*}{k}\} + f\Pi_I^*$$

$$-f\{1-(1-\bar{\beta}^*)^k\} - \frac{f}{k}\,\alpha^*$$

$$- \frac{f}{k}\bar{\beta}^{*2}\{\xi-\alpha^*\} - f\alpha^*\bar{\beta}^*(1-\bar{\beta}^*)^k\Big]$$

$$+ c_2\Big[fp - fp\Pi_I^+\gamma_s + f\alpha_s(\Pi_I^* - p\Pi_I^+)\Big]$$

using Theorem 2.2.2 and Theorem (4.1.4).

We shall find the value of k which minimizes the expected total cost, for the following cases;-

(i) large $\frac{\Delta}{\sigma}$ and small p,

(ii) large $\frac{\Delta}{\sigma}$ and arbitrary p.

<u>Theorem 4.1.8</u>

Assuming p, i.e., the a-priori probability of a factor to be defective to be small, and $\frac{\Delta}{\sigma}$ large, the size 'k' of the group-factor which minimizes the expected total cost 'C' in a step-wise group screening design with errors in observations is given by

$$k = \left[\frac{2c_1\{1-(2-\xi)(1-\alpha_s)p\}}{2(1-\alpha_I)(c_1+c_2\alpha_s)p-(1-\alpha_s)(1+\alpha^*)c_1p}\right]^{\frac{1}{2}},$$

subject to k being real, where α_s is the probability of declaring a non-defective factor defective in the subsequent steps, α_I is the probability of declaring a non-defective group-factor defective in the initial step, α^* is the proportion of group-factors declared defective at any step but due to errors in observations no factor from each such group-factor is declared defective on individual tests, c_1 is the cost of inspection per run and c_2 is the loss incurred per incorrect decision. The variable ξ takes the value 0 if $\alpha^*=0$ and the value 1 otherwise.

Proof

For large $\frac{\Delta}{\sigma}$ and small p, we have

$$E(R) \simeq h + \frac{f}{k} + f\alpha^* + f(1-\alpha_s)p\left\{1 - \frac{(2-\xi)}{k} - k + \frac{1}{2}(k+1)(1-\alpha^*)\right\}$$

$$+ f(1-\alpha_I)kp + f\alpha_I$$

using corollary 2.2.4, and

$$I \approx f\alpha_s\left[(\alpha_I - p) + (1-\alpha_I)pk\right]$$

using corollary 4.1.2.

The expected total cost thus becomes

$$C \approx c_1\left[h + \frac{f}{k} + f\alpha^* + f(1-\alpha_s)p\left\{1 - \frac{(2-\xi)}{k} - k + \frac{1}{2}(k+1)(1-\alpha^*)\right\}\right.$$

$$\left. + f(1-\alpha_I)kp + f\alpha_I\right]$$

$$+ c_2\left[f\alpha_s\{(\alpha_I - p) + (1-\alpha_I)kp\}\right] \qquad (4.1.27)$$

from (4.1.25).

Assuming continuous variation in k, the optimum size of the group-factor is obtained by solving the equation

$$\frac{dC}{dk} = 0.$$

This implies

$$- \frac{c_1}{k^2}\left[1-(2-\xi)(1-\alpha_s)p\right]$$

$$+ c_1\left[(1-\alpha_I)p-(1-\alpha_s)p+\frac{1}{2}(1-\alpha_s)(1-\alpha^*)p\right]$$

$$+ c_2(1-\alpha_I)p \quad = \quad 0$$

i.e.

$$k \approx \left[\frac{2c_1\{1-(2-\xi)(1-\alpha_s)p\}}{2(1-\alpha_I)(c_1+c_2\alpha_s)p-(1-\alpha_s)(1+\alpha^*)c_1p}\right]^{\frac{1}{2}}$$

(4.1.28).

The value of k in (4.1.28) is real if

$$p < \frac{1}{2(1-\alpha_s)(1-\frac{\xi}{2})} \qquad (4.1.29),$$

which is true if

$$p < \frac{1}{2} \qquad \left(\text{c.f. } (2.2.47)\right).$$

This value of k will be in the neighbourhood of the point of minimum of the expected total cost 'C' given in (4.1.27) if

$$\frac{d^2C}{dk^2} > 0$$

i.e. if

$$\frac{2}{k^3}\{1-(2-\xi)(1-\alpha_s)p\}c_1 > 0$$

i.e.

$$p < \frac{1}{2(1-\alpha_s)(1-\frac{\xi}{2})}$$

which is condition (4.1.29).

Therefore the value of k given in (4.1.28) is in the neighbourhood of the point of minimum of the expected total cost C given in (4.1.27).

This completes the proof of the Theorem.

The corresponding minimum value of 'C' is obtained by substituting this value of k in the expression for 'C' in (4.1.27).

The next case we are interested in is when $\frac{\Delta}{\sigma}$ is large and p arbitrary. Using corollary 2.2.2 and corollary 4.1.1, the expected total cost becomes

$$C \approx c_1\Bigg[h+\frac{2f}{k}+f-\frac{f(1-\alpha^*)}{k}\left\{\frac{1-\left\{1-(1-\alpha_s)p-\alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}^{k+1}}{(1-\alpha_s)p+\alpha_s\left(1-(1-\alpha_I)q^k\right)}\right\}$$

$$+f\{1-(1-\alpha_I)q^k\}-f\left\{1-\left\{1-(1-\alpha_s)p-\alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}^k\right\}$$

$$-\frac{f\alpha^*}{k}-\frac{f}{k}\left\{(1-\alpha_s)p+\alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}^2\{\xi-\alpha^*\}$$

$$-f\alpha^*\left\{(1-\alpha_s)p+\alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}\left\{1-(1-\alpha_s)p-\alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}^k$$

$$+f\left\{(1-\alpha_s)p+\alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}\left\{1-\frac{2-\xi}{k}-\frac{\alpha^*}{k}\right\}\Bigg]$$

$$+c_2\left[f\alpha_s\left(q-(1-\alpha_I)q^k\right)\right] \qquad (4.1.30).$$

The value of k that minimizes C given in previous page is a solution of the equation

$$\frac{dC}{dk} = 0$$

i.e.

$$c_1\left[\frac{1}{k^2}\left[-2+\alpha^*+\left\{(1-\alpha_s)p+\alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}(2-\xi+\alpha^*)\right.\right.$$

$$\left.+\left\{(1-\alpha_s)p+\alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}^2(\xi-\alpha^*)\right]$$

$$-(1-\alpha_I)q^k \ln q\left[-\alpha_s+\frac{(2-\xi)\alpha_s}{k}+\frac{\alpha_s\alpha^*}{k}+1\right]$$

$$+k\alpha_s(1-\alpha_I)q^k \ln q\left\{1-(1-\alpha_s)p-\alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}^{k-1}$$

$$\times\left[1-\alpha^*\left\{(1-\alpha_s)p+\alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}\right]$$

$$+\alpha_s(1-\alpha_I)q^k \ln q\left[\frac{2(\xi-\alpha^*)}{k}\left\{(1-\alpha_s)p+\alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}\right.$$

$$\left.-\alpha^*\left\{1-(1-\alpha_s)p-\alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}^k\right]$$

$$+\frac{(1-\alpha^*)}{k^2}\left[\frac{1-\left\{1-(1-\alpha_s)p-\alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}^{k+1}}{(1-\alpha_s)p+\alpha_s\left(1-(1-\alpha_I)q^k\right)}\right]$$

$$+\frac{\alpha_s(1-\alpha^*)}{k}\left[\frac{(k+1)\left\{1-(1-\alpha_s)p-\alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}^k}{(1-\alpha_s)p+\alpha_s\left(1-(1-\alpha_I)q^k\right)}\right]$$

$$\times(1-\alpha_I)q^k \ln q$$

$$- \frac{(1-\alpha^*)}{k}\left[\frac{\alpha_s(1-\alpha_I)q^k \ln q\left[1-\left\{1-(1-\alpha_s)p-\alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}^{k+1}\right]}{\left\{(1-\alpha_s)p+\alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}^2}\right]$$

$$- c_2\alpha_s(1-\alpha_I)q^k \ln q = 0 \qquad (4.1.31).$$

The value of k that minimizes C in (4.1.30) is nearer to the value of k given in Theorem 4.1.8 and can be obtained using Newton - Raphson's iterative method applied to equation (4.1.31).

4.2 SCREENING WITH UNEQUAL A-PRIORI PROBABILITIES

In this section, we will discuss the performance of a step-wise group-screening design with unequal a-priori probabilities of factors to be defective and with errors in observations.

4.2.1 Calculation of the expected number of incorrect decisions

Incorrect decisions arise in the following ways:-

(i) declaring defective factors as non-defective in the initial step,

(ii) declaring defective factors as non-defective in subsequent steps and

(iii) declaring non-defective factors as defective in subsequent steps.

Let $P^*_{k_i}(j)$ be the probability that exactly j factors from the i^{th} group-factor of size k_i that is declared defective in the initial step are declared defective in the subsequent steps. Then

$$P^*_{k_i}(j) = \begin{cases} 1 - \dfrac{1}{\pi^*_{Ii}} \{1-(1-\bar{\beta}^*_i)^{k_i}\} & j=0 \\ \dfrac{1}{\pi^*_{Ii}} \dbinom{k_i}{j} \bar{\beta}^{*\,j}_i (1-\bar{\beta}^*_i)^{k_i-j} & j=1,2,\ldots,k_i \end{cases}$$

as explained earlier in (3.2.17).

Let $E_{k_i}(j)$ denote the expected number of factors declared defective from the i^{th} group-factor that was declared defective in the initial step. Then

$$E_{k_i}(j) = \frac{1}{\Pi^*_{Ii}} k_i \bar{\beta}^*_i$$

$$= k_i \bar{\beta}^+_i \qquad (4.2.1)$$

where $\bar{\beta}^+_i$ is the conditional probability that a factor chosen at random from a group-factor of size k_i that is declared defective in the initial step is declared defective.

Lemma 4.2.1.

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 $(i=1,2,\ldots,g)$. Then

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

where γ_{si} is the probability that a defective factor from the i^{th} group-factor that is declared defective in the initial step is declared defective in the subsequent steps and Π^+_{Ii} is the probability that the i^{th} group-factor containing at least one defective factor is declared defective.

Proof

The total number of factors declared defective in the subsequent steps is

$$\sum_{i=1}^{g} k_i \bar{\beta}^+_i U_i \qquad \big(\text{c.f. } (4.2.1)\big),$$

where U_i is as already defined in (3.2.12).
The probability that a factor that is declared defective from the i^{th} group-factor that is declared defective is defective is equal to

$$1 - p_i^+ \qquad (4.2.2)$$

where

$$p_i^+ = \frac{\alpha_{si}(1-p_i')}{\bar{\beta}_i^+} \qquad (4.2.3)$$

using (4.1.5), and

$$p_i' = \frac{p_i \Pi_{Ii}^+}{\Pi_{Ii}^*} \qquad (4.2.4),$$

is the probability that a factor chosen at random from the i^{th} group-factor that is declared defective in the initial step is defective. Then

$$M_R = \sum_{i=1}^{g} k_i \bar{\beta}_i^+ (1-p_i^+) U_i$$

Hence,

$$E(M_R) = \sum_{i=1}^{g} k_i \bar{\beta}_i^+ (1-p_i^+) E(U_i)$$

$$= \sum_{i=1}^{g} k_i \left[\bar{\beta}_i^+ - \alpha_{si}(1-p_i')\right] \Pi_{Ii}^*$$

i.e.,

$$E(M_R) = \sum_{i=1}^{g} k_i \gamma_{si} p_i' \Pi_{Ii}^* \qquad (4.2.5),$$

replacing $\bar{\beta}_i^+$ by $\{\gamma_{si} p_i' + \alpha_{si}(1-p_i')\}$ as given in (3.2.16a).

Substituting the value of p_i' given in (4.2.4) in (4.2.5) we get

$$E(M_R) = \sum_{i=1}^{g} k_i p_i \pi_{Ii}^{+} \gamma_{si} \qquad (4.2.6).$$

This completes the proof of the lemma.

In the lemma that follows, we obtain an expression for the expected number of defective factors declared non-defective in the subsequent steps from all the group-factors that are declared defective in the initial step.

Lemma 4.2.2

The expected number of defective factors declared non-defective from all the group-factors that are declared defective in the initial step is given by

$$I_S = \sum_{i=1}^{g} k_i p_i \pi_{Ii}^{+} (1-\gamma_{si}).$$

Proof

The expected total number of defective factors in all the g group-factors in the initial step is equal to

$$\sum_{i=1}^{g} k_i p_i$$

The number of defective factors declared non-defective in the initial step is equal to

$$\sum_{i=1}^{g} (1-U_i) k_i p_i^{(o)} \qquad (4.2.7),$$

where $p_i^{(o)} = \dfrac{p_i(1-\Pi^+_{Ii})}{(1-\Pi^*_{Ii})}$ is the probability

that a factor chosen at random from the i^{th} initial step group-factor declared non-defective, is defective. Therefore

$$I_S = E\left[\sum_{i=1}^{g} k_i p_i - \sum_{i=1}^{g} (1-U_i) k_i p_i^{(o)} - M_R\right]$$

i.e.,

$$I_S = \sum_{i=1}^{g} k_i p_i - \sum_{i=1}^{g} k_i p_i^{(o)}(1-\Pi^*_{Ii}) - E(M_R)$$

$$= \sum_{i=1}^{g} k_i p_i \Pi^+_{Ii}(1-\gamma_{si}) \qquad (4.2.8)$$

using (4.2.6).

This completes the proof.

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

<u>Lemma 4.2.3</u>

$$I_I = \sum_{i=1}^{g} k_i p_i (1-\Pi^+_{Ii})$$

<u>Proof</u>

$$I_I = E\left[\sum_{i=1}^{g} (1-U_i) k_i p_i^{(o)}\right] \qquad \text{(c.f.(4.2.7))}$$

$$= \sum_{i=1}^{g} k_i p_i (1-\Pi^+_{Ii}) \qquad (4.2.9),$$

substituting for $E(U_i) = \Pi^*_{Ii}$ and simplifying.

This proves the lemma.

Let M_u be the number of non-defective factors declared defective in the subsequent steps. Then we have the following lemma.

Lemma 4.2.4

$$E(M_u) = \sum_{i=1}^{g} k_i \alpha_{si} (\Pi^*_{Ii} - p_i \Pi^+_{Ii})$$

Proof

The total number of factors declared defective in the subsequent steps is

$$\sum_{i=1}^{g} k_i \bar{\beta}^+_i U_i \qquad \left(\text{c.f.}(4.2.1)\right),$$

where U_i is as already defined in (3.2.12).

Thus

$$M_u = \sum_{i=1}^{g} k_i \bar{\beta}^+_i U_i p^+_i$$

which implies

$$E(M_u) = E\left[\sum_{i=1}^{g} k_i \bar{\beta}^+_i U_i p^+_i\right]$$

$$= E\left[\sum_{i=1}^{g} k_i U_i \alpha_{si} (1-p'_i)\right] \qquad \left[\text{c.f. } (4\cdot 2\cdot 3)\right]$$

i.e.,

$$E(M_u) = \sum_{i=1}^{g} k_i \alpha_{si} (\Pi^*_{Ii} - p_i \Pi^+_{Ii}) \qquad (4.2.10),$$

on replacing U_i by $E(U_i) = \Pi^*_{Ii}$ and p'_i by $\dfrac{p_i \Pi^+_{Ii}}{\Pi^*_{Ii}}$.

This completes the proof.

Theorem 4.2.2

Let I be the expected total number of incorrect decisions in a step-wise group screening design with g group-factors in the initial step such that the i^{th} group-factor of size k_i contains factors with a-priori probability p_i of being defective $(i=1,2,...,g)$.

Then

$$I = \sum_{i=1}^{g} k_i p_i - \sum_{i=1}^{g} k_i p_i \Pi_{Ii}^{+} \gamma_{si} + \sum_{i=1}^{g} k_i \alpha_{si} (\Pi_{Ii}^{*} - p_i \Pi_{Ii}^{+})$$

where Π_{Ii}^{*}, Π_{Ii}^{+} and γ_{si} are as defined earlier.

Proof

The expected total number of incorrect decisions is given by

$$I = I_I + I_S + E(M_u)$$

$$= \sum_{i=1}^{g} k_i p_i (1-\Pi_{Ii}^{+}) + \sum_{i=1}^{g} k_i p_i \Pi_{Ii}^{+} (1-\gamma_{si})$$

$$+ \sum_{i=1}^{g} k_i \alpha_{si} (\Pi_{Ii}^{*} - p_i \Pi_{Ii}^{+})$$

using lemmas 4.2.2, 4.2.3 and 4.2.4.

i.e.,

$$I = \sum_{i=1}^{g} k_i p_i - \sum_{i=1}^{g} k_i p_i \Pi_{Ii}^{+} \gamma_{si} + \sum_{i=1}^{g} k_i \alpha_{si} (\Pi_{Ii}^{*} - p_i \Pi_{Ii}^{+}) \tag{4.2.11}$$

This completes the proof of the Theorem.

Corollary 4.2.1

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

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

Proof

$$I = \sum_{i=1}^{g} k_i p_i - \sum_{i=1}^{g} k_i p_i \Pi_{Ii}^{+} \gamma_{si} + \sum_{i=1}^{g} k_i \alpha_{si} (\Pi_{Ii}^{*} - p_i \Pi_{Ii}^{+})$$

$$= \sum_{i=1}^{g} k_i p_i - E(M_R) + E(M_u).$$

Hence I will take its maximum value when $E(M_R)$ is minimum and $E(M_u)$ is maximum.

But

$$E(M_R) = \sum_{i=1}^{g} k_i p_i \Pi_{Ii}^{+} \gamma_{si} \qquad \left(\text{c.f. } (4.2.6)\right),$$

takes its minimum value when

$$\Pi_{Ii}^{+} = \sum_{s_i=1}^{k_i} \binom{k_i - 1}{s_i - 1} p_i^{s_i - 1} q_i^{k_i - s_i} \Pi_{Ii}(s_i \phi_{Ii}, \alpha_{Ii}) \text{ is}$$

replaced by $\Pi_{Ii}(\phi_{Ii}, \alpha_{Ii})$.

i.e.,

$$\text{Min } E(M_R) = \sum_{i=1}^{g} k_i \gamma_{si} p_i \Pi_{Ii}(\phi_{Ii}, \alpha_{Ii}) \qquad (4.2.12).$$

Next

$$E(M_u) = \sum_{i=1}^{g} k_i \alpha_{si} (\Pi^*_{Ii} - p_i \Pi^+_{Ii}),$$

takes its maximum value when Π^*_{Ii} is replaced by its maximum value and Π^+_{Ii} is replaced by its minimum value.

i.e., when

Π^*_{Ii} is replaced by $\left\{\alpha_{Ii} q_i^{k_i} + (1-q_i^{k_i}) \Pi_{Ii}(k_i \phi_{Ii}, \alpha_{Ii})\right\}$

and

Π^+_{Ii} is replaced by $\Pi_{Ii}(\phi_{Ii}, \alpha_{Ii})$.

Thus

$$\text{Max } E(M_u) = \sum_{i=1}^{g} k_i \alpha_{si} \left\{\alpha_{Ii} q_i^{k_i} + (1-q_i^{k_i}) \Pi_{Ii}(k_i \phi_{Ii}, \alpha_{Ii}) - p_i \Pi_{Ii}(\phi_{Ii}, \alpha_{Ii})\right\} \quad (4.2.13).$$

Using (4.2.12) and (4.2.13) in (4.2.11) we get

$$\text{Max } I = \sum_{i=1}^{g} k_i p_i - \sum_{i=1}^{g} k_i p_i \Pi_{Ii}(\phi_{Ii}, \alpha_{Ii}) \gamma_{si} + \sum_{i=1}^{g} k_i \alpha_{si} \left\{\alpha_{Ii} q_i^{k_i} + (1-q_i^{k_i}) \Pi_{Ii}(k_i \phi_{Ii}, \alpha_{Ii}) - p_i \Pi_{Ii}(\phi_{Ii}, \alpha_{Ii})\right\} \quad (4.2.14)$$

This proves the corollary.

Corollary 4.2.2

For large $\frac{\Delta_i}{\sigma}$ and arbitrary p_i's ,

$$\max I = \sum_{i=1}^{g} k_i \alpha_{si} \left[q_i - (1-\alpha_{Ii}) q_i^{k_i} \right].$$

Proof

Since $\Pi_{Ii}^{*}(s_i \phi_{Ii}, \alpha_{Ii}) \approx 1$ and $\gamma_{si} \approx 1$ for large $\frac{\Delta_i}{\sigma}$, the result follows from (4.2.14) by replacing

$\Pi_{Ii}(s_i \phi_{Ii}, \alpha_{Ii})$ by 1,

and

γ_{si} by 1.

Corollary 4.2.3

For large $\frac{\Delta_i\text{'s}}{\sigma}$ and small p_i's ,

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

Proof

The result is obtained by replacing $q_i^{k_i}$ by $1-k_i p_i$ in corollary 4.2.2.

4.2.2 Optimum sizes of the initial group-factors considering the expected total number of incorrect decisions and the expected total number of runs

Since we cannot minimize both max I and E(R)′at the same time, we will try to minimize one of them while fixing the value of the other.

We shall discuss the case when $\frac{\Delta_i\text{'s}}{\sigma}$ are large and p_i's small.

Under the above assumptions, the maximum expected total number of incorrect decisions is given by

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

as given in corollary 4.2.3.

The problem is to minimize

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

subject to the conditions

$$\text{(i)}\quad h+g+\sum_{i=1}^{g} k_i\alpha_i^*+\sum_{i=1}^{g}(1-\alpha_{si})k_ip_i\left\{1-\frac{2-\xi_i}{k_i}-k_i+\frac{1}{2}(1-\alpha_i^*)(k_i+1)\right\}$$

$$+\sum_{i=1}^{g}\alpha_{Ii}k_i+\sum_{i=1}^{g}k_i^2(1-\alpha_{Ii})p_i = \nu \quad \text{(fixed)}$$

$$\text{(ii)}\quad \sum_{i=1}^{g} k_i = f$$

$$\text{(iii)}\quad k_i > 0 \qquad i=1,2,\ldots,g.$$

Using the method of Lagranges multipliers, let

$$F(k_1, k_2,\ldots,k_g, \lambda_1, \lambda_2)$$

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

$$+ \lambda_1\left[h+g-\nu+\sum_{i=1}^{g} k_i\alpha_i^*\right.$$

$$+\sum_{i=1}^{g}(1-\alpha_{si})k_i p_i\left\{1-\frac{(2-\xi_i)}{k_i}-k_i-\frac{1}{2}(1-\alpha_i^*)(k_i+1)\right\}$$

$$+\sum_{i=1}^{g}\alpha_{Ii}k_i + \sum_{i=1}^{g}k_i^2(1-\alpha_{Ii})p_i\Bigg]$$

$$+\lambda_2\left[\sum_{i=1}^{g}k_i - f\right] \qquad (4.2.15).$$

For critical values,

$$\frac{\partial F}{\partial k_i}=0\ ,\ (i=1,2,\ldots,g)\ ,\ \frac{\partial F}{\partial \lambda_1}=0 \text{ and } \frac{\partial F}{\partial \lambda_2}=0\ .$$

These imply

$$\alpha_{si}\alpha_{Ii}-\alpha_{si}p_i+2(1-\alpha_{Ii})\alpha_{si}k_i p_i$$

$$+\lambda_1\Big\{\alpha_{si}^*+\alpha_{Ii}+\frac{1}{2}(1-\alpha_{si})(3-\alpha_i^*)p_i$$

$$-(1-\alpha_{si})(1+\alpha_i^*)k_i p_i+2(1-\alpha_{Ii})k_i p_i\Big\}$$

$$+\lambda_2=0 \qquad (4.2.16),$$

$$h+g-\nu+\sum_{i=1}^{g}k_i\alpha_i^*+\sum_{i=1}^{g}\alpha_{Ii}k_i$$

$$+\sum_{i=1}^{g}(1-\alpha_{si})k_i p_i\left\{1-\frac{(2-\xi_i)}{k_i}-k_i-\frac{1}{2}(1+\alpha_i^*)(k_i+1)\right\}$$

$$+\sum_{i=1}^{g}k_i^2(1-\alpha_{Ii})p_i=0 \qquad (4.2.17)$$

and

$$\sum_{i=1}^{g}k_i=f \qquad (4.2.18).$$

From (4.2.16) we get

$$k_i = -\frac{\alpha_{si}\alpha_{Ii}+\lambda_1(\alpha_i^*+\alpha_{Ii})+\lambda_2}{2(1-\alpha_{Ii})(\lambda_1+\alpha_{si})-(1-\alpha_{si})(1+\alpha_i^*)}\frac{1}{p_i}$$

$$-\frac{\frac{1}{2}(1-\alpha_{si})(3-\alpha_i^*) - \alpha_{si}}{2(1-\alpha_{Ii})(\lambda_1+\alpha_{si})-(1-\alpha_{si})(1+\alpha_i^*)}$$

$$= A_i\frac{1}{p_i} + B_i \text{ , say } (i=1,2,\ldots,g) \tag{4.2.19}$$

where A_i and B_i are unknown.

It is rather difficult to obtain the exact value of k_i by eliminating λ_1 and λ_2. We shall try to obtain k_i for the special case when $\alpha_{Ii} = \alpha_I$, $\alpha_{si} = \alpha_s$, $\alpha_i^* = \alpha^*$.

Theorem 4.2.3

If $\alpha_{Ii} = \alpha_I$, $\alpha_{si} = \alpha_s$ and $\alpha_i^* = \alpha$, then for large $\frac{\Delta_i\text{'s}}{\sigma}$ and small p_i's, the value of k_i which minimizes the maximum value of the total expected number of incorrect decisions for a fixed value ν, of the expected total number of runs is given by

$$k_i = A\frac{1}{p_i} + B ,$$

where

$$B = -\frac{a}{2} + \left[\frac{a^2}{4} + \frac{(d\sum_{i=1}^{g}\frac{1}{p_i} - f^2 - agf)}{\sum_{i=1}^{g}p_i\sum_{i=1}^{g}\frac{1}{p_i} - g^2}\right]^{\frac{1}{2}}$$

$$A = \frac{f - Bg}{\sum_{i=1}^{g} \frac{1}{p_i}}$$

$$a = \frac{\frac{1}{2}(1-\alpha_s)(3-\alpha^*)}{(1-\alpha_I)-\frac{1}{2}(1-\alpha_s)(1+\alpha^*)}$$

and

$$d = \frac{\nu-h-g-(\alpha^*+\alpha_I)}{(1-\alpha_I)-\frac{1}{2}(1-\alpha_s)(1+\alpha^*)} .$$

Proof:

When $\alpha_{Ii} = \alpha_I$, $\alpha_{si} = \alpha_s$ and $\alpha_i^* = \alpha^*$, $\xi_i = \xi$ $(i=1,2,...,g)$ so that (4.2.15) becomes

$$F(k_1,k_2,k_3,...,k_g,\lambda_1,\lambda_2)$$

$$= \alpha_s\alpha_I f - \alpha_s \sum_{i=1}^{g} k_i p_i + \alpha_s(1-\alpha_I) \sum_{i=1}^{g} k_i^2 p_i$$

$$+ \lambda_1 \left[h+g-\nu+(\alpha^*+\alpha_I)f+(1-\alpha_I) \sum_{i=1}^{g} k_i^2 p_i \right.$$

$$\left. + (1-\alpha_s) \sum_{i=1}^{g} k_i p_i \left\{ 1- \frac{2-\xi}{k_i} - k_i + \frac{1}{2}(1-\alpha^*)(k_i+1) \right\} \right]$$

$$+ \lambda_2 \left[\sum_{i=1}^{g} k_i - f \right] .$$

Thus (4.2.16) and (4.2.17) becomes

$$- \alpha_s p_i + 2\alpha_s(1-\alpha_I)k_i p_i$$

$$+ \lambda_1(1-\alpha_s)\left\{p_i - 2k_i p_i + (1-\alpha^*)k_i p_i + \frac{1}{2}(1-\alpha^*)\dot{p}_i\right\}$$

$$+ 2\lambda_1(1-\alpha_I)k_i p_i + \lambda_2 = 0 \qquad (4.2.20)$$

and

$$\lambda_1\left[h+g-\nu+\alpha^* f+(1-\alpha_s)\sum_{i=1}^{g} k_i p_i\left\{1-\frac{(2-\xi)}{k_i} - k_i + \frac{1}{2}(1-\alpha^*)(k_i+1)\right\}+\alpha_I f+(1-\alpha_I)\sum_{i=1}^{g} k_i^2 p_i\right] = 0$$

$$(4.2.21)$$

respectively.

From (4.2.20) we get

$$k_i = -\frac{\lambda_2}{2(1-\alpha_I)(\lambda_1+\alpha_s)-(1-\alpha_s)(1+\alpha^*)}\ \frac{1}{p_i}$$

$$-\frac{\frac{1}{2}(1-\alpha_s)(3-\alpha^*) - \alpha_s}{2(1-\alpha_I)(\lambda_1+\alpha_s)-(1-\alpha_s)(1+\alpha^*)}$$

$$= A\frac{1}{p_i} + B \qquad (4.2.22),$$

where A and B are constants to be determined.

Multiplying (4.2.22) by 1, p_i and $k_i p_i$ and summing each result over i we get

$$f = A\sum_{i=1}^{g}\frac{1}{p_i} + Bg \qquad (4.2.23)$$

$$\sum_{i=1}^{g} k_i p_i = Ag + B\sum_{i=1}^{g} p_i \qquad (4.2.24)$$

and

$$\sum_{i=1}^{g} k_i^2 p_i = Af + B\sum_{i=1}^{g} k_i p_i \qquad (4.2.25).$$

From (4.2.24) and (4.2.25) we obtain

$$\sum_{i=1}^{g} k_i^2 p_i = Af + BAg + B^2 \sum_{i=1}^{g} p_i \qquad (4.2.26).$$

But from (4.2.21),

$$\sum_{i=1}^{g} k_i^2 p_i + \frac{\frac{1}{2}(1-\alpha_s)(3-\alpha^*)}{(1-\alpha_I)-\frac{1}{2}(1-\alpha_s)(1+\alpha^*)} \sum_{i=1}^{g} k_i p_i$$

$$= \frac{\nu-h-g-(\alpha^*+\alpha_I)f}{(1-\alpha_I)-\frac{1}{2}(1-\alpha_s)(1+\alpha^*)}$$

$$= d, \text{ say,} \qquad (4.2.27).$$

Using (4.2.24) and (4.2.26) in (4.2.27), we get

$$Af + BAg + B^2 \sum_{i=1}^{g} p_i + aAg + a\,B \sum_{i=1}^{g} p_i = d$$

where

$$a = \frac{\frac{1}{2}(1-\alpha_s)(3-\alpha^*)}{(1-\alpha_I)-\frac{1}{2}(1-\alpha_s)(1+\alpha^*)} \qquad (4.2.28)$$

i.e.

$$Af + (B+a)Ag + B(B+a) \sum_{i=1}^{g} p_i = d \qquad (4.2.29).$$

From (4.2.23)

$$A = \frac{f - Bg}{\sum_{i=1}^{g} \frac{1}{p_i}} \qquad (4.2.30).$$

Substituting this value in (4.2.29) we get

$$\left(\sum_{i=1}^{g} p_i \sum_{i=1}^{g} \frac{1}{p_i} - g^2\right) B^2 + a\left(\sum_{i=1}^{g} p_i \sum_{i=1}^{g} \frac{1}{p_i} - g^2\right) B$$

$$+ f^2 + agf - d\sum_{i=1}^{g}\frac{1}{p_i} = 0 \qquad (4.2.31).$$

Therefore

$$B = -\frac{a}{2} \pm \left[\frac{a^2}{4} + \frac{\left(d\sum_{i=1}^{g}\frac{1}{p_i} - f^2 - agf\right)}{\sum_{i=1}^{g} p_i \sum_{i=1}^{g}\frac{1}{p_i} - g^2}\right]^{\frac{1}{2}} \qquad (4.2.32).$$

Substituting these values of B in (4.2.30), we get the corresponding values of A. We shall now show that the quantity within the square root in (4.2.32) is positive.

Now

$$d\sum_{i=1}^{g}\frac{1}{p_i} - f^2 - agf$$

$$= \left(\sum_{i=1}^{g} k_i^2 p_i + a\sum_{i=1}^{g} k_i p_i\right)\sum_{i=1}^{g}\frac{1}{p_i} - f^2 - agf$$

(c.f. (4.2.27) and (4.2.28)).

i.e.,

$$d\sum_{i=1}^{g}\frac{1}{p_i} - f^2 - agf$$

$$= \left[\sum_{i=1}^{g}(k_i + \frac{a}{2})^2 p_i - \frac{a^2}{4}\sum_{i=1}^{g} p_i\right]\sum_{i=1}^{g}\frac{1}{p_i} - f^2 - agf \qquad (4.2.33)$$

By Schwartz's inequality

$$\left(\sum_{i=1}^{g} p_i\right)\left(\sum_{i=1}^{g}\frac{1}{p_i}\right) \geq \left(\sum_{i=1}^{g}\sqrt{p_i}\,\frac{1}{\sqrt{p_i}}\right)^2 = g^2 \qquad (4.2.34)$$

and

$$\left[\sum_{i=1}^{g}(k_i+\frac{a}{2})^2 p_i\right]\left[\sum_{i=1}^{g}\frac{1}{p_i}\right] \geq \left(\sum_{i=1}^{g}(k_i+\frac{a}{2})\sqrt{p_i}\ \frac{1}{\sqrt{p_i}}\right)^2$$

$$= f^2 + agf + \frac{a^2}{4} g^2$$

(4.2.35).

Using (4.2.35) in (4.2.33) we get

$$d\sum_{i=1}^{g}\frac{1}{p_i} - f^2 - agf$$

$$\geq \frac{a^2}{4}(g^2 - \sum_{i=1}^{g} p_i \sum_{i=1}^{g}\frac{1}{p_i}) \qquad (4.2.36).$$

From (4.2.34) and (4.2.36) we conclude

$$\frac{a^2}{4} + \frac{\left(d\sum_{i=1}^{g}\frac{1}{p_i} - f^2 - agf\right)}{\sum_{i=1}^{g} p_i \sum_{i=1}^{g}\frac{1}{p_i} - g^2} \geq 0. \qquad (4.2.37).$$

For positive values of k_i, we get two sets of solutions. To check which set gives a minimum value, we examine

$$\text{Max } I = \alpha_s\alpha_I f - \alpha_s\sum_{i=1}^{g} k_i p_i + \alpha_s(1-\alpha_I)\sum_{i=1}^{g} k_i^2 p_i$$

which is to be minimized.

Substituting the values of k_i's given in (4.2.22) and noting that

$$A = \frac{f - Bg}{\sum_{i=1}^{g}\frac{1}{p_i}} \quad \text{as given in (4.2.30)}$$

we get

$$\text{Max } I = B\alpha_s\left((1-\alpha_I)\,B-1\right)\left[\sum_{i=1}^{g} p_i - \frac{g^2}{\sum_{i=1}^{g} \frac{1}{p_i}}\right]$$

$$+ \alpha_s\alpha_I f + \alpha_s(1-\alpha_I)f^2 \frac{1}{\sum_{i=1}^{g}\frac{1}{p_i}} - \frac{\alpha_s f g}{\sum_{i=1}^{g}\frac{1}{p_i}}$$

(4.2.38)

But

$$\sum_{i=1}^{g} p_i \geq \frac{g^2}{\sum_{i=1}^{g}\frac{1}{p_i}} \qquad \left(\text{c.f. } (4.2.34)\right)$$

i.e.

$$\sum_{i=1}^{g} p_i - \frac{g^2}{\sum_{i=1}^{g}\frac{1}{p_i}} \geq 0 \qquad (4.2.39).$$

From (4.2.38) we conclude max I is minimum when

$B\left[B(1-\alpha_I)-1\right]$ takes its minimum value.

But

$$B = -\frac{a}{2} \pm \left[\frac{a^2}{4} + \frac{\left(d\sum_{i=1}^{g} p_i - f^2 - agf\right)}{\sum_{i=1}^{g} p_i \sum_{i=1}^{g}\frac{1}{p_i} - g^2}\right]^{\frac{1}{2}}$$

$$= -\frac{a}{2} \pm t \ , \quad \text{say},$$

where $\frac{a}{2} > 0$ $\left(\text{c.f.} (4.2.28)\right)$.

When $B = -\frac{a}{2} + t$,

$$B\left[B(1-\alpha_I)-1\right] = \left(-\frac{a}{2} + t\right)^2(1+\alpha_I) + \frac{a}{2} - t$$

(4.2.40).

When $B = -\frac{a}{2} - t$,

$$B\left[B(1-\alpha_I)-1\right] = (-\frac{a}{2} - t)^2(1-\alpha_I) + \frac{a}{2} + t \tag{4.2.41}$$

But

$$(-\frac{a}{2} + t)^2(1-\alpha_I) - t < (-\frac{a}{2} - t)^2(1-\alpha_I) + t \tag{4.2.42}$$

since $a>0$ and $t>0$.

Therefore $B\left[B(1-\alpha_I)-1\right]$ takes its minimum value when

$$B = -\frac{a}{2} + t$$

i.e., when

$$B = -\frac{a}{2} + \left[\frac{a^2}{4} + \frac{(d\sum_{i=1}^{g}\frac{1}{p_i} - f^2 - agf)}{\sum_{i=1}^{g} p_i \sum_{i=1}^{g}\frac{1}{p_i} - g^2}\right] \tag{4.2.43}$$

This proves the theorem.

Interchanging the roles of max I and E(R) in Theorem 4.2.3, we obtain the following Theorem

<u>Theorem 4.2.4</u>

If $\alpha_{Ii} = \alpha_I$, $\alpha_{si} = \alpha_s$, $\alpha_i^* = \alpha^*$ and $\xi_i = \xi$, then for large $\frac{\Delta_i\text{'s}}{\sigma}$ and small p_i's the value of k_i which minimizes the expected total number of runs for a fixed value of the maximum total expected number of incorrect decisions, ω, say, is given by

$$k_i = Q\frac{1}{p_i} + T$$

where

$$T = \frac{-1}{2(1-\alpha_I)} - \left[\frac{1}{4(1-\alpha_I)^2} + \frac{(e \sum_{i=1}^{g} \frac{1}{p_i} - gf-(1-\alpha_I)f^2)}{(1-\alpha_I)\{ \sum_{i=1}^{g} p_i \sum_{i=1}^{g} \frac{1}{p_i} -g^2\}} \right]^{\frac{1}{2}}$$

$$Q = \frac{f - gT}{\sum_{i=1}^{g} \frac{1}{p_i}}$$

and

$$e = \frac{\omega - \alpha_s \alpha_I f}{\alpha_s} .$$

Proof:

The problem is to minimize

$$E(R) = h + g + (\alpha^* + \alpha_I)f + (1 - \alpha_I) \sum_{i=1}^{g} k_i^2 p_i$$

$$+ (1-\alpha_s) \sum_{i=1}^{g} k_i p_i \left\{ 1- \frac{2-\xi}{k_i} - k_i + \frac{1}{2}(1-\alpha^*)(k_i+1) \right\}$$

subject to the following conditions

(i) $\max I = \alpha_s \alpha_I f - \alpha_s \sum_{i=1}^{g} k_i p_i + \alpha_s(1-\alpha_I) \sum_{i=1}^{g} k_i^2 p_i$

$= \omega$ (fixed).

(ii) $\sum_{i=1}^{g} k_i = f$

(iii) $k_i > 0.$

Using the method of Lagranges multipliers, let

$$F(k_1,k_2,\ldots,k_g,\lambda_1,\lambda_2)$$

$$= h+g+(\alpha^*+\alpha_I)f+(1-\alpha_I)\sum_{i=1}^{g} k_i^2 p_i$$

$$+(1-\alpha_s)\sum_{i=1}^{g} k_i p_i\left\{1-\frac{2-\xi}{k_i}-k_i+\frac{1}{2}(1-\alpha^*)(k_i+1)\right\}$$

$$+\lambda_1\left[\alpha_s\alpha_I f-\omega-\alpha_s\sum_{i=1}^{g} k_i p_i+\alpha_s(1-\alpha_I)\sum_{i=1}^{g} k_i^2 p_i\right]$$

$$+\lambda_2\left(\sum_{i=1}^{g} k_i - f\right) \qquad (4.2.44).$$

For critical values,

$$\frac{\partial F}{\partial k_i} = 0 \quad (i=1,2,\ldots,g), \quad \frac{\partial F}{\partial \lambda_1} = 0 \text{ and } \frac{\partial F}{\partial \lambda_2} = 0 .$$

These imply

$$2(1-\alpha_I)k_i p_i + \frac{1}{2}(1-\alpha_s)(3-\alpha^*)p_i-(1-\alpha_s)(1+\alpha^*)k_i p_i$$

$$-\lambda_1\alpha_s p_i + 2\lambda_1\alpha_s(1-\alpha_I)k_i p_i + \lambda_2 = 0 \qquad (4.2.45)$$

$$\alpha_s\alpha_I f - \omega - \alpha_s\sum_{i=1}^{g} k_i p_i + \alpha_s(1-\alpha_I)\sum_{i=1}^{g} k_i^2 p_i = 0$$

$$(4.2.46)$$

and

$$\sum_{i=1}^{g} k_i = f \qquad (4.2.47)$$

From (4.2.45), we get

$$k_i = \frac{-\lambda_2}{2(1-\alpha_I)(1+\lambda_1\alpha_s)-(1-\alpha_s)(1+\alpha^*)}\ \frac{1}{p_i}$$

$$+\frac{\lambda_1\alpha_s - \frac{1}{2}(1-\alpha_s)(3-\alpha^*)}{2(1-\alpha_I)(1+\lambda_1\alpha_s)-(1-\alpha_s)(1+\alpha^*)}$$

$$= Q \frac{1}{p_i} + T \qquad (4.2.48)$$

where Q and T are constants to be determined.

Multiplying (4.2.48) by 1, p_i and $k_i p_i$ and summing each result over i, we get

$$f = Q \sum_{i=1}^{g} \frac{1}{p_i} + gT \qquad (4.2.49)$$

$$\sum_{i=1}^{g} k_i p_i = gQ + T \sum_{i=1}^{g} p_i \qquad (4.2.50)$$

and

$$\sum_{i=1}^{g} k_i^2 p_i = fQ + T \sum_{i=1}^{g} k_i p_i \qquad (4.2.51).$$

From (4.2.50) and (4.2.51) we obtain

$$\sum_{i=1}^{g} k_i^2 p_i = fQ + gQT + T^2 \sum_{i=1}^{g} p_i \qquad (4.2.52).$$

But from (4.2.46)

$$\sum_{i=1}^{g} k_i p_i + (1-\alpha_I) \sum_{i=1}^{g} k_i^2 p_i = \frac{\omega - \alpha_s \alpha_I f}{\alpha_s}$$

$$= e, \text{ say}, \qquad (4.2.53).$$

Using (4.2.50) and (4.2.52) in (4.2.53) we get

$$gQ + T \sum_{i=1}^{g} p_i + (1-\alpha_I) \left[fQ + gQT + T^2 \sum_{i=1}^{g} p_i \right]$$

$$= e \qquad (4.2.54).$$

From (4.2.49)

$$Q = \frac{f - gT}{\sum_{i=1}^{g} \frac{1}{p_i}} \qquad (4.2.55).$$

Substituting this in (4.2.54) we get

$$(1-\alpha_I)\left\{\sum_{i=1}^{g} p_i \sum_{i=1}^{g} \frac{1}{p_i} - g^2\right\}T^2 + \left[\sum_{i=1}^{g} p_i \sum_{i=1}^{g} \frac{1}{p_i} - g^2\right]T$$

$$+gf+(1-\alpha_I)f^2 - e\sum_{i=1}^{g} \frac{1}{p_i} = 0$$

$$(4.2.56).$$

Therefore

$$T = \frac{-1}{2(1-\alpha_I)} \pm \left[\frac{1}{4(1-\alpha_I)^2} + \frac{e\sum_{i=1}^{g} \frac{1}{p_i} - gf - (1-\alpha_I)f^2}{(1-\alpha_I)\left\{\sum_{i=1}^{g} p_i \sum_{i=1}^{g} \frac{1}{p_i} - g^2\right\}}\right]^{\frac{1}{2}}$$

$$(4.2.57).$$

Substituting these values of T in (4.2.55) we get the corresponding values of Q.

We shall now show that the values of T given in (4.2.57) are real. This is so if the quantity within the square root is positive.

Now

$$e\sum_{i=1}^{g} \frac{1}{p_i} - gf - (1-\alpha_I)f^2$$

$$= \sum_{i=1}^{g} k_i p_i \sum_{i=1}^{g} \frac{1}{p_i} + (1-\alpha_I) \sum_{i=1}^{g} k_i^2 p_i \sum_{i=1}^{g} \frac{1}{p_i} - gf - (1-\alpha_I)f^2$$

(c.f. (4.2.53))

$$= (1-\alpha_I) \sum_{i=1}^{g} \frac{1}{p_i} \sum_{i=1}^{g} \left[k_i + \frac{1}{2(1-\alpha_I)}\right]^2 p_i$$

$$- \frac{(1-\alpha_I)}{4(1-\alpha_I)^2} \sum_{i=1}^{g} \frac{1}{p_i} \sum_{i=1}^{g} p_i - gf - (1-\alpha_I)f^2$$

$$\geq (1-\alpha_I)f^2 + fg + \frac{g^2}{4(1-\alpha_I)} - \frac{1}{4(1-\alpha_I)} \sum_{i=1}^{g} \frac{1}{p_i} \sum_{i=1}^{g} p_i$$

$$-gf-(1-\alpha_I)f^2 \qquad (4.2.58),$$

using Schwartz's inequality (c.f.(4.2.34)).

i.e.,

$$\frac{e \sum_{i=1}^{g} \frac{1}{p_i} - gf - (1-\alpha_I)f^2}{(1-\alpha_I)\left\{\sum_{i=1}^{g} p_i \sum_{i=1}^{g} \frac{1}{p_i} - g^2\right\}} \geq \frac{\frac{1}{4(1-\alpha_I)}\left[g^2 - \sum_{i=1}^{g} \frac{1}{p_i} \sum_{i=1}^{g} p_i\right]}{(1-\alpha_I)\left\{\sum_{i=1}^{g} p_i \sum_{i=1}^{g} \frac{1}{p_i} - g^2\right\}}$$

$$= -\frac{1}{4(1-\alpha_I)^2} \qquad (4.2.59).$$

Thus the quantity inside the square root in (4.2.57) is positive.

Therefore we get two sets of solutions for k_i's. To check which set gives a minimum value, we examine

$$E(R) = h+g+(\alpha^*+\alpha_I)f+(1-\alpha_I)\sum_{i=1}^{g} k_i^2 p_i$$

$$+(1-\alpha_s)\sum_{i=1}^{g} k_i p_i \left\{ \left[1-\frac{2-\xi}{k_i} - k_i + \frac{1}{2}(1-\alpha^*)(k_i+1) \right] \right\}$$

$$= h+g+(\alpha^*+\alpha_I)f-(1-\alpha_s)(2-\xi)\sum_{i=1}^{g} p_i$$

$$+\frac{1}{2}(1-\alpha_s)(3-\alpha^*)\sum_{i=1}^{g} k_i p_i$$

$$+\left\{(1-\alpha_I)-\frac{1}{2}(1-\alpha_s)(1+\alpha^*)\right\}\sum_{i=1}^{g} k_i^2 p_i$$

which is to be minimized.

Substituting the values of k_i given in (4.2.48) and noting that

$$Q = \frac{f-gT}{\sum_{i=1}^{g}\frac{1}{p_i}} \qquad \left(\text{c.f.}(4.2.55)\right)$$

we get

$$E(R) = h+g+(\alpha^*+\alpha_I)f-(1-\alpha_s)(2-\xi)\sum_{i=1}^{g} p_i$$

$$+\frac{1}{2}(1-\alpha_s)(3-\alpha^*)\frac{fg}{\sum_{i=1}^{g}\frac{1}{p_i}} + \left\{(1-\alpha_I)-\frac{1}{2}(1-\alpha_s)(1+\alpha^*)\right\}\frac{f^2}{\sum_{i=1}^{g}\frac{1}{p_i}}$$

$$+\left(\sum_{i=1}^{g} p_i - \frac{g^2}{\sum_{i=1}^{g}\frac{1}{p_i}}\right)T\left[\frac{1}{2}(1-\alpha_s)(3-\alpha^*)\right.$$

$$\left.+\left\{(1-\alpha_I)-\frac{1}{2}(1-\alpha_s)(1+\alpha^*)\right\}T\right]$$

Noting that

$$\sum_{i=1}^{g} p_i - \frac{g^2}{\sum_{i=1}^{g} \frac{1}{p_i}} \geq 0 \qquad \left(\text{c.f.}(4.2.39)\right),$$

E(R) given above is minimum when

$$T\left[\tfrac{1}{2}(1-\alpha_s)(3-\alpha^*)+\left\{(1-\alpha_I)-\tfrac{1}{2}(1-\alpha_s)(1+\alpha^*)\right\}T\right]$$

takes its minimum value. For the two values of T in (4.2.57), the above expression is smallest when

$$T = \frac{-1}{2(1-\alpha_I)} - \left[\frac{1}{4(1-\alpha_I)^2} + \frac{e\sum_{i=1}^{g}\frac{1}{p_i} - gf - (1-\alpha_I)f^2}{(1-\alpha_I)\left\{\sum_{i=1}^{g} p_i \sum_{i=1}^{g}\frac{1}{p_i} - g^2\right\}}\right]^{\frac{1}{2}}$$

(4.2.60).

This proves the Theorem.

<u>4.2.3</u> <u>Optimum sizes of group-factors in the initial step in relation to the total cost</u>

In this section, we shall define the expected total cost as a linear function of the expected number of runs and the expected number of incorrect decisions and try to obtain the sizes of the group-factors so that the expected total cost is minimum.

Let the cost of inspection per run be c_1 and suppose that the loss due to incorrect decisions is proportional to the maximum expected number of incorrect decisions. Let c_2 be the loss per unit of the maximum expected number of incorrect decisions. Then the expected total cost C is given by

$$C = c_1 E(R) + c_2 \max I.$$

$$= c_1 \Bigg[h+f+2g- \sum_{i=1}^{g} \frac{1}{\bar{\beta}_i^*}(1-\alpha_i^*)\{1-(1-\bar{\beta}_i^*)^{k_i+1}\}$$

$$+ \sum_{i=1}^{g} k_i \bar{\beta}_i^* \left\{1- \frac{2-\xi_i}{k_i} - \frac{\alpha_i^*}{k_i}\right\} + \sum_{i=1}^{g} k_i \Pi_{Ii}^*$$

$$- \sum_{i=1}^{g} k_i \{1-(1-\bar{\beta}_i^*)^{k_i}\} - \sum_{i=1}^{g} \alpha_i^* + \sum_{i=1}^{g} (\alpha_i^*-\xi_i)\bar{\beta}_i^{*2}$$

$$- \sum_{i=1}^{g} k_i \alpha_i^* \bar{\beta}_i^* (1-\bar{\beta}_i^*)^{k_i} \Bigg]$$

$$+ c_2 \Bigg[\sum_{i=1}^{g} k_i p_i - \sum_{i=1}^{g} k_i p_i \gamma_{si} \Pi_{Ii}(\phi_{Ii}, \alpha_{Ii})$$

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

$$\left. - p_i \Pi_{Ii}(\phi_{Ii}, \alpha_{Ii}) \right\} \Bigg]$$

using Theorem (3.2.1) and corollary (4.2.1).

Theorem 4.2.5

For large values of $\frac{\Delta_i\text{'s}}{\sigma}$ and small values of p_i's ($p_i \leq p$, $i=1,2,\ldots,g$), the value of k_i which minimizes the expected total cost is given by

$$k_i = \left(f - \sum_{i=1}^{g} G_i H_i \right) \frac{G_i}{\sum_{i=1}^{g} \bar{G}_i} + G_i H_i$$

where

$$G_i = \frac{-1}{\left\{2(1-\alpha_{Ii})-(1-\alpha_{si})(1+\alpha_i^*)\right\} c_1 p_i + 2\alpha_{si}(1-\alpha_{Ii}) c_2 p_i}$$

and

$$H_i = c_1 \left\{ (\alpha_i^* + \alpha_{Ii}) + \frac{1}{2}(1-\alpha_{si})(3-\alpha_i^*) p_i \right\} - c_2 \alpha_{si} (\alpha_{Ii} - p_i) .$$

Proof

For large values of $\frac{\Delta_i\text{'s}}{\sigma}$ and small values of p_i's, the expected total cost is given by

$$C = c_1 \Bigg[h + g + \sum_{i=1}^{g} (\alpha_i^* + \alpha_{Ii}) k_i + \sum_{i=1}^{g} (1-\alpha_{Ii}) k_i^2 p_i$$

$$+ \sum_{i=1}^{g} (1-\alpha_{si}) k_i p_i \left\{ 1 - \frac{2-\xi_i}{k_i} - k_i + \frac{1}{2}(1-\alpha_i^*)(k_i+1) \right\} \Bigg]$$

$$+ c_2 \left[\sum_{i=1}^{g} k_i \alpha_{si} \left\{ (\alpha_{Ii} - p_i) + (1-\alpha_{Ii}) k_i p_i \right\} \right]$$

using corollaries (3.2.2) and (4.2.3).

We wish to minimize the expected total cost 'C' given above subject to the conditions

(i) $\sum_{i=1}^{g} k_i = f$

(ii) $k_i > 0 \qquad i=1,2,\ldots,g.$

Using the method of Lagranges multiplier let

$$F(k_1,k_2,\ldots,k_g,\lambda)=c_1\Bigg[h+g+\sum_{i=1}^{g}(\alpha_i^*+\alpha_{Ii})k_i+\sum_{i=1}^{g}(1-\alpha_{Ii})k_i^2p_i$$

$$+\sum_{i=1}^{g}(1-\alpha_{si})k_ip_i\Bigg\{1-\frac{2-\xi_i}{k_i}-k_i$$

$$+\frac{1}{2}(1-\alpha_i^*)(k_i+1)\Bigg\}\Bigg]$$

$$+c_2\left[\sum_{i=1}^{g}k_i\alpha_{si}\left\{(\alpha_{Ii}-p_i)+(1-\alpha_{Ii})k_ip_i\right\}\right]$$

$$+\lambda\left[\sum_{i=1}^{g}k_i-f\right] \qquad (4.2.61),$$

where λ is the Lagrange's multiplier.

Assuming continuous variations in k_i, the critical values of k_i are obtained from the equations

$$\frac{\partial F}{\partial k_i}=0$$

and $\qquad\qquad (4.2.62).$

$$\frac{\partial F}{\partial \lambda}=0$$

Conditions (4.2.62) imply

$$c_1\Big[(\alpha_i^*+\alpha_{Ii})+\frac{1}{2}(1-\alpha_{si})(3-\alpha_i^*)p_i+2(1-\alpha_{Ii})k_ip_i$$

$$-(1-\alpha_{si})(1+\alpha_i^*)k_ip_i\Big]$$

$$+c_2\Big[\alpha_{si}(\alpha_{Ii}-p_i)+2\alpha_{si}(1-\alpha_{Ii})k_ip_i\Big]+\lambda=0$$

$$(4.2.63)$$

and

$$\sum_{i=1}^{g} k_i = f \qquad (4.2.64).$$

From (4.2.63) we get

$$k_i = \frac{-\lambda - c_1\left\{(\alpha_i^* + \alpha_{Ii}) + \frac{1}{2}(1-\alpha_{si})(3-\alpha_i^*)p_i\right\} - c_2\alpha_{si}(\alpha_{Ii} - p_i)}{\left\{2(1-\alpha_{Ii}) - (1-\alpha_{si})(1+\alpha_i^*)\right\}c_1p_i + 2\alpha_{si}(1-\alpha_{Ii})c_2p_i}$$

(4.2.65).

Summing (4.2.65) over 'i' we get

$$f = -\lambda \sum_{i=1}^{g} \frac{1}{\left\{2(1-\alpha_{Ii}) - (1-\alpha_{si})(1+\alpha_i^*)\right\}c_1p_i + 2\alpha_{si}(1-\alpha_{Ii})c_2p_i}$$

$$- \sum_{i=1}^{g} \frac{c_1\left\{(\alpha_i^* + \alpha_{Ii}) + \frac{1}{2}(1-\alpha_{si})(3-\alpha_i^*)p_i\right\} - c_2\alpha_{si}(\alpha_{Ii} - p_i)}{\left\{2(1-\alpha_{Ii}) - (1-\alpha_{si})(1 + \alpha_i^*)\right\}c_1p_i + 2\alpha_{si}(1-\alpha_{Ii})c_2p_i}$$

(4.2.66).

Let

$$G_i = \frac{-1}{\left\{2(1-\alpha_{Ii}) - (1-\alpha_{si})(1+\alpha_i^*)\right\}c_1p_i + 2\alpha_{si}(1-\alpha_{Ii})c_2p_i}$$

(4.2.67)

and

$$H_i = c_1\left\{(\alpha_i^* + \alpha_{Ii}) + \frac{1}{2}(1-\alpha_{si})(3-\alpha_i^*)p_i\right\} - c_2\alpha_{si}(\alpha_{Ii} - p_i)$$

(4.2.68).

Then (4.2.66) becomes

$$f = \lambda \sum_{i=1}^{g} G_i + \sum_{i=1}^{g} H_iG_i$$

i.e.,

$$\lambda = \left(f - \sum_{i=1}^{g} G_i H_i \right) \frac{1}{\sum_{i=1}^{g} G_i} \qquad (4.2.69).$$

Using (4.2.67), (4.2.68) and (4.2.69) in (4.2.65) we obtain

$$k_i = \left(f - \sum_{i=1}^{g} G_i H_i \right) \frac{G_i}{\sum_{i=1}^{g} G_i} + G_i H_i$$

(4.2.70).

This completes the proof.

APPENDICES

The tables given in the following appendices, result from the theories developed in this thesis. For all practical purposes, the values of k's (i.e. the group-sizes) in all tables should be rounded to the nearest integers.

TABLES RESULTING FROM CHAPTER II

Table 1(a): Optimum group-sizes in the initial step and expected number of runs for selected a-priori probabilities for step-wise designs with f=100, and without errors in observations

$$k \approx \left(\frac{2-4p}{p}\right)^{\frac{1}{2}}$$

$$\text{Min } E(R) \approx 1 + \frac{3fp}{2} + f(2p)^{\frac{1}{2}}(1-2p)^{\frac{1}{2}}$$

(c.f. Theorem 2.1.3).

p	k	Min E(R)
0.001	44.68	5.62
0.002	31.56	7.61
0.005	19.90	11.70
0.010	14.00	16.50
0.015	11.37	20.31
0.020	9.80	23.60
0.025	8.72	26.54
0.045	6.36	36.37
0.060	5.42	42.50
0.080	4.58	49.66
0.100	4.00	56.00
0.150	3.06	69.33
0.200	2.45	79.99
0.250	2.00	88.50

Table 1(b): Relative performance of step-wise designs and corresponding two stage group-screening designs for f=100 and selected values of p and without errors in observations

In column 3,

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

[c.f.(2.1.32)].

In column 5,

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

[c.f.(2.1.47)]

	Step-wise group screening		Two stage group screening	
p	k	Min E(R)	k	Min E(R)
0.001	45	5.58	32	7.28
0.002	32	7.55	23	9.85
0.005	21	11.54	15	14.91
0.010	15	16.17	11	20.56
0.015	12	19.82	9	24.83
0.020	11	22.96	8	28.42
0.025	9	25.76	7	31.53
0.035	8	30.65	6	36.91
0.045	7	34.37	5	41.56
0.060	6	40.69	5	47.61
0.080	5	47.39	4	54.36
0.100	5	53.29	4	60.39
0.150	4	65.78	3	72.92
0.200	3	75.93	3	83.13
0.250	3	84.85	3	92.15

Remarks

(1) The integer value that minimizes E(R) is obtained using a computer search in both cases.

(2) The table indicates that for small values of p, step-wise designs are prefferable to corresponding two stage designs but for higher values of p, step-wise designs have distinct advantage over two stage designs.

Tables 1(c), 1(d), 1(e), 1(f), 1(g), 1(h), 1(i), 1(j), and 1(k):

Optimum group sizes in the initial step and expected number of runs for selected prior probabilities and with errors in observations for step-wise designs

The integer value of the group size 'k' that minimizes

$$E(R) = h + \frac{2f}{k} + f - \frac{f(1-\alpha^*)}{k}\left[\frac{1-\left\{1-\left[(1-\alpha_s)p-\alpha_s\left(1-(1-\alpha_I)q^k\right)\right]\right\}^{k+1}}{(1-\alpha_s)p+\alpha_s\left(1-(1-\alpha_I)q^k\right)}\right]$$

$$+f\left[(1-\alpha_s)p+\alpha_s\{1-(1-\alpha_I)q^k\}\right]\left[1-\frac{2-\xi}{k}-\frac{\alpha^*}{k}\right]$$

$$+f\left[1-(1-\alpha_I)q^k\right]-f\left[1-\left\{1-(1-\alpha_s)p-\alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}^k\right]$$

$$-\frac{f\alpha^*}{k}-\frac{f}{k}\left\{(1-\alpha_s)p+\alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}^2(\xi-\alpha^*)$$

$$-f\alpha^*\left\{(1-\alpha_S)p+\alpha_S\left(1-(1-\alpha_I)q^k\right)\right\}$$

$$\times\left[1-(1-\alpha_S)p-\alpha_S\left(1-(1-\alpha_I)q^k\right)\right]^k$$

(c.f. Corollary 2.2.2)

has been obtained using a computer search.

Table 1(c): $f=100$, $\alpha_I=0.005$, $\alpha^*=0.005$, $\alpha_s=0.002$

p	0.001	0.002	0.005	0.010	0.020	0.030	0.035	0.040	0.050	0.060	0.080	0.100	0.150	0.200
k	40	35	29	15	9	9	9	9	9	7	7	6	4	4
Min E(R)	6.44	8.31	12.70	17.00	24.16	29.34	31.85	34.31	39.07	43.35	50.99	58.11	71.52	83.42

Table 1(d): $f=100$, $\alpha_I=0.01$, $\alpha^*=0.01$, $\alpha_s=0.01$

p	0.01	0.02	0.03	0.035	0.04	0.05	0.06	0.07	0.08	0.09	0.10	0.15	0.20
k	15	14	9	9	9	9	7	7	7	7	6	4	4
Min E(R)	17.35	24.55	29.81	32.28	34.71	39.43	43.84	47.70	51.45	55.10	58.62	72.07	83.93

Table 1(e): $f=100$, $\alpha_I=0.05$, $\alpha^*=0.05$, $\alpha_s=0.02$

p	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10	0.15	0.20
k	29	14	9	9	9	9	7	7	7	7	4	4
Min E(R)	22.61	29.85	35.58	40.03	44.38	48.61	52.59	56.08	59.49	62.81	76.39	87.49

Table 1(f): $f=100$, $\alpha_I=0.05$, $\alpha^*=0.05$, $\alpha_s=0.05$

p	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10	0.15	0.20
k	31	15	14	9	9	9	7	7	7	7	4	4
Min E(R)	17.76	27.26	33.75	38.71	43.20	47.64	51.91	55.54	59.11	62.61	76.38	87.75

Table 1(g): $f=100$, $\alpha_I=0.10$, $\alpha^*=0.10$, $\alpha_s=0.05$

p	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10	0.15	0.20
k	33	15	14	9	9	9	9	7	7	7	4	4
Min E(R)	23.92	34.12	40.09	45.52	49.54	53.54	57.49	61.32	64.55	67.72	81.63	91.98

Table 1(h): $f=100$, $\alpha_I=0.10$, $\alpha^*=0.10$, $\alpha_s=0.10$

p	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10	0.15	0.20
k	29	15	14	9	9	9	9	7	7	7	4	4
Min E(R)	19.57	30.45	37.42	43.46	47.88	52.36	56.82	60.59	64.14	67.63	81.58	92.39

<u>Table 1(i):</u> $f=500$, $\alpha_I=0.005$, $\alpha^*=0.005$, $\alpha_s=0.002$

p	0.001	0.002	0.005	0.010	0.020	0.030	0.040	0.050	0.060	0.080	0.100	0.150	0.200
k	47	34	22	16	10	8	8	7	6	6	6	4	3
Min E(R)	27.7	37.5	57.6	81.1	116.6	143.3	166.5	187.7	207.6	243.0	276.6	345.6	404.9

<u>Table 1(j):</u> $f=500$, $\alpha_I=0.01$, $\alpha^*=0.01$, $\alpha_s=0.01$

p	0.01	0.02	0.03	0.035	0.04	0.05	0.06	0.07	0.08	0.09	0.10	0.15	0.20
k	16	12	8	8	8	7	7	6	6	6	6	4	3
Min E(R)	82.4	118.4	146.1	157.5	168.9	190.3	210.2	228.2	245.6	262.6	279.1	348.4	407.5

Table 1(k): $f=500$, $\alpha_I=0.05$, $\alpha^*=0.05$, $\alpha_s=0.05$

p	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10	0.15	0.20
k	26	16	10	8	8	7	7	6	6	6	4	4
Min E(R)	108.5	145.9	173.9	196.8	217.0	236.1	254.0	270.4	286.1	301.5	369.9	425.5

Remark:

Tables 1(c), 1(d), ..., 1(k) clearly indicate that with errors in observations, min E(R) increases with p and is higher than the corresponding value without errors in observations.

APPENDIX II

TABLES RESULTING FROM CHAPTER III

Tables 2(a), 2(b), 2(c) and 2(d): Optimum group sizes in the initial step and expected number of runs for selected unequal a-priori probabilities for f=100 and without errors in observations

$$k_i \approx \left(f+\frac{3}{2}g\right)\frac{1}{p_i \sum_{i=1}^{g}\frac{1}{p_i}} - \frac{3}{2}$$

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

$$\times \frac{1}{\sum_{i=1}^{g}\frac{1}{p_i}}$$

(c.f. Theorem 3.1.2)

Table 2(a): $p_i \leq p = 0.010 \quad g=7$

i	p_i	k_i
1	0.004	23.714
2	0.005	18.671
3	0.006	15.309
4	0.007	12.908
5	0.008	11.107
6	0.009	9.706
7	0.010	8.585
Total		100.000

Min E(R)=13.419.

For the corresponding two stage design,

$$\min E(R) = 17.127.$$

Table 2(b): $p_i \leq p = 0.015$, $g=9$

i	p_i	k_i
1	0.007	17.175
2	0.008	14.841
3	0.009	13.025
4	0.010	11.573
5	0.011	10.384
6	0.012	9.394
7	0.013	8.556
8	0.014	7.838
9	0.015	7.214
Total		100.000

$$\text{Min } E(R) = 17.109$$

For the corresponding two stage design,

$$\text{Min } E(R) = 21.518$$

Table 2(c): $p_i \leq p = 0.035$, $g=13$

i	p_i	k_i
1	0.008	17.097
2	0.009	15.031
3	0.010	13.378
4	0.013	9.944
5	0.015	8.418
6	0.017	7.251
7	0.020	5.939
8	0.022	5.263
9	0.025	4.451
10	0.027	4.010
11	0.030	3.459
12	0.033	3.008
13	0.035	2.751
	Total	100.000

Min E(R) = 22.064

For the corresponding two stage design,

min E(R) = 26.450

Table 2(d): $p_i \leq p = 0.1$ g=20

i	p_i	k_i
1	0.040	9.571
2	0.045	8.341
3	0.050	7.357
4	0.053	6.856
5	0.055	6.552
6	0.060	5.880
7	0.062	5.643
8	0.065	5.312
9	0.070	4.826
10	0.075	4.405
11	0.078	4.177
12	0.080	4.036
13	0.082	3.901
14	0.085	3.710
15	0.087	3.590
16	0.090	3.420
17	0.092	3.314
18	0.095	3.162
19	0.098	3.019
20	0.100	2.928
	Total	100.000

Min E(R) = 45.216

For the corresponding two stage design, min E(R) = 55.065

Remark: Tables 2(a), 2(b), 2(c) and 2(d) indicate that when screening with unequal prior probabilities, step-wise designs are prefferable to corresponding two stage designs.

Tables 2(e), 2(f), 2(g), 2(h), 2(i) and 2(j): Optimum group sizes in the initial step and the expected number of runs for selected unequal a-priori probabilities for step-wise designs with f=100 and with errors in observations

$$k_i \approx \left(f + \sum_{i=1}^{g} \frac{\left[\alpha_i^* + \alpha_{Ii} + \frac{1}{2}(1-\alpha_{si})(3-\alpha_i^*)p_i\right]}{\left[2(1-\alpha_{Ii})-(1-\alpha_{si})(1+\alpha_i^*)\right]p_i} \right)$$

$$\times \frac{1}{\left[2(1-\alpha_{Ii})-(1-\alpha_{si})(1+\alpha_i^*)\right]p_i \sum_{i=1}^{g} \frac{1}{\left[2(1-\alpha_{Ii})-(1-\alpha_{si})(1+\alpha_i^*)p_i\right]}}$$

$$- \frac{\left[\alpha_i^* + \alpha_{Ii} + \frac{1}{2}(1-\alpha_{si})(3-\alpha_i^*)p_i\right]}{\left[2(1-\alpha_{Ii})-(1-\alpha_{si})(1+\alpha_i^*)\right]p_i}$$

(c.f. Theorem 3.2.2).

$$E(R) \approx h+g+\sum_{i=1}^{g} k_i\alpha_i^* + \sum_{i=1}^{g}(1-\alpha_{si})k_ip_i\left\{1 - \frac{2-\xi_i}{k_i} - k_i + \frac{1}{2}(1-\alpha_i^*)(k_i+1)\right\}$$

$$+ \sum_{i=1}^{g} \alpha_{Ii} k_i + \sum_{i=1}^{g} k_i^2 (1-\alpha_{Ii}) p_i$$

(c.f. corollary 3.2.2)

$\xi_i = 1$ when $\alpha_i^* \neq 0$.

<u>Table 2(e)</u>: $h=1$, $g=7$, $p_i \leq p = 0.010$

$\alpha_i^* = \alpha^* = 0.005$, $\alpha_{Ii} = \alpha_I = 0.005$

$\alpha_{si} = \alpha_s = 0.002$

i	p_i	k_i
1	0.004	23.722
2	0.005	18.675
3	0.006	15.310
4	0.007	12.907
5	0.008	11.104
6	0.009	9.702
7	0.010	8.580
	Total	100.000

Min E(R) = 14.405

<u>Table 2(f):</u> $h=1$, $g=7$, $p_i \leq p = 0.010$

$\alpha_{Ii} = \alpha_I = 0.01$, $\alpha_i^* = \alpha^* = 0.01$,

$\alpha_{si} = \alpha_s = 0.01$

i	p_i	k_i
1	0.004	23.720
2	0.005	18.674
3	0.006	15.310
4	0.007	12.906
5	0.008	11.105
6	0.009	9.703
7	0.010	8.582
Total		100.000

Min E(R) = 15.365

Table 2(g): h=3, g=13, $p_i \leq p = 0.035$

$\alpha_{Ii}=\alpha_I=0.005$, $\alpha_i^*=\alpha^*=0.005$, $\alpha_{si}=\alpha_s=0.002$

i	p_i	k_i
1	0.008	17.111
2	0.009	15.042
3	0.010	13.386
4	0.013	9.948
5	0.015	8.419
6	0.017	7.251
7	0.020	5.936
8	0.022	5.259
9	0.025	4.446
10	0.027	4.005
11	0.030	3.453
12	0.033	3.001
13	0.035	2.743
Total		100.000

min E(R) = 25.239

Table 2(h): $h=3$, $g=13$, $p_i \leq p = 0.035$

$\alpha_{Ii}=\alpha_I=0.01$, $\alpha_i^*=\alpha^*=0.01$, $\alpha_{si}=\alpha_s=0.01$

i	p_i	k_i
1	0.008	17.107
2	0.009	15.039
3	0.010	13.384
4	0.013	9.947
5	0.015	8.419
6	0.017	7.251
7	0.020	5.937
8	0.022	5.260
9	0.025	4.447
10	0.027	4.006
11	0.030	3.455
12	0.033	3.003
13	0.035	2.745
	Total	100.000

min E(R) = 26.175

Table 2(i): $h=4$, $g=20$, $p_i \leq p = 0.100$

$\alpha_{Ii}=\alpha_I=0.005$, $\alpha_i^*=\alpha^*=0.005$, $\alpha_{si}=\alpha_s=0.002$

i	p_i	k_i
1	0.040	9.581
2	0.045	8.348
3	0.050	7.362
4	0.053	6.860
5	0.055	6.555
6	0.060	5.883
7	0.062	5.644
8	0.065	5.314
9	0.070	4.826
10	0.075	4.403
11	0.078	4.176
12	0.080	4.033
13	0.082	3.898
14	0.085	3.707
15	0.087	3.587
16	0.090	3.417
17	0.092	3.310
18	0.095	3.158
19	0.098	3.014
20	0.100	2.924
	Total	100.000

min $E(R) = 50.422$

Table 2(j): h=4, g=20, $p_i \leq p = 0.100$

$\alpha_{Ii}=\alpha_I=0.01$, $\alpha_i^*=\alpha^*=0.01$, $\alpha_{si}=\alpha_s=0.01$

i	p_i	k_i
1	0.040	9.578
2	0.045	8.346
3	0.050	7.361
4	0.053	6.858
5	0.055	6.554
6	0.060	5.882
7	0.062	5.644
8	0.065	5.313
9	0.070	4.826
10	0.075	4.404
11	0.078	4.176
12	0.080	4.034
13	0.082	3.899
14	0.085	3.708
15	0.087	3.588
16	0.090	3.418
17	0.092	3.311
18	0.095	3.159
19	0.098	3.016
20	0.100	2.925
	Total	100.000

min E(R) = 51.219

Remark: The tables 2(e), 2(f), ..., 2(j) clearly indicate that with errors in observations the value of minimum E(R) increases.

APPENDIX III

TABLES RESULTING FROM CHAPTER IV

Tables 3(a), 3(b), 3(c), 3(d) and 3(e): Step-wise group screening plans for f = 100 with selected prior probabilities 'p' and with a specified number of incorrect decisions 'ω'

In these plans,

$$k \approx \frac{\log(f\alpha_s q - \omega) - \log f\alpha_s(1-\alpha_I)}{\log q}$$

|c.f. (4.1.23)|

and

$$E(R) \approx h + \frac{2f}{k} + f - \frac{f(1-\alpha^*)}{k}\left[\frac{1-\left\{1-\left\{(1-\alpha_s)p-\alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}\right\}^{k+1}}{(1-\alpha_s)p+\alpha_s\left(1-(1-\alpha_I)q^k\right)}\right]$$

$$+\left[(1-\alpha_s)p+\alpha_s\{1-(1-\alpha_I)q^k\}\right]\left[1-\frac{2-\xi}{k}-\frac{\alpha^*}{k}\right]$$

$$+\left[1-(1-\alpha_I)q^k\right]-f\left[1-\left\{1-(1-\alpha_s)p-\alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}^k\right]$$

$$-\frac{f\alpha^*}{k}-\frac{f}{k}\left\{(1-\alpha_s)p+\alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}^2\{\xi-\alpha^*\}$$

$$-f\alpha^*\left\{(1-\alpha_s)p+\alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}$$

$$\times\left[1-(1-\alpha_s)p-\alpha_s\left(1-(1-\alpha_I)q^k\right)\right]^k$$

[c.f. (4.1.22)]

Table 3(a): $\alpha_I = 0.05$, $\alpha^* = 0.05$, $\alpha_s = 0.05$, $\omega = 1$

p	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10	0.15	0.20
k	18.35	9.76	6.90	5.47	4.61	4.04	3.63	3.33	3.09	2.90	2.34	2.06
min E(R)	21.72	30.18	37.12	43.01	48.23	53.95	59.28	61.27	66.97	67.43	81.87	95.00

Table 3(b): $\alpha_I = 0.05$, $\alpha^* = 0.05$, $\alpha_s = 0.05$, $\omega = 2$

p	0.06	0.07	0.08	0.09	0.10	0.11	0.12	0.13	0.14	0.15	0.20
k	9.13	8.04	7.23	6.60	6.09	5.68	5.34	5.05	4.81	4.60	3.88
min E(R)	47.81	54.49	55.80	58.01	64.07	65.00	66.82	72.53	74.15	75.69	87.25

Table 3(c): $\alpha_I = 0.10$, $\alpha^* = 0.10$, $\alpha_s = 0.05$, $\omega = 1$

p	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10	0.15
k	12.97	7.08	5.12	4.14	3.55	3.16	2.89	2.68	2.51	2.39	2.00
min E(R)	32.99	41.35	51.03	57.44	62.98	67.88	69.26	75.21	79.82	81.13	94.46

Table 3(d): $\alpha_I = 0.10$, $\alpha^* = 0.10$, $\alpha_s = 0.05$, $\omega = 2$

p	0.06	0.07	0.08	0.09	0.10	0.11	0.12	0.13	0.14	0.15	0.20
k	8.26	7.30	6.58	6.02	5,58	5.22	4.92	4.67	4.45	4.27	3.63
min E(R)	56.62	58.25	60.63	65.83	67.85	69.74	75.50	77.16	78.71	80.17	94.31

<u>Table 3(e):</u>. $\alpha_I = 0.10$, $\alpha^* = 0.10$, $\alpha_s = 0.10$, $\omega = 3$

p	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10	0.15	0.20
k	26.44	13.87	9.69	7.60	6.34	5.51	4.92	4,47	4.12	3.85	3.03	2.63
min E(R)	22.56	30.95	39.69	46.78	52.97	55.55	61.70	63.52	69.06	70.37	85.37	95.45

Tables 3(f), 3(g), 3(h), 3(i): Step-wise group screening plans for f = 100 which minimize expected total cost for selected prior probabilities 'p'

In these plans, the integer value of k that minimizes the expected total cost

$$C = c_1 E(R) + c_2 I$$

where

c_1 = the cost of observing a run

c_2 = the cost of an incorrect decision

$$E(R) \approx h + \frac{2f}{k} + f - \frac{f(1-\alpha^*)}{k}\left[\frac{1-\left\{1-\left[(1-\alpha_s)p-\alpha_s\left(1-(1-\alpha_I)q^k\right)\right]\right\}^{k+1}}{(1-\alpha_s)p+\alpha_s\left(1-(1-\alpha_I)q^k\right)}\right]$$

$$+f\left[(1-\alpha_s)p+\alpha_s\{1-(1-\alpha_I)q^k\}\right]\left[1-\frac{2-\xi}{k}-\frac{\alpha^*}{k}\right]$$

$$+f\left[1-(1-\alpha_I)q^k\right]-f\left[1-\left\{1-(1-\alpha_s)p-\alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}^k\right]$$

$$-\frac{f\alpha^*}{k}-\frac{f}{k}\left\{(1-\alpha_s)p+\alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}^2\{\xi-\alpha^*\}$$

$$-f\alpha^*\left\{(1-\alpha_s)p+\alpha_s\left(1-(1-\alpha_I)q^k\right)\right\}$$

$$\times\left[1-(1-\alpha_s)p-\alpha_s\left(1-(1-\alpha_I)q^k\right)\right]^k$$

(c.f.(4.1.22))

and

$$I \approx f\alpha_s\{q-(1-\alpha_I)q^k\}, \quad \text{(c.f. Corollary 4.1.1)}$$

has been obtained using a computer search.

The minimum C given in column 5 is a relative figure using c_1 i.e., the cost of observing a run as the unit. [i.e. it indicates the values of C/c_1].

<u>Table 3(f):</u> $\alpha_I=0.05$, $\alpha^*=0.05$, $\alpha_s=0.05$

$c_2 : c_1 = 1 : 5$

p	k	E(R)	I	min C
0.01	30	17.762	1.436	18.049
0.02	15	27.256	1.392	27.535
0.03	14	33.753	1.749	31.103
0.04	9	38.708	1.510	39.010
0.05	9	43.200	1.756	43.552
0.06	9	47.644	1.978	48.039
0.07	7	51.907	1.792	52.266
0.08	7	55.541	1.950	55.931
0.09	7	59.109	2.095	59.528
0.10	7	62.605	2.228	63.050
0.11	6	65.972	2.089	66.390
0.12	4	69.170	1.551	69.481
0.13	4	71.607	1.629	71.933
0.14	4	74.012	1.702	74.352
0.15	4	76.384	1.770	76.738
0.16	4	78.724	1.835	79.091
0.17	4	81.031	1.896	81.410
0.18	4	82.171	1.925	83.694
0.19	4	85.543	2.005	85.944
0.20	4	87.748	2.054	88.159

Table 3(g): $\alpha_I = 0.05$, $\alpha^* = 0.05$, $\alpha_s = 0.05$

$c_2 : c_1 = 3 : 5$

p	k	E(R)	I	min C
0.01	30	17.762	1.436	18.624
0.02	15	27.256	1.392	28.092
0.03	14	33.753	1.749	34.802
0.04	9	38.708	1.510	39.614
0.05	9	43.200	1.756	44.254
0.06	9	47.644	1.978	48.831
0.07	7	51.907	1.792	52.982
0.08	7	55.541	1.950	56.711
0.09	7	59.109	2.095	60.367
0.10	7	62.605	2.228	63.942
0.11	6	65.972	2.089	67.225
0.12	4	69.170	1.551	70.101
0.13	4	71.607	1.629	72.584
0.14	4	74.012	1.702	75.033
0.15	4	76.384	1.770	77.447
0.16	4	78.724	1.835	79.825
0.17	4	81.031	1.896	82.168
0.18	4	83.304	1.952	84.475
0.19	4	85.543	2.005	86.746
0.20	4	87.748	2.054	88.980

Table 3(h): $\alpha_I = 0.1$, $\alpha^* = 0.1$, $\alpha_s = 0.05$

$c_2 : c_1 = 1 : 5$

p	k	E(R)	I	min C
0.01	32	23.930	1.688	24.267
0.02	15	34.116	1.576	34.431
0.03	14	40.094	1.912	40.476
0.04	9	45.523	1.684	45.860
0.05	9	49.539	1.914	49.922
0.06	9	53.538	2.122	53.962
0.07	9	57.486	2.308	57.948
0.08	7	61.317	2.090	61.735
0.09	7	64.546	2.225	64.991
0.10	7	67.721	2.348	68.191
0.11	7	70.837	2.460	71.329
0.12	7	73.887	2.561	74.400
0.13	6	76.767	2.399	77.247
0.14	4	79.489	1.838	79.857
0.15	4	81.634	1.901	82.014
0.16	4	83.754	1.960	84.146
0.17	4	85.849	2.014	86.252
0.18	4	87.918	2.065	88.331
0.19	4	89.961	2.113	90.383
0.20	4	91.976	2.157	92.408

Table 3(i): $\alpha_I = 0.1$, $\alpha^* = 0.1$, $\alpha_s = 0.05$

$c_2 : c_1 = 3 : 5$

p	k	E(R)	I	min C
0.01	32	23.930	1.688	24.942
0.02	15	34.116	1.576	35.062
0.03	14	40.094	1.912	41.241
0.04	9	45.523	1.684	46.533
0.05	9	49.539	1.914	50.688
0.06	9	53.538	2.122	54.810
0.07	9	57.486	2.308	58.871
0.08	7	61.317	2.090	62.571
0.09	7	64.546	2.225	65.881
0.10	7	67.721	2.348	69.130
0.11	7	70.837	2.460	72.312
0.12	6	74.011	2.310	75.397
0.13	6	76.767	2.399	78.207
0.14	4	79.489	1.838	80.592
0.15	4	81.638	1.901	82.775
0.16	4	83.754	1.960	84.930
0.17	4	85.849	2.014	87.058
0.18	4	87.918	2.065	89.157
0.19	4	89.961	2.113	91.228
0.20	4	91.976	2.157	93.270

Tables 3(j), 3(k), 3(ℓ) and 3(m):

Optimum group-sizes obtained by minimizing the maximum value of the total expected number of incorrect decisions I for a specified value of the expected total number of runs, ν, when $\alpha_{Ii} = \alpha_I$, $\alpha_i^* = \alpha^*$, $\alpha_{si} = \alpha_s$ and $f = 100$ for selected unequal a-priori probabilities.

In these plans,

$$k_i = A\frac{1}{p_i} + B$$

where

$$B = -\frac{a}{2} + \left[\frac{a^2}{4} + \frac{(d\sum_{i=1}^{g}\frac{1}{p_i} - f^2 - agf)}{\sum_{i=1}^{g} p_i \sum_{i=1}^{g}\frac{1}{p_i} - g^2}\right]^{\frac{1}{2}}$$

$$A = \frac{f - Bg}{\sum_{i=1}^{g}\frac{1}{p_i}}$$

$$a = \frac{\frac{1}{2}(1-\alpha_s)(3-\alpha^*)}{(1-\alpha_I) - \frac{1}{2}(1-\alpha_s)(1+\alpha^*)}$$

and

$$d = \frac{\nu - h - g - (\alpha^* + \alpha_I)}{(1-\alpha_I) - \frac{1}{2}(1-\alpha_s)(1+\alpha^*)}$$

(c.f. Theorem 4.2.3)

for

$$h = 4\left[\frac{g}{4}\right] - g$$

$$\text{Minmax } I = \alpha_s \alpha_I f - \alpha_s \sum_{i=1}^{g} k_i p_i + \alpha_s (1-\alpha_I) \sum_{i=1}^{g} k_i^2 p_i$$

(c.f. Corollary 4.2.3)

<u>Table 3(j):</u> $h=3$, $g=13$, $\nu=24$, $\alpha_I=0.05$

$\alpha^*=0.05$, $\alpha_s=0.05$, $p_i \leq p = 0.035$

i	p_i	k_i
1	0.008	15.380
2	0.009	13.691
3	0.010	12.340
4	0.013	9.533
5	0.015	8.286
6	0.017	7.332
7	0.020	6.259
8	0.022	5.706
9	0.025	5.043
10	0.027	4.682
11	0.030	4.232
12	0.033	3.863
13	0.035	3.653
Total		100.000

MinMax I = 0.760

Table 3(k): $h=3$, $g=13$, $\nu=25$, $\alpha_I=0.05$,

$\alpha^*=0.05$, $\alpha_s=0.05$, $p_i \leq p = 0.035$

i	p_i	k_i
1	0.008	10.338
2	0.009	9.757
3	0.010	9.292
4	0.013	8.326
5	0.015	7.897
6	0.017	7.568
7	0.020	7.199
8	0.022	7.009
9	0.025	6.780
10	0.027	6.656
11	0.030	6.501
12	0.033	6.375
13	0.035	6.302
Total		100.000

MinMax I = 0.813

Table 3(ℓ): $h=4$, $g=20$, $\nu=49$, $\alpha_I=0.05$,

$\alpha^*=0.05, \alpha_s=0.05$, $p_i \leq p = 0.100$

i	p_i	k_i
1	0.040	8.614
2	0.045	7.641
3	0.050	6.863
4	0.053	6.467
5	0.055	6.227
6	0.060	5.696
7	0.062	5.508
8	0.065	5.247
9	0.070	4.863
10	0.075	4.529
11	0.078	4.350
12	0.080	4.237
13	0.082	4.131
14	0.085	3.980
15	0.087	3.885
16	0.090	3.752
17	0.092	3.667
18	0.095	3.546
19	0.098	3.434
20	0.100	3.363
Total		100.000

MinMax I = 1.528

Table 3(m): h=4, g=20, ν=50, α_I=0.05,

α^*=0.05, α_s=0.05, $p_i \leq p$ = 0.100

i	p_i	k_i
1	0.040	6.142
2	0.045	5.835
3	0.050	5.589
4	0.053	5.464
5	0.055	5.388
6	0.060	5.220
7	0.062	5.161
8	0.065	5.078
9	0.070	4.957
10	0.075	4.850
11	0.078	4.795
12	0.080	4.759
13	0.082	4.725
14	0.085	4.678
15	0.087	4.648
16	0.090	4.605
17	0.092	4.579
18	0.095	4.540
19	0.098	4.505
20	0.100	4.482
	Total	100.000

MinMax I = 1.564

Tables 3(n), 3(o), 3(p) and 3(q):

Optimum group-sizes which minimize expected total cost C, when $\alpha_{Ii}=\alpha_I$, $\alpha_i^*=\alpha^*$ and $\alpha_{si}=\alpha_s$, for f=100 and for selected unequal a-priori probabilities

In these plans,

$$k_i = \left(f - \sum_{i=1}^{g} G_i H_i\right) \frac{G_i}{\sum_{i=1}^{g} G_i} + G_i H_i$$

where

$$G_i = \frac{-1}{\left\{2(1-\alpha_I)-(1-\alpha_s)(1+\alpha^*)\right\}c_1 p_i + 2\alpha_s(1-\alpha_I)c_2 p_i}$$

and

$$H_i = c_1\left\{(\alpha^*+\alpha_I)+\frac{1}{2}(1-\alpha_s)(3-\alpha^*)p_i\right\} - c_2\alpha_s(\alpha_I - p_i).$$

$$C = c_1 E(R) + c_2 I$$

where

$$E(R) = h+g+(\alpha^*+\alpha_I)f-(1-\alpha_s)\sum_{i=1}^{g} p_i+(1-\alpha_I)\sum_{i=1}^{g} k_i^2 p_i$$

$$+ \frac{1}{2}(1-\alpha_s)(3-\alpha^*)\sum_{i=1}^{g} k_i p_i$$

$$- \frac{1}{2}(1-\alpha_s)(1+\alpha^*)\sum_{i=1}^{g} k_i^2 p_i$$

$$I = \alpha_s\alpha_I f - \alpha_s\sum_{i=1}^{g} k_i p_i + \alpha_s(1-\alpha_I)\sum_{i=1}^{g} k_i^2 p_i$$

(c.f. Theorem 4.2.5)

The minimum C given is a relative figure using c_1 i.e., the cost of observing a run as the unit. [i.e. it indicates the value of C/c_1] .

Table 3(n): $h=3$, $g=13$, $\alpha_I=0.05$, $\alpha^*=0.05$

$\alpha_s=0.05$, $c_2:c_1=1:5$, $p_i \leq p = 0.035$

i	p_i	k_i
1	0.008	17.129
2	0.009	15.056
3	0.010	13.397
4	0.013	9.952
5	0.015	8.421
6	0.017	7.250
7	0.020	5.933
8	0.022	5.254
9	0.025	4.440
10	0.027	3.998
11	0.030	3.445
12	0.033	2.992
13	0.035	2.733
Total		100.000

$E(R) = 33.577$

$I = 0.771$

Min $C = 33.731$

Table 3(o): h=3, g=13, α_I=0.05, α^*=0.05

α_s=0.05, c_2:c_1=3:5 $p_i \leq p$ = 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) = 33.577

I = 0.770

Min C = 34.039

Table 3(p): $h=4$, $g=20$, $\alpha_I=0.05$, $\alpha^*=0.05$

$\alpha_s=0.05$, $c_2:c_1=1:5$, $p_i \leq p = 0.100$

i	p_i	k_i
1	0.040	9.593
2	0.045	8.357
3	0.050	7.368
4	0.053	6.865
5	0.055	6.559
6	0.060	5.886
7	0.062	5.644
8	0.065	5.315
9	0.070	4.826
10	0.075	4.402
11	0.078	4.174
12	0.080	4.032
13	0.082	3.895
14	0.085	3.704
15	0.087	3.583
16	0.090	3.412
17	0.092	3.305
18	0.095	3.153
19	0.098	3.009
20	0.100	2.918
	Total	100.000

$E(R) = 56.421$

$I = 1.546$ Min $C = 56.731$

Table 3(q): h=4, g=20, α_I=0.05, α^*=0.05,

α_s=0.05, $c_2:c_1$=3:5, $p_i \leq p$ = 0.100

i	p_i	k_i
1	0.040	9.565
2	0.045	8.337
3	0.050	7.354
4	0.053	6.853
5	0.055	6.550
6	0.060	5.880
7	0.062	5.642
8	0.065	5.313
9	0.070	4.827
10	0.075	4.403
11	0.078	4.179
12	0.080	4.037
13	0.082	3.902
14	0.085	3.712
15	0.087	3.592
16	0.090	3.423
17	0.092	3.316
18	0.095	3.164
19	0.098	3.020
20	0.100	2.931
	Total	100.000

E(R) = 56.422 ; I = 1.545 ; Min C = 57.349

CONCLUDING REMARKS

The usual sampling inspection plan, consists of drawing a sample or samples from the population. All the items in the sample(s) are then examined. If the proportion of defective items in the sample(s) is small, then they are replaced by good ones and all the items in the population are accepted. In such a case, some items are passed without being inspected. In group screening designs however, every item is subject to inspection either in groups or individually. Group screening designs are thus some kind of 100% sampling inspection plans.

In this thesis, a class of group screening designs which we have called "the step-wise designs" are studied. The step-wise group screening design requires fewer runs than the corresponding two stage group screening design, for all prevalence rates of defectives for which a two stage group screening design has fewer runs than a single stage design. The two stage group-screening design and consequently the step-wise group screening design has fewer runs, than an s-stage ($s \geq 3$) group screening design for prevalence rates of defective greater than 0.09. The step-wise group screening design has fewer runs than the single stage design for a wider range of prevalence rates of defective items than an s-stage ($s \geq 2$) group screening design.

Group screening techniques can be used in industries in sorting out defective items from non defective ones with substantial saving in cost of inspection and time. In chemical industry, the technique has been used for example in (i) classifying an unknown chemical element, (ii) selecting the best catalyst for a chemical reaction from a large number of compounds which are possible candidates. Group screening techniques have also been applied in Biological experiments.

In this thesis, we have assumed that the direction of the effect of a defective factor is known or is correctly assumed a-priori. Further work could be done in step-wise design by relaxing this assumption. We could allow the possibility of cancellation of effects.

BIBLIOGRAPHY

[1] Connor, W.S. "Developments in the Design of Experiments - Group screening Designs"; The proceedings of the sixth conference on the Design of experiments in Army Research Development and Testing (1961).

[2] Curnow, R.N. "A note on G.S. Watson's paper, A study of the group screening method"; Technometrics 7, 444-446, (1965).

[3] Dorfman, R. "The detection of defective members of large populations"; Annals of Mathematical Statistics, 14, 436-440, (1943).

[4] Finucan, H.M. "The blood testing problem"; Applied Statistics, 13, 43-50, (1964).

[5] Garey, M.R. and Hwang, F.K. "Isolating a single defective using group testing"; J.A.S.A. 69, No. 345 151-153 (1974).

[6] Graff, L.E. and Roeloffs, R. "Group testing in the presence of test error; an extension of Dorfman's Procedure; Technometrics, 14, 113-122, (1972).

[7] Kleijnen, J.P.C. "Screening Designs for poly-factor experimentation"; Technometrics, 17, 487-493, (1975).

[8] Li, C.H. "A sequential method for screening experimental variables"; J.A.S.A. 57, 455-477, (1962).

[9] Mauro, C.A. and Smith, D.E. "The performance of two-stage Group screening in Factor screening Experiments"; Technometrics 24, 325-330, (1982).

[10] Mauro, C.A. "On the performance of two stage group screening designs"; Technometrics vol.26, No.3, 255-264, (August 1984).

[11] Mauro, C.A. and K.C. Burns, "A comparison of Random balance and Two stage Group screening designs: A case study"; Comm. stat. Theory and Methods vol. 13, No. 21, 2625-2647, (1984).

[12] Odhiambo, J.W. "Group-Screening designs with more than Two stages"; Ph.D. Thesis submitted to the University of Nairobi,(1982).

[13] Patel, M.S. "On the performance of the group screening method"; Research Triangle Institute, Technical Report No. 4, May (1961).

[14] Patel, M.S. "Group screening with more than two stages"; Technometrics 4, 209-217, (1962).

[15] Patel, M.S. "A note on Watson's Paper"; Technometrics, 5, 397-398, (1963).

[16] Patel, M.S. "On selecting the factors for experimentation"; Kenya Journal of science and Technology, Series A, vol. 1, 99-101, (1980)

[17] Patel, M.S. "On some aspects of a two stage group screening method"; Paper read at the First Eastern African Symposium, (June 1981).

[18] Patel, M.S. "A critical look at two stage group screening method"; Kenya Journal of science and Technology series A, vol. 4, No. 2, 63-88 (1983).

[19] Patel, M.S. and Ottieno J.A.M. "Optimum two stage Group-Screening designs"; Comm. stat. Theor. Meth. vol.13, No. 21, 2649-2663, (1984).

[20] Patel, M.S. and Ottieno J.A.M. "Two-stage group-screening designs with equal prior probabilities and no errors in decisions"; Comm. Stat. Theor. Meth. vol. 13, No.9 1149-1160, (1984).

[21] Patel, M.S. and Ottieno J.A.M. "Two-stage group-screening designs with unequal a-priori probabilities"; Comm. stat. Theor. Meth. vol.13, No. 6, 761-779, (1984).

[22] Plackett, R.L. and J.P. Burman "The design of optimum multi-factorial experiment"; Biometrica, 33, 305-325, (1946).

[23] Samuels, S.M. "The exact solution to the two-stage group screening problem"; Technometrics 20, 497-500, (1978).

[24] Sobel, M. and P.A. Groll "Group-testing to eliminate efficiently all defectives in a binomial sample"; The bell system Journal, 38, 1179-1252, (1959).

[25] Sobel, M. and P.A. Groll "Binomial Group testing with an unknown proportion"; Technometrics 8, 631-656, (1966).

[26] Sterret, A. "On the detection of defective members of large populations"; Annals of Mathematical statistics 28, 1033-1036, (1957).

[27] Thompson, K.H. "Estimation of the proportion of vectors in a natural population of insects"; Biometrics, 18, 568-578, (1962).

[28] Watson, G.S. "A study of group screening method"; Technometrics 3, 371-388, (1961).

[29] William, G. Hunter and Reiji Mezaki "Catalyst selection by group screening"; Industrial and Engineering Chemistry, vol. 56, No. 3, 38-40, (1964).