On Selecting the Factors for Experimentation

by

M. S. Patel

Purdue University

Mimeograph Series Number 4

Department of Statistics

April 1963

On Selecting the Factors for Experimentation[1]

by

M. S. Patel
Purdue University

1. Introduction and Summary. When it is presumed that a large number of factors contribute to an observed response in an experiment, of which only a few are really effective in changing the response, Connor [1], Watson [4] and the author [2] himself have suggested experimental technique known as the Group-screening method. Designs used in this technique have been called Group-screening designs. In these designs, the factors are tested in groups initially and then at later stages, the factors within the significant group factors are tested in smaller group sizes and finally the significant factors are detected. It was assumed that a factor is effective with a priori probability p. The method was developed to detect the true effects with as few runs as possible. In this paper, a different approach is suggested. For instance, it is pointed out that instead of starting an experiment with a large number of factors, it is preferable to proceed with a relatively small number of factors with probability of being effective different from 1/2. It is shown that the presence of such factors in the experiment tend to minimize

[1] Part of the work presented in this paper was done by the author while he was at the Research Triangle Institute, Durham, North Carolina, under Contract No. DA-01-009-ORD-816, Office of Ordnance Research Project No. 2579.

the maximum number of correct decisions that can be made from the information available. Obviously the number of runs is reduced and the number of correct decisions is made optimum.

2. Assumptions, Terminology and Notation. For simplicity, it will be assumed that (i) all factors have, independently the same a priori probability p of being effective, (ii) effective factors have the same effect $\Delta > 0$, (iii) none of the factors interact, (iv) the errors of observations are independent $N(0, \sigma^2)$ with $\sigma^2 \neq 0$ and known. A factor will be said to be effective if it produces a non-zero change in the mean value. The word effective will also be used whenever a factor is found to be significant. If the effective factors have different effects, then Δ in (ii) will denote the minimum amount of these effects. Only orthogonal designs, i.e., the designs in which the estimates of the effects are orthogonal will be used in this paper.

Let $[f/4]$ denote the least integer greater than $f/4$ except that $[f/4] = 0$ for $f = 0$. Each factor has two levels, the upper level and the lower level. The experiment is run with $4[f/4]$ runs so as to estimate the main effects orthogonally (cf. Plackett and Burman [3]). The variance of the estimates is $\sigma^2/4[f/4]$. Test of hypothesis that the main effect is zero may be made using a normal deviate test η. The power function of this test will be denoted by

$$\pi\left(\sqrt{2[f/4]}\,\frac{\Delta}{\sigma}, \alpha\right)$$

where α is the level of significance.

3. Criteria for Optimum Design. The distribution of η, under the alternative hypothesis, is normal with mean equal to the non-centrality parameter. Hence,

(1) $$\pi\left(\sqrt{2[f/4]}\,\frac{\Delta}{\sigma}, \alpha\right) = \int_{\eta(\alpha)}^{\infty} (2\pi)^{-1/2} \exp - (\eta-\delta)^2/2 \; d\eta$$

$$= \int_{\eta(\alpha)-\delta}^{\infty} (2\pi)^{-1/2} \exp - z^2/2 \; dz \; ,$$

where

(2) $$\alpha = \int_{\eta(\alpha)}^{\infty} (2\pi)^{-1/2} \exp - z^2/2 \; dz = 1 - F[\eta(\alpha)]$$

and the non-centrality parameter is

(3) $$\delta = \sqrt{4[f/4]}\,\frac{\Delta}{\sigma} \; .$$

Let E and $\bar{E}$ denote the number of effective factors declared effective, and the number of ineffective factors declared effective respectively. Then the expected value of E is

(4) $$f\ p\ \pi(\sqrt{2[f/4]}\ \frac{\Delta}{\sigma},\alpha),$$

and the expected value of $\bar{E}$ is $fq\alpha$ where $q = 1-p$. Hence, the expected number of correct decisions is given by

(5) $$N = f\ p\ \pi(\sqrt{2[f/4]}\ \frac{\Delta}{\sigma},\alpha) + f\ q(1-\alpha).$$

Critical values of N occur at values of α satisfying the equation

(6) $$\frac{dN}{d\alpha} = fp\ \frac{d\pi}{d\alpha} - fq = 0.$$

From (1),

(7) $$\frac{d\pi}{d\alpha} = \left\{-(2\pi)^{-1/2} \exp - [\eta(\alpha) - \delta]^2/2\right\} \frac{d\eta(\alpha)}{d\alpha},$$

and from (2) follows

(8) $$1 = -F'[\eta(\alpha)]\ \frac{d\eta(\alpha)}{d\alpha},$$

i.e.,

(9) $$\frac{d\eta(\alpha)}{d\alpha} = -1/f(\eta(\alpha)),$$

where $f(z)$ is the normal ordinate at z. Substituting in (7), we get

(10) $$\frac{d\pi}{d\alpha} = \exp\ [\eta(\alpha)\delta - (\delta^2/2)]\ .$$

Therefore (6) implies

(11) $$\eta(\alpha) = [\log_e(q/p) + \delta^2/2]/\delta \ .$$

Since

(12) $$\frac{d^2N}{d\alpha^2} = -fp\delta \exp [\eta(\alpha)\delta - (\delta^2/2)]/f(\eta(\alpha)).$$

which is negative for all α, the value of α given by (11) maximizes N. Substituting this value of α in (5), we get

(13) $$\max_{\alpha} N = f\ p\ F[\{-\log_e(q/p) + \delta^2/2\}/\delta] + f\ q\ F[\{\log_e(q/p) + \delta^2/2\}/\delta].$$

Denoting $\max_{\alpha} N$ by N^*, it is easy to show that, at $p = 1/2$

(14) $$\frac{dN^*}{dp} = 0,$$

and

(15) $$\frac{d^2N^*}{dp^2} = 8f.\ f(\delta/2)/\delta$$

which is positive. Hence, N^* has a minimum at $p = 1/2$. This leads us to conclude that, from the point of view of achieving the maximum number of correct decisions in an experiment, the optimum design should exclude those factors for which the a priori probability of significance is 1/2 or is in the neighborhood of 1/2. It also provides us with some information about the factors before an elaborate experiment is actually planned.

REFERENCES

[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.

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

[3] Plackett, R. L. and Burman, J. P. (1946). The design of optimum multifactorial experiments. Biometrika. 33. 305-325.

[4] Watson, G. S. (1961). A study of the group-screening method. Technometrics, 3. 371-388.

(*) exact information will be provided in the revised copy.