TWO-SAMPLE DECISION PROBLEM— DETERMINATION OF AMOUNT OF EXPERIMENTATION

By A. H. Manwani

Institute of Agricultural Research Statistics (I.C.A.R.), New Delhi

1. Introduction

Suppose we are required to take a decision whether or not a given new process of production P_1 should be adopted in preference to an old process say P_0. While planning an experiment for taking a decision to this end, we have to keep two propositions in view. At the first instance, the experiment should be precise enough to enable us to arrive at a right type of decision by detecting the smallest possible gain or loss due to the adoption of the new process. Secondly, the cost of experiments leading to the decision should not be so large as to make the experiment uneconomical; in other words, the cost incurred on the experiment should be justifiable when judged from the point of view of the expected gain accruing from the correct decision to be arrived at. These two propositions are contradictory to one another, since the amount of precision is directly related to the size of the experiment. Hence, there arises a need for planning the experiment so as to achieve an optimum balance between the two propositions. In other words, having realized the importance of experimental results and the size of gain or loss due to right or wrong decision, it is required to allocate a reasonable amount of expenditure and the experimental resources which may give results with a reasonable degree of precision.

Assuming that the cost of the experiment is proportional to the number of observations to be taken for the experiment, the problem at hand reduces to the determination of the optimum number of observations to be taken for deciding between two alternatives, *i.e.*, the acceptance of P_1 or P_0. For this, the procedure suggested here, is to define a suitable risk function involving the cost of experimentation and the expected gain which will result if the production process P_1 is rightly accepted. The total risk will be defined as the cost of experimentation *minus* the expected gain. The risk function so defined,

will naturally involve certain unknown parameters as well as certain other constants depending on the external economic factors of the decision process. The latter economic factors, can be determined with sufficient accuracy from experience and knowledge of the situations under which the experiment is being carried out. As far as the estimation of the parameters is concerned, it has been suggested that a preliminary experiment of a small size say n_1 be carried out at the first instance and the estimates of the parameters obtained from this small experiment may be properly used in defining the risk function. The risk function so defined, has been called the Integral risk function. This function is now mainly a function of the additional number of observations, say n_2 required to be determined, and the estimates of the parameters based on the set of values of n_1 observations in the first sample. The problem now, is to see whether or not the risk will be reduced in case, the experiment is continued one stage further and the proposed decision is taken on the basis of the results obtained from the two experiments together, *i.e.*, on the basis of $n_1 + n_2$ observations. In case, the risk involved in continuing the experiment one stage further is less than that in stopping the experiment at the first stage, *i.e.*, taking decision on the basis of first n_1 observations only, the value of n_2, *i.e.*, the size of the experiment at the second stage has to be determined in such a manner as to possess certain well-defined optimum properties. Once the optimum number of observations for the experiment has been determined, it will invariably fix the optimum amount of precision attached with the decision process, since the two are directly related. Following this approach, Grundy and others (1956) have determined the amount of experiments for various cost situations in the case when the observations in the experiment are distributed normally with a known variance (restriction imposed in Grundy's method). The present work deals with the case when the observations are taken from a normal population with unknown variance.

2. Decision Rule for Choosing between Two Alternatives

Let us suppose that $x_{11}, x_{12} \cdots x_{1n_1}$ are the values of n_1 observations in the first experiment of size n_1. Each x_i denotes the value of net gain per unit of production accruing from adopting the new process P_1 over the old process P_0. Let $\bar{x}_1$ be the mean and s_1^2 the mean square based on n_1 observations. Throughout, it will be assumed that all x_i's are distributed normally with the same unknown mean and the variance σ^2. Clearly, $\bar{x}_1$ and s_1^2 are respectively the estimates of the average net gain η and variance σ^2. Let us say, the first experiment

is not found to be sufficient enough to decide whether or not it is worthwhile to adopt the new process P_1 in preference to P_0. Hence, another independent experiment of size n_2 is carried out. This experiment gives values say, x_{21}, x_{22}, x_{23}, $\cdots$ x_{2n_2} of the net gain. Let $\bar{x}_2$ be an independent estimate of η. The combined estimate of η from two experiments will be:

$$\bar{x} = \frac{n_1\bar{x}_1 + n_2\bar{x}_2}{n_1 + n_2}.$$

The decision rule whether the P_1 should be accepted or rejected is given by:

accept P_1 if $\bar{x} > 0$,

reject P_1 *i.e.*, retain P_0 if $\bar{x} \leq 0$.

The question, whether the first n_1 observations are adequate or not for taking the decision and in case they are not, how many more observations should be taken in order to arrive at a definite decision has been discussed in the following paragraph.

3. Definition of the Risk Function

Let $P(n_2, \bar{x}_1, \eta, \sigma^2)$ denote the conditional probability of accepting P_1 on the basis of the decision rule given in (2) for given values of $\bar{x}_1$ and n_2. It can easily be seen that:

$$P(n_2, \bar{x}_1, \eta, \sigma^2) = P(\bar{x} > 0/n_2, \bar{x}_1, \eta, \sigma^2)$$
$$= \phi\left[\frac{(n_1\bar{x}_1 + n_2\eta)}{\sqrt{n_2}\sigma}\right]$$

where

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-v/2}\, dv. \tag{3.1}$$

The conditional risk function $R = R(n_2, \bar{x}_1, \eta, \sigma^2)$ is defined as,

$$R = kn_2 - M\eta P(n_2, \bar{x}_1, \eta, \sigma^2) \tag{3.2}$$

where 'k' is the cost of taking single observation and 'M' is a constant measuring the scale of production or the units on which P_1 will be recommended. In the case, where P_1 represents the application

of fertilizers or other improved practices, 'M' may be taken as units, say Hectares on which these practices will be applied if the experiment so recommends. In the case where P_1 represents some manufacturing process, 'M' represents the number of units which will be manufactured through P_1 over a specified time. It is obvious that the risk function as defined above, involves the parameters η and σ^2 of which no knowledge is available except their respective estimates $\bar{x}_1$ and s_1^2 as obtained from the initial small experiment of size n_1. There are two ways by which the risk function can be made independent of the unknown parameters. At the first instance, R can be averaged over an *a-priori* distribution of η and σ^2, if such distribution can be determined or guessed. Secondly, R can be integrated over the joint fiducial distribution of η and σ^2 (Fisher). As it is not easy to choose an optimum *a-priori* distribution of η and σ^2, we may adopt the second course.

The fiducial distribution $f(\eta, \sigma^2)$ of η and σ^2 is given by,

$$df(\eta, \sigma) = \sqrt{\frac{n_1}{2\pi}} \frac{1}{\sigma} \{e^{-n(\eta - \bar{x}_1)^2/2\sigma^2}\} \times \frac{1}{\left|\frac{n_1 - 1}{2}\right.} \left(\frac{n_1 - 1}{\sigma^2} s_1^2\right)^{(n_1-1)/2} e^{-(n_1-1)/2\, s_1^2}\, d\eta d\sigma.$$

Hence, the integral risk function of the expected value of R as averaged over $df(\eta, \sigma)$ is given by,

$$\bar{R} = kn_2 - M \int_{-\infty}^{\infty} \int_{\infty}^{\infty} \eta \phi \left(\frac{n_1 x_1 + n_2 \eta}{\sqrt{n_2}\sigma}\right) df\ (\eta, \sigma). \tag{3.3}$$

Integrating over η first we get,

$$\bar{R} = kn_2 - M\left[\sqrt{\frac{N}{(N+1)\, n_1}}\, A_1 + \bar{x}_1 A_2\right]$$

where

$$A_1 = \frac{1}{\left|\left(\frac{n_1 - 1}{2}\right)\right.} \int_0^{\infty} \sigma^2 \phi'\left(\frac{l}{\sigma}\right) e^{-K^2/\sigma^2} \left(\frac{K^2}{\sigma^2}\right)^{(n_1-1)/2} \frac{2d\sigma}{\sigma}$$

$$= s_1 \sqrt{\frac{(n_1 - 1)}{(n_1 - 3)}}\, T_{n_1-3}\left[\sqrt{\frac{(N+1)\,(n_1-3)}{N(n_1-1)}}\, X\right],$$

$$N = \frac{n_2}{n_1}, \quad l = \frac{n_1(N+1)}{N}, \quad K = \sqrt{\frac{n_1 - 1}{2}}\, s_1,$$

$$X = \eta_1^{\frac{1}{2}} \frac{\bar{x}_1}{s_1^2}.$$

$T_{n_1-3}(x)$ being the ordinate of 't' distribution with $n_1 - 3$ degree of freedom at x point and

$$A_2 = \frac{1}{\Gamma\left(\frac{n_1-1}{2}\right)} \int_0^\infty \phi\left(\frac{l}{\sigma}\right)\left(\frac{K^2}{\sigma^2}\right)^{(n_1-1)/2} e^{K^2/\sigma^2} \frac{2d\sigma}{\sigma}$$

$$= \int_V \int_{<(l/K)U} \chi^2_{n_1-1}(U)\ \phi'(V)\, dU dv.$$

'U' being the chi-square variate with $\chi^2_{n_1-1}(U)$ distribution and V the normal variate with the distribution $\phi(V)$. Or,

$$A_2 = T_{n_1-1}\left(\sqrt{\frac{N+1}{N}}\, X\right).$$

where

$$T_{n_1-1}(x) = \frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi(n-1)}\,\Gamma\left(\frac{n-1}{2}\right)} \int_{-\infty}^{x} \left\{\frac{1}{\left(1 + \frac{t^2}{n-1}\right)^{n/2}}\right\} dt.$$

Hence,

$$\bar{R} = kn_2 - M\left[\sqrt{\frac{N(n_1-1)}{(N+1)(n_1-3)}}\, s_1 T'_{n_1-3}\left(\sqrt{\frac{(N+1)(n_1-3)}{N(n_1-1)}}\, X\right)\right.$$

$$\left. + \bar{x} T_{n_1-1}\left(\sqrt{\frac{N+1}{N}}\, X\right)\right] \qquad (3.4)$$

or we may preferably put it in the form,

$$\bar{Q} = N - \lambda\sqrt{\frac{N}{N+1}}\left[\theta T_{n_1-1}(\theta)\right.$$

$$\left. + \sqrt{\frac{n_1-1}{n_1-3}}\, T'_{n_1-3}\left(\sqrt{\frac{n_1-3}{n_1-1}}\, \theta\right)\right] \qquad (3.5)$$

where

$$\bar{Q} = \frac{\bar{R}}{kn_1}, \quad \lambda = \frac{M}{k} n_1^{-3/2} s_1 \quad \text{and} \quad \theta = X\sqrt{\frac{N+1}{N}}.$$

It will be seen that the Integral risk function $\bar{Q}$ as defined above is a generalized form of the risk function given by Grundy and others in the case when σ^2 is known, *viz.*,

$$\bar{Q} = N - \lambda \sqrt{\frac{N}{N+1}} \left[\theta\phi(\theta) + \phi'(\theta)\right] \tag{3.6}$$

where

$$\lambda = \frac{M}{k} n_1^{-3/2} \sigma, \quad \theta = X \sqrt{\frac{N+1}{N}}, \quad X = n^{1/2} \frac{\bar{x}}{\sigma}.$$

4. Study of the Integral Risk Function

Let us denote,

$$\bar{Q}_0 = \operatorname*{Lim}_{N \to 0} \bar{Q}$$

then,

$$\frac{\triangle \bar{Q}}{N} \equiv \frac{\bar{Q} - \bar{Q}_0}{N}$$

$$= 1 - \lambda N^{-1/2} (N+1)^{-1/2} \left[\sqrt{\frac{n_1 - 1}{n_1 - 3}} T'_{n_1 - 3}\left(\sqrt{\frac{n_1 - 3}{n_1 - 1}}\, \theta\right) - |\theta| T_{n_1 - 1}(-|\theta|)\right] \tag{4.1}$$

and

$$\frac{N^{3/2} (N+1)^{1/2} \frac{\partial}{\partial N}\left(\frac{\triangle \bar{Q}}{N}\right)}{\sqrt{\frac{n_1 - 1}{n_1 - 3}} T'_{n_1 - 3}\left(\sqrt{\frac{n_1 - 3}{n_1 - 1}}\, \theta\right)}$$

$$= \frac{2N+1}{2N+2} - \frac{|\theta| T'_{n_1 - 1}(-|\theta|)}{\sqrt{\frac{n_1 - 1}{n_1 - 3}} T'_{n_1 - 3}\left(\sqrt{\frac{n_1 - 3}{n_1 - 1}}\, \theta\right)} \tag{4.2}$$

also

$$\frac{\partial \bar{Q}}{\partial N} = 1 - \frac{\lambda}{2} N^{-1/2} (N+1)^{3/2} \sqrt{\frac{n_1 - 1}{n_1 - 3}} T'_{n_1 - 3}\left(\sqrt{\frac{n_1 - 3}{n_1 - 1}}\, \theta\right) \tag{4.3}$$

and

$$\frac{4N^{5/2}(N+1)^{5/2}}{\lambda T'_{n_1-1}(\theta)} \cdot \frac{\partial^2 \bar{Q}}{\partial N^2}$$

$$= (4N+1)\, N \frac{n_1 - 1}{n_1 - 2} - X^2 (N+1) \left(1 - \frac{4N+1}{n_1 - 3}\right) \quad (4.4)$$

From 4.4, it can be seen that for $n_1 \geqq 4$, $\partial^2 \bar{Q}/\partial N^2$ is negative in the neighbourhood of $N = 0$, and continuously increases with N, becoming positive after passing the value zero. Hence $\partial \bar{Q}/\partial N$ initially decreases from unity and after attaining some minimum value, increases and attains unity value. Therefore, the shape of $\bar{Q}$ could be one of the form given in the following three figures:

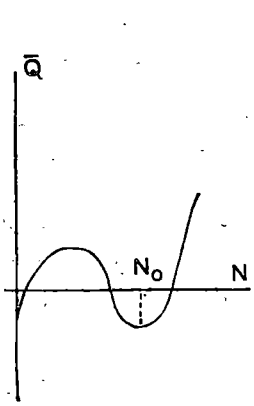

FIG. 1

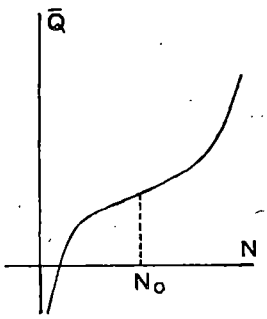

FIG. 2

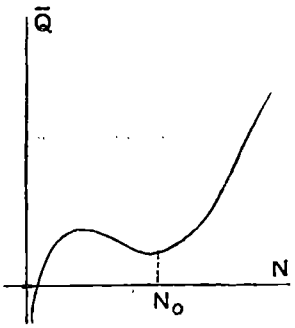

FIG. 3

In case, the shape of $\bar{Q}$ is of the form as given in Fig. 1, the minimum value exists at $N = N_0$ and at this point the risk $\bar{Q}$ is less than the initial risk at $N = 0$. Hence, in this case it will be worthwhile to conduct additional experiment of size $n_2 = n_1 N_0$. In case, the shape be of the form as given in Fig. 2, the minimum value does not exist hence no further experimentation need be carried out. Similarly, Fig. 3 shows that although the minimum value exists at some $N = N_0$, however, at this point, the risk involved is more than that the one at $N = 0$. Hence, in both these latter cases it will not pay to carry out experimentation beyond the first stage, *i.e.*, the decision based on the first n_1 observations will be as good as that taken on the basis of any additional amount of experimentation.

Now, referring to Fig. 1 and from (4.1) and (4.2), it can be seen that the minimum value of 'N' given by the equation $\partial\bar{Q}/\partial N = 0$, *i.e.*,

$$N^{1/2}(N+1)^{3/2} = \frac{\lambda}{2}\sqrt{\frac{n_1-1}{n_1-3}}\,T'_{n_1-3}\left(\sqrt{\frac{n_1-3}{n_1-1}}\,\theta\right) \qquad (4.5)$$

will be an absolute minimum if the inequality,

$$\frac{2N_0+1}{2N_0+2} > \frac{|\theta|\,T_{n_1-1}(-|\theta|)}{\sqrt{\left(\frac{n_1-1}{n_1-3}\right)}\;T'_{n_1-3}\left(\sqrt{\frac{n_1-3}{n_1-1}}\,\theta\right)}$$

is atisfied, where

$$\theta = X\sqrt{\frac{(N_0+1)}{N_0}}\,.$$

'N_0' is the larger value obtained by solving the Equation (4.5).

The explicit solution of the Equation (4.5) is rather difficult. However, one can easily get a graphical solution of this equation by drawing a Nomogram for various values of n_1. The Nomograms given in appendix have been drawn for $n_1 = 4$, 5 and 6. With the help of these Nomogram the value of n_2 could be determined for given values of $X = n_1^{1/2}X_1/s_1$ and $\lambda = M/K\,n_1^{-3/2}\,s_1$ corresponding to values of $n_1 = 4$, 5 and 6.

5. To Show that the Rule given above is Admissible

We have

$$R(x_1,\ s_1,\ \eta,\ \partial,\ n_2) = kn_2 - k'\eta P(\bar{x} > 0 \mid n_2,\ x_1,\ \eta,\ \partial).$$

Hence, the unconditional risk attached with the decision process can be put as,

$$R(\eta,\ \partial,\ \psi) = \int_{-\infty}^{\infty}\int_{0}^{\infty} R[x_1,\ s_1,\ \eta,\ \sigma,\ \psi(x,\ s_1)]\,df(x_1 s_1)$$

where $f(x_1, s_1)$ is the joint density function of s_1 and x_1 and $\psi(x_1, s_1) = n_2$; it being assumed that ψ is the function which defines 'n_2' in terms of x_1 and s_1.

Let us now assume that $\psi(x_1, s_1)$ is measurable and $E[\psi(x_1, s_1)]$ is finite.

Hence $E(P)$ and $E(n_2)$ are continuous functions.

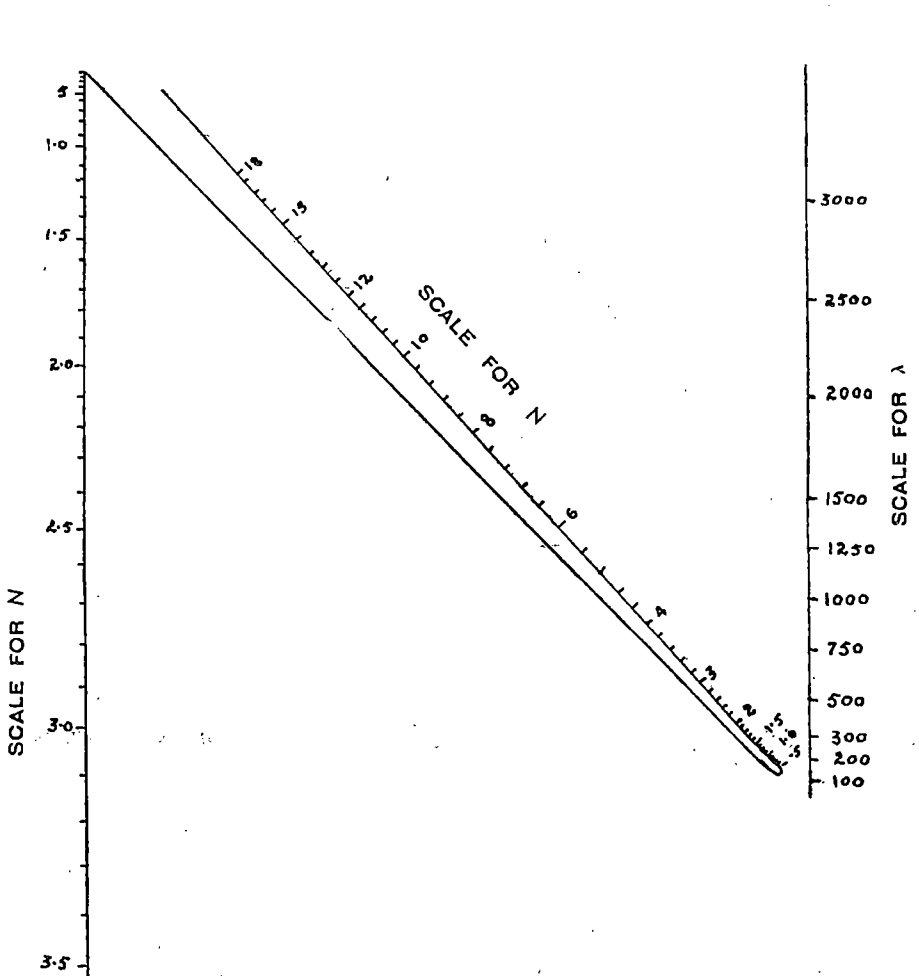

FIG. 4. Nomogram for X, λ and N when $n_1=4$

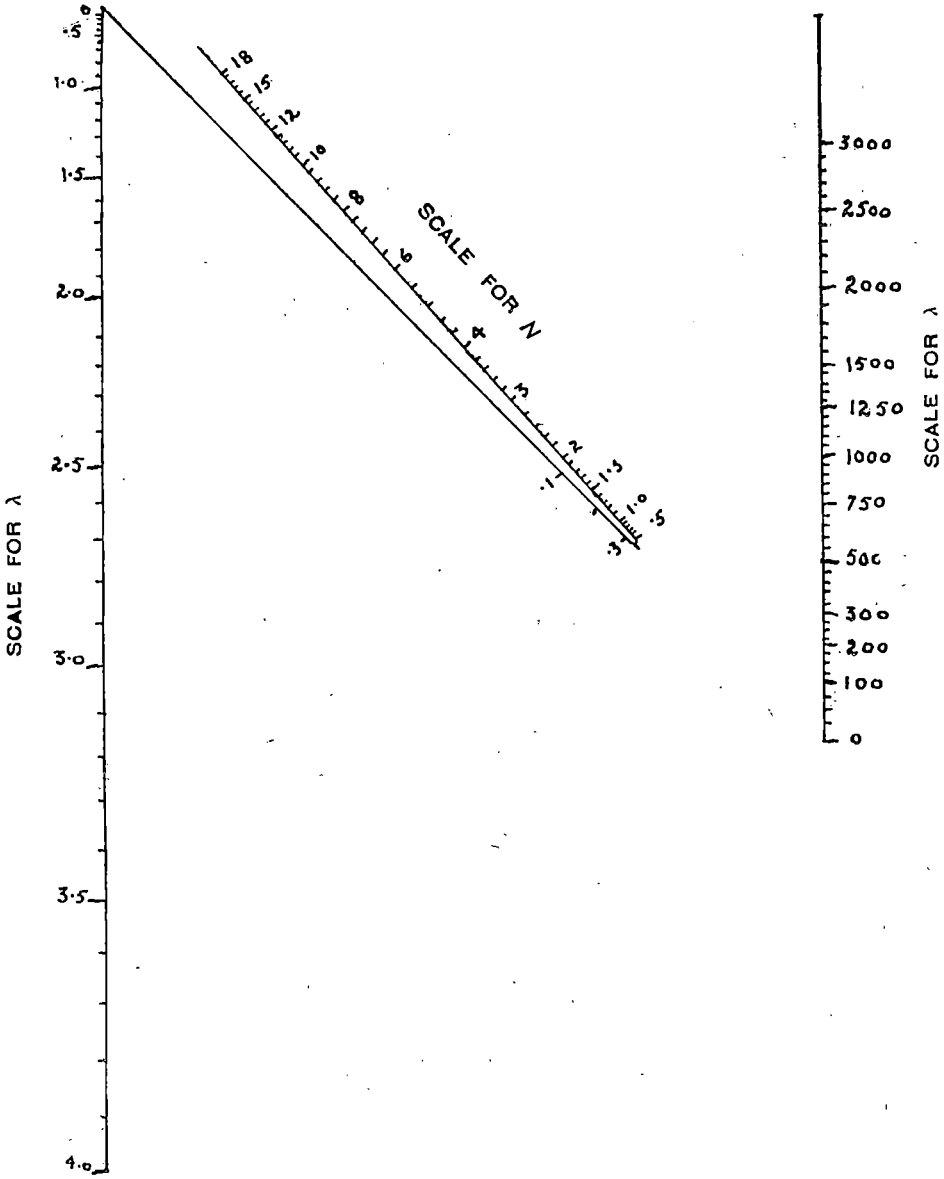

FIG. 5. Nomogram for X, λ and N when $n_1=5$

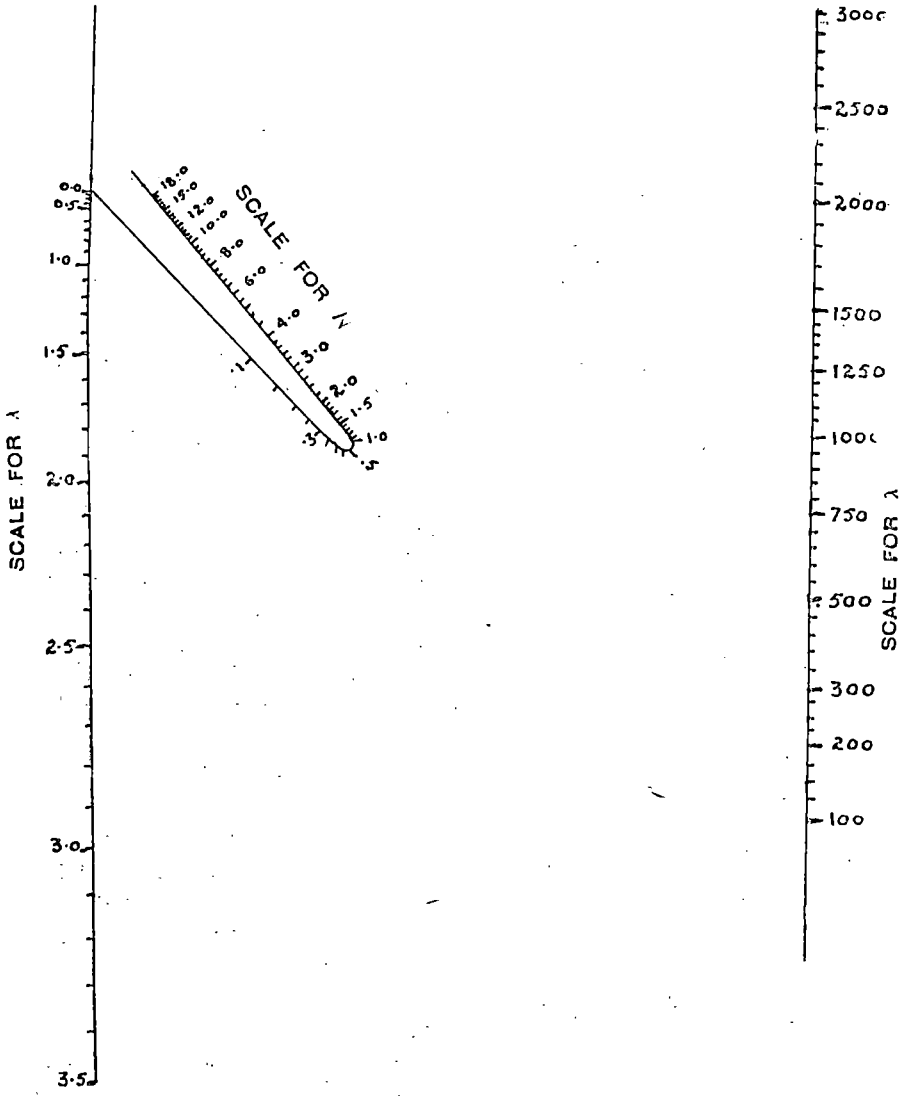

FIG. 6. Nomogram for X, λ and N when $n_1=6$

If we define,

$$R_0(\eta, \sigma, \psi) - R(\eta, \sigma, \psi) = R_0(\eta, x_1, s_1, n_2) - R(\eta, x_1, s_1, n_2)$$
$$= +M\eta \qquad \text{if} \quad \eta > 0$$
$$= 0 \qquad \text{if} \quad \eta \leqslant 0,$$

then by expressing

$$\iiint R_0 [\eta,\ \sigma,\ x_1,\ s_1,\ \psi(x_1,\ s_1)]\ df(x_1,\ s_1)\ d\eta d\sigma$$

as a repeated integral in two possible ways (since R_0 is a continuous function, the inversion of the order of integral is permissible), we obtain the identity

$$\int_{-\infty}^{\infty}\int_{0}^{\infty} R_0(\eta,\ \sigma,\ \psi)\ d\eta d\sigma = \int_{-\infty}^{\infty}\int_{0}^{\infty} R_0(x_1,\ s_1,\ n_2)\ dx_1 ds_1 \equiv \Psi(\psi).$$

If ψ is so chosen that $\Psi(\psi)$ is minimum for all ψ, the integral on the left-hand side is also minimised. Hence, there exists no $\psi(n_1,\ s_1)$ which has the risk $R_0(\eta,\ \sigma,\ \psi)$ as small as the one given by the rule given here for all η and σ, and smaller for some η and σ. Hence, the rule given here is admissible.

6. Concluding Remarks

From theoretical point of view, the decision rule discussed above raises two important issues. Firstly, the rule for rejecting or accepting the hypothesis is based on the positive or negative sign of the mean. This procedure seems to be sound on intuitive grounds, but one does not know how does it compare with the more powerful procedure based on 't' test. The optimum properties of sign rule have been considered by K. Matusita (1951). However, since the alternative procedure based on 't' criterion by no means yields to an easy solution, the procedure suggested here could serve as a guiding principle for the choice of the amount of experimentation. Secondly, while deriving the solution, the controversial idea of fiducial probability has been used. This, however, has been done under the circumstances where it is difficult or rather impossible to suggest any suitable *a-priori* joint distribution of mean and the variance of the underlying normal population. But whatever be the objection to the use of fiducial distribution, it has been shown that the procedure based on this distribution is an admissible one. Because of this property, the decision rule given here could be regarded as an optimum one.

From practical point of view, it may be hard to determine values of the constants like 'M' and 'k'. Again, since the process will be retained for a certain period of time, it is difficult to say for what exact period the value of M should be determined. It may be safer to take a reasonably longer period so that even if there is small gain from the change over to new process, this becomes substantial when the extent of application is fairly large and is repeated over a course of time.

Whatever the case may be, the importance of determining the optimum amount of experimentation could hardly be minimized especially in situations involving exhorbitant cost of experimentation. And the discussion given above serves as a guiding principle.

The method of determining the amount of experimentation for the second stage has been explained in the examples below:

Example.—Under simple fertilizer trials in cultivators' fields on paddy crop during the year 1962–63, the following responses were obtained to 20 lb./acre doze of nitrogen in each of the initial five experiments conducted at one of the centres in Lucknow District of U.P. The total area under paddy in the centre is reported to be 20 thousand acres. It has been estimated that the cost of conducting single experiment comes to Rs. 60. This includes the cost of fertilizer, salary of the field staff and cost of statistical analysis. The additional cost of fertilizer alone comes to Rs. 18 per acre; however, if the cost of applying the fertilizer is also included, the additional cost of fertilizer application comes to Rs. 24 per acre. Hence, unless the additional yield from the fertilizer application exceeds Rs. 24 per acre, it will not be worthwhile to recommend the fertilizer. The price of paddy in the district has been reported to be Rs. 12 per maund.

Experiment No.	Response in md./acre
1	1·52
2	1·82
3	1·88
4	1·25
5	3·15

Average response = 1·92 md./acre.

Mean square of the response = 0·5340 (md./acre)^2.

In this case, the values of different quantities occurring in x and λ are as follows:

$n_1 = 5$, $M = 20{,}000$ acres, $k =$ Rs. 60,

$\bar{x}_1$ = Price of paddy × average response − additional cost of fertilizer application.

= 12 × 1·92 − 24 in case, the cost of spraying the fertilizer in the field is taken into account.

= − 0·96 Rs./acre,

or

= 12 × 1·92 − 18 in case, the cost of fertilizer alone is taken into account.

= 5·04 Rs./acre.

and

$s_1 = \sqrt{0 \cdot 5340} \times$ price of paddy:

$= 0 \cdot 73 \times 12 = 8 \cdot 75$ Rs./acre.

Hence

$$X = \frac{n^{1/2}\bar{x}_1}{s_1}$$

$$= -0 \cdot 24 \quad \text{or} \quad +1 \cdot 28$$

and

$$\lambda = \frac{n^{-3/2} M s_1}{k}$$

$$= 262 \text{ Rs.}$$

The values of N for $n_1 = 5$ as seen from the table, are 5·6, when $X = -0 \cdot 24$ and 2·4, when $X = 1 \cdot 28$. For both these values of N, it can be seen that the inequality,

$$\frac{2N+1}{2N+2} > \frac{\theta T_{n_1-1}(-|\theta|)}{\sqrt{\frac{n_1-1}{n_1-3}}\, T_{n_1-3}\left(\sqrt{\frac{n_1-3}{n_1-1}}\,\theta\right)}$$

is satisfied. Hence the number of experiments in the second stage should be 12 or 28.

7. Acknowledgements

The author is indebted to Dr. A. R. Roy, Professor and Head of Department of Statistics, Lucknow University, Lucknow, for his guidance and constant encouragement given in the investigation underlying this paper. It is also a pleasure to thank Dr. B. V. Sukhatme,

Professor of Statistics, Institute of Agricultural Statistics, New Delhi, for giving many valuable suggestions during the writing of this paper.

8. REFERENCES

1. Finney, D. J. .. "The statistician and the planning of experiments. *J.R.S.S.*, 1956, Series A, Part I.

2. Yates, F. .. "Principles governing the amount of experimentation in developmental work, " *Nature*, 1952, **171**, 391.

3. Grundy, P. M., Healy, N. J. R. and Rees, D. D. "Decision between the two alternative—How many experiments," *Biometric*, 1954, **10**, 317.

4. ———— .. "Economic choice of the amount of experiments," *J.R.S.S.*, 1956, **18**, Series B, No. 1.

5. Fisher, R. A. .. "Fiducial argument in statistical inference," *Contribution to Mathematical Statistics*, Paper No. 25, p. 395.

6. Matusita, K. .. "On testing statistical hypothesis," *Annals of the Institute of Statistical Mathematics* (*A.I.S.M.*), Japan, 1952, **6** (2).

7. Manwani, A. H. .. "*On Economics of Experimentation*," Unpublished thesis submitted as a partial fulfilment of the Research Course in Agricultural and Animal Husbandry Statistics at the Institute of Agricultural Research Statistics (I.C.A.R.), New Delhi, 1959.

8. Pearson, E. S. and Hartley, H. O. *Biometrica Tables for Statisticians*, **1**, Table 9, p. 132.

APPENDIX

As shown in (4.5), the amount of additional experimentation $n_2 = Nn_1$ will be determined by the larger root 'N' of the equation,

$$N^{1/2}(N+1)^{3/2} = \frac{\lambda}{2}\sqrt{\frac{n_1-1}{n_1-3}}\, T'_{n_1-3}\left(\sqrt{\frac{n_1-3}{n_1-1}}\,\theta\right)$$

where,

$$\lambda = \frac{M}{k}\, n_1^{-3/2}\, s_1 \quad \text{and} \quad \theta = \sqrt{\frac{N+1}{N}}\, X.$$

$$X = \frac{n^{1/2}\, x_1}{s_1}.$$

TABLE I

Values of additional number of experiments relative to $n_1 = 4$

X \ λ	100	250	500	750	1000	1250	1500	2000	2500	3000
0·2	4·2	7·2	10·5	12·6	14·7	16·8	18·8	21	23	25
0·4	4·0	6·8	9·6	12·0	14·4	16·0	17·5	20	22	24
0·6	3·6	6·0	8·7	11·2	13·2	14·8	16·2	19	21	23
0·8	3·0	5·2	8·0	10·0	11·8	13·2	14·6	17·4	19·0	20·5
1·0	2·6	4·6	7·1	8·9	10·7	12·0	13·4	15·6	17·4	18·8
1·2	2·2	4·2	6·2	7·8	9·4	10·8	12·0	14·1	15·6	17·0
1·4	1·8	3·8	5·6	7·1	8·4	9·6	10·6	12·6	14·1	15·6
1·6	1·5	3·4	5·2	6·4	7·6	8·6	9·6	11·4	13·0	14·3
1·8	1·2	3·0	4·4	5·6	6·8	7·8	8·6	10·2	11·6	12·9
2·0	0·5	2·6	3·9	5·1	6·1	7·1	7·8	9·2	10·5	11·7
2·2	0·5	2·2	3·6	4·4	5·7	6·4	7·0	8·4	9·6	10·8
2·4	0·3	2·0	3·4	4·4	5·3	6·0	6·6	7·8	8·4	9·9
2·6	0	1·7	3·0	3·9	4·8	5·6	6·0	7·2	8·2	9·2
2·8	0	1·5	2·8	3·6	4·3	5·0	5·7	6·6	7·6	8·4
3·0	0	1·4	2·4	3·4	3·9	4·5	5·2	6·1	7·0	7·9

TABLE II

Values of additional number of experiments relative to $n_1 = 5$

X \ λ	100	250	500	750	1000	1250	1500	2000	2500	3000
·20	3·2	6·0	8·4	10·8	12·6	14·4	16·2	18·6	21·0	23·5
·40	2·8	5·2	8·0	10·0	11·8	13·6	15·2	18·0	20·0	22 0
·60	2·4	4·6	7·2	9·2	10·4	12·0	13·8	16·2	18·0	19·5
·80	1·9	3·9	6·2	8·0	9·2	10·4	11·6	13·8	15·6	17·8
1·0	1·3	3·1	5·3	6·8	8·0	9·0	10·0	11·6	13·4	15·0
1·2	0·7	2·6	4·3	5·6	6·6	7·6	8·4	10·0	11·2	12·2
1·4	0	1·9	3·4	4·8	5·6	6·4	7·2	8·0	9·6	10·7
1·6	0	1·4	2·8	3·9	4·8	5·6	6·0	7·2	8·2	9·2
1·8	0	1·0	2·2	3·2	3·9	4·6	5·2	6·0	7·0	8·0
2·0	0	0	1·8	2·7	3·3	3·9	4·5	5·3	6·1	6·9
2·2	0	0	1·3	2·0	2·7	3·2	3·7	4·6	5·4	5·9
2·4	0	0	0·9	1·7	2·2	2·7	3·1	3·8	4·6	5·2
2·6	0	0	0	1·2	1·9	2·3	2·7	3·3	3·9	4·2
2·8	0	0	0	0·8	1·4	1·8	2·2	2·8	3·3	3·8
3·0	0	0	0	0	1·1	1·5	1·8	2·4	2·9	3·3

$T_{n-1}(x)$ being the ordinate of 't' distribution at x. The value of N given from the equation does not depend upon the sign of X value. The graphic solution of the above equation can be obtained by drawing nomograms for different values of n_1. The nomograms have been drawn for $n_1 = 4$, 5 and 6. Each nomogram consists of two straight line scales each corresponding to 'X' and λ and one curved scale corresponding to N. For obtaining the solution, a straight line is to be drawn joining the given values of X and λ on the respective scales. The point where this line cuts the scale for N reads the values of optimum N in addition to the inequality

TABLE III

Values of additional number of experiments relative to $n_1 = 6$

X \ λ	100	250	500	750	1000	1250	1500	2000	2500	3000
·20	3·4	6·2	9·8	11·8	13·2	15·0	16·2	18·4	20·5	22·0
·40	3·0	6·0	8·8	10·8	12·0	13·6	15·0	18·0	19·5	21·0
·60	2·0	4·4	6·4	8·4	10·0	11·2	12·0	14·6	16·2	17·4
·80	1·2	3·4	5·2	6·8	8·0	9·0	10·0	11·6	13·2	15·0
1·0	0	2·4	4·0	5·2	6·4	7·6	8·4	9·6	11·0	12·0
1·2	0	1·5	3·0	4·0	4·8	5·6	6·4	7·8	8·7	9·6
1·4	0	0	2·0	3·0	3 8	4·4	4·8	6·0	6·9	7·9
1·6	0	0	1·0	1·9	2·6	3·2	3·7	4·4	5·2	6·0
1·8	0	0	0	1·1	1·8	2·3	2·8	3·4	4·0	4·5
2·0	0	0	0	0	0	1·5	1·9	2·3	3·1	3·7
2·2	0	0	0	0	0	0	1·1	1·8	2·3	2·9
2·4	0	0	0	0	0	0	0	1·1	1·7	2·1
2·6	0	0	0	0	0	0	0	0	0	1·4
2·8	0	0	0	0	0	0	0	0	0	0
3·0	0	0	0	0	0	0	0	0	0	0

$$\frac{2N+1}{2N+2} > \frac{\theta\, T_{n_1-1}(-|\theta|)}{\sqrt{\frac{n_1-1}{n_1-3}}\, T_{n_1-3}\left(\sqrt{\frac{n_1-3}{n_1-1}}\,\theta\right)}$$

being satisfied. Otherwise the value of N will be taken to be zero.

With the help of these nomograms, tables giving the values of N, have been prepared for $X = 0{\cdot}20$ to $3{\cdot}0$ and $\lambda = 100$ to 3,000.