SEQUENTIAL EXPERIMENTATION*

R. A. FISHER

Department of Genetics
Cambridge

THE PRESENT USE of the term sequential is intended to be of a broader import than the formal use of the word as associated with the systematic procedure known as sequential analysis. The experimenter does not regard his material as wholly passive but instead looks to what may be learnt from it with a view to the improvement and extension of the enquiry. This willingness to learn from it how to proceed is the essential quality of sequential procedures. Wald introduced the sequential test, but the sequential idea is much older. For example, what is the *policy* of a research unit? It is that in time we may learn to do better and follow up our more promising results. The essence of sequential experimentation is a series of experiments each of which depends on what has gone before. For example, in a sample survey scheme, as explained by Yates, a pilot survey is intended to supply a basis for efficiently planning the subsequent stages of a survey. Again, successively smaller units may be chosen at various stages and the chance of what is chosen at a particular stage depends on what has gone before.

These ideas are illustrated by two concrete examples which naturally arise in genetic work. One of these is quite near to Wald's sequential testing. We are indebted to Wald for laying down a detailed procedure in advance specifying how we shall react in all cases, so giving to the sequences the same unity as a single act of sampling possesses, and enabling the ordinary notions of estimation, efficiency and sufficiency to be applied to sequential experimentation. The chief interest in the first problem will be in its comparison with and slight deviations from Wald's procedure.

Suppose that there is evidence of a lethal factor in heterozygotes linked with a visible factor such as black or brown color in the house mouse. Let us represent

B black

b brown

l lethal

and $+$ absence of lethal factor.

Then in the mating

$$B + /bl = B + /bl$$

the brown mice occur with a frequency much less than one quarter. There is another gene "misty", about ten units from the black/brown locus, and we may wish to ascertain the location of the lethal gene relative to brown and misty. Omitting the genetical details, a useful mating would be

$$BlM/b + m = blm/b + m,$$

which gives two brown mice to one black if linkage is close, and it will be necessary to test the $Bl/b +$ mice before use. A sequential test may be used to find such animals soon enough, and only males would be tested as females cannot be tested sufficiently in time. The other test mating is

$$B + /b + = bl/b +$$

which gives equal numbers of black and brown, and we wish to eliminate such males. This differs from the usual sequential test in that our object is the practical one of securing material of a particular kind, rather than the theoretical one of determining the proportion of each kind which is present. I will suppose that we have an ample supply of mice to be tested, so that we need have no hesitation about discarding any individual, and there is no interest in minimising the probability of rejecting a mouse of the desired kind. The only limitations will be on the number of cages which have been provided for the females.

If at some stage we have counted x brown and y black mice, the probability of reaching that stage depends on whether we have used a $2 : 1$ or $1 : 1$ mouse. The probability in the case of a $2 : 1$ mouse is

$$P_1 = C(\tfrac{2}{3})^x(\tfrac{1}{3})^y,$$

where C is the number of paths to reach that stage, and is like a binomial coefficient, but not equal to it for in the sequential test certain paths are blocked by elimination or acceptance.

The probability for a $1 : 1$ mouse is

$$P_2 = C(\tfrac{1}{2})^x(\tfrac{1}{2})^y.$$

Taking logarithms

$$L_1 = \log C + x \log(\tfrac{2}{3}) + y \log(\tfrac{1}{3})$$

$$L_2 = \log C + x \log(\tfrac{1}{2}) + y \log(\tfrac{1}{2})$$

$$L_1 - L_2 = x \log(\tfrac{4}{3}) - y \log(\tfrac{3}{2}).$$

Lines of equal likelihood ratio would be parallel lines if x and y were continuous variates. Actually we can take cells (x,y) as shown in the figure and the acceptance or rejection contours will be broken lines like those shown in the figure.

Suppose that the initial frequency of 1 : 1 and 2 : 1 mice was 9 : 1. Then the frequency at the point x,y will be $9e^{L_2}$: e^{L_1} and the ratio of good to bad will be

$$e^{L_1}/9e^{L_2} = \tfrac{1}{9}e^{L_1-L_2}.$$

When $L_2 > L_1$ the quality is worse than in the original material, and we can reject it right away. In the diagram, the solid line divides the positive likelihood region from the negative likelihood region.

Suppose we want the ratio of the desired to the undesired type to be 5 : 1, in the selected animals. Then

$$L_1 - L_2 = \log 45,$$

shown by the double line. Mice below this line corresponding to $L_1 - L_2 > \log 45$ would be accepted as being at least 45 times as likely to be desirable as the original ones. We could choose any other ratio, depending on our anxiety regarding the use of the wrong genotype.

If the young were examined one at a time, the number of paths to a particular cell in the diagram can be written down in a manner similar to that used by Fermat and Pascal to write the "arithmetical triangle". This would provide the value of C for any cell. We could find the frequencies with which the two types are distributed over the area, and they would be crossing over the lower boundary line at all points in the ratio of about 5 : 1. In addition, we could divide the intermediate area into zones with lines of equal likelihood. This provides the possibility of testing more intensively those more probably desirable and hence of hastening the finding of acceptable mice in a mass of unacceptable material.

I shall now consider a different type of sequential experimentation. Geneticists are often expected to advise on the improvement of dairy cattle. The naive geneticist may think that he can accomplish this merely by selecting high producing cows, though it is really much more difficult than that. Now it is possible to visualise a situation in which one big cow is equal in all respects to two little cows, including food consumption. Then it would be possible for our geneticist with 50 million dollars of public funds to expend over a 20 year period to develop a cow which produces twice as much milk as the animals from which he started. The geneticist would then say that his breeding program has been successful in doubling the yield of milk, and "if only farmers were

keeping as many cows as they used to, the milk supply would be doubled. Unfortunately, the farmers are now keeping only half as many cows." This is an extreme case with a fairly obvious moral.

A Sequential Selection Chart for Mice
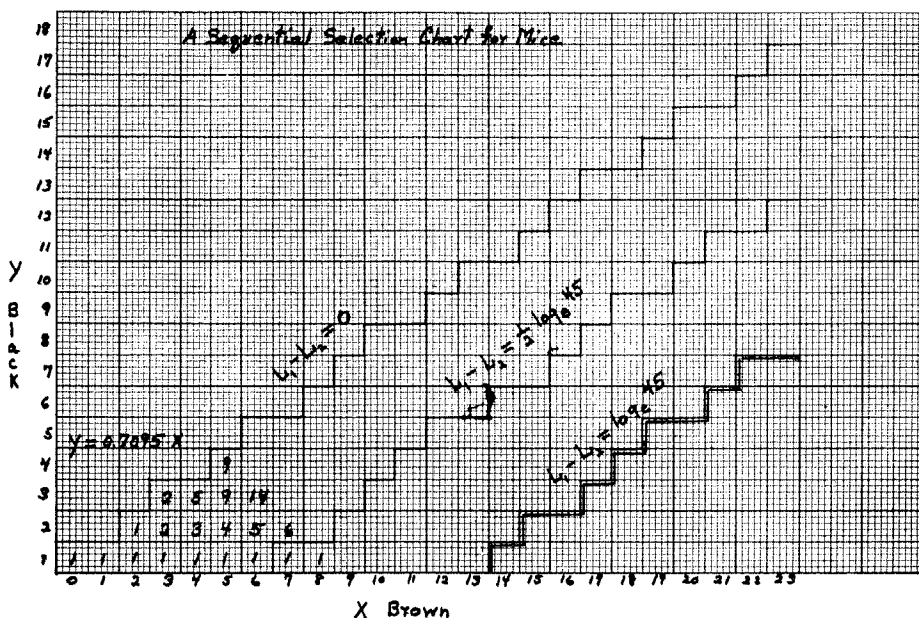
If we select a type only for milk yield and at random for other characters then we are likely to injure it in other respects. When you select for one character you affect everything else. You need to select on all valuable characters to avoid injuring the stock, for example, selecting for milk yield alone may decrease longevity.

What do we mean by measuring the milk yield of a cow? For a long time now milk yield records have been available. Such records may have been advantageous, but in many cases certainly they have been abused. There have even been cases where the milk produced has been fed back to the cows with a view to obtaining record yields.

Many physiologists believe that it is possible to make a clean separation of Science and Economics, in that they can describe the reactions of an animal in physical terms, and leave it to the farmer to judge what uses of it are profitable. I do not believe such a clean separation can be made. Where economics comes in is that for any machine there will be conditions of favorable working and unfavorable working which are within the choice of the operator of the machine. To choose among machines we must test each when operated in the most profitable way. In feeding cows there will be more or less profitable operations open to

the farmer and the scientific problem becomes definite when these alternatives are recognized. This becomes technically very difficult. How can we ensure that an individual cow is tested in conditions which are most profitable for that individual? How can we induce the cow to tell us how she ought to be fed to give us the maximum profit? This can be attempted in an experimental situation by using the accepted feeding standards to establish a basic rate of feeding for the individual animal taking account of the shape of the lactation curve, and the maintenance requirement of the animal. This provides a feeding curve to which may be added several parallel curves both above and below it, separated by intervals corresponding to differences in feed which are the economic equivalent of about one gallon of milk a week.

We can make the following index table for the individual cow, where a is the overhead cost and z is the cost per unit of milk.

Week	Food x (in cents)	Milk y	Cost per unit $(x + a)/y = z$
1	x_1	y_1	z_1
2	x_2	y_2	z_2
3	x_3	y_3	z_3
4	x_4	y_4	z_4
5	x_5	y_5	z_5
6	x_6	y_6	z_6
..	..	..	..

If the cost per unit in the second week is less than the average in the first and third weeks, then in the following week we should go back to the same feeding level as in the second week. Comparisons and decisions may be made by looking at successive sets of three records on z. If cost per unit for the third week is more than the average for the second and fourth weeks we again raise the feeding level. No comparison is then made at the end of the fifth week and the sixth week feeding level is taken to be that of the fourth. We may go wrong in a procedure of this kind, but it should be tried and scrutinized. There should be some such method of using animal reactions to guide experiments in milk yield.

I would call this a sequential test of milk yield. It has the advantage that we do get cost at the end and that we can then give preference to those animals which produce most cheaply. I suggest that the use of identical twins may be of valuable assistance in estimating the precision of the trial.