RANDOM SELECTION IN THE THERAPEUTIC TRIAL

BY

R. F. WRIGHTON

Department of Medical Statistics, University of Birmingham

I. INTRODUCTION

This paper completes the formal part of an inquiry into the credentials of the statistical method as related to the evaluation of therapeutic measures. I have summarized the general conclusions of this inquiry in an earlier communication (Wrighton, 1955). What follows is an attempt to elucidate concepts which, whilst of less immediate relevance than those I have dealt with earlier (Wrighton, 1953), are, in the final analysis, of equally vital importance.

In employing concepts not peculiar to his own discipline, the biologist can often draw on a vocabulary refined and tested in the laboratory so as to be for all practical purposes quite unequivocal. This is regrettable not so in the case of statistics. The relevant aspect of physical reality is seemingly no-one’s professional responsibility, unless it be the mathematician’s. But mathematicians have never approached unanimity on the subject and, while only too ready to suggest technical procedures owing their origins to unstated and dubious assumptions, have in the main seemed unprepared to admit that fundamental inquiry is necessary or profitable.

The laboratory situation we need to examine in this context is properly remote from the practical problems which are our ultimate concern. We have an urn containing unknown numbers of balls of various specifications.* We may then ask: first, what precise meaning can we have in requiring that a number of balls should be drawn from the urn at random? Secondly, assuming the first question answered, what kind of information about the contents of the urn can we legitimately derive from a sample thus drawn? The second is the question I have already attempted to answer. Semantic confusion is there associated with words such as significance, confidence, and likelihood, and the notion of inverse probability. The first is the question I am concerned to answer here. As I have suggested above, it is of more than technical interest. Indeed it must, strictly speaking, be answered before the other can be approached with complete propriety. The semantic issues here relate to the notions of chance event, random procedure, and direct probability.

The method of random selection in common use requires the numbering of the balls in the urn, and uses published lists of random digits to effect the withdrawal. There are two types of objections we can make to such a procedure. We can object that the use of an established table of random numbers is illegitimate, however the table has been compiled, or we can object specifically to the mode of structure of the table, viewed in itself as a dynamic randomizing device. In considering such objections, I must slightly anticipate discussion in the body of the paper and use loosely terms whose meanings I intend later to make more precise.

With any randomizing procedure we anticipate that successive results will be discordant, yielding, on appropriate tabulation, a chaotic array in which no pattern can be detected. The compiler of a table of random digits can be judged after the event only by his success in producing such an appearance. To do so, however, he may or may not invoke what we can refer to as a disorderly procedure. In the fact of the existence of this choice lies an important objection to the method of random digits. Suppose, for example, we have equal numbers of black and white counters filling the squares of a chess-board in a systematic arrangement. We may use the disorderly procedure of shuffling in order to redistribute the counters amongst the squares; and we can suppose, if our shuffling is vigorous, that almost certainly, no new pattern will appear. We may, however, proceed quite differently in order to obtain the same type of...
result, for we may, though with great labour, systematically enumerate the possible arrangements of counters on the board and select from amongst these one which passes all of a number of pre-assigned and systematically compiled tests for freedom from pattern.* We are led, then, to distinguish sharply between the concept of a static chaos and the concept of a chaotic procedure. The latter may be sufficient, but it is not necessary for the production of the former. With our present preoccupation we are concerned only with chaotic procedures. If we use a static chaos we cannot hope to ensure that the appearance of a particular sample is in the strictest sense unanticipated.

A further aspect of the failure of the static random digit table to fulfill our need may be briefly mentioned. Suppose, on the one hand, the compiler of such a table uses a disorderly procedure to effect his end, and produces, against long odds, a table with markedly systematic features. If he does not wish to pass this misfortune on to the user, he will either suppress the table or give warning of the regularities present.† In either case he admits that the calculus of probability is not strictly applicable when the table is used. If, on the other hand, no undesirable regularities appear, the compiler is no better off; if subsequent calculations are to be strictly valid, he must bear in mind the whole range of possible outcomes of the random procedure and must examine his conscience to discover whether any result conceivable could have caused his behaviour to vary. We must conclude, therefore, that the use of random number tables is quite inconsistent with the end we have in view.

We consider now objections to currently suggested methods of compiling random number tables, viewed as randomizing devices in their own right. Some methods we must regard as a fortiori invalid on grounds already discussed. These include any method which uses printed or otherwise pre-determined data, such as end digits in mathematical tables and census reports. There remain ad hoc and infinitely protractable procedures. One such method takes as random numbers the middle digits of arbitrarily chosen large numbers multiplied together on a modern digital computer. This method, and others like it, we must reject out of hand, since our ideas can be little clarified in replacing the word random by arbitrarily chosen. Another method utilizes so-called stochastic phenomena met with in physics. This too we must reject, since in principle at least we must attend to the possibility of fresh experimental inquiry vitiating any assumption we make. There remain methods (such as that of Kendall and Babington Smith) which make direct use of apparatus similar to that used in familiar games of chance. We shall see in what follows that it is possible to formulate a satisfactory approach on this basis; but to do so we shall have to make quite clear that a game of chance is and, furthermore, make a vital emendation of the procedure which takes over unqualified the results which flow directly from a chance apparatus used therein. In doing this we shall in effect be providing new supports to the foundations of a part of probability theory which is at least of great historical importance.

Since we are not concerned here in developing a theory descriptive in aim, we shall conveniently adopt a genetic approach, finding accordingly that in laying the foundations for a formal development we can at the outset provide ourselves with the means of recognizing those situations in the real world to which the system must be regarded as applicable. In doing so we shall avoid the paradoxes encountered by those who seek, as it were after the event, to make the domain of application of probability theory commensurate with a wide and vaguely defined domain of indeterminate behaviour. It follows that, while we shall make claim to associate an unequivocal meaning with the term probability in the limited context of situations in which stochastic events are purposively evolved, we shall not pretend that this directly clarifies the sense in which the word is or should be used, for example, in physics, in genetics, in general statistics, or in logic.

II. THE CONCEPT OF FORTUTTY

The following is the problem which we seek to resolve: given N distinguishable objects, to derive an unambiguous procedure by which we can select n objects from amongst them, so that, immediately before the act of selection, every such combination of objects may be said to possess the same opportunity of being withdrawn. We refer to this as the lottery problem.

We can effect a casual resolution by matching the objects concerned with numbered tickets, shuffling these together more or less thoroughly and calling upon a disinterested person to withdraw from
RANDOM SELECTION IN THE THERAPEUTIC TRIAL

amongst them a batch of the requisite size. In the simplest situation, where \( N=2 \) and \( n=1 \), we may adopt a more clearly defined procedure. We may toss a coin and, following a pre-assigned convention, withdraw the object labelled 1 or the object labelled 2 according as the coin lands with heads or tails uppermost. Here we may reasonably allow a potential beneficiary to operate the mechanism employed. Since the chance events concerned derive from a procedure for the following of which clear-cut instructions, seemingly admitting of no wilful misinterpretation, can be given, we approach a situation in which impartiality is strictly ensured and in which the notion of chance acquires objectivity. It is characteristic of such procedures that it is possible to verify their efficacy quantitatively. The validity of the coin-tossing procedure in particular is attested to by our experience that, with ordinary precautions and an approximately symmetrical coin, no operator whatever his efforts, can succeed in deflecting the proportion of heads observed in a large number of trials from the neighbourhood of \( \frac{1}{2} \). With this sort of situation in mind we can proceed to develop a general approach to the problem we have under consideration.

In discussing the generalized gaming device, we must distinguish sharply between the physical apparatus employed—for example a coin, the medium in which it is tossed, and the surface upon which it lands—and the instructions, in practice often ill-defined, which dictate the actions of an operator as he manipulates the apparatus. More precisely, in such situations, we have to bear in mind three entities—the apparatus employed, the operator using the apparatus, and the instructions issued to the operator. We shall require first to examine the limitations imposed by unavoidable imperfections in the apparatus employed, and secondly to establish the minimum precise formulation necessary to make the operator's instructions adequately explicit. We must in particular ensure that instructions for the evolution of stochastic sequences shall contain no word—for example shuffle, haphazard, erratic, non-purposive—whose inclusion would imply that we allow the responsibility for the impartiality hoped for in the operator's behaviour to depend in some way on his own interpretation of the purpose behind the procedure. To succeed in this will be to find a definition of the term probability which gives the word an important part of its connotation in the classical mathematical theory, but which anticipates charges of circularity and equivocation.

Let us ask at this point what it is that gaming devices of seemingly proven adequacy have in common. It will be seen that:

\( (a) \) A routine action is prescribed which may have one of a specified number of distinguishable outcomes, each associated with one facet of a symmetrical apparatus.

\( (b) \) The instructions given to the operator of the device are such that, whoever he is, it lies quite beyond his powers of sensory discrimination to predetermine any particular one of these outcomes.

We examine these properties in a chance mechanism which is conveniently simple.

Consider a teetotum formed from a disk transfixed at its centre by a style and divided into \( r \) equal sections, so that, when the operator spins it, any one of \( r \) outcomes may result according to what point on the circumference of the disk comes to rest in contact with the surface on which it is spun. We may plot the results of successive spins upon a straight line which is marked off at intervals of \( 2\pi \) units from an origin representing the fixed point from which the top is set in motion. We allow the operator to be any human being capable of carrying out our instructions. We instruct him to twist the top so that it spins completely round at least \( n \) times. Let us consider the simplest case, \( r=2 \), the disk being divided by a diameter into two sections, one coloured black, the other white. We may distinguish three cases:

\( (i) \) \( n \) small.—The operator may acquire such skill that he can predetermine whether the outcome is black or white. We may refer to sequences of labelled events derived in this way as determinate.

\( (ii) \) \( n \) intermediate.—The operator is still able to exercise his choice but not with invariable success; fluctuations in the labelled sequence of events occur and are attributable to intrinsic factors, such as fatigue, or to seemingly trivial external circumstances distracting the operator or subverting the train of events he has set in motion. Resulting sequences we may refer to as statistical or indeterminate. With this type of instruction we can devise a game of skill between operators of comparable aptitude.

\( (iii) \) \( n \) large.—The operator is unable to influence the outcome of a trial in any way. Resulting sequences we may refer to as fortuitous. With this type of instruction we can devise a true game of chance.

Following this distinction, we may regard a line upon which the outcomes of individual trials are
mapped, as divided into three domains: of determinacy, indeterminacy, and fortuity. Vagueness in the definition of the boundaries of these domains need not trouble us. In serious application we shall wish to deal exclusively with sequences of events we can regard beyond all question as either determinate or fortuitous. Statistical series, as defined above, will not submit themselves to formal manipulation.

The principal assumption we shall need to make is that domains of fortuity exist, that is to say, that fortuitous behaviour exists, as distinct from what we have called indeterminate behaviour. Clearly a direct method of demonstrating this is inconceivable. Any evidence we may adduce in its favour must be incomplete insofar as it will relate to protracted sequences of trials. A statistical theory is not invalidated if occasional events within a sequence supposedly obeying probabilistic laws are, according to a particular criterion, determinate; it may merely, and perhaps to an insignificantly small extent, be inadequately descriptive. But the lottery problem can admit of no resolution at all if anomalously determinate events are allowed to intrude into the procedure of selection used; for here a repetition of the procedure, within which an occasional anomaly of this type might be seen as unimportant, need not occur. We wish to be able to say, even if in the course of a lifetime we undertake one sampling experiment and one only, or organize one lottery and one only, that in doing so we can act rationally.

We must re-emphasize that the assumption we make presupposes a three-fold system of apparatus, instructions, and operator, and that in consequence fortuity in the sense and circumstances in which we shall use the word will be regarded as relative to the sensory equipment of the operator. We shall regard, for example, the outcome of the tossing of a coin as fortuitous in an unqualified sense, only insofar as we believe that a task (the inducing of the result heads, say) is implicitly set which lies quite beyond the ability of all potential operators to fulfill. If an extraterrestrial visitor were to show himself able to spin a coin several feet in the air so that it invariably landed with head uppermost we should reasonably admire his ability. We should not be led to deprecate such laws of chance as we can now formulate.

By the above analysis we see that a variety of simple actions can provide the means for evolving stochastic sequences. We require only that the task of inducing a determinate outcome can be made exacting to any desired extent. We are released from the necessity for relying on traditional, and possibly inconvenient, devices. Further and more important, with an understanding of the rationale of any device chosen, we shall be able to see how its necessary inadequacies are remediable.

III. Probability Rules assuming Perfect Apparatus and Absolute Fortuity

A chance apparatus may prove deficient in two respects. In the first place, and trivially, physical difficulties may intervene before the boundary of the domain of fortuity is reached. For example, we can conceive of a coin too heavy to be used for our present purposes. In the second place we must recognize that there are no means by which the apparatus used can be made physically perfect. All dice are slightly asymmetrical, and all tops a little unbalanced. In discussing the fundamental rules of probability, however, we shall find it convenient to assume provisionally that the apparatus used is ideal.

With ideal apparatus we can in all cases, as in the example of the top, map the outcomes of individual trials on a line which is divided at equal intervals. When outcomes with a particular specification are distinguished a repeated pattern is imposed on this line. Thus Fig. 1 will depict a section of the domain of fortuity corresponding to the chance procedure in which the apparatus used is a symmetrical top divided into six equal sections, one of which is painted black, the outcome being scored $\phi$ if the black sector comes to rest in contact with the surface on which the top is spun, the angle through which the top passes before it comes to rest being $\phi$.

In this example, the mapping is linear. Extending the notions discussed, we can conceive of situations in which the mapping is on a plane, and in which a repeated pattern is imposed upon a meshwork of equal rectangular cells. In such situations we issue a two-fold instruction to the operator, who will have, in effect, to compound two stochastic actions. For example (Fig. 2) we may instruct the operator to spin a needle of length $2a$ across an area laid out in parallel strips each $l$ units wide, imparting to it a minimum angular and a minimum linear momentum; or alternatively—and the effect will be the same—we can require the operator first to project the needle directly across the surface and then to spin it, pivoted at the point at which its centre has earlier come to rest. The cells of the meshwork on which outcomes are to be plotted will have sides of lengths $l$ and $\pi$, and the repetitive pattern will be of the form exhibited in the diagram, $\theta$ being the angle through which the needle is spun and $x$ the distance through which the centre of the needle travels before it comes to rest. This affords the celebrated problem of Buffon's needle.
The meshwork of the mapping of any two-dimensional fortuitous behaviour must comprise identical rectangular cells. This follows because the primitive notion of fortuity we have prescribed requires that the picture in the domain of fortuity should be unaltered when the skeleton and pattern is shifted a whole cell-length along either axis. Accordingly we define the probability of a fortuitous event unambiguously, if we say (using the two-dimensional case in illustration) that it is the ratio of the area representing that event in a single cell of a rectangular mapping to the total area of the cell. Thus in Fig. 1, where the mapping is linear,

$$\text{Prob(event)} = \frac{1}{6};$$

whereas in Fig. 2,

$$\text{Prob(event)} = \frac{2}{7\ell} a \int \sin \theta \, d\theta = \frac{4a}{7\ell}.$$  

This definition, in conjunction with the primitive axiom we have adopted, at once establishes the probability of an event which is certain as unity. It also yields directly the addition rule for compound events, the rule establishing probabilities in conditional sequences, and the product rule for successive trials. For example, the probability that a head results from adequate tossing of an ideal coin is $\frac{1}{2}$. In order to see that the probability of two heads occurring in successive tosses is $\frac{1}{4}$ we need to note that, since the individual outcome is fortuitous, we may alter the labelling of the coin between tosses without altering any essential feature of the procedure. Accordingly, the outcomes HT, HH, TH, and TT are of precisely equal standing, appear as equal squares in the mapping of the outcome of the double trial, and are each, therefore, to be taken to occur with probability $\frac{1}{4}$.

It must be emphasized that, in taking up this position, we advance beyond the classical notion of geometrical probability, completely avoiding its ambiguities in making our definition ultimately depend on a precise specification of the instructions issued to the operator. Equally, we avoid the obverse difficulty which results from the adoption of a frequency definition.* Furthermore, other

---

* See in particular the discussion of the needle problem by von Mises (von Mises, 1931).
IV. VERIFIABLE CONSEQUENCES—BERNOULLI’S THEOREM

Let us suppose that an experiment is repeated in a laboratory and that circumstances are so controlled that a particular result can be produced with unfailing success. Let us suppose further that, unknown to the experimenter, an assistant adopts the following procedure. At each repetition of the experiment he tosses a coin (which we may suppose to be ideal) ten times; if ten heads result he alters the conduct of the experiment so as to produce a contradictory result. If we can reasonably suppose that the possibility of deception thus introduced will have little effect on any honest appraisal of the results of the experiments—and this would surely be difficult to deny—we can reasonably accept, as effectively determinate, results in a probabilistic experiment which are predicted as occurring with probability very near to unity. One proviso only is needed; we are obliged faithfully to record the results of any empirical investigation of a chance mechanism, with no selection subsequent to the actual performance.

With this understanding we can derive verifiable and non-trivial consequences of the fortuity axiom. Consider the fortuitous event $E$ with probability $p$ produced by means of ideal apparatus. By virtue of the addition and multiplication rules, it follows from a known proposition concerning the terms of binomial series that, if $n$ trials are performed, at each of which the event $E$ may occur, then given $\delta$ and $\varepsilon$, it is possible to find a number $N_0$ such that for all $n > N_0$,

$$\text{Prob} \left\{ \left| \frac{m}{n} - p \right| > \delta \right\} < \varepsilon,$$

where $m$ is the number of occurrences of $E$. This is commonly known as Bernoulli’s theorem. It is only at this point that we need to invoke it. If we can agree on an acceptably small value for $\varepsilon$ we can then say that, in a large number of trials, using ideal apparatus, the fortuitous event $E$, whose probability is $p$, will occur with a proportionate frequency approximately equal to $p$, and in a particular inquiry we can define numerically what we mean by the term approximately and what we mean by the term large.

V. NON-IDEAL SYSTEMS

We proceed to examine the consequences of using imperfect apparatus for the elaboration of random sequences. If a biased die is repeatedly thrown, we shall in general still be able to assert that sequences evolved are random, since it may still be quite beyond the operator’s powers to induce determinate outcomes. But we shall no longer be able to associate the probability $\frac{1}{6}$ with the occurrence of any particular outcome. We must find out just what type of statement can be made in a situation such as this.

The simplest conceivable deviation from the ideal result occurs in situations where all predicted probabilities are deflected by fixed amounts from their true values. Thus the probability of throwing heads with a biased coin might be supposed to become $\frac{1}{2} + \varepsilon$ instead of $\frac{1}{2}$. Have we, however, any right to make the assumption that a physical bias is in this way unambiguously representable by a numerical probability value? Clearly not, for to do so is merely to replace one ideal probabilistic system by another. In any particular situation, such as we envisage here, the reasons for rejecting the notion are especially cogent. The departures from theoretical perfection we are considering will be very small, since we shall be supposing that we have made every possible attempt to ensure the symmetry of the apparatus used. Such departures, which as we shall show later may well be important, we must assume will remain concealed after any attempt to verify their existence in trials of realizable length. In these circumstances we shall have no right to postulate anything less than the most general specification of departure from the ideal set-up.

This conclusion is suggested by a consideration of the behaviour of stochastic sequences the apparatus for the manufacture of which is knowingly biased. Consider, for example, the case of coin-tossing. The coin will usually come to rest within a second or two of landing on a level surface. But occasionally, especially if the axis round which the coin is made to
spin whilst in the air is nearly vertical, the coin will spin topwise for a while after touching the surface. Normally this will not matter, but if we bias the coin (a penny) by soldering to the tail-side a sixpence, then, although the frequency of tails will be little affected in ordinary throws, in the sub-sequence of throws in which surface-spinning occurs the outcome will almost always be tails. If we impart a bias by attaching a disk rather lighter than a sixpence, we might suppose the situation to be represented roughly as follows:

In normal throws, \( \text{Prob}(\text{tails}) = \frac{1}{2} \);

In throws resulting in surface - spinning, \( \text{Prob}(\text{tails}) = \frac{1}{2} + \varepsilon \).

Now the operator will alter his method of throwing from time to time and eventually he may learn how to ensure that a throw with associated probability \( \frac{1}{2} + \varepsilon \) results. Unless we alter our instructions to the operator, the most we can say of this system is that the probability of the result labelled tails alternates at perhaps unknown intervals between \( \frac{1}{2} \) and \( \frac{1}{2} + \varepsilon \). Within these limits the operator may be able consciously to impose on the system any binary pattern of underlying probabilities he wishes. Even if he does not possess the clue which would tell him how to vary the probability at will, we must still allow that such patterns may result. We have here an indication of what sort of departure from the ideal we should properly seek to account for.

The most general behaviour of this type we can envisage will occur when the probabilities concerned vary unpredictably from trial to trial but remain always within certain specified limits. Let us suppose that in the system with which we are concerned each of \( r \) outcomes \( A_1, A_2, \ldots, A_r \), should ideally result with probability \( \frac{1}{r} \). Then, at the \( j^{th} \) trial of a sequence, we shall assume that:

\[
\text{Prob}(A_i) = \frac{1}{r} + \delta_i, \quad \text{where} \quad \sum_i \delta_i = 0
\]

and \( |\delta_i| < \delta \), for all \( i \) and \( j \).

We shall refer to \( \delta \) as the maximum proportional bias of the system.

In application to the lottery problem, we must be able to assign an unequivocal value to the maximum proportional bias of the system which is specified by the physical apparatus, instructions, and operator employed. Such an assessment must be based on an examination of observed frequencies. However, this in no way contradicts our initial requirement that the primary grounds for supposing a sequence to be stochastic must be \( a \ priori \). If at each assay the operator could perform an infinite number of trials, we should be able to assert that in no case could he succeed in educing limiting ratios greater than \( \frac{1}{r} + \delta \) or less than \( \frac{1}{r} - \delta \).

With a finite but large number of trials, at each assay we should be able with almost complete assurance to say, by reason of Bernoulli’s theorem, that he will not succeed in educing limiting ratios outside a range

\[
\left[ \frac{1}{r} + \delta + \Delta, \frac{1}{r} - \delta - \Delta \right].
\]

It follows that, if we can feel confident that the operator cannot educ in a fixed number of trials a proportionate frequency outside the range \( \frac{1}{r} \pm \Delta \), we are justified in accepting the maximum proportional bias of the system as less than \( r \Delta \).

In the case of coin-tossing sufficient is known to suggest that we are dealing with a truly fortuitous procedure. Assuming it implicitly specified that a minimum initial angular momentum should be imparted to the coin, the \( a \ priori \) requirements for the production of stochastic events are satisfied, and on the basis of experience we can reasonably say that in batches of 2,000 tosses no operator will succeed in obtaining a proportion of heads outside the range \( \frac{1}{2} \pm \frac{1}{10} \).

If we can accept this, we have an adequate specification of a simple stochastic system whose maximum proportional bias is \( \frac{1}{5} \).

We can thus resolve the simplest form of the lottery problem confident that the error we commit is less than 20 per cent. This does not immediately appear as a very striking achievement. But we shall proceed to show that, by using so crude and elementary a finding, we can resolve the lottery problem in its most general form.

VI. THE INVOLUTION* OF A STOCHASTIC SEQUENCE

In the ideal system, the operator cannot alter any long-run numerical property of a derivative sequence. But, if we admit the possibility of "bias", we allow him a limited opportunity for exercising a choice, or

* The simplest case of what we here call involution occurs when two coins are tossed, the labelling of the sides of one as heads or tails being switched if the other shows up, e.g. as heads. This procedure was suggested by Laplace. It has been pointed out to me that the procedure has been re-introduced recently, being referred to as filtering or filtration. It will be clear from what follows why the term involution is preferred. There exists no real analogy with harmonic analysis in the present context.
admit that he can in some way unthinkingly influence the course of events. It is natural to inquire whether we may use the element of fortuity necessarily present in the sequences evolved to vitiate such activity. In other words, we may anticipate that it is possible so to turn a sequence in upon itself as to rid it of irregularities imposed on it by the physical imperfections of the apparatus used.

Let us suppose, as above, that we have a stochastic system for which at every trial each of \( r \) events ideally occurs with probability \( \frac{1}{r} \), and in which the maximum proportional bias is \( \delta \). We must allow that, within the limits afforded by the specification of \( \delta \), the operator has the knowledge to impose any underlying stochastic pattern he wishes upon the sequence which results from a set of ordered trials. We may divide these patterns into two classes according to whether we do or do not impose one condition on the circumstances under which the trials are carried out.

(a) We may suppose the trials to be undertaken with no restrictions imposed beyond those implicit in the specification of the apparatus used and the primitive instructions given to the operator.

(b) We may require that the operator should be kept in ignorance at each stage of the outcome of earlier trials, results being noted down by an independent recorder.

This distinction is sharp. We can conveniently refer to the two types of sequence respectively as observed and blind. In the case of blind sequences, any pattern imposed by the operator will have been so imposed independently of any knowledge he possesses of the outcome of any single trial within the sequence under consideration. Using the word in this special sense, we can, therefore, refer to trials in a blind sequence as independent.

Consider now the case \( r = 2 \). Let us suppose the two possible outcomes of the individual trial to be labelled 0 and 1. We can consider a sequence of trials to be condensed into a sequence of pairs of events. Each pair will be labelled either 0,0; 0,1; 1,0; 1,1. Let us first consider the case where the initial sequence is blind. We cannot ensure that for a particular pair the operator will not behave so that

\[
\text{at the first trial, } \quad \text{Prob}(0) = \frac{1}{2} + \frac{\delta}{2},
\]

and at the second trial, \( \text{Prob}(0) = \left( \frac{1}{2} - \frac{\delta}{2} \right) \).

In this case, by the product rule,

\[
\begin{align*}
\text{Prob}(0,0) &= \frac{1}{4} - \frac{\delta^2}{4}; \\
\text{Prob}(0,1) &= \frac{1}{4} + \frac{\delta}{2} + \frac{\delta^2}{4}; \\
\text{Prob}(1,0) &= \frac{1}{4} - \frac{\delta}{2} + \frac{\delta^2}{4}; \\
\text{Prob}(1,1) &= \frac{1}{4} - \frac{\delta^2}{4}.
\end{align*}
\]

Equally we cannot ensure that he will not behave so that, at each individual trial,

\[
\text{Prob}(0) = \frac{1}{2} + \frac{\delta}{2},
\]

so that:

\[
\begin{align*}
\text{Prob}(0,0) &= \frac{1}{4} + \frac{\delta}{2} + \frac{\delta^2}{4}; \\
\text{Prob}(0,1) &= \frac{1}{4} - \frac{\delta^2}{4}; \\
\text{Prob}(1,0) &= \frac{1}{4} - \frac{\delta^2}{4}; \\
\text{Prob}(1,1) &= \frac{1}{4} - \frac{\delta}{2} + \frac{\delta^2}{4}.
\end{align*}
\]

These are the two extreme cases. We can say therefore that the maximum proportional bias relating to paired events is approximately doubled.

Consider now the expression

\[
\text{Prob}(0,0) + \text{Prob}(1,1)
\]

in the blind sequence. If \( \delta_1 \), and \( \delta_2 \) are the effective biases at the two loci in a representative pair \( \{ | \delta_1 |, | \delta_2 | < \delta \} \),

then Prob(0,0) + Prob(1,1)

\[
\begin{align*}
&= \left( \frac{1}{2} + \frac{\delta_1}{2} \right) \left( \frac{1}{2} + \frac{\delta_2}{2} \right) + \left( \frac{1}{2} - \frac{\delta_1}{2} \right) \left( \frac{1}{2} - \frac{\delta_2}{2} \right) \\
&= \frac{1}{2} + \frac{\delta_1 \delta_2}{2},
\end{align*}
\]

which is a maximum when \( \delta_1 = \delta_2 = \delta \), and a minimum when \( \delta_1 = -\delta_2 = \delta \).

Thus, \( \text{Prob}(0,0) + \text{Prob}(1,1) = \frac{1}{2} \pm \frac{\delta^2}{2} \).

That is to say, if we relabel like pairs 0 and unlike pairs 1, then, if \( \delta \) is the maximum proportional bias of the original sequence, we obtain a fresh stochastic sequence in which the ideal probabilities of the two possible events are \( \frac{1}{2} \), and the maximum proportional bias of which is reduced to \( \delta^2 \). We may repeat this procedure using groups of 3, 4, . . . elements in order to produce stochastic sequences.
RANDOM SELECTION IN THE THERAPEUTIC TRIAL

... with maximum proportional bias as small as we please. With the sequence divided into groups of \( n \), the maximum proportional bias of the condensed sequence becomes \( \delta^a \).

We are now in a position to see why it is essential to distinguish between blind and observed sequences. For, if we suppose that at the first position in a pair the operator employs a proportionate bias \( \delta \),

\[
\text{Prob}(0) = \frac{1}{2} + \frac{\delta}{2},
\]

and that at the second position he employs a proportionate bias of \( +\delta \) if the outcome of the preceding trial has been 0, and \( -\delta \) if the outcome of the preceding trial has been 1, then,

\[
\begin{align*}
\text{Prob}(0,0) &= \frac{1}{4} + \frac{\delta}{2} + \frac{\delta^2}{4}; \\
\text{Prob}(0,1) &= \frac{1}{4} - \frac{\delta^2}{4}; \\
\text{Prob}(1,0) &= \frac{1}{4} - \frac{\delta}{2} + \frac{\delta^2}{4}; \\
\text{Prob}(1,1) &= \frac{1}{4} - \frac{\delta^2}{4}.
\end{align*}
\]

Whence

\[
\text{Prob}(0,0) + \text{Prob}(1,1) = \frac{1}{2} + \frac{\delta}{2}.
\]

It follows that the involuted sequence will have the same maximum proportionate bias as the initial sequence.

Consider now the more general case. Let us suppose, using \( \pm \) as indicated above, that

\[
\text{Prob}(A_i) = \frac{1}{r} \pm \frac{\delta}{r};
\]

\[i = 1, 2, \ldots, r.\]

Let us again consider a pair of trials, supposing that the operator has to deal with maximum proportionate biases \((\varepsilon_1, \ldots, \varepsilon_r)\) at the first trial, and \((\eta_1, \ldots, \eta_r)\) at the second trial, where

\[
\sum \varepsilon_i = \sum \eta_i = 0; \quad |\varepsilon| < \delta, \quad |\eta| < \delta; \quad i = 1, \ldots, r.
\]

We then have:

\[
\text{Prob}(A_i A_j) = \left( \frac{1}{r} + \frac{\varepsilon_i}{r} \right) \left( \frac{1}{r} + \frac{\eta_j}{r} \right).
\]

Let us sum this over \( r \) values of the pair \((i, j)\) so that each value of \( i \) and each value of \( j \) is represented once and only once. The terms linear in \( \varepsilon \) and \( \eta \) will vanish so that:

\[
\sum_{r \text{ values}} \text{Prob}(A_i A_j) = \frac{1}{r} + \frac{1}{r^2} \sum_{r \text{ values}} \varepsilon_i \eta_j.
\]

Now the selection of groups of pairs in this way may be represented by the laying out of a Latin square with rows and columns labelled according to the numbering of the first and second members of the pair. We select from all possible pairs those corresponding to a particular labelling of cells within the square.

The maximum value of

\[
\sum_{i=1}^{n} x_{ij} y_{ij}
\]

subject to

\[
\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i = 0;
\]

\[
|x_i| \leq X, \quad |y_i| \leq Y
\]

is \( kXY \), where \( k = n \) if \( n \) is even, and \((n - 1)\) if \( n \) is odd.

It follows that, in the general symmetrical stochastic sequence with probability parameters

\[
\frac{1}{r} \pm \frac{\delta}{r},
\]

involvation by Latin square in pairs yields a stochastic sequence with probability parameters

\[
\frac{1}{r} \pm \frac{\delta}{r} \text{ or } \frac{1}{r} \pm \frac{r-1}{r^2} \delta,
\]

according as \( r \) is even or odd. Involuting in groups of \( s \) by means of an \( s \)-dimensional Latin structure, a condensed sequence results with a maximum proportionate bias of at most \( \delta^s \).

The legitimacy of the involvation procedure, insofar as it relates to the minute biases which are our primary concern, depends on the theoretical argument advanced above. But when the bias is great its effectiveness can easily be observed in the behaviour of observable frequencies.

In a sequence of 3,000 spins of a badly balanced top, with two segments, one labelled \( X \), the author observed 1,799 (59.9 per cent.) \( X \)s.

In 3,000 double spins with involvation, he observed 1,480 (49.3 per cent.) effective \( X \)s.

In 3,000 triple spins with involvation, he observed 1,496 (49.9 per cent.) effective \( X \)s.

In 3,000 quadruple spins with 4-fold involvation he observed 1,512 (50.4 per cent.) effective \( X \)s.

VII. RESOLUTION OF THE LOTTERY PROBLEM IN THE GENERAL CASE

Let us suppose that we have initially a stochastic sequence in which \( r \) events \((A_1, A_2, \ldots, A_r)\) ideally occur with equiprobability and in which the maximum proportionate bias is \( \delta \). We may conveniently
regard $r$ as even. In using such a sequence to select $n$ objects at random from amongst $N$ we need to note:

(a) That the reduced sequence, derived from the original sequence by deleting the occurrences of $(\delta - s)$ of the fundamental events, comprises a stochastic sequence of $s$ events, say $(A_1, \ldots, A_s)$ in which the maximum allowable proportionate bias is $\delta$.

(b) That the involuted sequence, obtained by in-voluting the fundamental sequence in groups of $k$, comprises another sequence of $r$ ideally equiprobable events with maximum allowable proportionate bias $(1 + \delta)^k - 1$.

(c) That the expanded sequence, obtained by labelling separately distinguishable groups of $k$ events, comprises a stochastic sequence with $r^k$ ideally equiprobable events and maximum proportionate bias $(1 + \delta)^k - 1$.

Any sequence required in sampling may be derived from the fundamental sequence by appropriate reduction, involution and expansion. Reduction will be most conveniently carried out at the last stage. It is necessary to inquire in what order involution and expansion should be carried out.

Where two-fold involution is followed by two-fold expansion, the initial sequence being divided into fours, the resulting sequence comprises equiprobable events with probabilities:

$$\frac{1}{r^2} \pm \frac{2\delta^2 + \delta^4}{r^2}.$$  

Where expansion is followed by involution, we have to consider in more detail the probabilities associated with four successive trials.

Suppose the effective biasses at these trials are $\varepsilon_{ik}$, where $i = 1, \ldots, r; \; k = 1, \ldots, 4$; and

$$\sum_i \varepsilon_{ik} = 0$$

for all values of $k$.

Then

$$\text{Prob}(A_i, A_m, A_u, A_v) = \left(\frac{1}{r} + \varepsilon_{i1}\right)\left(\frac{1}{r} + \varepsilon_{m2}\right)\left(\frac{1}{r} + \varepsilon_{u3}\right)\left(\frac{1}{r} + \varepsilon_{v4}\right),$$

and the probability of any individual event in the sequence obtained in following expansion by involution is this expression summed over $r^4$ terms in which each pair $(i, m)$ is represented once and once only, and each pair $(u, v)$ once and once only, that is to say over cells with a particular labelling in a Latin square containing $r^4$ cells. Now in any Latin square rows and columns may be interchanged so that one set of similarly labelled cells lies along the principal diagonal of the square. If we re-label the biases and call the event associated with the principal diagonal $E$, then

$$\text{Prob}(E) = \sum_{i,j} \text{Prob}(A_i, A_j, A_k, A_l)$$

$$= \frac{1}{r^2} + \frac{1}{r^2} \sum_{i,j} \varepsilon_{i1} \varepsilon_{j2} \varepsilon_{k3} \varepsilon_{l4} + \sum_{i,j} \varepsilon_{i1} \varepsilon_{j2} \varepsilon_{k3} \varepsilon_{l4},$$

on expanding and using the relations between the $\varepsilon_{ij}$.

Now, if we put $\varepsilon_{ij} = (-)^{i-j} \frac{\delta}{r}$, we get:

$$\text{Prob}(E) = \frac{1}{r^2} + \frac{2\delta^2}{r^2} + \frac{\delta^4}{r^2}.$$  

This indicates that we shall lose nothing vis à vis our present problem if we involute first and expand afterwards.

Consider now the problem of selecting one object at random from amongst $N$. Let us suppose the initial stochastic sequence we have to use is binary, with probability parameter $\frac{1}{2} \pm \frac{\delta}{2}$, and, for convenience, that the proportionate accuracy required in the final outcome is also $\delta$. If $k$ is the degree of expansion required and $r$ is the degree of involution, then $k$ is the smallest integer for which $2^k > N$, and $r$ is the smallest integer for which $(1 + \delta)^k - 1 < \delta$.

That is to say, $r$ is the smallest integer greater than

$$\frac{\log \left[ (\delta + 1)^{1/k} - 1 \right]}{\log \delta}.$$  

For example, for $N=1,000$: 

$$\delta = \frac{1}{10}, \; k = 10, \; r = 2,$$

and for $N=50,000,000$:

$$\delta = \frac{1}{10}, \; k = 26, \; r = 3.$$  

Finally, consider the general problem of selecting $n$ objects at random from amongst $N$, using the same initial sequence and requiring of the result the same degree of accuracy.

We wish to ensure that each of the $\binom{N}{n}$ possible choices is made with probability $\frac{1}{N} \left(1 \pm \delta\right)$.

With $N$ large, it would be impracticable to label each combination and employ the same method as for selecting a single object.

Accordingly we label the population $1, \ldots, N$ as before, and select by means of a sequence with probability parameter $\frac{1}{N} \pm \varepsilon$, successively reducing it as items are withdrawn (we are supposing all along
that sampling is without replacement. The ideal probability of getting a particular \( n \)-fold sample can be expressed in the form,
\[
\frac{n}{N} \left( \frac{n-1}{N-1} \right) \cdots \left( \frac{1}{N-n+1} \right),
\]
so as to correspond with the procedure employed.
Possibly associated with each term of this product is a term \((1 + \varepsilon)\) corresponding to the largest bias allowed. Accordingly, \( \varepsilon \) must be chosen so that
\[
(1 + \varepsilon)^n - 1 < \delta.
\]

For the present purpose we can take \( \varepsilon = \frac{\delta}{n} \).

It follows, \( k \) being defined as before, that, for the degree of involution required in the initial sequence, we have:
\[
r = \text{smallest integer for which } (1 + \delta)^k - 1 < \frac{\delta}{n}.
\]

For example, if we take \( \delta = \frac{1}{10} \),
\[
N = 1,000, \quad n = 50; \quad \text{then } k = 10, r = 4.
N = 1,000, \quad n = 500; \quad \text{then } k = 10, r = 5.
N = 50,000,000, \quad n = 1,000; \quad \text{then } k = 26, r = 6.
N = 50,000,000, \quad n = 25,000,000; \quad \text{then } k = 26, r = 10.
\]
The values for \( r \) we obtain are therefore sufficiently small for us to take it that for practical purposes the lottery problem is satisfactorily resolved.

VIII. General Considerations

In conclusion we may briefly review the approach we have taken. It is customary to classify theories of probability which have any pretense to direct application as deriving either from a subjective or from a frequency notion of probability. Both of these we have implicitly repudiated. In spite of this, however, and, although we do not seek in any way to release ourselves from the discipline of numerical appraisal which a frequency theory imposes, the starting point of our inquiry has essentially been in the subjectivist camp. In this context, however, subjective is perhaps an unfortunate term. What the non-frequency schools have in common is that they all relate probability in some way to ignorance; and since ignorance is a quality commonly unequally distributed amongst individuals, theories result which are subjective in application.

What we have sought to do is to discover an objective form of ignorance and to relate the concept of chance to it accordingly. Keynes (1921) remarks:

When we attribute a coincidence to objective chance, we mean not only that we do not actually know a law of connection, but, speaking roughly, that there is no

law of connection to be known. And when we say that the occurrence of one alternative rather than another is due to chance, we mean not only that we know no principle by which to choose between the alternatives, but also that no such principle is knowable.

The individual’s limited capacity for sensory discrimination merely affords us the only means possible of establishing a communicable meaning for the phrase “necessary ignorance” implied in the above quotation.

The position of frequency theories vis à vis our own is, however, of some interest. A frequency theory of probability (in other words a statistical theory) has the crippling disadvantage, as far as our present purpose is concerned, that it cannot provide any constructive clue to guide us in initially setting about producing a random sequence. The suggestion it will offer is that one should, by trial and error, seek to find such a sequence as it occurs in nature, testing sequences one by one for stability of proportions/frequencies in long series of trials. Such a procedure may well be regarded as sufficiently distasteful to warrant the rejection of any theory which proposes it. But suppose that we do provisionally accept it and that we succeed in obtaining a method of randomization which yields proportionate frequencies approximately in accord with theoretical requirements. We are then in the same position as we found ourselves at the end of Section V (above); namely, we can solve approximately the simplest form of the lottery problem, but no more. But since we shall not have admitted any a priori notions which would enable us to make the necessary distinction between observed and blind sequences we cannot, as we did in Section VI, proceed to invoke the procedure of involution to advance matters further. A sequence established as random in the methodology of naive empiricism, we must presume, be taken to have no qualities beyond those established by the direct scrutiny of routine enumeration. Moreover, we cannot obtain an alternative resolution of the difficulty by extrapolation. We might, working on the basis of empirical inquiry, find ourselves able to say that the “probability” of a coin landing heads is approximately \( p \). But, clearly, we can have no right to infer from this that the probability, for example, of a run of ten heads conditional on the occurrence of some run of ten identical throws is approximately \( p \) also. An empirical theory of this type can summarize, but it cannot, in any strict sense of the word, predict.

Finally, since in the context of the therapeutic trial the lottery problem is only of interest insofar as it must be resolved before we can discuss the inverse sampling problem, we shall mention briefly the
relation borne by our present findings to the second question propounded at the beginning of this communication. If we admit the necessary existence of an assessable experimental error in any sampling procedure, we must require of an inverse theory that it shall be able to take such error into account. In my suggested approach to inverse problems (Wrighton, 1953), we associate with the design of any statistical experiment an appropriately small probability value. In the simple situation of single-stage sampling, the maximum proportionate error of this probability value will be identical with the maximum proportionate bias of the sequence used for the selection of the sample after it has been adjusted by involution. It follows that, if the probability of drawing a particular sample is

\[
\frac{1}{N} (1 \pm \delta)
\]

and the theoretical uncertainty level is \( \varepsilon \), the “true” uncertainty level of the experiment is

\[
\varepsilon (1 \pm \delta).
\]

**SUMMARY**

(1) The object of this paper is to elucidate the notion of probability as associated with games of chance and with the procedure of random sampling as used in a therapeutic trial.

(2) We see the notion of probability as arising in situations where an operator performs a task for which his limited capacity for sensory discrimination precludes a determinate outcome.

(3) We note the significance of unavoidable imperfections in the physical apparatus used and put forward a method for rendering them ineffective.

(4) On the basis of the particular meaning we thus attach to the term *probability*, we advance a procedure for dividing \( N \) individuals at random into groups of \( n \) and \( N-n \) members.

(5) The *probability* of any particular allocation will be assignable as

\[
\frac{1}{N} (1 \pm \varepsilon),
\]

where \( \varepsilon \) is small and prescribed and represents the maximum absolute error we are prepared to allow to the procedure.

(6) The full details of the proposed procedure are given in the text. Here it suffices to emphasize in what way the approach is fundamentally different from current recipes for random choice.

(7) Any fixed table of random numbers is the record of a unique sample ostensibly specified by a randomizing procedure; it cannot therefore be the basis for a continuing randomizing process.

(8) Methods hitherto proposed to effect such an end, though immune to this objection, fail to guarantee the accuracy of the probability values they lead to, and in many cases do not admit of complete specification.

(9) By circumscribing the notion of chance employed in this context, the procedure here discussed claims to be free from the foregoing defects.

**REFERENCES**


