PART I

ON EXPECTATION AND UNCERTAINTY

I

PROBABILITY AND UNCERTAINTY*

In his novel *The Widows of the Magistrate*, Keith West tells how certain Chinese officials once plotted rebellion against their Emperor. The brief passage that I am going to reproduce describes the thoughts of a certain sentry, who had to decide whether to obey his immediate superior, the treacherous Captain of the Guard, or to stand alone against the rebels in loyal defence of the Emperor's representative, the Lady Hibiscus:

> In the room above, where the great drum stood, the sentry named Kwong Hui was testing the stacked bows of mulberry wood and setting the arrows in order.
>
> 'I am a man who seizes opportunity', he told the admiring women and the sleeping children.
>
> 'If I obey the Captain of the Guard, two things may happen. Either the rebellion succeeds, and I remain a soldier in the guard, or the rebellion fails, when I lose my head. Whereas if I obey the Lady Hibiscus, two things may happen. Either the rebellion succeeds, and I lose my head, or the rebellion fails, when I shall receive rewards quite beyond my imagination to conceive. Now of these four possibilities, the last only attracts me. So I shall strive to hold this tower unentered, as long as is possible, until the arrival of help from elsewhere. That is the course of wisdom, as well as the course of courage, and I am deficient in neither wisdom nor courage.'

This eminently wise and sensible decision, reached with such incisive logic, might not have been so readily attained had the sentry been acquainted with the theory of probability. For then he might have argued thus: 'I find in the record of history a thousand cases similar to my own, wherein the person concerned decided upon treachery, and in only four hundred of these cases the rebellion failed and he was beheaded. On balance, therefore, the advantage seems to lie with treachery, provided one does it often enough.'

Having one's head cut off is, for the person concerned, rather final. Had the sentry decided to support the rebellion, he might have had time, just before the axe fell, to reflect that he would never, in fact, be able to repeat his experiment a thousand times, and that thus the guidance given him by actuarial considerations had proved illusory.

* *Metroeconomica*, vol. 1 (1949), pp. 161–73. Trieste.

When some kind of performance, such as the tossing of a coin or the throwing of a pair of dice, has been many times repeated in suitably uniform circumstances, we can establish for each possible result of such performance the approximate number of times it will occur in a given number of repetitions of the performance. If then some value is assigned to each possible result, so that with a tossed coin we count, for example, a head as worth 1 and a tail as worth 0, or with dice we value any throw by the number appearing on the upper faces of the dice, and if for each possible result we multiply its valuation by the number of times it will occur in some suitably large number of repetitions of the performance, we shall get a valuation of this series of performances considered as a whole. Provided our frequency-ratios are correct (it will be remembered that we have postulated 'suitably uniform' circumstances and a 'large' number of repetitions: I do not propose here to probe further into these terms), this valuation of the series of performances considered as a whole has nothing whatever to do with ignorance or uncertainty: it is knowledge. If each of the possible courses of action open to me in some situation consists in such a series of performances, and if the values assigned to various possible outcomes of any one type of performance properly reflect my tastes, then my choice can fall without hesitation on that series whose valuation is highest.

What conditions must be satisfied in order that this sort of knowledge may be applied when an individual decision-maker is faced with a choice (as every one of us is, every day and every hour) between a number of rival courses of action? The conditions can be epitomized by two statements:

(1) The frequency-ratios, unless derived *a priori*, must have been obtained from a set of performances sufficiently uniform and sufficiently numerous. This condition may be satisfied in varying degrees and will accordingly give answers of varying precision. We must notice that the experience or the set of performances from which the frequency-ratios are obtained need not have been suffered or made by the person who proposes to use the knowledge.

(2) The experiment, to the valuation of whose outcome the frequency-ratios are to be applied, must be what I shall call a *divisible* experiment.

It is this second condition which, in a great and fundamentally important class of cases, cannot be satisfied. The first condition may often, of course, be only very poorly met, and the valuation assigned to an experiment will not then fully deserve to be called knowledge. True uncertainty will have entered; not because frequency-ratios

find a place in the calculation, but, on the contrary, because the strict conditions which justify their use are absent. Sometimes, as when a new invention is to be exploited, it is impossible in the nature of things to get strictly relevant frequency-ratios. Still we must remember that other people's experience, gathered at many times and places, may be as good as a man's own for determining frequency-ratios, and the difficulty of obtaining them is the lesser strand in my argument. That argument is really concerned with the question whether and when frequency-ratios have meaning and applicability for a contemplated future experiment, and I want, therefore, to attempt a classification of experiments from this point of view.

When an experiment is such that its result will be obtained by adding together the results of a series of separate performances, and when these performances are going to be sufficiently uniform in their circumstances, and sufficiently numerous, for frequency-ratios derived from a past series of similar performances to be applied, I shall call the experiment, which consists in the totality of the contemplated series of performances considered all together as one whole, a *divisible* experiment. Now it is plain that not every experiment is divisible in this sense. But when confronted with a non-divisible experiment, can we not handle the difficulty by seeking other experiments sufficiently like this one, sufficiently numerous, and able to be all performed at dates near enough in future time to provide us with the bricks for building up a divisible experiment? Can we not, in other words, treat any non-divisible experiment, not as something having an importance in its own right, as a particular, individual event, but as a mere anonymous contribution to a series or aggregation of performances, a series essentially similar to the tossing of a coin a thousand times? Can we not, in brief, treat it as what I shall call a *seriable* trial? It is plain that the Chinese sentry would have come to grief had he relied on such a solution, and his case gives us a clue. It does not matter whose past experience we use as a source of frequency-ratios, but the fears and hopes attached to future experiments only concern us if these fears and hopes are our own, only if we ourselves are going to be the gainers or losers by those experiments. There are two ways in which this condition can be met: either a great number of people must agree in advance to pool the results of the experiments which they as individuals will separately make; or else each individual must be able to feel sure that he himself as potential gainer or loser will in fact be able to repeat his own experiment many times. The former method is called insurance, and its scope as we know is very far from covering all the contingencies of life. Nor should we wish it

to do so if it could, for to insure against failure would, in this hard and competitive world, mean also insuring against success. The second method is logically excluded from all that type of cases which I shall call *crucial* experiments.

By a crucial experiment I mean one where the person concerned cannot exclude from his mind the possibility that the very act of performing the experiment may destroy for ever the circumstances in which it was performed. We must remember that an essential part of these circumstances is the individual's own stock of experience and mental attitude. This does not, indeed, prevent a given act from being many times performed by him in relevantly uniform external, objective circumstances, but it does prevent the significance for him of the separate performances being uniform, for in strictness it is plain that every single thing we do changes our stock of experience. Part of the satisfaction we get by imagination of any contemplated act is the thrill of its success, if that should come. But this thrill cannot be the same on any later occasion as it was the first time.

Such thoughts may seem far from the pedestrian objectivity of business. So let us turn to a crucialness arising in a different way. Some experiments, like a chess move, must by their nature, whatever their outcome, inevitably change the whole course of relevant future events for the individual. If I stake my whole capital on a scheme for gold-prospecting, I may win a fortune that will transform my whole attitude and objective situation, or I may lose everything, so that by the time I have built up fresh savings I shall be an older and a different man with different ideas living in a different world. A general who adopts one scheme of tactics in a critical battle cannot expect to have his opportunity presented again should things go wrong. If I choose the wrong career when I am young, I cannot be young again. A political party which adopts a given 'platform' in an election campaign has made a choice that may well turn the course of history; there can be no repetition in relevantly identical conditions.

Let me return once more to my borrowed parable of the sentry. Here was a case where the decision-maker could not feel sure that he would be prevented from repeating his experiment many times. Had he supported the rebellion, it might have succeeded, and so he might have lived to change his allegiance again and again as the tide of affairs suggested. But does this mean that he could have applied actuarial principles on the first or any other occasion when this choice between loyalty and treachery was presented? Plainly not. Frequency-ratios give us foreknowledge of the outcome of an experiment which is going to consist of the aggregation of many separate

but similar acts. But this phrase 'is going to' puts us in danger of begging the whole question. The application of frequency-ratios only makes sense if the individual can feel sure that there will be many repetitions, that there will be a divisible experiment in which the immediate act, with which he is now concerned, will be swallowed up. Thus what we may call a 'contingently crucial' experiment, that is, an act which, for all the decision-maker can tell, can have the consequence of making further similar acts impossible for him, is for all practical purposes as crucial as one where the impossibility of repetition is logically certain.

A crucial or even contingently crucial experiment must be treated by the decision-maker as in effect unique and never-to-be-repeated. But even when an experiment is not such that repetition is logically impossible, the individual may believe that some of the circumstances which affect the result will change away from the state of affairs prevailing at his viewpoint and never be restored. Speculation in building-land in the hope of a great migration to a supposed new goldfield is perhaps an example. Such opportunities may recur only at intervals of decades. Even with such long intervals, however, you may think we ought not to use the word 'unique'. If so, let us speak of *isolated* experiments. When a man must look down a long vista of future years to find any prospect of repeating the experiment he is about to make, such distant possible repetitions must be discounted, not only because he can feel no assurance that he will really be able to make them, but also because it is not in human nature to see a distantly future event so vividly, to give it a significance as great, as a like event in the immediate future. Whether the lack of assurance and the lack of vividness are two facets of the same thing I am not sure. But if experiments spaced widely out in future time are, as it were, a rapidly convergent series, we cannot by taking them all together make a divisible experiment.

What we have achieved so far is the distinction between an experiment of the kind I have called divisible, whose outcome is the result of adding the outcomes of many separate performances all in certain respects uniform, and an experiment of the kind we may call non-divisible non-seriable, which, so far as the individual is concerned who stands to gain or lose by it, can be neither itself broken down into a number of uniform additive parts nor treated as part of a divisible experiment. Now it is plain, I think, that for a non-divisible non-seriable experiment no frequency-ratio can have any meaning or relevance. Perhaps it is not difficult to see why the opposite view has sometimes been held. When any one of the separate

performances of which a divisible experiment is composed is considered by itself as an experiment in its own right, as a particular event which for the time being is thought of in isolation from other performances, so that we are interested in the result of this performance for its own sake and not merely as an unidentified drop in an aggregative ocean, then this performance is, by assumption, non-seriable, and there is plainly an overwhelming presumption that it is, in that case, also non-divisible (for if it consisted itself of a series of more elementary performances, why cannot this latter series be prolonged?). Let us agree that even an instance of the tossing of a coin, by a man who has been tossing the same coin steadily for hours and intends to continue, can be looked on in this light as a non-divisible non-seriable experiment. But there will be in such a case a temptation to confuse the particular, mentally isolated performance with the divisible experiment in which it could be included, and to treat the 'known' outcome of the latter as though it were a guide to that of the isolated performance. The fact that out of 600 throws a die has shown an ace 100 times tells us that out of 6000 throws it will show an ace about 1000 times. But what does it tell us about the *next* throw, the 6001st? Suppose the captains in a Test Match have agreed that instead of tossing a coin for choice of innings they will decide the matter by this next throw of the die, and that if it shows an ace Australia shall bat first, if any other number, then England shall bat first. Can we now give any meaningful answer whatever to the question 'Who will bat first?' except 'We do not know'? For a non-divisible non-seriable experiment the concept of frequency-ratios is wholly irrelevant. Yet because a seriable performance, one which is going to be treated as a mere item contributing to a divisible experiment, resembles a non-divisible non-seriable experiment in the fact that we can hold a number of different hypotheses as to what its own outcome will be, and because, when each of these contingencies is given a valuation, the process of forecasting the valuation of the divisible experiment consists precisely in taking each of these hypotheses in turn, multiplying its valuation by a frequency-ratio, and adding the various products together, we are tempted to think that this same additive procedure can be applied to the rival, mutually exclusive hypotheses concerning the non-seriable experiment. Such a transfer of method is plainly illegitimate, it does not make sense, and its result will be logically meaningless.

What basis is there left on which a person, faced with a set of rival courses of action each of which is a non-divisible non-seriable experiment, can make his choice? If we ask what in such a case it is

rational to do there is no answer, if rationality means choosing the most preferred amongst a set of attainable ends. For he does not know what ends are attainable with the means at his disposal, indeed, in the face of this ignorance, his powers of action are not properly described as 'means' at all. But if we ask what it is sensible and natural for him to do, the example of the Chinese sentry again gives us the clue. The sentry saw two possible experiments, and for each of them there were two hypothetical outcomes; and for each experiment he saw no reason whatever to rely on one of its hypotheses rather than the other. For each experiment one of its two hypothetical results was agreeable and the other disagreeable, and it happened also that the two disagreeable hypotheses, one for each experiment, were identical. When he compared the two favourable hypotheses, however, one was incomparably more enjoyable to imagine than the other. Thus his choice was made.

Now I can imagine my readers saying: But this is a very special case. Two very special circumstances were present, which enabled the decision-maker to make his choice by a simple comparison of the pure *content* of the two agreeable hypotheses, leaving quite out of account, first, the question, which might arise most insistently in another case, of the relative *credibility* of the two hypotheses, and second, the whole matter of the disagreeable hypotheses. Am I saying, you may ask, that when we try to peer down our vista of future time we see nothing but the sharp clear-cut black lines of sure knowledge and the pure blank white spaces of sheer ignorance? Do our expectations resemble thus an engraving rather than a painting? Are there no greys, no half-tones? Those who may happen to have seen my *Expectation in Economics** will know that I am very far from neglecting that twilight zone between belief and disbelief without which the words 'hope and fear' would scarcely have any meaning. There is not time now to explain the nature of the measure of true uncertainty, the measure of acceptance of a hypothesis, that I there tried to adumbrate. Its chief characteristics are, first, that it is a measure of disbelief, of doubt, rather than of belief or positive confidence; and secondly, that it can be assigned to the various members of an exhaustive set of rival suggested answers to some question in degrees which are independent of each other. Thus in contrast to numerical probability, which is expressed in proper fractions which, taken over the entire set of possible contingencies, must sum to unity, the degrees of 'potential surprise' respectively assigned to the various members of a 'subjectively complete' set of

* Cambridge, 1949.

rival hypotheses need not sum to any particular total; the degree assigned to any one hypothesis can be altered without affecting the degree assigned to others; new additional hypotheses can be formed by splitting up what I have called in my book the 'residual hypothesis' (which is a mere formal recognition that, in the complex cases of real life, not all possibilities have necessarily been thought of) into additional explicit hypotheses, and the total number of rival hypotheses composing the exhaustive set can thus be increased, without the need to assign higher degrees of 'suspectness', i.e. of potential surprise, to the other, initially explicit hypotheses. These aspects of potential surprise, however, need not trouble us now. For the present purpose we are simply concerned with a variable which, when its value (in the algebraic sense) is high, makes the hypothesis to which it is assigned dimmer, less real, less interesting. I may perhaps be allowed to invoke an astronomical analogy. The brightness of the planet Venus, as seen from the earth, depends both upon its distance and upon its phase.* When the planet is at the full it is far away from us on the other side of the sun, and is therefore relatively dim. When it is at its nearest to us it lies between us and the sun, and so its visible face is wholly dark or illuminated only in a thin crescent. The planet appears to us brightest at an intermediate position, when a thick crescent is illuminated at a phase corresponding to that of the 5-day-old moon. Now these two variables which affect the brightness of the planet have their counterparts in the two variables which, I think we may say, affect the brightness, the interest, of a hypothesis. Those two variables are, first, the pure content of the hypothesis, and secondly, the degree of potential surprise assigned to it, the difficulty which the individual has in banishing from his mind, when contemplating in imagination a happy outcome, the thought that, after all, it may not come true. Upon these two variables there depends what we may call the degree of stimulus that a given hypothesis can impart.

If what I have been saying seems reasonable thus far, this is the position we have reached: a person faced with a set of rival courses of action amongst which he must choose will find that for each of them he has in mind a set of mutually exclusive hypotheses about what its outcome would be. For any one course of action, the hypotheses differ amongst themselves in their power to arouse his interest and release his mental energies: some, though exciting by their content, are too unplausible to be seriously interesting; others, though perfectly credible, are commonplace and dull. But from whatever

* These variables are of course not independent of each other.

combination of these two sources it springs, each has in some degree a power to stimulate the decision-maker's mind. If the content is agreeable, the stimulus, so far as it can influence the decision-maker, will pull him towards the course of action, if disagreeable it will push him away from it. Here we have apparently a situation of great complexity: whole terms of hypotheses are pulling and pushing the individual towards or away from each possible experiment, one might think his position would resemble that of a referee trying to manage ten games of football all being played simultaneously on the same pitch. How can he find for each experiment some single measure of its attractiveness?

Possible answers to this question fall, I think, into two classes. We can suppose that the decision-maker allows every one (or every member of some large subset) of the rival hypotheses, to which he concedes any plausibility, concerning the outcome of a course of action, to contribute something to his assessment of the merits of that course. Or we can suppose that a very few particular hypotheses will seem to him to subsume the significance of all the others, and will therefore select themselves as representatives of the entire set of rival hypotheses. Each of these two classes of answer can be exemplified by constructions using what I call potential surprise. We may call these two constructions respectively the integrative and the focus-values solutions.

In the conception that I have been outlining, each of the rival hypotheses that an individual has in mind concerning the outcome of some particular experiment can, if we like, be thought of as a very tiny plot of ground whose dimensions we shall not consider. Upon this tiny plot let us imagine a column to be erected, of a height which will represent the power of the hypothesis to stimulate the individual's mind. The little plots on which such columns are to be erected can be imagined, if you like, as forming a huge chequer-board with not just 64 squares but a very large number. As we move from left to right across this board we shall consider that we are moving towards hypotheses of successively more attractive content. As we move from the nearer to the farther edge of the board we shall consider that we are moving towards hypotheses carrying successively higher degrees of unplausibility, of potential surprise. Now it will be clear, I think, that every one of the entire set of hypotheses which the individual entertains in regard to some experiment will be able to find its appropriate, determinate place on such a board. Thus we shall see, stretching across the board, a collection of columns of varying height. How will these columns be arranged? Is it not plausible that the