PROBABILITY, RARITY, INTEREST, AND SURPRISE

WARREN WEAVER

Dr. Weaver (Ph.D., University of Wisconsin, 1921) taught at his alma mater from 1920 to 1932, when he resigned as chairman of the Department of Mathematics to become director of the Division of Natural Sciences of the Rockefeller Foundation. During World War II he served on the National Defense Research Committee, the Naval Research Advisory Committee, and as a member of the War Department Research Advisory Panel.

THE unusual is often interesting. Moreover, the frequency definition of probability, which is the one consciously or tacitly used in all applications to experience, makes it clear that an improbable event means precisely a rare event. Hence, an improbable event is often interesting. But is an improbable event always interesting? We shall see that it is not. If an event actually occurs, and if its probability, as reckoned before its occurrence, is very small, is the fact of its occurrence surprising? The answer is that it may be, or it may not be.

Suppose one shuffles a pack of cards and deals off a single bridge hand of thirteen cards. The probability, as reckoned before the event, that this hand contain any thirteen specified cards is 1 divided by 635,013,559,600. Thus the probability of any one specified set of thirteen cards is, anyone would agree, very small.

When one hand of thirteen cards is dealt in this way there are, of course, precisely 635,013,559,600 different* hands that can appear. All these billions of hands are, furthermore, equally likely to occur; and one of them is absolutely certain to occur every time a hand is so dealt. Thus, although any one particular hand is an improbable event, and so a rare event, no one particular hand has any right to be called a surprising event. Any hand that occurs is simply one out of a number of exactly equally likely events, some one of which was bound to happen. There is no basis for being surprised at the one that did happen, for it was precisely as likely (or as unlikely, if you will) to have happened as any other particular one.

So we see that an event should not necessarily be viewed as surprising, even though its a priori probability is very small indeed. If you ever happen to be present at the bridge table when a hand of thirteen spades is dealt, remember that what you ought to say is this: "My friends, this is an improbable and a rare event: but it is *not* a surprising event. It is, however, an interesting event."

For it remains stubbornly true that, under the circumstances described, we are sometimes interested in the outcome, and sometimes uninterested. It seems clear that this occurs because the hands, although equally likely, are just not equally interesting. Thirteen spades *is* an interesting hand, and indeed a Yarborough with no cards above the ten, is interestingly poor; whereas millions of the hands are dull and wholly uninteresting messes.

On the other hand, think of a round disk of metal, like a coin except with perfectly plane faces and no milling on the edge. If this coin-disk were tossed, and landed in an almost perfectly vertical position, with practically no spin, it would remain standing on its edge. One could easily adjust the thickness and the mode of spinning so that the probability of its coming down and standing on its edge would be, say, one in a billion.

When this metal disk is tossed, there are, then, three possible outcomes. It may come up heads (for which the probability differs from one half by one half of a billionth), or it may come down tails (which is exactly as likely as heads), or it may stand on its edge (for which the probability is one one-billionth).

Either heads or tails would be probable, not rare, uninteresting, and not surprising. Standing on edge would be improbable, rare, interesting, and surprising. But this surely surprising event is more than 635 times as probable as would be an uninteresting hand in our previous illustration. So standing on edge is not surprising just because it has a low probability. Why is it surprising?

It is surprising, of course, not because its probability is small in an absolute sense, but rather because its probability is so small as compared with the probabilities of any of the other possible alternatives. Standing on edge is half a billion times as unlikely as the only other two things that can happen, namely, heads or tails: whereas the dull hand is precisely as likely (or unlikely) as any of the possible alternative outcomes of the dealing process.

* Order of appearance of cards in the dealing process is not taken into account: only the final constitution of the hand. Two hands are, of course, "different" if they differ in one or more cards.

Thus one concludes that *probability* and *degree of rarity* are essentially identical concepts: that a *rare* event is *interesting* or not depending on whether you consider it interesting or not† and that an event is *surprising* only providing its probability is very small as compared with the probabilities of the other accessible alternatives. This requires that a *surprising* event be a *rare* event, but it does not at all require that a *rare* event be a *surprising* event.

It is easy to make this idea more precise, and to define a Surprise Index (S. I.) for an event. To do this, it is necessary to say a preliminary word about what is called, in probability theory, mathematical expectation, or simply expectation.

If a probability experiment can result in any one of n ways, W_n, and if the a priori probabilities of these various ways are $p_1, p_2, \ldots p_n$, then any quantity Q which takes on the values $Q_1, Q_2, \ldots Q_n$ as the experiment eventuates in its possible ways W_n is said, relative to this experiment, to have the expected value

$$E(Q) = p_1Q_1 + p_2Q_2 + \ldots + p_nQ_n.$$

This definition sounds a little formidable, but the essence of the matter is really very simple. For, since the probabilities can be thought of as the relative frequencies of occurrence of the various outcomes, the formula gives a weighted average of Q, weighting each possible value of Q proportionately to the frequency with which it may be expected to occur. That is to say, the mathematical expectation of any quantity Q is simply the average value that Q may be expected to tend toward as the number of experiments becomes large.

Now, when a probability experiment is carried out, one of the ways W necessarily occurs—that is, in fact, what we mean by "doing the experiment." This way W (the one that actually happens) had a certain a priori probability p that it would occur; and, since it has occurred, we can say that we have realized, in this experiment, a probability p. How much probability can one expect on the average to realize in a given experimental setup? We have just seen that mathematical expectation answers precisely this kind of question. We must calculate the expected value of the probability. Thus

$$E(p) = p_1^2 + p_2^2 + \ldots + p_n^2$$

is the average amount of probability we can expect to realize per trial of the experiment in question. And the way, say, W_i that the experiment actually comes out, is to be considered surprising, according to whether the Surprise Index

$$(\text{S.I.})_i = \frac{E(p)}{p_i} = \frac{p_1^2 + p_2^2 + \ldots + p_n^2}{p_i}$$

is large compared with unity or is not. The S. I., as so defined, obviously measures whether the probability realized, namely, p_i, is small as compared with the probability that one can expect on the average to realize, namely, $E(p)$. If this ratio is small and S. I. correspondingly large, then one has a right to be surprised.

For the card-hand illustration, all the p_i are equal, and at once it turns out that the Surprise Index for any such hand is equal to unity. Thus in this experiment no outcome, however rare it may be, however interesting or uninteresting, is in the least surprising.

For the coin illustration, however, if we say that W_1 denotes a head, W_2 a tail, and W_3 standing on edge, then it is at once calculated that $(\text{S.I.})_1$ and $(\text{S.I.})_2$, which are equal, differ from unity only in the eleventh decimal place; whereas $(\text{S.I.})_3$ is within one of being a half million. Standing on edge is, therefore, decidedly surprising, as well as being improbable and interesting; whereas neither a head nor a tail is at all surprising.

It was remarked, earlier, that a hand of thirteen spades is rare and interesting, but should not be viewed as surprising. How about two hands of solid trumps in one evening or, what is entirely equivalent, any two identical hands in one evening? That is a very different matter! If an experiment consists of dealing just two hands, then this experiment can come out in either one of two ways. It may, way W_1, result in two unlike hands, or it may, way W_2, result in two identical hands. Since W_2 has one chance in many billions of occurring, it is easily calculated that the Surprise Index for W_1 is essentially unity, whereas the S. I. for W_2 is an extremely large number. Thus, two unlike hands in succession are not in the least surprising, whereas two alike in succession are very surprising indeed.

This calculation also indicates that there is in fact a viewpoint which justifies surprise at a single hand of thirteen spades. Suppose, for example, that one is so much interested in a perfect hand that he lumps together all imperfect ones. That is, for this person there are only two sorts of hands, perfect and imperfect. For him the deal of a hand comes out in only two ways. The calculation of the Surprise Index corresponding to these two ways is, of course, exactly that given in the preceding paragraph. The S. I. for an imperfect hand

† It may perfectly well (alas for after-dinner conversation) be interesting to you, say, because it happened to you before, and not at all interesting to anyone else.

is essentially unity, whereas the S. I. for a perfect hand is more than 635,000,000,000. It is therefore necessary to modify the general statements made previously and say that when an improbable event is so interesting that all its alternatives are lumped together as events so dull as to be indistinguishable, then the interesting event may thereby become a surprising event.

For those who play with statistics it may be worth while to point out that it seems to be difficult indeed to calculate, in general terms, the expectation $E(p)$ of the probability for either the binomial or the Poisson case. The resulting series involves squares of factorials, which are most unpleasant.‡

For the normal case, it is trivially easy to show that the expectation in the probability density, as expressed in terms of the standard deviation σ, is given by $1/2\sigma\sqrt{\pi}$. The Surprise Index for a deviation x from the mean is found at once to be equal to $\sqrt{2}e^{-x^2/2\sigma^2}$. Thus the Surprise Index is large compared with unity provided x is large compared with σ; which is of course just what one would expect.

This last calculation, because of long experience with, and feel for, standard deviation, gives one a little notion as to what would be a sensible agreement, throughout this discussion, as to what "large" ought to mean. A Surprise Index of 3 or 5 is surely not large: one of 10 begins to be surprising: one of 1,000 is definitely surprising: one of 1,000,000 or larger is very surprising indeed: one of 10^{12} would presumably qualify as a miracle.

Also, for those who have a more professional interest in statistics, it may be interesting to mention that the ideas here presented have some rather remote connection with the concept of "likelihood" as introduced more than twenty years ago by R. A. Fisher. The simplest case to which Fisher's notions apply is similar to the cases here considered in that a probability event has occurred and one wishes to draw certain inferences from this fact of occurrence. But, in the Fisher case, the probability of occurrence depends upon some unknown parameter, and the main point is that one is attempting to infer some information concerning this unknown parameter. In the cases to which the ideas of the present article apply, there is no such unknown parameter, it being assumed that all the necessary facts are at hand to permit a precise calculation of the probabilities in question. In actual experiments, it is seldom if ever true that all the facts for an absolutely precise calculation of the probabilities are indeed available. Thus, in a card experiment, one could not be entirely certain that the shuffling process is perfect. Indeed, one could have quite an extensive argument as to what "perfect" means. Thus, in actual experiments, it is quite likely to be true, at least in a precise sense, that certain unknown parameters are present: and therefore the Fisher theory is, if one wishes to be quite precise, necessary. I do not think that this destroys the interest in the above remarks, which do apply to those many instances in which one is quite willing to simplify the situation by agreeing that the facts concerning the probabilities are known.

‡ I have spent a few hours trying to discover that someone else had summed these series and I spent substantially more trying to do it myself; I can only report failure, and a conviction that it is a dreadfully sticky mess.

The excuse for making these rather obvious, but still not trite, remarks to a general audience is that all scientists have to be concerned with probability; and it is not altogether rare, although surely surprising, to find scientists who are surprised to find improbable things occurring. They always have a right to be interested, but only seldom do they have a right to be surprised.

It is hoped, then, that the considerations here presented will give some pause to those who say too glibly and too simply: "But the theory of this experiment *must* be accepted, for the probability of the results occurring by pure chance is only 10^{-20}. Surely nothing as improbable as that could ever occur." The mere fact that a probability is low should not of itself lead to incredulity. In fact, life consists of a sequence of extremely improbable events. One must always, before concluding that a certain event is surprising, examine the probability of the alternatives. For there are circumstances, and an evening of hands of bridge is a very good example, under which the occurrence of an over-all event of fantastically small probability is, in fact, inevitable.