V.—CRITICAL NOTICES.

Wahrscheinlichkeit, Statistik, und Wahrheit. R. von Mises. Second Edition. Wien: Julius Springer, 1936. Pp. viii, 282. M. 16.

THE first edition of this work was published in 1928. It now reappears, in a considerably enlarged form, as Vol. III of the series *Schriften zur wissenschaftlichen Weltauffassung*, edited by Prof. Frank of Prague and the late Prof. Schlick of Vienna. The author is a very distinguished mathematician, formerly of Berlin and now professor at Istanbul. So far as I can discover, the first edition was not noticed in MIND. For this reason, and because the book contains an extremely clear and able statement of one form of the Frequency Theory of Probability by an acknowledged expert in the technique of the subject, I propose to review it in some detail.

The book consists of six divisions. Each is called a "lecture"; but they must have been considerably expanded from their original length. The first four of them contain the statement and explanation and defence of von Mises' theory; the other two are accounts of the application of the calculus of probability to statistics and the errors of observations and to physics. The first lecture deals with the definition of "probability", and the second with the elements of the calculus of probability. In the third von Mises considers critically alternative views to his own, and tries to deal with the arguments of opponents and the alleged improvements suggested by half-converted friends. Plainly there is some close connexion between a frequency-theory of probability and those theorems which may be grouped together under the head of "the laws of great numbers". In the fourth lecture von Mises considers carefully the meaning of these theorems and their precise relation to his frequency-definition of "probability".

The essential points in Lecture I. are the following: (i) The word "probability" may be compared, *e.g.*, with the word "work", in so far as it is undoubtedly used in many different senses in ordinary life, and it is hopeless to look for a definition of either which will both cover all these senses and mark out something capable of measurement and mathematical treatment. The proper course is to begin by attending to those regions in which the word "probability" is admittedly used in a sense in which it can be and has been made the subject of a calculus. These are games of chance, insurance, and certain mechanical and physical problems. This clear central nucleus is surrounded by a penumbra of borderline cases, such as the credibility of witnesses. In the outer darkness, and explicitly excluded by von Mises from consideration, come such usages as the "probability" of an historical narrative, the "inner probability" of a work of art, and so on. I suppose that he would therefore exclude from consideration the "probability" of an alleged law of nature, such as the conservation of energy, and the "probability" of a scientific theory, such as Einstein's theory of gravitation or the nebular hypothesis.

(ii) We must next try to discover what is common and peculiar to the cases that fall within the clear central nucleus. According to von Mises we find two such characteristics, one fairly obvious and the other less so. (*a*) We have a certain clearly delimited class of observable phenomena, *e.g.*, throws with a die, which are very numerous and can be conceived to become indefinitely more numerous as time goes on. Each member of this class must manifest *some one*, and cannot manifest *more than one*, of a certain set of alternative characteristics. *E.g.*, in the case of the fall of a die, the characteristics are a 1 uppermost or a 2 uppermost or . . . a 6 uppermost. The relative frequencies with which these various alternatives have been manifested can be determined at any moment, and it is conceived that each of them would approach indefinitely nearly to a certain limiting value as the total number of observed members of the class was indefinitely increased. (*b*) The frequencies with which the various alternatives are manifested among the members of a class might approach to limiting values in the way described above, and yet the following situation might exist. There might be one or more general rules for choosing infinite sub-classes out of the original infinite class, such that in these selected sub-classes the limiting frequency of a given alternative would be *different from* its limiting frequency in the original class. Now the second condition is that this possibility must be ruled out. The original class must be such that the limiting frequency for any alternative is *the same* for the class as a whole and for any infinite selection from it, provided only that the question whether a given individual does or does not fall into the selected sub-class is independent of the particular alternative which *it* manifests. Von Mises refers to this second condition as the "principle of Excluded System" or the "principle of Indifference to Ordinal Selection" (*Stellungauswahl*). We will call it "Randomness". He defines a "Collectivity" (*Kollektiv*) as any class which answers to these two conditions.

(iii) "Probability", as used by von Mises, has a meaning only in reference to collectivities. The minimum intelligible statement predicating a probability is of the form "The probability of the occurrence of alternative *a* in the collectivity C_A is *p*". And this means what is meant by the statement "C_A is a collectivity every member of which is a manifestation of some one of the alternatives *a*, *a′*, *a″* . . .; and the limiting value of the ratio of the number of its members which manifest *a* to the total number of its members is *p* ".

If we consider a certain particular die, *e.g.*, the relevant collectivity will be the past, the present, and all the possible future throws with *that* die. With von Mises' definition it would be sensible to ask "What is the probability of throwing a 6 with *that* die ? " But, so far as I can see, it would be meaningless to ask "What is the probability that I, who am now just about to throw that die, shall throw a 6 with it on this occasion ? " The case of vital statistics would seem to be somewhat different, since each man can die but once. Here the collectivity might be, *e.g.*, Englishmen reaching the age of 40 during 1937, considered in respect of the two alternatives of surviving or not surviving their 41st birthday. It seems to me that the notion of a collectivity, answering to von Mises' two conditions which both involve infinity and limits, can hardly be regarded as a legitimate extrapolation from the observable class in this case, even if it can in the case of throwing a die.

Be this as it may, it would be meaningless, on von Mises' definition, to ask "What is the probability of *Mr. Smith*, who became 40 in 1937, surviving to *his* 41st birthday ? " This is admitted and asserted by von Mises, but he uses an argument which is really relevant to a different point. The argument is that Mr. Smith, beside being an English *man*, is an English *human being*, is a *European* human being, and so on. Now the statistics for persons of 40 surviving to their 41st birthday are different for all these different classes, and Mr. Smith is equally a member of all of them. Why single out the statistics for one of them, *viz.*, the class of *English men*, rather than the statistics for another of them, as "the probability that Mr. Smith will survive to his 41st birthday" ? If you answer that it is unreasonable to take the statistics of a less determinately delimited class when you can get those of a more determinately delimited class, why stop at the class of English men ? Mr. Smith may be a Yorkshireman, an Etonian, and a Plymouth Brother, besides being an English man. If you go far in this direction, you will define a class of which he is the only known member, and then the notion of limiting frequency will be completely inapplicable. My comment on this argument is twofold. In the first place, it is not needed in order to show that it is meaningless to talk of the probability of a particular event on von Mises' definition of "probability". This is immediately obvious from the definition. Secondly, if a person does attach a meaning to "probability" as applied to particular events, all that the argument will teach him is what he knew already, *viz.*, that he must never talk of *the* probability of an event without qualification, but must always talk of its probability with respect to such and such data. There is nothing in the argument to prevent such a person from saying that the probability of Mr. Smith surviving to his 41st birthday, relative to the datum that he is an Englishman of 40 and to that alone, is measured by the frequency with which Englishmen of 40 have been found to survive to 41.

(iv) The fact that there are collectivities, in von Mises' sense, is an empirical fact. The evidence for the existence of limiting frequencies in games with dice, cards, etc., is provided by the experience of gamblers, proprietors of casinos, governments holding lotteries, and so on. The evidence that these limiting frequencies are the same for all selections which fulfil von Mises' conditions is provided by the failure of all gambling "systems". On p. 16 von Mises says that the probability of a certain die throwing a 6, as defined by him, is "a physical property of the die, of the same kind as its weight, its thermal conductivity, etc." I think it is plain that these assertions are highly questionable; but I shall defer consideration of them until we have seen what von Mises has to say about the laws of Great Numbers, which are likely to be relevant in this connexion.

We can now pass to the second Lecture, which is concerned with the objects and methods of the Calculus of Probability. The general problem of the calculus may be stated as follows: "You are given the probabilities for the various alternatives in certain collectivities. You are asked to infer the probabilities for the various alternatives in certain other collectivities *derived from* the former". It is no part of the business of the calculus to provide the original probabilities; these must be supplied by observation or postulated hypothetically. To think otherwise is to make a mistake about the calculus of the kind which a person would make who confused geometry with mensuration. In every probability-calculation both the premises and the conclusion are statements of probabilities. Lastly, we must remember that the probabilities 0 and 1, on von Mises' theory, do not mean "certainly not" and "certainly", respectively. They mean only that the frequency of a certain alternative in a certain collectivity tends to 0 or to 1, respectively, as the number of terms is indefinitely increased.

The question that remains is "What is meant by *deriving* a collectivity from other collectivities, and how is it done?" Von Mises says that the process of derivation has four and only four fundamental forms, and that any particular case can be reduced to a single or a repeated application of one or more of these four procedures. He calls them *Selection*, *Mixture*, *Division* and *Combination*. I will now explain what he means by them.

(i) *Selection.* This consists in selecting an infinite class, in accordance with some rule, from the members of a collectivity, and considering the probability of the *same* alternatives within the selected class. It follows from the definition of a collectivity that the probabilities are unchanged.

(ii) *Mixture.* Here we consider the same set of terms as before, but we take as a single alternative a disjunction of several of the original alternatives. Thus the original collectivity might be the throws of a certain die, considered in respect of the six alternatives 1, 2, 3, 4, 5, or 6 uppermost, and the probabilities of these might be p_1, p_2 . . . p_6, respectively. The derived collectivity might be the throws with the same die, considered in respect of the two alternatives odd or even uppermost. The first of these is a disjunction of the original alternatives 1, 3, and 5; and the second is a disjunction of the original alternatives 2, 4, and 6.

The rule for calculating the new probabilities in such cases is, of course, the Addition Rule. Von Mises remarks that this rule is often carelessly formulated. It is often said that the probability of (p or q) is equal to the sum of the probabilities of p and of q, provided that p and q are mutually exclusive. He points out that the probability of dying in one's 40th year or getting married in one's 41st year is not the sum of the probabilities of dying in one's 40th year and getting married in one's 41st year, although the alternatives are mutually exclusive. The condition which must be added is that one and the same *collectivity* is under consideration throughout. (On the Keynes-Johnson theory the corresponding condition would be that one and the same *datum*, *e.g.*, the proposition h, must be considered throughout.)

(iii) *Division*. The essential point of this may, I think, be put most clearly as follows. Suppose that, in your original collectivity, a certain set of n mutually exclusive and collectively exhaustive alternatives, a_1, a_2, . . . a_n was considered. Form a new collectivity by excluding from consideration every member of the original collectivity which manifests any of the alternatives a_{m+1} . . . a_n, and consider this in respect of the limiting frequencies of the remaining alternatives a_1, a_2, . . . a_m. The rule here is that, if p_1, p_2, . . . p_m be the original probabilities for the alternatives a_1, a_2, . . . a_m, respectively, then the new probability for each will be got by dividing its old probability by the sum $p_1 + p_2 + . . . + p_m$.

It seems to me rather futile to offer this as a fundamental procedure in the calculus. It is easy to show that the rule is a consequence of applying the general principle of inverse-probability to a certain very simple special case. And the rule of inverse-probability is itself an immediate consequence of the rule of multiplication, which von Mises introduces later in connexion with what he calls "Combination". The proof of these statements is as follows. The multiplicative rule, stated in the Keynes-Johnson notation, is $(x . y)/h = (x/h)(y/xh) = (y/h)(x/yh)$. From this there immediately follows the rule of inverse-probability, *viz.*, $x/yh = (x/h)(y/xh) \div (y/h)$. In order to get von Mises' rule of "Division" we have merely to substitute the disjunction $a_1 \vee a_2 \vee$ . . . $\vee a_m$ for y and to substitute a_1, *e.g.*, for x. Then $x/h = p_1$; y/xh obviously $= 1$, since it is the probability of an alternative proposition given that one of the alternants is true; and $y/h = p_1 + p_2 + . . . + p_m$, since the alternatives are by hypothesis mutually exclusive and are being considered with respect to the same datum. So von Mises' rule of Division follows at once.

(iv) *Combination.* Suppose we start with two collectivities, C_A and C_B, for which the alternative possibilities are respectively $a_1, a_2, \ldots a_n$ and $b_1, b_2, \ldots b_m$. Let R be any relation which correlates the terms of C_A and C_B with each other in pairs. Consider the class each member of which is a *pair* of correlated terms, one from C_A and the other from C_B. Since the C_A-constituent of any such pair has the n alternatives $a_1 \ldots a_n$ open to it, whilst the C_B-constituent of the same pair has the m alternatives $b_1 \ldots b_m$ open to it, and each of the former could be combined with each of the latter, the terms of our new class can be considered in respect of the $n \cdot m$ conjunctive alternatives of the form $a_1b_1, a_2b_1 \ldots a_nb_1$; $a_1b_2, a_2b_2, \ldots a_nb_2$; $\ldots a_1b_m, a_2b_m, \ldots a_nb_m$. (An example would be if C_A were the collectivity whose members are the throws of a certain die, considered in respect of turning up 1, 2, . . . 6; if C_B were the collectivity whose members are the throws of a certain penny considered in respect of turning up heads or tails; and if R were the relation of simultaneity between a throw with the die and a throw with the coin.) The problem here is to infer the limiting frequencies of each of the $n \cdot m$ conjunctive alternatives in the new class from the limiting frequencies of each of the n alternatives in C_A and the limiting frequencies of each of the m alternatives in C_B.

The reader may have noticed that I have spoken of forming a new *class*, and not of forming a new *collectivity*, by this method. The reason is that it is not necessary that a class formed in this way out of two collectivities should be itself a collectivity. Certainly it will have one of the two defining properties of a collectivity, *viz.*, that the frequencies with which each of the alternative possibilities is manifested by its members has a limiting value. But it need not have the other property, *viz.*, randomness, *i.e.*, the indifference of these limiting frequencies to ordinal selection from the class. Now, unless the class formed by combination be itself a collectivity, the limiting frequencies with which the various alternatives are manifested by its members will not be "probabilities", as defined by von Mises. He gives the following example of two collectivities which are not combinable into a collectivity. Suppose that C_A consists of the measured values of a certain meteorological phenomenon at a certain place at 8 a.m. on successive days; suppose that C_B consists of the measured values of the same phenomenon at the same place at 8 p.m. on successive days; and suppose we make a new class each member of which is the values at 8 a.m. and 8 p.m. on the same day. It might be that at every full-moon a certain value of one causally necessitates the same value of the other. The new class would then not be a collectivity, and the limiting frequencies of the alternatives in it would not be probabilities as defined by von Mises.

Assuming that the two correlated collectivities C_A and C_B are such that they can be combined to form a collectivity $C^R{}_{AB}$, there are still two different possibilities to be considered. C_A and C_B

may either be or not be "mutually independent". Suppose that a bag is known to contain red, white, and blue counters and no others, and that two counters are drawn in immediate succession on a great many occasions and their colours noted. Let C_A be the class of "first drawings" and let C_B be the class of "second drawings" from this bag. If the rule of the game is that the first counter is to be replaced on each occasion before the second is drawn, then C_A and C_B are independent. If, on the other hand, the rule is that the first counter drawn is to be kept out on each occasion until after the second has been drawn and that the two are then to be replaced, then C_A and C_B are not independent. Von Mises gives a rather complicated definition of "independence" on p. 62. It amounts to the following: Let C_A and C_B be two collectivities whose terms can be correlated one-to-one. We say that C_B is "independent of" C_A if, and only if, the following condition is fulfilled. Select in any way that you like an infinite class from C_A. Consider the terms of C_B which are correlated with the terms of this selected sub-class. Select from them, in any way that you like, an infinite sub-class. Then the limiting frequency of each alternative within this latter sub-class must be the same as the limiting frequency of the same alternative within the whole class C_B.

According to von Mises the only way in which you can tell whether two collectivities are independent or not is by experiment. If they are independent, the limiting frequency for any alternative *ab* in the combined collectivity $C^R{}_{AB}$ will be equal to the product of the limiting frequency for *a* in C_A and the limiting frequency for *b* in C_B. Otherwise there will not be this equality. The only way in which such a question can possibly be decided is by carrying out a long enough series of observations.

Von Mises ends the lecture by working out in elaborate detail, in terms of his four processes of derivation, the simple problem in dice-throwing which the Chevalier de Méré set to Fermat, whose solution of it was the beginning of the calculus of probability. (There is a bad misprint in von Mises' solution on p. 73. In the last equation on that page the reader should substitute 5/6 for 1/6 and 35/36 for 1/36.)

The third Lecture is entitled "Critique of the Foundations". Von Mises first criticises certain alternative theories, and then considers criticisms on his own theory and proposed modifications of it. The most important points are the following.

(i) The classical definition of "the probability of an event" originated with Laplace and has been handed down in successive mathematical text-books ever since. He defined the "probability of an event" as the ratio of the number of cases favourable to it to the total number of cases both favourable and unfavourable to it, all these being assumed to be "*equally possible*". Von Mises fastens on the last proviso. He has little difficulty in showing that "equally possible " can mean only " equally likely to happen ". So Laplace's statement is undoubtedly circular, if taken as a *definition* of the " probability of an event ". The only way to avoid this charge of circularity is to say that Laplace takes the notion of " *equally* probable " as indefinable, and then proceeds to define the statement that the probability of an event is so-and-so in terms of this notion.

Now anyone who takes this view will be in difficulties whenever he has to deal with a case, such as a loaded die, where the probabilities of the various alternatives are not equal. He will have to try to split up the unequally probable alternatives into disjunctions of different numbers of more fundamental alternatives all of which are equally probable ; or else to admit that the theorems of the calculus of probabilities cannot be applied. Now von Mises makes the following criticisms at this point. (*a*) This kind of analysis, even if it can be performed, is extremely artificial in the case of loaded dice, insurance problems, etc. (*b*) Yet no one hesitates to apply the theorems of probability to the limiting frequencies which are found by observation in these cases. And (*c*) in point of fact the equal probabilities, in the case of a die which is fair, have to be established in precisely the same empirical way as the unequal probabilities in the case of a die which is loaded. In each case they are simply the limiting frequencies with which the various alternatives present themselves in a collectivity of throws. If the limiting frequencies for the various alternatives 1, 2, . . . 6 are all 1/6 in the case of die A and are, *e.g.*, 1/21, 2/21, . . . 6/21, respectively, in the case of die B, there is no rational ground for regarding the latter set of unequal probabilities as any less fundamental than the former set of equal probabilities.

Von Mises suggests that people have thought that equi-probability is fundamental, because they have thought that there are cases in which they could tell *a priori* that the alternatives are *equally* probable, whilst no-one imagines that he can tell *a priori* what are the probabilities of the various alternatives when they cannot be seen to be equi-probable. He is referring, of course, to the so-called " Principle of Indifference ". He argues, quite successfully in my opinion, that in any actual case the evidence for equi-probability is always empirical, though it does not always take the form of carrying out a series of trials with the particular object under consideration. In dealing with any particular die or penny, we know that dice are generally deliberately made as " fair " as possible, that pennies are generally made with a head and a tail and not with two heads or two tails, and so on. Again, suppose we did know *a priori* that an *accurately* cubical object, made of *perfectly* homogeneous material, would be equally likely to fall with any of its faces uppermost if *fairly* thrown. How could we possibly apply this *a priori* knowledge in any particular case ? As a matter of fact we know quite well that a die is *not* a perfectly symmetrical object, since it has different numbers of spots on different faces. How can we tell, except empirically, that this difference is irrelevant to the frequency with which these variously spotted faces will fall uppermost ?

(ii) The Laplaceans profess to find in Bernoulli's and Bayes's theorems, *i.e.*, in the laws of Great Numbers, a "bridge" by which they can pass safely to and fro between their definition of "probability" and the frequency-theory. Von Mises holds that this view is fallacious; but the point must be deferred until we consider his account of these laws.

(iii) Von Mises uses the well-known paradoxes and contradictions, which arise when the Principle of Indifference is employed to determine the probabilities of a *continuous* set of alternatives, in order to reinforce his contention that the Principle is worthless and that probabilities are always limiting frequencies based either on direct observation or postulated hypothetically and tested by observable consequences. In this connexion he criticises von Kries's "*Spielraum*"-theory of probability.

It seems to me that von Mises' criticisms on alternative theories are highly damaging; it remains to be seen what he has to say in answer to attacks on his own theory.

(i) We may defer his answer to the contention that there is a contradiction between the frequency-definition of "probability" and the result of Bernoulli's theorem.

(ii) It may be objected that, according to von Mises, probabilities are defined as the *limits* to which observed frequencies within a class approach indefinitely near as the number of members of the class is indefinitely increased, and that nevertheless he describes them as physical properties discoverable by observation. To this the only answer that I can find is the retort that mechanics makes use of the notions of points, material particles, etc.; and that the notions of density, velocity, etc., in physics all involve proceeding to a limit and yet are determined by experiment and observation.

(iii) Two objections may be made in respect of the "randomness", which is an essential part of von Mises' definition of a "collectivity". (*a*) It might be contended that the phrase "infinite class for which there is no intrinsic rule of construction", which is what von Mises' definition of a "collectivity" seems to involve, is simply meaningless. To this von Mises' answer is that the Formalist school of mathematicians need not object, provided that phrase is not self-contradictory, and that the Intuitionist school need not object, provided that a series answering to this description could be constructed by a procedure which they admit in other cases to be valid. Now the phrase has not been shown *to be* self-contradictory, though it has also not been shown *not* to be so. And Intuitionists do admit series for which the only rule of construction is to throw a die continually and note what turns up on each occasion.

(*b*) It might be objected that, even if the two factors in von Mises' definition of a "collectivity" are severally intelligible, yet they are mutually inconsistent. It might be said that, unless there is a law connecting position of a term in a series with the alternative which it manifests, it is meaningless to talk of the frequency with which that alternative is manifested within the series as having a limiting value. Yet the condition of "randomness" just is the condition that there is no such law. To this von Mises makes the following answer. (α) There are plenty of series which *are* given by an intrinsic rule, where, nevertheless, we *cannot* say whether the frequency of a certain alternative has a limiting value or not. (An example is the following. Suppose you take the series of digits in the endless decimal which expresses the square-root of π, and substitute a 0 for each even digit and a 1 for each odd digit. The series is constructed according to a rule; but there is no answer to the question whether the frequency with which 1's occur in it has a limiting value.) This appears to me to be interesting, but quite irrelevant to the objection under discussion. (β) He says that, unless there is something in the description of a series which positively *excludes* the possibility of the frequency of an alternative in it having a limit, you are at liberty to suppose that there is such a limit and to work out the consequences. (γ) If it is objected that this reduces the whole calculus to a game, he points to the practical applications of the theory in physics and social statistics.

I think it must be admitted that the objections which we have been considering are highly plausible, and that von Mises' answers to them are not very convincing. But I think that we can go further. These objections may be called "logical", in the sense that they raise doubts as to whether any clear meaning can be attached to the statement that there are "collectivities" and "probabilities" as defined by von Mises. But, even if these logical difficulties could be removed, a serious epistemological question would remain. How are we justified in passing from the empirical premise that the frequency with which a certain die has fallen with 6 uppermost in the N times which, so far as we know, it has been thrown is so-and-so, to the conclusion that, if it were thrown infinitely many times, the frequency would approach indefinitely near to the limiting value so-and-so? Again, how can we establish empirically the very sweeping universal negative proposition there there is *no* way of selecting an infinite sub-class from the original class of throws which would have a different limiting frequency for the same alternative? If we have any rational ground for believing such conclusions on such evidence, must it not involve principles of "probability" in some important sense of "probability" not contemplated by von Mises? This would not necessarily be any objection to von Mises' definition; for he is admittedly confining his attention to "probability" in the sense in which it can be measured and made the subject of a calculus. But it would show that we should have no reason to believe any propositions about probability, in his sense, unless there are logical principles of probability, in another sense.

The rest of Lecture III is devoted to writers who agree in the main with von Mises but propose a less rigid condition in defining collectivities than that of complete randomness. The least rigid of these suggested conditions is that the series must be " Bernoullian ". Suppose that p is the probability of a certain alternative being manifested by a term in the series, and suppose that we take as the terms of a new series the first n, the second n, . . . and so on, terms of the old series. Then the Bernoullian condition is that the probability of any term in the new series being any particular ordered sequence of r occurrences and n-r non-occurrences of the given alternative must be $p^r(1 - p)^{n-r}$ for all values of n and r. Other writers, such as Popper and Reichenbach, have proposed a more rigid condition, which includes the Bernoullian condition and another besides. Von Mises claims to show that series can be constructed which answer to these conditions and yet have limiting frequencies for the occurrence of certain alternatives which no-one in his senses would admit to be the *probabilities* of those alternatives. Hence a more rigid condition is needed in order to demarcate collectivities whose limiting frequencies shall be what are commonly taken as the probabilities of such and such alternatives. He mentions the American mathematician Copeland as one who has come nearest to defining conditions which are sufficient and yet are less sweeping than his own condition of complete randomness.

Finally, on pp. 120 to 122, von Mises gives a sketch of the work of the mathematician Dörge, who has tried to construct an axiomatic theory on von Mises' lines which shall avoid the criticisms brought against the theory in its original form. This looks very interesting, but it is too technical to be summarised here.

We can now pass to the fourth Lecture, which deals with the Laws of Great Numbers, *i.e.*, with Bernoulli's, Poisson's, and Bayes's theorems, and with later extensions and polishings of these. Von Mises' discussion of these questions seems to me to be extremely valuable and illuminating.

Let us take Bernoulli's theorem and Bayes's theorem as typical, since the former is simple to state and the latter is, in a certain sense, the " inverse " of it. We will begin with Bernoulli's theorem. I think that the essential points in von Mises' discussion of it may be stated as follows.

(i) Whatever meaning we may attach to the word " probability ", both the premises and the conclusion of Bernoulli's theorem are in terms of probability.

(ii) The correct statement of the theorem is as follows. Suppose that the probability of a certain alternative being realised on any one occasion of a certain kind is p. (Take, *e.g.*, the probability of throwing a 6 in any one throw with a certain die.) Consider a set of n such occasions; *e.g.*, n successive throws with this die. Let ϵ be any fraction, *e.g.*, one-millionth. Let $\pi_{n,\epsilon}$ be the probability that this alternative will be manifested not less than $pn - n\epsilon$ times and not more than $pn + n\epsilon$ times in such a set of n occasions. Then, no matter how small ϵ may be, the probability $\pi_{n,\epsilon}$ will approach indefinitely near to 1 as n is indefinitely increased.

(iii) We must now interpret this proposition when "probability" is defined in terms of limiting frequency. I shall state it in my own way, but I shall be giving what is in fact von Mises' interpretation of it. Consider, *e.g.*, a series each member of which is *a single throw* with a certain die. Let N be the total number of times it has been thrown, and let N(6) be the total number of these which have turned up 6. It is assumed that the ratio N(6)/N approaches indefinitely near to a certain limit p as N is indefinitely increased. And it is assumed that this series is "random". Now consider a new series each term of which is *a set of* n *throws* with the same die. Let ϵ be any fraction, *e.g.*, one-millionth. Let N′ be the total number of such *sets* that have occurred, and let $N'(pn \pm n\epsilon)$ be the number of such sets which contain not less than $pn - n\epsilon$ and not more than $pn + n\epsilon$ 6's in each. Then (*a*) the new series is "random". (*b*) The ratio $N'(pn \pm n\epsilon)/N'$ approaches indefinitely near to a certain limiting value $\pi_{n,\epsilon}$ as N′ is indefinitely increased. And (*c*) no matter how small ϵ may be, this limiting ratio $\pi_{n,\epsilon}$ will approach indefinitely near to 1 as n, the number of terms in each set, is indefinitely increased. This conclusion may be summed up more colloquially as follows. However small ϵ may be, if you increase the number of *terms in each set* and the *number of sets* sufficiently, an overwhelming majority of the sets will contain a proportion of 6's which differs from p by less than ϵ.

(iv) It is sometimes objected that, if the frequency-theory of probability were true, Bernoulli's theorem would consist in laboriously proving what is already asserted in the premise that the probability of a certain alternative being realised on any one occasion is p. It is quite evident from the interpretation of the theorem given above that this objection is mistaken.

(v) On the other hand, it is sometimes objected that the frequency-theory assumes something to be *certain* which the Bernoulli theorem proves to be only *very probable*. In the case of a die, *e.g.*, the frequency-theory assumes that the ratio N(6)/N has a certain exact limiting value p when N is indefinitely increased. But the Bernoulli theorem, it is alleged, shows that we have no right to assert more than that N(6)/N is very unlikely to differ by more than a certain pre-assigned small amount from p if N be made large enough. A glance at the accurate statement of the theorem above will show that this objection is invalid. The conclusion of the theorem, in our notation, is not about the limiting value of N(6)/N in the original series of *single throws* as N is indefinitely increased. It is about the limiting value of $N'(pn \pm n\epsilon)/N'$ in the series of *sets of* n *throws* when both n and N′ are indefinitely increased.

(vi) The notion that Bernoulli's theorem could act as a "bridge" between "probability" in the Laplacean sense and "probability" in the frequency sense is a complete delusion. In whatever sense "probability" is used in the premises it must be used in that sense in the conclusion. Let us take a concrete example to illustrate this. Bernoulli's theorem shows that, if the probability in the Laplacean sense of throwing a head with a certain coin is ½, then the probability in the Laplacean sense of getting between 49 per cent. and 51 per cent. of heads in a set of 10,000 throws with this coin is approximately ·95. It also shows that the probability in the Laplacean sense of getting between 49 per cent. and 51 per cent. of heads in a set of 100 throws with this coin is approximately ·16. Now in cases like the first, where the Laplacean probability is nearly 1, there is a strong tendency to pass surreptitiously from the Laplacean probability to assertions about limiting frequency. There is a strong tendency to state the conclusion in the form that in almost all sets of 10,000 throws the percentage of heads will fall between 49 and 51. But would a Laplacean be prepared to make a similar identification of Laplacean probability with limiting frequency in the second case, and to say that in 16 per cent. of sets of 100 throws the percentage of heads will fall between 49 and 51? If it is justifiable to identify *high* Laplacean probabilities with limiting frequencies of nearly 100 per cent., surely it must be equally justifiable to identify any lower Laplacean probability with a correspondingly lower limiting frequency. The plain fact is this. You cannot legitimately draw any conclusion about the limiting frequency with which a *certain proportion of heads* will occur in a series of *sets of* n *throws* unless you start with a premise about the limiting frequency with which *a head* will occur in a series of *single throws*. And, beside this premise, you will need the further premise that the occurrence of heads is "randomly distributed" in the original series of single throws, in the sense explained above.

Having, as I hope, made von Mises' position about Bernoulli's theorem and its relation to the frequency theory quite clear, I can deal much more briefly with Bayes's theorem. I shall again state von Mises' view in my own way. In order to be as concrete as possible I will again talk in terms of dice.

Suppose you have a set of N dice, each of which has been thrown n times and has given the *same* number $n(6)$ of sixes. Let $N(p)$ be the number of these dice which, if thrown an indefinitely large number of times, would turn up 6 with the limiting frequency p. (Of course p is a proper fraction capable of having any value from 0 to 1 inclusive). Then (*a*) for every possible value of p the corresponding ratio $N(p)/N$ has a characteristic limiting value as N is indefinitely increased. (*b*) Let ϵ be any fraction, *e.g.*, one-millionth, and let $N[n(6)/n \pm \epsilon]$ be the number of these dice which, if thrown for an indefinitely large number of times, would turn up 6 with a limiting frequency not less than $n(6)/n - \epsilon$ and not greater than $n(6)/n + \epsilon$. Then the ratio of $N[n(6)/n \pm \epsilon]$ to N approaches indefinitely near to 1 as limit when both n and N are indefinitely increased, no matter how small ϵ may be. The conclusion may be summed up more colloquially as follows. Suppose you have a very large number of dice, each of which has been thrown a great many times and has given the same proportion of 6's. Let ϵ be any fraction. Then, if only the dice be numerous enough and you throw each of them long enough, an overwhelming majority of them will give 6's in a proportion which differs by less than ϵ from the observed proportion, no matter how small ϵ may be.

It is obvious that this theorem is of the utmost importance for the practical application of the frequency theory of probability. For the essential point of it is the following. It enables you to start with the *observed* frequencies in a number of similar series, and to conclude that the *limiting* frequencies in the great majority of these series differ very little from the observed frequencies.

Lecture IV concludes with a fascinating account of the extensions of Bernoulli's and Bayes's theorems which have been made in recent years by Polya and Cantelli, and with an introduction to the notion of Statistical Functions.

I shall touch very lightly on the two remaining lectures, although they are of extreme interest. In Lecture V von Mises explains and deals with Marbe's problem of the expectant father who hopes that his child will be a boy and studies the recent birth-statistics; with Polya's treatment of the statistics of epidemics; and with Lexis's notion of normal, sub-normal, and super-normal dispersion. The fundamental problem of statistics, according to von Mises, is to discover whether a given set of observations can be regarded either (*a*) as a finite part of a certain collectivity, or (*b*) if not, can be regarded as following by certain assignable processes from certain collectivities. He compares the whole procedure to Kepler's observations leading first to Newton's laws of planetary motion and these leading in turn to the calculation of the actual complex and not truly elliptical paths of the planets. The lecture ends with a discussion of the theory of errors of observation, illustrated by the device known as Galton's Board.

The sixth and last lecture deals with the applications of probability in physics. It treats of the classical kinetic theory of gases; the theory of Brownian movement; the theory of radio-active discharge; the more recent developments of gas-theory by Einstein, Bose, and Fermi; and the Uncertainty Principle in quantum mechanics.

It would be difficult to recommend this book too highly. It is written with admirable clearness; it presupposes no advanced mathematical knowledge; it is full of the most interesting examples; and it provides at intervals admirable summaries of the argument and the conclusions. It is very much to be hoped that it will be translated into English.

C. D. BROAD.