VI.—CRITICAL NOTICES.

A Treatise on Probability. By J. M. Keynes, Fellow of King's College, Cambridge. London: Macmillan & Co. Ltd., 1921. Pp. xi, 466.

Mr. Keynes's long awaited work on Probability is now published, and will at once take its place as the best treatise on the logical foundations of the subject. The present reviewer well remembers going over the proofs of the earlier parts of it in the long vacation of 1914 with Mr. Keynes and Mr. Russell. From these innocent pleasures Mr. Keynes was suddenly hauled away on a friendly sidecar to advise the authorities in London on the *moratorium* and the foreign exchanges. Mr. Russell (like the foreign exchanges) received a shock, from which he has never wholly recovered, in learning that the logic books had been deceiving him by their reiterated assertions that "man is a rational animal"; and the *Treatise on Probability* was held up till this year.

The present treatise is essentially philosophical rather than mathematical, although it contains a fair amount of mathematics. It is divided into five parts. The first defines probability and discusses how far it can be measured. The second gives the fundamental theorems of probability in strict logical form. This part owes a great deal to Mr. W. E. Johnson, to whose magnificent work on this subject Mr. Keynes acknowledges his great obligations. Indeed the Muse of Probability seems to have fixed her seat at King's College, Cambridge, of which both Mr. Keynes and Mr. Johnson are fellows. The third part deals with the logical principles of inductive and analogical generalisation; and the fifth with the connected, but more complex, problem of inductive correlation or statistical inference. In between these two is sandwiched Part IV., which is entitled "Some Philosophical Applications of Probability". This is concerned with a number of historically interesting problems, and in particular with the application of probability to ethics. At the end of the work Mr. Keynes provides an admirable bibliography of books and articles on probability and kindred subjects.

In this review I shall try to give an outline of Mr. Keynes's theory. I shall not have many serious criticisms to make, because I am substantially in agreement with him, and where I am not persuaded by his arguments the subject is so difficult that I have little of value to suggest as an alternative to his views.

The fundamental thesis of the book is that probability is a relation between propositions, which may be compared with implication. When p implies q the belief that p is true justifies an equally strong belief in q. But there are numberless cases where a belief in p justifies a certain degree of belief in q, but does not justify so strong a belief in q as we have in p. In such cases there is a certain logical relation between p and q, and this relation is of the utmost importance for logic. But it is not the relation of implication. It is this other relation with which probability is concerned. This probability relation is capable of degree, since it may justify a more or a less confident belief in q. The typical probability statement is of the form "p has to q a probability relation of degree x". Implication may perhaps be regarded as the strongest probability relation, or better as a limit of all possible probability relations.

There is however a very important difference, which is not merely one of degree, between the implicative and the probability relations. There is nothing corresponding to the Principle of Assertion in probability. If one proposition implies another and we know that the first is true we are justified by the Principle of Assertion in going on to believe the second by itself, and in dropping all reference to the first. We can never do this in probability. We can never get beyond statements of the form "p has such and such a probability with respect to the datum q". Propositions are true or false in themselves, though we may need to know their relations to other propositions in order to *know* whether they are true or false. But probability is of its very nature relative. When we talk of *the* probability of a proposition this phrase is always elliptical, as when we say that the distance of London is 120 miles. We simply assume that the person to whom we are speaking will supply from his own mind the same data as we are taking. Two important consequences flow from this. In the first place, a proposition may be highly probable with respect to certain data and yet be false. Its turning out to be false makes no difference whatever to the fact that it is highly probable with respect to these data. Secondly, one and the same proposition may have many different probabilities at the same time, so long as the data are different in each case. In particular a proposition may be highly probable with respect to a certain set of data and highly improbable with respect to another set of data which includes the first set as a part. Thus, if the only fact that you know about a man is that he has recently swallowed arsenic, it is highly probable with respect to these data that he will be dead in the next half hour. If you afterwards get the additional piece of information that he has taken an emetic, the probability that he will die in the next half hour, on the combined data, is much smaller. Neither probability is in any way more "correct" than the other. This essential relativity of probability is absolutely fundamental, and most previous expositions have suffered by failing to grasp it.

To express these facts Mr. Keynes takes over a useful symbol from Mr. Johnson. He writes $q/p = x$ for "the probability of q with respect to the datum p is of magnitude x". Two questions at once arise: (1) Can probability always be measured? and (2) Why do we commonly prefer a probability with respect to wider data to a probability with respect to narrower data? These questions are dealt with by Mr. Keynes in two chapters in the first part.

(1) Mr. Keynes argues that there is no reason to suppose that all probabilities fall into a single scale. All indeed lie between certain truth and certain falsehood, but there may be innumerable series leading from the one to the other. It is only probabilities that lie in the same course that can be directly compared. Two different courses may cut each other at one or more points, *i.e.*, there may be certain probabilities which are common to several different series. When this happens there is a possibility of indirectly comparing two probabilities in different series by comparing both with one that is common to the two series. But, even when we confine ourselves to the probabilities of a single series, there is no guarantee that we shall be able to set up a consistent system of numerical measures for them. Not every series of comparable magnitudes is measurable. The mathematicians have naturally exaggerated the amount of numerically measurable probability in the world; and, when they came across probabilities that were not comparable, or, if comparable, not numerically measurable, they passed by and "thanked God that they were rid of a rogue". Probabilities are only measurable in the comparatively rare cases where we have a field of possibilities which can be split up disjunctively into exhaustive, exclusive, and equiprobable alternatives. This does happen in games of chance and in the "bag" problems in which mathematicians exercise themselves, but not in many other cases.

It must be noticed that this view of Mr. Keynes's is much more radical than the view that all probabilities are theoretically measurable, but that in most cases the practical difficulties are insuperable. Mr. Keynes points out that there is one and only one theory of probability on which the latter view is plausible. This is the Frequency Theory, which he proceeds to discuss.

There is something bluff and Anglo-Saxon about the Frequency Theory, which no doubt accounts for its extreme popularity with the Island Race in general and with Prof. Whitehead in particular. Moreover there is a real but rather complex connexion between probability and frequency by way of Bernoulli's Theorem; and the very narrow limits within which that theorem and its converse can be applied have been overlooked by most people, as Mr. Keynes points out in the later parts of the present work. Thus there are many excuses for accepting the Frequency Theory. Mr. Keynes has little difficulty in showing that, in the simple-minded form in which it appears in Venn's *Logic of Chance*, it is unsatisfactory, and that Venn tacitly assumes in many places a sense of probability other than that which is laid down in his definitions. Prof. Whitehead's form of the theory, as might be expected, is a good deal more subtle. Unfortunately it is not easy to make out exactly what it is. Mr. Keynes states it in the way in which he has understood it from private correspondence, but admits that he may be mistaken about Whitehead's meaning. It is therefore hardly profitable for a third person to discuss this form of the theory. But it is open to a reviewer to point out what seems to him to be a fallacy in Mr. Keynes's arguments against the theory. Keynes argues that Whitehead's form of the theory shares with Venn's the defect that it cannot satisfactorily explain the fundamental axiom connecting the probability of a disjunctive proposition with the probabilities of its separate parts, *i.e.*, the proposition

$$(p\vee q)/h = p/h + q/h - pq/h.$$

On the Frequency Theory, as interpreted by Mr. Keynes, the datum h determines a certain class α of propositions of which p is a member, a certain class β of which q is a member, and certain classes γ and δ of which the propositions pq and $p\vee q$ are respectively members. The probability of p with respect to h is then defined as the ratio of the number of true propositions in the class α to the total number of propositions in this class. Similar definitions apply, *mutatis mutandis*, to the probabilities of q, pq, and $p\vee q$, respectively. He then points out, quite truly, that the question whether the fundamental addition-theorem mentioned above will hold at all depends entirely on what particular classes, α, β, γ and δ, the datum h does determine for the four propositions in question. So far I quite agree, and think that this is a very serious difficulty in the way of the theory in question. But Mr. Keynes then proceeds to tell us what must be the values of the classes, α, β, γ, and δ, if the equation is to hold. He says that δ must be the class of propositions of the form $p\vee q$, where p is a member of α and q of β; and that γ must be the class $\alpha\beta$ of propositions. It is very easy to make up simple concrete examples to disprove this; *i.e.*, to make up examples in which the fundamental theorem does not hold even when the classes of reference are determined in this particular way. But it is better to disprove it quite generally. It can be shown that the number of propositions in Mr. Keynes's class δ is $(\alpha)(\beta) - (\alpha\beta)/2 - (\alpha\beta)^2/2$. It can also be shown that the number of *true* propositions in this class is :—

$$(\alpha_r)(\beta) + (\beta_r)(\alpha) - (\alpha_r)(\beta_r) - (\alpha\beta)_r\,(\alpha\beta) + (\alpha\beta)_r^2/2 - (\alpha\beta)_r/2$$

where (α) = the number of propositions in the class α; (α_r) means the number of true propositions in the class α; and similar meanings attach to the other symbols. If the fundamental equation is to hold, the ratio of the second expression to the first must be equal to—

$$\frac{(a_\tau)}{(a)} + \frac{(\beta_\tau)}{(\beta)} - \frac{(a\beta)_\tau}{(a\beta)}.$$

It is quite obvious that this will not in general be true; and therefore that either Mr. Keynes or I have made some blunder in the algebra of classes. I am pretty certain that Mr. Keynes is wrong, but of course I may be wrong too. However this may be, the real force of Mr. Keynes's general criticism is not diminished, even if he has made an algebraical slip here.

If the measurement and comparison of probabilities be possible only in a few specially favourable cases it is peculiarly important to be sure what those cases are. This leads to the question: When may we judge two probabilities to be equal? And this leads us at once to one of the *cruces* of the Theory of Probability, *viz.*, the famous *Principle of Non-Sufficient Reason*, or, as Mr. Keynes prefers to call it, the *Principle of Indifference*. In the negative and critical part of this chapter Mr. Keynes found most of the work already done for him by Von Kries, one of the few writers on the philosophical side of probability who are really worth reading. Von Kries had already pointed out the absurd results which a light-hearted use of the Principle of Indifference had led to. He did indeed attempt to base on these a positive statement of the proper limits of the Principle; but I am relieved to notice that Mr. Keynes finds the precise upshot of Von Kries's positive theory as hard to grasp as I have always done myself.

By studying the cases where the uncritical use of the Principle of Indifference ends in absurdities Mr. Keynes elicits the following conditions which must be fulfilled if it is to be applicable. (1) The various alternatives under consideration must be capable of being put into the same form, *i.e.*, they must simply be different instances of a single propositional function ϕ. This cuts out the wild applications of the Principle to pairs of contradictory alternatives in which Jevons habitually indulged. The two alternatives "x is red" and "x is not red" are not of the same form. The first means that x has the colour red. The second certainly does not mean that x has the colour "non-red," for non-red is not a colour. (2) The alternatives must not be sub-divisible into other alternatives of the same form as themselves. Given that x is an inhabitant of Europe it follows that he lives either in Great Britain or in France or in Germany or. . . . These alternatives are of the same form, and so far all is well. But each of them is divisible into sub-alternatives of the same form as itself. The alternative that x lives in Great Britian is divisible into such alternatives as that he lives in England, that he lives in Scotland, etc. . . . It is by ignoring this condition that mathematicians who treat of geometrical probability so often reach different solutions of the same problem.

Subject to these two conditions Mr. Keynes states the Principle as follows. The alternatives $\phi(a)$ and $\phi(b)$ are equally probable with respect to the data h, provided that h can be written in the form $f(a)\ f(b)\ h'$, where $f(a)$ and $f(b)$ are logically independent, h' is absolutely irrelevant to both alternatives, and $f(a)$ and $f(b)$ are the only parts of h that are relevant to $\phi(a)$ and $\phi(b)$ respectively. (There is a puzzling mistake in Mr. Keynes's symbolism on p. 60, § 21. He says: "It might be the case that . . . $\phi(x) = x$ is the only propositional function common to all of them" (*i.e.*, the alternatives). He cannot possibly mean this, for it is sheer nonsense that $\phi(x)$ which is a proposition about x should ever be identical with x itself. What he really means is simply that $\phi(x)$ might be nothing but $x = a . \vee . x = b . \vee . x = c . \vee .$ . . . where a, b, c, . . . are just proper names or other designations of the alternatives. Such a ϕ will not do. His real point therefore is that the alternatives must be members of a class which is defined intensively, and not by a mere enumeration of its members.)

It will be seen then that all judgments of indifference involve judgments of irrelevance. We have to know what part of h is irrelevant to both $\phi(a)$ and $\phi(b)$ before we can see whether h does fall into the form required for the Principle of Indifference. These judgments of irrelevance are of fundamental importance in Probability, and no rules can be given for making them. In the end we have to come down to direct insight, just as we have to do in the end in judging the validity of any deductive argument.

Mr. Keynes makes one very important observation here on the dangers of symbolism. So long as we are dealing with mere a's and b's all that we know about them is that they are both instances of some ϕ. But the moment you substitute something definite, like Socrates, for a, and something else definite, like Plato, for b, you can no longer assume that the conditions for the Principle of Indifference still hold. The moment you know, not merely that you are dealing with a ϕ, but also know *which* particular one of the ϕ's you are dealing with, you may have fresh relevant information.

Having treated the conditions under which two or more probabilities may be judged to be equal Mr. Keynes turns to the question: "Under what conditions can one probability be judged to be greater or less than another?" Such comparisons can only be made directly when either (a) we have the same data, and one of the propositions whose probability is sought is a conjunctive containing the other proposition as a part; or (b) when the proposition whose probability is sought is the same in both cases, but the datum in one is a conjunctive which includes the data of the other as a part. Into the exact refinements that are needed here I will not enter. Mr. Keynes shows that, by combining cases (a) and (b), we can sometimes indirectly compare probabilities which do not fall under either rubric.

(2) The prolegomena to the measurement of probability are now completed, and we can turn to another most important question

which has already been mentioned. If there is nothing to choose in point of correctness between the probabilities of a proposition with respect to a wider and to a narrower set of data why do we prefer the former probability to the latter? Why do we attach more weight to the low probability of the patient who is known to have taken both arsenic and an emetic dying in the next half hour than to the much higher and equally correct probability of the same event relative to the narrower data that he has taken arsenic? This extremely puzzling question is attacked by Mr. Keynes in a chapter on the *Weight of Arguments*. I do not know of any other writer who has raised it except myself in the chapter on Causation in *Perception, Physics, and Reality*; though I do not doubt that Mr. Johnson has an elaborate treatment of it up his sleeve. Roughly speaking, any increase in the amount of relevant evidence increases the weight of an argument, though it may leave the probability unchanged or may decrease it. We have already seen an example of the latter; let us now consider the former. Suppose we start with a probability a/h. A new piece of evidence k may arise, and k may consist of two parts k_1 and k_2, one of which is favourably and the other unfavourably relevant to a/h. In that case it is possible that $a/hk = a/h$. Nevertheless the weight of a/hk is greater than that of a/h. Mr. Keynes discusses various cases in which weights can be compared; and he considers the relation between weight and what is called "probable error" in statistics. In general a big probable error is a sign of scanty observations, and therefore of a low weight for one's result. But this correlation is not absolutely invariable. I wish that Mr. Keynes had discussed why we feel it rational to prefer an argument of greater weight to one of less weight. I think that our preference must be bound up in some way with the notion that to every event there is a finite set of conditions relative to which the event is certain to happen or certain not to happen. So long as the evidence is scanty a high probability with respect to it does not make it reasonable to act as if we knew that the event would happen, because it is reasonable to suppose that we have only got hold of a very small selection of the total conditions and that the missing ones may be such as to be strongly relevant in an unfavourable direction. If the probability remains high relative to a nearly exhaustive set of data we feel that there is less danger that the missing data may act in the opposite direction. In fact, what we assume is that a high probability with respect to a wide set of data is a sign of certainty with respect to the *complete* set of relevant data.

This exhausts the main features of Part I. Part II. is largely the formal development of the fundamental axioms of probability. Much of it could be accepted by a person who rejected Mr Keynes's view as to what probability really is. The most exciting theorems in this part are due to Mr. Johnson, whose valuable conception of "Coefficients of Dependence" is introduced and explained. It is worth while to mention a very plausible fallacy in probable reasoning which is detected and dealt with mathematically by Mr. Johnson's methods. It seems plausible to hold that if k is favourably relevant to m/h and m is favourably relevant to x/h then k must be favourably relevant to x/h. It is shown here that this is not in general true; and the two conditions under which alone it is true are elicited. It is fairly easy to illustrate part at least of this fallacy by an example. The fact that a man is a doctor increases the probability that he will have visited smallpox patients, and the fact that a person has visited smallpox patients increases the probability that he will get smallpox. It by no means follows that the fact that a man is a doctor increases the probability that he will get smallpox. For this fact also increases the probability that he is properly vaccinated and that he will take reasonable precautions. And this of course decreases the probability that he will get smallpox. Thus we see that it is not enough that k shall be favourably relevant to something that is favourably relevant to x. It is also necessary that k shall not be favourably relevant to anything that is unfavourably relevant to x. The second condition is more subtle, and I cannot at the moment think of any simple example that would illustrate it. As an example of the power of the Keynes-Johnson methods the reader is advised to look at Chapter XVII, in which Mr. Keynes solves in a few lines problems over which Boole spent pages of algebra, arriving as often as not at results which are certainly wrong.

To the mathematician I should imagine that the most interesting thing in this part would be Mr. Keynes's beautiful treatment of Laws of Error, and his general solution of the problem: What form must the law of error take in order that the most probable value of a measured variable shall be represented by the arithmetic, the geometric, the harmonic, and other means, of the observed values? I know of no treatment of this subject which approaches Mr. Keynes's for clearness and generality. To most readers of MIND, however, the chapters of greatest interest will be the earlier ones on the notions of *Groups* and *Requirement*.

Both these notions were first devised by Mr. Johnson to deal with such problems in deductive reasoning as are raised by Mill's attack on the Syllogism and by the apparent paradox about a false proposition implying all propositions and a true proposition being implied by all propositions. Mr. Keynes first explains the applications of the theory, and then proceeds to give his own extension of it to the case of probable reasoning.

A group, so far as I can understand, consists of a set of propositions which must contain some formal principles of inference, and includes in addition all propositions that follow from the fundamental set by the principles which are contained in that set. A group is said to be real if the set of propositions which determine it are all known to be true, otherwise it is said to be hypothetical. It is of course possible for the same group to be determined by several alternative sets of propositions, though a given set necessarily determines a single group. Mr. Keynes and Mr. Johnson are both persuaded of the extreme importance of the theory of groups in the logic of inference. I agree with them to this extent, that the *facts* that the theory of groups takes into account are of vital importance. But it does seem to me that they can all be stated much more simply in other terms; and I have failed to find anything specially important that follows from the group notation and would not have been discovered without it. Possibly I am only exhibiting my ignorance. The essential point that the group theory is meant to bring out is the distinction between what Johnson calls the Logical and the Epistemic factors in inference. The latter is the question of the order in which we get our knowledge. *E.g.*, p implies q provided that either p is false or q is true. So far it is irrelevant how we came to know that this disjunction holds. But when we say "if p then q" we mean something more than this. We mean that it is possible to know that p is false or q is true without having to know that p is false or having to know that q is true. And the only way in which we can know such a thing is by seeing that the disjunction is an instance of some formally true hypothetical such as "if S*a*P then $\bar{\mathrm{P}}a\bar{\mathrm{S}}$". Again, if we want to infer q from p it is obviously necessary to be able to know that p is false or q true before you know whether q is true or not. All this can be and is expressed by Mr. Keynes in terms of the theory of groups; and my only doubt is whether it becomes any clearer or leads to anything further when so expressed.

A proposition has a probability with respect to a set of data h when neither it nor its contradictory falls into the group determined by h. Does this really enlighten us any more than to know (what is equivalent to it) that neither the proposition nor its contradictory must follow logically from the *premises* mentioned in h by the known *formal principles* of deductive logic? On page 131 Mr. Keynes has a formidable definition in terms of groups of the statement that "the probability of p does not require q within the group determined by h". When this definition is unpacked it seems to me to amount to no more than this: You can make a selection h' out of h such that no part of h outside h' will alter the probability p/h' when added to h'; and some part of h outside h' when added to h' will alter the probability of q/h'. If this be the right interpretation, it is far easier to grasp than Mr. Keynes's definition in terms of groups.

Not only am I doubtful of the fruitfulness of the group theory, I am also not satisfied that Mr. Keynes's treatment of hypothetical groups is adequate. All groups must, so far as I can see, include in their fundamental set formal principles of inference as well as premises. I quite understand that the premises may be hypothetical. But can we really allow the generating principles to be hypothetical also? Mr. Keynes does not discuss this point, which seems to me to be a very important one for a person who is going to admit hypothetical groups.

Let us next turn to Mr. Keynes's theory of inductive generalisation, which is contained in Part III. It is peculiarly gratifying to me to find how nearly Mr. Keynes's view of the nature and limits of induction agrees with that put forward quite independently by me in two articles in MIND. We both agree that induction cannot hope to arrive at anything more than probable conclusions, and that therefore the logical principles of induction must be the laws of probability. We both agree that, if induction as applied to nature is to lead to results of reasonably high probability, nature must fulfil certain conditions which there is no logical necessity why it should fulfil. Finally, we agree as to the nature of those conditions, in general outline at any rate. In some way the amount of ultimate variety in nature must be limited, if induction is to be practically valuable; the infinite variety of nature, as we perceive it, must rest on combinations of a comparatively few ultimate differences. But of course Mr. Keynes's theory is far more detailed and subtle than anything of which I am capable; and it is, so far as I know, the only account of the logic of this process which a self-respecting logician can read with any satisfaction.

The problem of induction boils down to this: We examine n things. They have the r properties $p_1 \ldots p_r$ in common; this is called their total positive analogy. There is also a set of properties $q_1 \ldots q_s$ such that each is present in some of the things and none is present in all of them; this is called the total negative analogy. Both the positive and the negative analogies in any actual case are pretty certain to be greater than the *known* positive and negative analogies, which form the only basis of our argument. Our object is to prove some proposition of the form that everything which has the properties $p_1 \ldots p_m$ has the properties $p_{r-t} \ldots p_r$. It is obvious that this can only be possible if some part of the known analogy is irrelevant. *E.g.*, all the examined instances agree in the fact that we have examined them, that they are confined to certain limits of space and time, and differ from all unexamined instances in these respects. Whenever this part of the known analogy is relevant to the attempted generalisation, it is clear that the attempt is doomed to fail. Thus an essential factor in all inductive generalisations is judgments of irrelevance. Many of them no doubt depend on past experience, but Mr. Keynes holds that there must be a residuum which is *a priori*. The only importance of the Uniformity of Nature is that it is a general principle of irrelevance, which asserts that *mere* differences of date and position are irrelevant. Mr. Keynes raises the question in a note whether this is affected by the Theory of Relativity; but he does not answer his own question. However this may be, it seems to me that the Uniformity of Nature, thus defined, is a mere pious platitude; since—whether space and time be absolute or relative—no two objects or events ever do differ merely in date or place.

Such differences always involve their being in intimate spatio-temporal relations with different sets of objects or events, and these differences cannot be assumed to be irrelevant.

Our generalisation always refers to much less than the known positive analogy. When we argue that all swans are white our generalisation only concerns whiteness and those few properties by which we define a swan. But all the examined swans were known to have many other common properties beside these, and we do not know that these are all irrelevant. All that we positively know to be irrelevant at this stage is the properties in the known negative analogy. We can reduce the dangers thus involved by seeking other instances which increase the known negative analogy. For this purpose mere *number* is unimportant. One instance which is known to differ from the previously examined ones in many of those properties which the generalisation assumes to be irrelevant is of more importance than dozens of instances which are exactly like those already examined. But there remains a danger due to the fact that the total analogy is almost certain to be greater than the known positive analogy. The extra and unknown analogies may be relevant; and, since we do not know what they are, we do not know where to look for negative analogies which will prove them to be irrelevant. In this case the only course is to increase the *number* of instances, trusting that, even though they do not differ in any known respects from those that have already been examined, they will probably between them differ in many of the unknown points of positive analogy from the examined instances. All this however only tells us how to diminish the objections to an inductive generalisation. It does not tell us that any inductive generalisation will possess a reasonable degree of probability, even when we have carried out these processes to the utmost. Something more is clearly needed if inductive generalisation is to be trustworthy.

The extra factor is dealt with in the chapter on Pure Induction. It is easy to prove that an hypothesis becomes more and more probable the more mutually independent consequences of it are verified. It is also easy to prove that, if it starts with a finite probability, sufficient verification of mutually independent consequences will make its probability approach as near as we please to unity. The problem that remains is: What justifies us in ascribing a finite antecedent probability to any inductive generalisation? To this Mr. Keynes answers that we are only justified if we assume that all the variety of perceptible properties springs from a comparatively small number of generating properties.

To each generating property there corresponds a large group of perceptible qualities, but we must admit the possibility that the class of perceptible qualities corresponding to ϕ_1 and the class corresponding to ϕ_2 may partially overlap. If so the group common to the two will not tie us down to a single generator. Setting this possibility aside for the moment, we see that if a group α of perceptible qualities is found to be accompanied by a group β there is a finite probability that the complete group $\alpha\beta$ corresponds to a single generator, or that the generators of α include among them the generators of β. If this is so α will not be able to occur without β, and there is thus a finite antecedent probability of the generalisation, on which induction can build. If we allow that a group of perceptible qualities may have a plurality of possible generators this argument breaks down; but if we assume that the plurality of possible generators for every set is finite we can still assign a finite antecedent probability to inductive *correlations*, which assert that the next S, or at least a certain proportion of the S's, will be P.

Mr. Keynes seems to me to be right here; and it is true that this is the kind of assumption that does lie at the back of all our scientific reasoning. I have only two remarks to make. (1) Does the theory of generators add anything to the facts? Would it not be enough to assume that perceptible qualities do tend to occur in bundles? This is the whole cash-value of the assumption, and the doctrine of generators seems to be nothing more than a hypothetical explanation of our assumption. (2) Mr. Keynes holds that there is no circle in saying *both* that no inductive generalisation can acquire a finite probability without this assumption, *and* that the results of induction may make this assumption progressively more and more probable.

It is therefore not necessary that the fundamental inductive assumption should be certain. It is enough if it ever had a finite probability; for all subsequent experience has tended to support it. What Mr. Keynes means is, I think, this: If the world is a system with a finite number of generating properties we might expect to find a good deal of regularity and repetition in it. Now, up to the present, we have found more and more regularity and repetition the more carefully we have looked for them. Thus the actual course of experience has been such as to increase the probability of the inductive hypothesis, provided that it started with any finite probability. This works out in practice to the result that a large part of the confidence that we now feel in any inductive generalisation is due, not to the special evidence for it, but to the enormous and steadily increasing amount of regularity that we have found in other regions. There is, I think, no circle in this. Thus the one fundamental assumption of induction is that we can know somehow that the inductive hypothesis that nature is fundamentally finite has a finite antecedent probability. Mr. Keynes admits that it is very difficult to see how we can know this. It is certainly not an *a priori* principle, self evident for all possible worlds, that every system must depend on a finite number of generators. We can only suppose that in some way we can see directly that this has a finite probability for the actual world. But the epistemology of this is at present wrapped in mystery.

In Part IV. many interesting problems are discussed; but I must only glance at them. Mr. Keynes ranges from Psychical Research to *Principia Ethica*, and from the *Argument from Design* to the Petersburg Problem; and he has something illuminating to say about all of them. From the point of view of pure probability the most important thing in this part is the definitions of an objectively chance event and of a random selection. The former is very important in connexion with statistical mechanics, the latter in connexion with most statistical reasoning. A chance event is not one which is supposed to be undetermined. Nor is it always one whose antecedent probability is very small. To throw a head with a penny is a chance event, but its probability is $\frac{1}{2}$. An event may be said to be a matter of chance when no increase in our knowledge of the laws of nature, and no practicable increase in our knowledge of the facts that are connected with it, will appreciably alter its probability as compared with that of its alternatives.

Part V. deals with the principles of statistical inference. It is too technical for me to give any complete account of it, so I will confine myself to a very short summary of the most important points in it. (1) Mr. Keynes considers the conditions under which Bernoulli's theorem holds, and shows that they are so restricted that we can seldom in practice count on their being fulfilled. (2) He severely criticises Laplace, and particularly his famous Rule of Succession. This occurs in connexion with the attempted inversion of Bernoulli's theorem. I agree with Mr. Keynes about this rule, but it seems to me that he is a little unfair to it in one respect. He assumes that it always deals with cases where what is drawn is replaced before the next drawing. On that supposition it is true, as he points out, that the formula only holds as the number of drawings tends to infinity. But the same formulæ hold without this restriction when the objects drawn are not replaced. And surely, if the Rule claims to have the slightest application to our investigations of nature, the latter is the right alternative. For we cannot observe the same event twice over, any more than we can draw a counter twice out of a bag if we do not replace it. (3) On all these subjects Mr. Keynes prefers Bortkiewicz, Tschuproff, Tchebycheff, and Lexis to the classical French school. I am afraid that, with the exception of Lexis, these names are mere sternutations to most English readers; but I suppose we may look forward to a time when no logician will sleep soundly without a Bortkiewicz by his bedside. (I must remark in passing that the beginning of Mr. Keynes's sketch of Tchebycheff's theorem seems to the uninitiated to commit precisely the same kind of fallacy which Mr. Keynes himself points out in Maxwell's deduction of the law for the distribution of molecular velocities in a gas. This is on page 353, where it is said that "the probability that the sum $x + y + z \ldots$ will have for its value $x_\kappa + y_\lambda + z_\mu \ldots$ is $p_\kappa q_\lambda r_\mu \ldots$". Surely this forgets that a sum of this value could be made up in a great number of different ways by taking suitably chosen values of the variables. Why should not $x_\alpha + y_\beta + z_\gamma \ldots$ have the same value as $x_\kappa + y_\lambda + z_\mu \ldots$? In that case the probability will be much greater than $p_\kappa\, q_\lambda\, r_\mu \ldots$.) (4) About past statisticians Mr. Keynes makes a remark which exactly hits the nail. They never have clearly distinguished between the problem of *stating* the correlations which occur in the observed data, and the problem of *inferring* from these the correlations of unobserved instances. There is nothing inductive about the former; but, as it involves considerable difficulties, the statistician has been liable to suppose that, when he has solved these, all is over except the shouting. Thus the inductive theory of statistical inference practically does not exist, save for beginnings in the works of Lexis and Bortkiewicz. These beginnings Mr. Keynes describes and tries to extend.

There are several misprints in the book beside those that are mentioned in the list of errata. On page 170 the various kinds of h's have got mixed up in the course of the argument. On page 183 it is said that "we require a/ah_2h_1," when we really want a/ah_1h_2. On page 207 substitute $\phi(z)$ for $\phi(x)$ on the left-hand side of the equation. In the formula at the bottom of page 386 read f' for f in the second factor of both numerator and denominator. On page 395 in the first line after the equation read p_1 for the second p in the line.

I can only conclude by congratulating Mr. Keynes on finding time, amidst so many public duties, to complete this book, and the philosophical public on getting the best work on Probability that they are likely to see in this generation.

C. D. BROAD.

The Analysis of Mind. By BERTRAND RUSSELL, F.R.S. London: George Allen & Unwin, Ltd., 1921. Pp. 310.

"TRAVELLING, whether in the mental or the physical world, is a joy, and it is good to know that, in the mental world at least, there are vast countries still very imperfectly explored."

Many will feel that in those words, which occur early in his latest book, Mr. Russell has aptly summed up his own attitude to philosophy. For there has seldom been a bolder traveller in those realms than Mr. Russell, and seldom one who had more power to charm his readers by the accounts of his discoveries, or to communicate to them something of the zest he himself finds in such adventures. Almost every successive work which has come from his pen represents a new voyage of discovery, and most of his readers must at times have found it difficult to keep track of his rapid progress. But never before, I think, has he made so venturesome a journey as the present one, or covered in his survey so large a stretch of country.

We have already, of course, had preliminary reports of this