Concepts of Randomness
Author(s): Peter Kirschenmann
Source: *Journal of Philosophical Logic*, Vol. 1, No. 3/4 (Aug., 1972), pp. 395-414
Published by: Springer
Stable URL: http://www.jstor.org/stable/30226051
Accessed: 21-07-2016 07:51 UTC

PETER KIRSCHENMANN

CONCEPTS OF RANDOMNESS

I. INTRODUCTION

The notion of randomness has always been rather perplexing. Although it is frequently used in natural and social science, both technically and informally, it seems to have been somewhat neglected by philosophers of science ever since the discussion of the foundations of the so-called frequency theory of probability, in which it was assigned a basic role, has faded. Yet this discussion is of such significance that any attempt at clarifying the notion of randomness will have to relate to it. After a few preliminary remarks on some of the problems and puzzles of randomness, I shall, therefore, expound and discuss a concept of random distribution of a property in classes and sequences, defined in terms of relative frequencies and their limits. Because of certain shortcomings of this concept it appears advisable to turn to probabilities, in terms of which a quite different concept, viz., that of random conjunction of properties, can readily be defined as stochastic independence. This concept still has features clashing with the ordinary sense of 'randomness' which become manifest in cases where certain probabilities assume extreme values. However, when we take measures defined in information theory as measuring the degree of randomness, to which purpose they lend themselves particularly well, we find that these seemingly troublesome cases are rather harmless. A by-product of the discussion of measures of randomness is the concept of primitive randomness. The conclusion points out some further problems.

II. PRELIMINARY CONSIDERATIONS

The puzzling character of the notion of randomness can be brought out by means of some intriguing problems which have arisen in its context. Their discussion will lead to preliminary distinctions and clarifications.

Journal of Philosophical Logic **1** (1972) 395–414.

II. 1. *Apparent Self-Contradiction.*

A randomizer is supposed to generate random sequences of certain elements. In repeated tossings of a fair coin, for instance, one expects to get a sequence like

(1) 10110100011...,

where '1' stands for heads, and '0' for tails. Since however, at each toss heads and tails are equally likely to occur, the following sequence is as likely as sequence (1):

(2) 1111111111...

As one would not call this sequence 'random', one would have to say that a randomizer at least sometimes generates a non-random sequence, which seems contradictory. For similar reasons G. Spencer Brown speaks of "the ultimate self-contradiction of the concept of randomness" (1957, p. 57), ultimate, because as the sequences become longer some of the features he considers become more conspicuous.

Obviously, the seeming contradiction evaporates when we distinguish between random generation and random arrangement. Any *order* exhibited by the generated sequence is compatible with the fact that the generating mechanism is a *chance* mechanism.

Besides, sequence (2) can be said to be as likely as sequence (1) only when we confine our comparison to the chances of heads and tails in each single toss. As soon as we consider other features as well, like chances for runs of a certain length, sequence (2) becomes less likely than (1).

II. 2. *The π-puzzle.*

Usually, a sequence is regarded as random if it does not possess any regular pattern, or if the constituent elements occur in it according to no rule or law. The sequence formed by the decimal expansion of the transcendental number π,

(3) 314159265358979323846...,

not only appears to be random, but has also passed all statistical tests for randomness so far applied to it (cf. Pathria, 1962; also Stoneham, 1965). On the other hand, there are well-known mathematical formulas that can be used to compute the numerical value of π. Thus, sequence (3) can be said to be both random and governed by a rule, which seems paradoxical (cf. Venn, 1888, pp. 112f.).

The apparent paradox dissolves on realizing that again two different points of view are involved: the computational and the statistical. The formulas employed to compute the value of π do not consecutively specify each digit, nor can they be used to determine the frequencies of digits or other statistical features of the expansion of π. This is why mathematicians have to undertake what they call 'empirical' studies of this expansion. What the mathematical rules yield is rather a series of decimal expansions, infinite themselves, which converges to the value of π. On the other hand, unless we draw upon our familiarity with the expansion of π, none of the computational rules can be inferred from sequence (3). In general, any statistical structure of the generated sequence can be compatible with the fact that a generating rule yields just one determinate sequence (see, however, Section IV.2).

II. 3. *The Paradox of Random Selection.*

We customarily call the selection of an object from an aggregate of objects 'random' if each object has an equal chance of being chosen. Furthermore, we would think that the chances for an object to be so selected should not depend on how the objects are arranged in the aggregate. Yet, consider the question of the chances that a natural number picked at random will be a prime. When the natural numbers are taken in their natural order, i.e., when the aggregate in question is

(4) 1 2 3 4 5 6 7 8 9 10 11 12...,

the chance, as given by the relative frequency, of picking a prime tends to be zero. However, if we rearrange the natural numbers in the following way:

(5) 1 2 4 3 5 6 7 11 8 13 17 9 ...,

so that the aggregate consists of pairs of primes interspersed with single non-primes, then the chance of getting a prime is $\frac{2}{3}$. Accordingly, depending on the arrangement, the chances that a natural number selected at random will be a prime can vary from zero to one (cf. Russell, 1948, pp. 366f.).

The example shows that it is ambiguous to speak without further qualification of 'random selection' when the selection concerns infinite sets, be they denumerably infinite or not. An instance of the latter case is the familiar 'Bertrand's paradox' (cf., e.g., Kneale, 1949, pp. 184 ff.). The specification of the random selection involves a specification of the selection procedure together with a specification of the relevant set and its ordering (cf. Cannavo, 1966, pp. 136ff.).

II. 4. *The Puzzle of Inferences from Randomness.*

There are many successful applications of probability calculus and its laws to random or chance occurrences. In such applications statements are derived which are well corroborated. But 'random' means lawless and incalculable. How can one draw calculable conclusions from incalculabilities? (cf. Popper, 1968, p. 150)

(This question has several variants. On a subjectivist view, randomness is an expression of our ignorance. This raises the question of how we can conclude anything from what we do not know. Other variants concern ontological rather than logical relationships: How can there be chaos on a molecular level together with order on a molar level? Or, how can random mutations be at the origin of the orderly functioning and the well-adapted behavior of organisms?)

Clearly, one cannot draw conclusions from anything in random occurrences that is not amenable to logical treatment or calculation. Conclusions can be drawn, however, from hypotheses about probabilities or probability distributions, hypotheses which are assumed to hold for the occurrences in question. The statistician, on the other hand, is interested not so much in the random nature of occurences as, e.g., in their stable relative frequencies, which can be compared with hypothesized or calculated probabilities (cf. Kendall, 1941, p. 5).

II. 5. *The Problem of Collectives.*

R. von Mises took probability theory to be an emprical science, the science of 'collectives', i.e., mass phenomena, repetitive events, or potentially indefinitely long sequences of observations, which were to satisfy two conditions (see, e.g., von Mises, 1957). One of them is the principle of convergence, or the so-called limit-axiom. The other is the principle of irregularity, or the axiom of randomness, also called the 'principle of the excluded gambling system'.

The problem of collectives can be taken to include several questions, one being whether the principle of convergence makes any sense with respect to collectives. The main questions, however, have concerned the existence of collectives: Can they be shown, possibly by a constructive proof, to exist mathematically? Or, at least, can it be shown that the two conditions are compatible? Mathematicians, philosophers, and logicians have scrutinizingly dealt with these questions, without finding answers that would have complied with von Mises' original intentions or justified his approach. Since there exists an excellent review of this discussion by P. Martin-Löf (1969), I shall not go into any detail here (see also Section IV.2). With the advent of Kolmogorov's axiomatic approach to probability theory (Kolmogorov, 1933) interest in the problem of collectives has declined considerably. Now it is customary to characterize randomness in terms of probability rather than conversely (see also Section VI.3).

II. 6. *Undefinability of Randomness.*

There are conflicting views concerning the definition of a concept of randomness. R. von Mises' controversial principle of irregularity is a definition of absolute randomness in sequences. Several concepts of restricted randomness have been defined in the context of the problem of collectives. Random distribution and random conjunction of properties will be defined and discussed below. However, there are mathematicians and philosophers who have various doubts with regard to definitions of randomness. It has been said that it seems impossible to give a precise definition of what we mean by 'random', and that the sense of this term is better conveyed by examples, like tosses of a coin, combinations of genes, and death occurrences (Cramér, 1946, pp. 138ff.); or, that the concept of randomness is not extensionally definable, because it does not refer to particular elements but to the generating process for sequences (Rescher, 1961, p. 8); or, that any attempt at defining disorder in a formal way will lead to a contradiction, and that one has to appeal to something real like the urn model in order to say what disorder or random choice is (Freudenthal, 1968, p. 9).

Yet, nothing seems to preclude definitions of certain types of randomness. They may, of course, not coincide with a preconceived sense of 'randomness'. Illustrating randomness by means of examples cannot be our only resort. What is random in such examples should at least be susceptible to characterization, if not definition. What is true, though, is that a formal definition of absolute randomness, in the sense of disorder, of infinite sequences is impossible (see Section IV.2). But in this case, examples are of no avail either.

III. RANDOM DISTRIBUTION IN CLASSES AND SEQUENCES

In the following, I shall sketch a theory of random distribution drawing upon ideas elaborated by Kendall (1941), von Wright (1951, esp. ch.8), and Cannavo (1966). It possesses several features showing R. von Mises' influence. Its basic idea can easily be conveyed by an example. Take a well-specified group of people, e.g., the registered students of a college. Some of them have black hair, others do not. Now select from the group all those with blue eyes. If, in the selected subgroup, the proportion of students with black hair is the same as in the original group, the property of having black hair will be said to be randomly distributed in the original group, relative to the property of having blue eyes.

In greater detail and with some ramifications, the idea can be stated as follows. Let H be a property defining a denumerable class H'; H will be called the 'reference property', and H' the 'reference class'. If H' is finite, its members may be ordered or not. If H' is infinite, let its members be linearly ordered, in a definite way, so as to form a sequence (because of the problems discussed in Section II. 3). Let A be a property which some of the members of H' do have, i.e., $A'.H' \neq \emptyset$, where A' is the class defined by A. We shall call A the 'conjectural property', and A' the 'conjectural class'.

Def. 1. The *relative frequency* of A, given H – $F(A/H)$ for short – is defined as follows:

(a) If H' has a finite number N of members, and if n_A members of H' have property A, then $F(A/H) = n_A/N$.

(b) If H' is infinite; if n_A of the n first members of H' have property A; and if the limit $\lim_{n\to\infty} n_A/n$ exists: then $F(A/H) = \lim_{n\to\infty} n_A/n$.

Def. 2. Let S be a property. S is said to be a *selection* with respect to H' if and only if the subclass of H' consisting of the members of H' which have property S, is not empty, i.e., $H'.S' \neq \emptyset$; and in case H' is infinite, the subclass $H'.S'$ is also infinite.

If M is a material property, like the one in the introductory example, and a selection with respect to H', we shall speak of a 'material selection'. If F is a formal property and a selection with respect to H', we shall call it a 'formal selection'. Formal properties are to be attributed only to members of ordered reference classes, i.e., sequences. Formal selections refer to the places of the elements in a sequence. These places can be uniquely specified by the ordinal numbers of the elements. An example of a formal property is that of being the successor of three elements with the property A, or that of having an ordinal number exactly divisible by seven.

Since the subclass $H'.S'$ can also be considered as a reference class, the relative frequency of A, given $H.S$, i.e., $F(A/H.S)$, is defined according to Def. 1. The same holds for the relative frequency $F(S/H)$, or $F(A.S/H)$, when S, or $A.S$, is regarded as a conjectural property.

Def. 3. A is said to be *randomly distributed* in H' defined by H, *relative* to S, if and only if the relative frequency of A, given $H.S$, is equal to the relative frequency of A, given H; or symbolically

$$RD(A, S, H) \leftrightarrow F(A/H.S) = F(A/H),$$

where $F(A/H.S) = F(A/H) = f$ with $0 \leqslant f \leqslant 1$, according to Def. 1.

We shall speak of 'material randomness' or 'formal randomness' according to whether the selection is a material of a formal one. $\mathrm{RD}(A, S, H)$ can also be expressed by saying that S is irrelevant to the distribution of A in H'. In case $F(A/H.S) > F(A/H)$, S is favorably relevant to the distribution of A in H'; and in case $F(A/H.S) < F(A/H)$, S is unfavorably relevant to the distribution of A in H'.

Def. 3 can be generalized to cases with more than one selection. Let Φ be a family of properties which are selections with respect to H'. A is said to be randomly distributed in H', relative to Φ, if and only if (S) $[S \in \Phi \rightarrow \mathrm{RD}(A, S, H)]$. Φ can be called the 'domain of invariance' of the relative frequency of A, given H. Cannavo presupposes that the members of Φ are logically independent of each other. This condition eliminates redundancies in the domain of invariance, which then comprises only properties taken to be atomic.

Def. 4. The distribution of A in H' is said to be *absolutely random* if and only if there is no property S logically independent of A such that $F(A/H.S)$ is different from $F(A/H)$; or symbolically, with Ψ as the family of all properties which are logically independent of A, and selections with respect to H',

$$\mathrm{ARD}(A, H) \leftrightarrow (S)\,[S \in \Psi \rightarrow \mathrm{RD}(A, S, H)].$$

The requirement that S be logically independent of A, i.e., that S logically entails neither the occurrence nor the non-occurrence of A, is indispensable. For otherwise Def. 4 would be void in all but extreme cases. For example, if S logically entails A, RD(A, S, H) cannot hold except in those extreme cases where all, or none, of the members of H' have property A, that is, $A' \cdot H' = H'$, or $\emptyset$. (However, S may entail A by virtue of a theory holding for the kind of things which have property H, and may have properties S and A. If this is the case, then, by Def. 4, the distribution of A in H' is just not absolutely random.)

IV. DISCUSSION OF THE NOTION OF RANDOM DISTRIBUTION

Although M. G. Kendall, who treats only of formal randomness, speaks of 'a theory of randomness', the theory sketched above is not a theory in the proper sense. It does not contain any specific primitive concept regulated by a specific axiom. The definitions given presuppose nothing beyond ordinary analysis and an algebra of classes. There are some other points which deserve more detailed consideration.

IV. 1. *Random Distribution in Finite Classes and Sequences*

A consequence of Def. 3 is that, roughly speaking, adding an element, whatever its properties, to a finite random sequence destroys the randomness. In greater detail, we have:

Thm. 1. Let K' be a reference class or sequence obtained from H' by adding or subtracting one element (with reference property H), or by replacing one member with one differing in the property A or S. Then, except in four extreme cases, if A is randomly distributed in H', relative to S, it is not so in K', and conversely.

The proof is one of simple arithmetic, using Def. 1 (a). The four extreme cases are those in which $F(A/H) = F(A/K) = 0$ or 1, and $F(S/H) = F(S/K) = 0$ or 1. The reason for the great sensitivity of randomness in finite classes or sequences is, of course, the fact that the relative frequencies are proper fractions, which themselves are very sensitive to such operations on the reference class.

This sensitivity is unobjectionable as long as we are interested in exact proportions, in the composition of definite classes, or in the structure of definite linear arrangements. Adding one element to H' would, in this case, simply mean that we are considering a different class K' defined by a different property K. However, this sensitivity becomes a defect if the defined concept of random distribution is to apply to the randomness in the results of repeated trials like coin tossings. In that case we should assume that randomness once found to be present in the trial sequence is not bound to vanish with the performance of another trial.

Several ways of eliminating this defect may be proposed. (a) The practical-minded statistician will tend to weaken Def. 3 by demanding only that $F(A/H.S)$ and $F(A/H)$ have more or less the same value (Spencer Brown, 1957, p. 52). Although such an idea of approximate randomness can be given a precise formulation (cf. Kolmogorov, 1963), yielding a concept of randomness in a finite sequence different from the one defined above, it seems that it should concern the application of the concept of randomness rather than the concept itself (cf. Kendall, 1941, pp. 7f.). (b) One may take all classes and sequences to be potentially infinite. This case will presently be discussed. (c) One may think of using probabilities instead of relative frequencies for a definition of randomness. This means, however, not so much eliminating the defect mentioned as defining a different concept of randomness.

IV. 2. *Random Distribution in Infinite Sequences*

The condition, in Def. 1(b), that the limit $\lim_{n\to\infty} n_A/n$ exists, restricts the applicability of the notion of random distribution to sequences which possess such limits. For this to be the case, a sequence has to be given by a mathematical rule which specifies each element as a function of its place (ordinal number), whereas the sequences one is interested in when discussing randomness are just those which are not given by such a mathematical rule.

This is one of the shortcomings of R. von Mises' notion of collective. When requiring irregularity of his collectives, von Mises presupposed what we would call 'absolute formal randomness' of the distribution of every conjectural property. More particularly, the objection has been as follows. If there is such a rule as stated above, then it is easy to find a formal property which will yield a subsequence with relative frequencies different from those of the original sequence (cf., e.g., von Wright, 1940, p. 272). Thus, randomness in a collective can never be absolute.

As a way out of this difficulty it has been proposed that one deal with empirical sequences alone (Cannavo, 1966); but this would bring us back to finite sequences. A way of circumventing the difficulty is again to work with probabilities instead of relative frequencies and their limits. Then it is possible, as done in any textbook of probability theory, to define notions of convergence different from the concept of limit in ordinary analysis. These notions, however, do not apply to the frequentist's and empiricist's ideal of a single actual sequence, but only to whole families of sequences of a kind (see also Section VI. 3).

IV. 3. *Extreme Cases*

In the infinite sequence

(6) $AOAOAOAOAOAOAO\ldots,$

for which the limits of the relative frequencies exist, the property A is randomly distributed, relative to an infinite set of formal selections, viz., those which pick out the elements having an ordinal number that is exactly divisible by a given odd number. Furthermore, in the sequence

(7) $AAAAAAAAAAAAAA\ldots$

the distribution of A is absolutely random. This shows that the defined concept of random distribution does not coincide with the ordinary sense of 'randomness' of a sequence, which may or may not be taken as a shortcoming of the definition. What one has to conclude, however, is that the domain of invariance should not be regarded as indicating something like the degree of randomness.

Sequence (7) has already been labeled an 'extreme case', One could exclude this case by requiring $A'.H' \subset H'$, which would be parallel to the requirement $A'.H' \neq \emptyset$ made for conjectural properties in order to exclude the extreme case $A'.H' = \emptyset$.

The same holds for properties serving as selections. In every sequence, the conjectural properties are randomly distributed relative to at least one selection property, viz., the reference property, or one for which $H'.S' =$

H'. One could also exclude this extreme case by requiring $H'.S' \subset H'$, which would be parallel to the requirement $H'.S' \neq \emptyset$ made for selections.

When dealing with finite reference classes, the extreme cases can be equivalently characterized by $F(X/H) = 0$ or 1 (where 'X' stands for A or S). This is not possible with respect to infinite sequences, since $F(X/H) = 0$ (or 1) does not, in this case, imply $H'.X' = \emptyset$ (or H'), although the converse still holds.

IV. 4. *Statistical Independence*

Since the conditions $F(A/H.S) = F(A/H)$ and $F(S/H.A) = F(S/H)$ are equivalent we have:

Thm. 2. If A is randomly distributed in H', relative to S, then S is randomly distributed in H', relative to A, and conversely. Or, symbolically:

$$(A)(S)(H)\,[\mathrm{RD}(A, S, H) \leftrightarrow \mathrm{RD}(S, A, H)].$$

The two equivalent conditions above can be regarded as defining statistical independence of two properties in a reference class (although 'statistical independence' is mostly used as a synonym for 'stochastic independence' which is defined in terms of probabilities). The symmetry stated in Thm. 2 is brought out explicitly by another formulation of statistical independence, viz. $F(A.S/H) = F(A/H).F(S/H)$. In short, random distribution as here defined boils down to statistical independence.

As Thm. 2 holds also for formal selections we will have somewhat strange cases where a formal property is randomly distributed relative to another property. Besides, it has not been excluded that the conjectural property is itself a formal property. Thus, we have allowed for even stranger cases where one formal property is randomly distributed relative to another formal property. In particular, if both formal properties refer to the ordinal number of the elements of an infinite sequence, the randomness of this distribution does not at all depend on the structure of the sequence itself. Such cases are essentially of the same kind as that of sequence (6), as can be seen by taking the conjectural property, in this case, to be that of having an ordinal number which is odd. In sum, it is especially the formal properties and selections, with the concomitant requirement that the reference class be ordered, which lead to rather undesirable features of the notion of random distribution.

V. RANDOM CONJUNCTION OF PROPERTIES

In several places of the foregoing discussion it was suggested that a notion of randomness be defined in terms of probabilities instead of relative frequencies and their limits. In view of the relationship between the relative randomness of a distribution and statistical independence, I shall dedefine random conjunction of properties as stochastic independence. The definition presupposes the calculus of probability, in particular that of conditional probability, in its usual axiomatic formulation; the so-called algebra of events, however, is taken to be an algebra of properties. There will be no reference to formal properties in the sense of the preceding section; we shall no longer be concerned with sequences, nor directly with classes.

Let H be again a reference property, and A, B, ... conjectural properties which may or may not occur together with the reference property H. The non-occurrence of a property A will be designated by '$\bar{A}$'. Let the conditional probabilities $P(A/H)$ – i.e., the probability of A, given H –, $P(A/H.B)$, $P(A.B/H)$, etc. be defined as usual.

Def. 5. A and B are said to be *randomly conjoined* on H if and only if A and B are stochastically independent on H; or symbolically:

$$\mathrm{RC}(A, B, H) \leftrightarrow P(A.B/H) = P(A/H).P(B/H).$$

As can be seen from the relationships between probabilities, the following theorems hold:

Thm. 3. $(A)(B)(H)\,[\mathrm{RC}(A, B, H) \leftrightarrow \mathrm{RC}(B, A, H)]$,

Thm. 4. $(A)(B)(H)\,[\mathrm{RC}(A, B, H) \leftrightarrow \mathrm{RC}(\bar{A}, B, H) \leftrightarrow$
$\leftrightarrow \mathrm{RC}(A, \bar{B}, \mathrm{H}) \leftrightarrow \mathrm{RC}(\bar{A}, \bar{B}, H)]$,

Thm. 5. $(A)(B)(H)\,[P(B/H) \neq 0 \rightarrow (\mathrm{RC}(A, B, H) \leftrightarrow$
$\leftrightarrow P(A/H.B) = P(A/H))]$.

In analogy to Def. 4 we can say that A is absolutely random on H if and only if A is randomly conjoined with any other property logically independent of A. The relation $\mathrm{RC}(A, B, H)$ can also be expressed by saying that A and B are, on H, irrelevant to each other. In case $P(A.B/H) >$ $> P(A/H).P(B/H)$, they are said to be favorably relevant to each other; in case $P(A.B/H) < P(A/H).P(B/H)$, they are unfavorably relevant to each other.

VI. REMARKS ON THE NOTION OF RANDOM CONJUNCTION

Corresponding to the definition above, von Wright has defined concepts of relative and absolute chance. He calls $P(A/H)$ the chance, relative to G, that a positive instance of H will be a positive instance of A if $P(A/H.G) = P(A/H)$ (von Wright, 1951, pp. 226f.). There are others who regard stochastic independence as characterizing, or being a criterion of, randomness or chance (cf., e.g., Keynes, 1921, p. 287; Bohm and Schützer, 1955, p. 1039).

VI. 1. *Relation Between Random Distribution and Random Conjunction*

From the discussion in Section II. 1 it is clear that the concept of random distribution concerns randomness in the sense of disorder, whereas that of random conjunction concerns randomness in the sense of chance. Still von Wright claims that "(relative and absolute) random distribution implies (relative and absolute) chance.... An assertion of random distribution is, however, somewhat stronger than an assertion of chance.... 'normally', *i.e.*, this extreme case [the case $P(\bar{G}/H) = 1$] being excluded, (relative and absolute) chance implies random distribution" (1951, pp. 229f.). These claims are based on the unjustified belief that the so-called frequency interpretation of probability provides the means for unambiguously inferring probabilities from relative frequencies and their limits, and, with the exception of extreme cases, also vice versa. The extreme case excluded by von Wright concerns the equivalence of two conditions for random conjunction, as expressed in Thm. 5, rather than the transition from relative frequencies to probabilities, or from chance to random distribution. The actual relation between random distribution and random conjunction is rather methodological than logical. A random distribution may lead us to suppose random conjunction; and a hypothesis of random conjunction may be tested by checking distributions of properties for randomness (cf. also Bunge, 1956).

VI. 2. *Series of Reference Properties*

In analogy to Def. 5, concepts of random conjunction can be defined in other cases of stochastic independence as well, cases which are dealt with in any textbook of probability theory. A simple example is the case of what are called 'repeated trials under identical conditions' (cf. Feller, 1957, pp. 118f.). These form a series of repetitions of the same reference property H, together with the same set of mutually exclusive conjectural properties A_j, and the same conditional probabilities $P(A_j/H)$. Consider, for the sake of brevity, a series of only two instances of this kind; the reference property of the series as a whole will then be the ordered pair (H, H), associated with the conjectural properties (A_j, A_k), and the probabilities $P\{(A_j, A_k)//(H, H)\}$. For this case, we have:

Def. 6. A_j in the first instance and A_k in the second instance are said to be randomly conjoined on (H, H) if and only if $P\{(A_j, A_k)/(H, H)\} = P(A_j/H) . P(A_k/H)$.

One speaks of 'independent trials' if this condition is satisfied by all pairs (A_j, A_k). Thus, the possible outcomes of independent trials are randomly conjoined. As mentioned in the preceding section, this does not mean that the sequence of actual outcomes will exhibit random distribution of conjectural properties.

VI. 3. *Collectives Reconsidered*

With respect to infinite series of independent trials (i.e., not with respect to their outcomes alone), two theorems can be proven which are reminiscent of R. von Mises' conditions for collectives. The one is the familiar strong law of large numbers, which says, roughly speaking, that the possible relative frequencies of conjectural properties in unlimited series of independent trials under identical conditions converge almost always (or, with probability one) to the probabilities of conjectural properties in one trial. The possible relative frequencies are themselves defined in terms of the latter probabilities. The theorem does not concern the behavior of relative frequencies in actual sequences. Furthermore, the notion of convergence almost always is quite different from the concept of limit in ordinary analysis.

The other theorem is a modified version of the principle of the impossibility of a successful gambling system (cf. Feller, 1957, pp. 185f.). The systems considered are essentially such that whether or not a bet is made at a certain trial is stochastically independent of that trial. For this reason, probabilities have to be assigned to the occurrences of bets, which shows that gambling systems of this kind are different from the formal selections considered by von Mises. The theorem itself states that under any such system the successive trials betted at form a series of trials with unchanged probability for success. For all subseries of this kind the law of large numbers holds as well. W. Feller concludes that "the two theorems together describe the fundamental properties of randomness which are inherent in the intuitive notion of probability and whose importance was stressed with special emphasis by von Mises" (1957, p. 191). Aside from the fact that the two theorems are, as I have indicated, fundamentally different from von Mises' principles, the notion of random conjunction is evidently more basic than what is expressed by the two theorems, since both of them are derived under the assumption of stochastic independence.

VI. 4. *Extreme Cases*

The concept of random conjunction also has some apparently undesirable features, analogous to those of the concept of random distribution, which are again linked to extreme cases.

*Thm.*6. If $P(A/H)=0$ or 1, then A is randomly conjoined with any other conjectural property on H(i.e., A is absolutely random on H).

The feeling that this consequence is quite inappropriate might be alleviated by the observation that there is no contradiction involved in saying that A is absolutely random on H, but nevertheless determined by another property, which may be H itself. We shall presently see that there is a more stringent way of rendering the absolute randomness of extreme cases harmless.

Obviously, Thm. 6 implies that every conjectural property is randomly conjoined on H with any property entailed by H, hence with H itself. The latter can be taken to express the common idea that an assumption of some kind of randomness (which will be called 'primitive randomness' below) is implied whenever we speak of probabilities, i.e. not only when we speak of stochastic independence.

VII. MEASURES AND SOME FURTHER NOTIONS OF RANDOMNESS

Some of the measures defined in information theory can be taken as measures of randomness, as concerns both conjunctions of properties on a single reference property and conjunctions of properties in series of reference properties. Only the former case will be considered here.

Moreover, their mathematical properties and interrelations suggest that they be regarded as measuring degrees of randomness, not only in cases of random conjunction of properties, but also in cases where conjectural properties are not randomly conjoined, and even with respect to the probability distributions of single conjectural properties.

I shall keep with the stipulations preceding Def.5. Furthermore, '$\mathscr{A}$' will designate the probability distribution $p_X =_{\text{df}} P(X/H)$ associated with the variable X, where $X \in \{A, \bar{A}\}$. '$\mathscr{A}^*$' will designate the special case, in which $p_A = p_{\bar{A}} = \frac{1}{2}$. '$\mathscr{B}$' will stand for the probability distribution $p_Y =_{\text{df}} P(Y/H)$ associated with $y \in \{B, \bar{B}\}$. Other abbreviations are $p_{X/Y} =_{\text{df}} P(X/H.\ Y)$, and $p_{XY} =_{\text{df}} P(X.\ Y/H)$, with $p_{X/Y} . p_Y = p_{XY}$. The following notions are defined in information theory (cf. Rényi, 1966, pp. 440ff.).

Defs. 7.
(a) (amount of) information $I(\mathscr{A}) =_{\text{df}} \sum_{X \in \{A, \bar{A}\}} p_X . \log_2 (1/p_X)$
(b) joint information $I((\mathscr{A}, \mathscr{B})) =_{\text{df}} \sum_{X \in \{A, \bar{A}\}} \sum_{Y \in \{B, \bar{B}\}} p_{XY} . \log_2 (1/p_{XY})$
(c) conditional information $I(\mathscr{A}/\mathscr{B}) =_{\text{df}} I((\mathscr{A}, \mathscr{B})) - I(\mathscr{B})$
(d) relative information $I(\mathscr{A}, \mathscr{B}) =_{\text{df}} I(\mathscr{A}) + I(\mathscr{B}) - I((\mathscr{A}, \mathscr{B}))$.

All of these functions are non-negative. Def. 7(a) entails that $I(\mathscr{A}^*) = 1$. The following theorems, proved in information theory (cf. Rényi, 1966, pp. 447ff.), are of particular interest in the present context.

Thms. 7. For every $\mathscr{A}$ and $\mathscr{B}$:

(i) $I(\mathscr{A}) \leqslant I(\mathscr{A}^*)$
(ii.a) $I((\mathscr{A}, \mathscr{B})) \leqslant I(\mathscr{A}) + I(\mathscr{B})$
(ii.b) $I(\mathscr{A}/\mathscr{B}) \leqslant I(\mathscr{A})$
(ii.c) $0 \leqslant I(\mathscr{A}, \mathscr{B})$
(iii) $I(\mathscr{A}, \mathscr{B}) \leqslant \min(I(\mathscr{A}), I(\mathscr{B}))$.

VII.1. *Random Conjunction of Properties Redefined.*

In Thms. 7(ii), the equality signs hold precisely in the special case when $\mathscr{A}$ and $\mathscr{B}$ are stochastically independent on H, i.e., by virtue of Thm. 4, when A and B are randomly conjoined on H. Hence we can replace Def.5 by any one of the following definitions.

Defs. 5'. (a) $\text{RC}(A, B, H) \leftrightarrow I((\mathscr{A}, \mathscr{B})) = I(\mathscr{A}) + I(\mathscr{B})$
(b) $\text{RC}(A, B, H) \leftrightarrow I(\mathscr{A}/\mathscr{B}) = I(\mathscr{A})$
(c) $\text{RC}(A, B, H) \leftrightarrow I(\mathscr{A}, \mathscr{B}) = 0$.

VII.2. *Conditional and Joint Randomness.*

The range of the conditional information $I(\mathscr{A}/\mathscr{B})$ is given by $0 \leqslant I(\mathscr{A}/\mathscr{B}) \leqslant I(\mathscr{A})$. It takes on its maximum in case $\mathscr{A}$ and $\mathscr{B}$ are stochastically independent, i.e. A and B randomly conjoined. It takes on its minimum when $p_{A/B}=0$ or 1, and $p_{A/\bar{B}}=0$ or 1. (Note that two of these four cases, viz. $p_{A/B}=p_{A/\bar{B}}=0$ or 1, are cases of stochastic independence). $I(\mathscr{A}/\mathscr{B})$ can be said to measure the degree of conditional randomness, i.e., the degree of randomness in $\mathscr{A}$, given $\mathscr{B}$. This means, firstly, that the degree of conditional randomness involved in random conjunctions of properties is measured by $I(\mathscr{A}/\mathscr{B})_{\mathrm{RC}}=I(\mathscr{A})$. Secondly, this implies that we shall also speak of conditional randomness, as measured by $I(\mathscr{A}/\mathscr{B})$, in cases where $\mathscr{A}$ and $\mathscr{B}$ are not stochastically independent. In these cases, the conditional randomness can be said to be only partial, whereas in cases of stochastic independence it may be said to be complete. It should be clear that the distinction between complete and partial conditional randomness is not one in terms of degrees of conditional randomness. The degree of complete conditional randomness, i.e. $I(\mathscr{A}/\mathscr{B})_{\mathrm{RC}}$, can vary itself. Its range is the same as that of $I(\mathscr{A})$, which is given by $0 \leqslant I(\mathscr{A}) \leqslant I(\mathscr{A}^*)$. It is smallest in case $p_{A/B}=0$ or 1, and greatest for $p_{A/B}=\frac{1}{2}$ (where, as in all cases of stochastic independence, $p_{A/B}=p_{A/\bar{B}}=p_A$). In sum, the degree of conditional randomness itself is greatest in case of stochastic independence plus equiprobability.

The joint information $I((\mathscr{A}, \mathscr{B}))$ has properties analogous to those of the conditional information, and can be said to measure the degree of joint randomness. Its range is given by $0 \leqslant I((\mathscr{A}, \mathscr{B})) \leqslant I(\mathscr{A}) + I(\mathscr{B})$. As it is a symmetrical function with respect to $\mathscr{A}$ and $\mathscr{B}$, the degree of complete joint randomness, i.e. $I((\mathscr{A}, \mathscr{B}))_{\mathrm{RC}} = I(\mathscr{A}) + I(\mathscr{B})$, can be considered the most appropriate measure of the randomness involved in random conjunctions of two properties.

If one would not allow for partial randomness, as measured by $I(\mathscr{A}/\mathscr{B})$ or $I((\mathscr{A}, \mathscr{B}))$, then so-called random processes would not be random at all. Since the stages of processes depend on one another, they are, except in limiting cases, not stochastically independent, and hence not randomly conjoined. Nevertheless, processes can exhibit degrees of partial conditional, or joint, randomness.

VII.3. *Stochastic Dependence.*

The relative information $I(\mathscr{A}, \mathscr{B})$, whose range is given by $\min(I(\mathscr{A}), I(\mathscr{B})) \geqslant I(\mathscr{A}, \mathscr{B}) \geqslant 0$, has a peculiar status, because it is minimum in case of stochastic independence, and increases when conditional and joint information decrease. It is a measure of the stochastic dependence of $\mathscr{A}$ and $\mathscr{B}$, or, as it were, of the degree of non-randomness of $\mathscr{A}$ and $\mathscr{B}$ relative to one another. It is maximum when one of the properties A and B is determined by the other. In particular, $I(\mathscr{A}, \mathscr{B}) = I(\mathscr{A})$ if $p_{A/B} = 0$ and $p_{A/\bar{B}} = 1$, or $p_{A/B} = 1$ and $p_{A/\bar{B}} = 0$.

VII.4. *Primitive Randomness.*

The fact that the amount of information occurs in such relationships with the other measures of information as were stated and used in the foregoing suggests that it be also taken as measuring the degree of some kind of randomness. It certainly does so in cases of stochastic independence, where, as we have found, the degree of conditional randomness is measured by $I(\mathscr{A}/\mathscr{B})_{\mathrm{RC}} = I(\mathscr{A})$. But as it has been deemed plausible to speak of randomness also in cases other than those of stochastic independence, we may regard $I(\mathscr{A})$, in general, as measuring the degree of primitive randomness in $\mathscr{A}$. Its range is given by $0 \leqslant I(\mathscr{A}) \leqslant I(\mathscr{A}^*)$. It is minimum for $p_A = 0$ or 1; and it is maximum when the alternatives A and $\bar{A}$ are equiprobable. The latter matches the common view that the occurrence of a property is most random when its occurrence and its non-occurrence are equally likely.

VII.5. *Extreme Cases Reexamined.*

The cases where the degree of primitive randomness is minimum are precisely the extreme cases discussed in Section VI.4. In those cases, the property A was seen to be randomly conjoined with any other property B on H, or absolutely random on H, which seemed to clash with the ordinary understanding of randomness. Now this seeming oddity disappears insofar as, in these cases, the degree of primitive randomness $I(\mathscr{A})$ as well as the degree of conditional randomness $I(\mathscr{A}/\mathscr{B})$ are zero. In short, although the randomness is absolute, it is of vanishing degree. It is only the joint randomness whose degree will, in general, not be zero. However, it reduces, in these cases, to nothing but the primitive randomness in $\mathscr{B}$.

VIII. CONCLUSION

I have discussed a concept of random distribution of properties in classes and defined a concept of random conjunction of properties; I have also discussed measures of various kinds of randomness. In concluding, I shall only mention some further problems which await treatment. Both the concept of random conjunction and the measures of randomness rest upon the notion of probability, which was not explicitly dealt with in this paper. Since, however, assumptions of randomness are frequently brought forward as justifications for working with probabilities, the relationship of randomness and probability should be examined in detail. Another task is to show which concepts of randomness are relevant in the sciences. It seems that the concept of random distribution of properties in classes, though interesting in itself and of relevance to statistics, has no use in theories of the empirical sciences. However, the assumption of randomness in the sense of stochastic independence, very often combined with the notion of randomness in the sense of equiprobability, or maximum primitive randomness, undoubtedly plays a role in various fields of science. I do not know of any case where the measures of randomness are directly employed; the idea of stochastic dependence and that of probabilities other than equiprobabilities are, of course, widely used. Finally, some of the assumptions of random conjunction of properties made in the sciences seem to be justifiable in terms of an actual independence of the properties. This raises the question as to whether, and to what extent, the notion of an actual independence is fundamental to concepts of randomness.

ACKNOWLEDGEMENTS

This paper was written while I was on a leave of absence from Wayne State University and had a grant from the Canada Council. I should like to thank Prof. M. Bunge, director of the Foundations and Philosophy of Science Unit at McGill University, for many fruitful discussions.

McGill University, Montreal[1]

NOTE

[1] Presently at Monteith College, Wayne State University, Detroit, U.S.A.

BIBLIOGRAPHY

Bohm, D. and Schützer, W., 1955, 'The General Statistical Problem in Physics and the Theory of Probability', *Nuovo Cimento Suppl.* **2**, ser. 10, 2nd sem., 1004–47.

Bunge, M., 1956, 'A Critique of the Frequentist Theory of Probability', *Congresso Internacional de Filosofia*, São Paulo, Vol. **3**, 787–92.

Cannavo, S., 1966, 'Extensionality and Randomness in Probability Sequences', *Philosophy of Science* **33**, 134–46.

Cramér, H., 1946, *Mathematical Methods of Statistics*, Princeton University Press, Princeton.

Feller, W., 1957, *An Introduction to Probability Theory and Its Applications*, **1**, John Wiley, New York, London.

Freudenthal, H., 1968, 'Realistic Models in Probability', In I. Lakatos (ed.), *The Problem of Inductive Logic*, North Holland, Amsterdam, pp. 1–14.

Kendall, M. G., 1941, 'A Theory of Randomness', *Biometrika* **32**, 1–15.

Keynes, J. M., 1921, *A Treatise on Probability*, MacMillan, London.

Kneale, W., 1949, *Probability and Induction*, Clarendon, Oxford.

Kolmogoroff, A. N., 1933, *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Springer, Berlin.

Kolmogorov, A. N., 1963, 'On Tables of Random Numbers', *Sankhya* **25**, 369–76.

Martin-Löf, P., 1969, 'The Literature on von Mises Kollectivs Revisited', *Theoria* **35**, 12–37.

von Mises, R., 1957, *Probability, Statistics and Truth*, MacMillan, New York.

Pathria, R. K., 1962, 'A Statistical Study of Randomness Among the First 10000 Digits of π', *Mathematics of Computation* **16**, 188–97.

Popper, K. R., 1968, *The Logic of Scientific Discovery*, Harper Torchbook, New York.

Rényi, A., 1966, *Wahrscheinlichkeitsrechnung*, Deutscher Verlag der Wissenschaften, Berlin.

Rescher, N., 1961, 'The Concept of Randomness', *Theoria* **27**, 1–11.

Russell, B., 1948, *Human Knowledge*, Simon and Schuster, New York.

Spencer Brown, G., 1957, *Probability and Scientific Inference*, Longmans, London, New York, Toronto.

Stoneham, R. C., 1965, 'A Study of 60000 Digits of the Transcendental 'e'', *American Mathematical Monthly* **72**, 483–500.

Venn, J., 1888, *The Logic of Chance*, Chelsea, New York.

von Wright, G. H., 1940, 'On Probability', *Mind* **49**, 265–83.

von Wright, G. H., 1951, *A Treatise on Induction and Probability*, Routledge and Kegan Paul, London.