

A New Approach to Probability Theory with Reference to Statistics and Statistical Physics

Nijalingappa Umakantha

Department of Physics, Karnatak University, Dharwad, India

Email: numakantha131@yahoo.co.in

Received 26 February 2016; accepted 24 May 2016; published 27 May 2016

Copyright © 2016 by author and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

A new approach to probability theory is presented with reference to statistics and statistical physics. At the outset, it is recognized that the “average man” of a population and the “average particle” of a gas are only objects of thought, and not real entities which exist in nature. The concept of average (man) is generalized as a new concept of *represental* (man) whose epistemological status is intermediate between those of the *particular* (the man) and the *universal* (a man). This new concept has become necessary as a result of emergence of statistics as a new branch of human knowledge at the beginning of the nineteenth century. Probability is defined with reference to the represental. The concept of probability is the same in probability theory and in physics. But whereas in statistics the probabilities are *estimated* using random sequences, in statistical physics they are *determined* either by the laws of physics alone or by making use of the laws of probability also. Thus in physics we deal with probability at a more basic level than in statistics. This approach is free from most of the controversies we face at present in interpreting probability theory and quantum mechanics.

Keywords

Probability, Statistics and Statistical Physics, Monadic and Collectivistic Probabilities

1. Introduction

The theory of probability and the theory of quantum mechanics are well developed disciplines. They share two features. Both make use of the concept of probability and both have remained controversial regarding the interpretation of their basic concepts. In probability theory, the main issue has been whether probability, the basic

concept of the theory, is relevant to a single event (such as a single tossing of a coin) or to a large number of events (such as repeated tossings of the coin) [1]-[5]. Similarly in quantum mechanics the main issue has been whether the state function, the basic concept of the theory, is relevant to a single system or to a large number of systems [6]-[13]. The two problems are not just similar but are identical. They are due to an inadequacy in epistemology.

From the days of Plato it is well known in epistemology that particulars and universals form two levels of cognition [14]. For instance, Mr. Jones is a particular man (an object of perception) with whom we can associate a well defined space-time region and well defined properties like height, weight, number of children, etc.; in contrast, a man is a universal (an object of thought) with whom we cannot associate any space-time region and any well defined values of height, weight, number of children, etc. It is only from the beginning of the nineteenth century have scientists been dealing with large amounts of quantitative data relevant to agricultural yield, industrial products, and human population, etc. This led to a *new* branch of knowledge called statistics [15]-[17]. In statistics, we are not concerned with which particular (man) has which particular properties (like height, weight, number of children, etc.) but deal with their average values. Thus statistical analysis involves a *new* level of cognition. As early as 1857 Quetelet introduced the concept of “average man” for dealing with statistical data relevant to a population of men [18]. The average man is an object of thought with whom we *cannot* associate any space time region but *do associate* well defined (average) values of height, weight, number of children, etc. Evidently, the concept of average man belongs to a new level of cognition.

In this article, the concept of average man is given a firmer epistemological status by introducing the concept of represental man and a new approach to probability theory is developed. At the outset a distinction is made between the disordered sequences (to which the laws of probability are not applicable) and the random sequences (for which they are applicable) and the laws of probability are established. A clear distinction is made between monadic probabilities (which are relevant to single systems) and collectivistic probabilities (which are relevant to a large number of systems). From general considerations, it is shown that monadic probabilities should necessarily be independent of time, a result whose importance in quantum mechanics cannot be overemphasized. The role of probability in statistics and in statistical physics is made clear. For the sake of completeness (and the benefit of students of probability theory) the limitations of subjective application of probability to a particular (entity or an event) [19] [20] and also differences between the present approach and von Mises’ approach [21] [22] are mentioned.

2. Disordered Sequences and Random Sequences

For the sake of clarity, we first distinguish between the disordered sequences and the random sequences. Let us choose a long English text (a novel or an epic), ignore all grammatical signs, replace all capital letters by small letters, and write successive letters as a sequence giving each letter its serial number. For instance the sequence may start as follows (taken from Dante’s *The Divine Comedy*, translated into English by D.L.Sayers, Penguin Books, Harmondsworth, 1949):

$$\begin{array}{l} \text{Serial No: } 1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7 \quad 8 \quad 9 \quad 10 \quad 11 \quad 12 \quad 13 \quad \dots \\ \text{Letters: } \quad m \quad i \quad d \quad w \quad a \quad y \quad t \quad h \quad i \quad s \quad w \quad a \quad y \quad \dots \end{array} \quad (S2.1)$$

Thus we get a sequence of letters a, b, c, \dots, z . In this sequence there is no relation between the serial number and the letter recorded at the serial number in the sense that by making explicit use of only the serial number and the corresponding letter of any finite number of entries (the serial numbers being selected in any manner whatsoever), we cannot evolve a procedure (or a mathematical method, or a computer program) which can yield correctly the letters at all the serial numbers.

If in a long segment of this sequence containing N number of letters, the letter a occurs N_a number of times, the quantity $\omega(a) = N_a/N$ is called the relative frequency of the letter a in the segment. Similarly we can find the relative frequencies of all the other 25 letters. Evidently, some letters like e occur more often than the others like z . It is found that if N is sufficiently large, of the order of 10,000 or more, each letter has *approximately* a constant relative frequency independent of the actual number N in the segment and also independent of from which part of the sequence the segment is chosen. A sequence having these two properties is called a disordered sequence. In English literature $\omega(e)$ has the maximum value of 0.13 and $\omega(z)$ the minimum value of 0.001.

Next we may consider pairs of two successive letters, 1^{st} and 2^{nd} ; 3^{rd} and 4^{th} ; and so on; we get the pairs $mi, dw, ay, th, is, \dots$ from (S2.1). If there are N_{ab} number of pairs having the letters ab out of a total number of $N/2$ pairs

in a segment, then $\omega(ab) = N_{ab}/(N/2)$ is called the relative frequency of the pair of letters ab in the segment. We can find the relative frequencies of all the possible pairs of letters. Naturally we may expect some relation to hold good approximately between the relative frequency $\omega(ab)$ and the relative frequencies $\omega(a)$ and $\omega(b)$. Can we expect $\omega(yz)$ to be related to $\omega(y)$ and $\omega(z)$ by the same relation? In English language (as in any other language) words are so formed that they can be pronounced without much difficulty. We see that ab stands for a sound that can easily be pronounced but yz cannot be pronounced even with much difficulty. So we expect $\omega(yz)$ not to be expressible in terms of $\omega(y)$ and $\omega(z)$ by the same relation that expresses $\omega(ab)$ in terms of $\omega(a)$ and $\omega(b)$ even approximately. Or in general, the relative frequencies of all the possible (26×26) pairs of letters cannot be expressed (even approximately) in terms of the relative frequencies of the individual letters in the pair by the same single mathematical relation. Because of these properties the sequence under consideration is said to be a disordered (but not a random) sequence.

Next let us consider a sequence which we would later identify as a random sequence for developing the theory of probability. The theory of probability is normally developed with reference to a (random) sequence obtained by tossing of a coin repeatedly. But many subtle points become clear if human individuals are considered. Let us define the "state of parenthood" as the number of children an adult man has and denote the possible states by C_0, C_1, \dots, C_5 , corresponding respectively to the number of children being 0, 1, ... 5; for the sake of simplicity we have limited the maximum number to 5. Evidently, any particular man (like Mr. Jones) has, over a duration of time, either no children, or 1 child, or 2 children, ..., or 5 children; or in other words, a particular man exists, over a duration of time, in only one of the possible states C_0, C_1, \dots, C_5 .

We may collect the data about the states of parenthood of the adult men in a large population by approaching one particular man after the other in a systematic manner (by going from house to house, street to street, and city to city) and record as follows the states in a sequence giving each entry a serial number:

$$\begin{array}{l} \text{Serial No.: } 1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7 \quad 8 \quad 9 \quad 10 \quad \dots \\ \text{State: } \quad C_2 \quad C_0 \quad C_3 \quad C_3 \quad C_2 \quad C_1 \quad C_0 \quad C_5 \quad C_4 \quad C_1 \quad \dots \end{array} \quad (\text{S2.2})$$

In any homogeneous population where no stigma is attached to men for having a large number or a small number of children, any particular man (like Mr. Jones) would not take into consideration, for selecting his place of residence, the number of children he has or his future neighbors happen to have; similarly, in deciding how many children he should have any particular man residing in a location would not take into account the number of children his neighbors have. Or in other words, occurrence of a state (such as C_2) at a particular serial number (such as 5) in the sequence does not depend on the states at other serial numbers; this is a direct consequence of the absence of any relation between the place of residence of a man and his state of parenthood. Evidently, the sequence S2.2 is a disordered sequence, characterized by a fairly well defined value of the relative frequency $\omega(n)$ of the state C_n where n is 0, 1, ..., 5 corresponding to the states C_0, C_1, \dots, C_5 respectively. Similarly, we may expect groups of successive states such as $C_m C_n, C_l C_m C_n, \dots$, also to have fairly well defined values of the relative frequencies $\omega(mn), \omega(lmn), \dots$.

The sequence (S2.2) is, *by definition*, said to be a random sequence if it satisfies the following conditions: 1) In any long segment of the sequence all the relative frequencies $\omega(n), \omega(mn), \omega(lmn), \dots$, are *approximately* independent of the total number N of the elements in the segment and independent of from which part of the sequence the segment is chosen; 2) the relative frequency $\omega(mn)$ is approximately related to the relative frequencies $\omega(m)$ and $\omega(n)$ by the same mathematical formula for every possible pair of values of m and n ; 3) similarly, $\omega(lmn)$ is approximately related to $\omega(l), \omega(m)$, and $\omega(n)$ by a single formula for every possible values of l, m , and n ; and so on. Here we are *only* accepting that such formulae exist; we are *not* assuming any definite expressions for the formulae.

In the sequence of states under consideration, the state entered against each serial number belongs to one particular man in the population who has one definite state, C_0, C_1, \dots, C_5 . For instance, the state C_2 at the 5th serial number belongs to Mr. Jones who has two children. But in determining the relative frequency $\omega(2)$ of the state C_2 , we do not take cognizance of the fact that the 5th entry belongs to Mr. Jones. We just find the number of entries which correspond to the state C_2 and divide it by the total number of entries in the segment. Thus all the relative frequencies $\omega(n), \omega(mn), \omega(lmn), \dots$ are independent of such details as 1) the number of children different *particular men* in the population have and 2) the *places of residence* of these particular men. *An analysis of the data in a sequence in which (analysis) no cognizance is taken of which particular state belongs to which particular entities, is called statistical analysis.*

There is no unique way of selecting city after city, street after street, and house after house for collecting the

data; the data can be collected in many different ways. Thus we can have many different random sequences of the same states C_0, C_1, \dots, C_5 belonging to the same population. These random sequences are independent of one another in the sense that given one random sequence (*i.e.*, the serial numbers and the corresponding states in the sequence) we cannot derive the other random sequences by using a mathematical method or a computer program. Nevertheless, an important *empirical* fact which makes statistical analysis meaningful at all (as a general method) is that all the random sequences belonging to the same population are statistically equivalent in the sense that all the relative frequencies $\omega(n), \omega(mn), \omega(lmn)$, etc, have approximately the same set of values independent of the random sequence from which they are estimated. All these relative frequencies are characteristic of the population (though they can be thought of only with reference to a random sequence). In statistical analysis we are concerned only with the possible states and their relative frequencies, with no regard for particulars and their space-time regions.

We may rewrite the random sequence (S2.2) in the reverse order (which is equivalent to collecting the data by going from house to house in reverse order). The new sequence (S2.2R) is also a random sequence belonging to the same population. So the relative frequency $\omega(mn)$ has approximately the same value in both the sequences. The group of two ordered states $C_n C_m$ in (S2.2R) corresponds to the group $C_m C_n$ of (S2.2). So, $\omega(nm)$ of (S2.2R) is equal to $\omega(mn)$ of (S2.2) which is approximately equal to $\omega(mn)$ of (S2.2R); that is, in (S2.2R), $\omega(nm)$ is approximately equal to $\omega(mn)$. This should be so in every random sequence which belongs to the population. So the explicit expression giving $\omega(mn)$ in terms of $\omega(m)$ and $\omega(n)$ should be symmetric with respect to interchange of m and n for each pair of values of m and n . Similarly, the expression for $\omega(lmn)$ should be symmetric with respect to interchange of l and n (with fixed m) for all the possible values of l and n ; and so on.

Evidently, the relative frequencies of the states and the groups of states in any long segment of states satisfy the following mathematical equations *exactly*.

$$\Sigma \omega(n) = 1; \Sigma \Sigma \omega(mn) = 1; \Sigma \Sigma \Sigma \omega(lmn) = 1; \dots \tag{2.1}$$

(By way of clarification we may state that these equations are true of a disordered sequence also. The significance of the distinction between the disordered sequences and the random sequences would be mentioned later in Sec.4).

We wish to develop a mathematical theory for dealing with the statistical data characterized by the relative frequencies $\omega(n)$'s, $\omega(mn)$'s, $\omega(lmn)$'s, ..., in spite of the fact that these relative frequencies (evaluated from different long segments of different random sequences, or even from those of the same random sequence) have only approximately the same set of numerical values. Now the problem is: can we develop a (rigorous) *mathematical* theory to deal with quantities which are inherently *approximate*? A little reflection would show that there is nothing new in this problem. For instance, when measured with sufficient accuracy, the different measurements of the diameter D of a (real) solid spherical object lead to *approximately* the same values; this is so with the measurements of the surface area A , as well as of the volume V . Yet, the mathematically rigorous laws of geometry are relevant to these approximate values because the mean values of these quantities are interrelated approximately in accordance with the laws of geometry. In geometry we treat the radius r as an unspecified constant and in applying the laws of geometry to a (real) spherical object we (have to) choose an *appropriate* value of r based on the measured values. Historically, the laws of geometry were formulated first and later applied to spherical objects (like domes of buildings). *But in the present case we begin with approximate relative frequencies and then formulate a mathematically rigorous method for dealing with them.*

We have seen that the relative frequencies $\omega(n)$'s are associated with the states C_n 's as such and are free from references to the individual men in the population. So we can dispense with explicit references to the particular men and to the random sequences, and develop a new method of statistical analysis involving only the states C_n 's and the numerical values of $\omega(n)$'s associated with these states. Now *the most crucial epistemological* question is: if the states C_0, C_1, \dots, C_5 are free from references to the particular men (like Mr.Jones, Mr.Smith and others) in the population, then *whose* states do they denote in the new method of statistical analysis? After all, a state (*i.e.*, an attribute or a set of attributes) cannot exist without some object possessing it. Grin cannot exist without the Cheshire cat!

In order to find a solution to this epistemological problem, let us consider the typical statistical quantity, namely the average or the mean number of children "a man" in the population has. It is given by

$$\bar{n} = \omega(0)0 + \omega(1)1 + \dots + \omega(5)5. \tag{2.2}$$

Here the numbers 0, 1, ..., 5 are the possible numbers of children the particular men in the population can have; and $\omega(0), \omega(1), \dots, \omega(5)$ are respectively the fractions of the total number of the particular men in the population who have these numbers of children. Being a statistical quantity \bar{n} need not be an integer and can have a non-integral value, such as 2.718. Evidently, no particular man, like Mr. Jones, can have 2.718 children. We can only regard \bar{n} as the number of children of the “average man” of the population. The concept of average man was first introduced into statistics by Quetelet as early as 1869 [18]. Historically, this concept was popular among statisticians for many decades but slowly waned as its epistemological significance was not recognized by the philosophers of science.

As no particular man in the population can be the average man, the average man cannot be an object of perception and can only be an object of thought. A man, the universal, is also an object of thought but we cannot say a man (in the population) has 2.718 number of children. As mentioned earlier, the average man of statistics is a new kind of object of thought belonging to the new level of cognition. But the concept of average has limited connotation. For instance, in statistical analysis of the eye color of the men in a population (which may be black, blue, brown, green, grey, etc.) we cannot think of the “average” eye color of the average man; and further, in statistical analysis, apart from the average values we deal with distributions also. So, we introduce a more general concept than the average, called represental.

We develop a new method for dealing with statistical data with reference to the represental. In this method no explicit reference to any random sequence of the states and their relative frequencies is made. We deal with only a set of possible states of the represental and a set of numbers associated with these states. This is achieved in an abstract and mathematically rigorous manner in the theory of probability.

3. Theory of Probability

With reference to the states of parenthood of men in a population, we *define* the represental man as “one man” whose state of parenthood C is given by

$$C = P(0)C_0 + P(1)C_1 + \dots + P(5)C_5. \quad (3.1)$$

That is, the state of parenthood C is C_0 to the relative extent $P(0)$, C_1 to the relative extent $P(1)$, ..., and C_5 to the relative extent $P(5)$. Or in other words, the represental man exists in all the possible states C_0, C_1, \dots, C_5 to the relative extents $P(0), P(1), \dots, P(5)$ respectively; we can as well say (as it is more convenient), all these states coexist in the represental man with respective relative extents. In the theory these relative extents are treated as *unspecified* positive numbers, satisfying *exactly* the basic condition

$$P(0) + P(1) + \dots + P(5) = 1. \quad (3.2)$$

These unspecified numbers $P(n)$'s are best described as the *relative extents* to which the states C_n 's coexist in the represental man. In literature they are called the *probabilities* of the states C_n 's.

The probabilities are associated with the possible states *per se* (of the represental man) and not with the represental man. The represental man exists in all the possible states and hence if under any circumstances the probability of any one of these states is meaningful, then under the same circumstances the probabilities of all the other states are also equally meaningful; the probability of one possible state cannot exist without the probability of every other possible state. In all that follows we refer to these possible states simply as the states.

Some scope for confusion can be avoided if we appreciate the distinction between the concept of probability and the numerical value of probability. Coexistence of the states (in the represental) is the essence of the concept of probability; the relative extents of their coexistence are merely the numerical values of the probabilities. The concept of probability is independent of, and unaffected by the changes in, the numerical values of probabilities. For instance, socio-economic factors affect in a complicated manner the number of children different men in a population would like to have; such factors affect only the numerical values of the probabilities and do not affect the concept (or more aptly, the definition) of probability. Situation here is similar to the distinction between the concept of length and the numerical value of length; when an elastic string is stretched, only the numerical value changes but the concept (or more aptly, the definition) of length is unaffected.

Before going further, we may summarize the relations the states C_n 's bear with respect to a particular man, a man (the universal), and the represental man. Every particular man (an object of perception, like Mr. Smith) has only one definite state of parenthood, either C_0 or $C_1, \dots, \text{or } C_5$; *these states are mutually exclusive in a particu-*

lar man. State of parenthood is relevant (or meaningful with reference) to a man, the universal; but *no definite state can be attributed to a man*, neither C_0 , nor C_1 , ..., nor C_5 . All the states, C_0 , and C_1 , ..., and C_5 are attributed to the represental man (with probabilities $P(0)$, $P(1)$, ..., $P(5)$ respectively); *all the states coexist in the represental man*.

It is important to note that the concept of represental is distinct from the concept of class. A class of men is not a man, but the represental man is "one man". A flock of birds is not a bird, but the represental bird (defined with reference to some property such as weight) is "one bird".

We refer to the probabilities $P(n)$'s of the (basic) states C_n 's as the *primary* probabilities and treat their numerical values as unspecified constants. Other probabilities which are expressed in terms of these, are called the derived probabilities. The relations between the derived probabilities and the primary probabilities are mathematically exact. These relations are governed by the laws of probability theory. These laws should be consistent with the relations (mentioned in Sec.2) the relative frequencies (of the states in a random sequence) approximately satisfy by definition; otherwise the probability theory would have no relevance to random sequences (and hence to statistics). As these laws are well known, we just restate them with reference to the represental man.

1) If we consider the state of parenthood of Mr. Jones, it would be C_2 ; if we consider it again it would again be C_2 . Similarly, if we consider the state of parenthood of the represental man it would be C_0, C_1, \dots, C_5 with probabilities $P(0), P(1), \dots, P(5)$ respectively; in the second consideration also it would be C_0, C_1, \dots, C_5 with the probabilities $P(0), P(1), \dots, P(5)$ respectively. So in two considerations 6×6 pairs of states coexist; they are $C_0C_0, \dots, C_0C_5; C_1C_0, \dots, C_1C_5; \dots; \dots; \dots; C_5C_0, \dots, C_5C_5$ with respective probabilities $P(00), \dots, P(05); P(10), \dots, P(15); \dots; \dots; \dots; P(50), \dots, P(55)$. The derived probability $P(mn)$ of the ordered pair of states C_mC_n depends on the primary probabilities $P(m)$ and $P(n)$. According to the law of product of probabilities, $P(mn) = P(m)P(n)$. Similarly, in three considerations, the probability of the ordered group of states $C_lC_mC_n$ is $P(lmn) = P(l)P(m)P(n)$.

The law of product of probabilities: In two (or three, or four, ...) considerations of the represental, the derived probability of any ordered group of two (or three, or four, ...) states is the product of the primary probabilities of the two (or three, or four, ...) individual states in the ordered group.

2) We may treat the basic states C_0, C_2, C_4 as one single state C_e , and treat the basic states C_1, C_3, C_5 as another single state C_d (depending on the number of children being even or odd). The derived probabilities $P(e)$ and $P(d)$ of the states C_e and C_d are related to primary probabilities by:

$$P(e) = P(0) + P(2) + P(4); \quad P(d) = P(1) + P(3) + P(5). \quad (3.3)$$

The law of addition of probabilities: If some states are treated as a single new state, in one consideration of the represental the derived probability of this new state is the sum of the primary probabilities of the individual states included in the new state.

If the probability of a new state is to be meaningful there should be at least one other state which may be either one of the basic states or another new state. Every possible state (such as C_n) should be either treated as a state or included in only one new state.

3) In some cases we may consider only a few states such as C_0, C_2, C_4 , and ignore the other states. Under these conditions, only these states coexist in the represental man with the derived probabilities

$$P(0/e) = P(0)/P(e); \quad P(2/e) = P(2)/P(e); \quad P(4/e) = P(4)/P(e), \quad (3.4)$$

where the sign $/e$ indicates that we are concerned with only the even states. These are called conditional probabilities (because we are imposing the condition that we are concerned with only some and not all states).

The law of conditional probabilities: When only some states are retained for consideration, in a single consideration of the represental the probability of a retained state is given by the primary probability of the state divided by the sum of the primary probabilities of all the retained states.

Collectivistic and Monadic Probabilities

It is important from the point of view of physics to distinguish between the collectivistic probabilities and the monadic probabilities. If, in a random sequence of states, the successive states belong to different particular entities (or systems), the corresponding probabilities are called collectivistic probabilities. If the successive states

belong to the same entity (or system) at different successive instants of time, the corresponding probabilities are called monadic probabilities. Evidently, to be regarded as a random sequence, the relative frequencies belonging to different segments of the sequence (which correspond to different gross durations of time) should have approximately the same set of values. This means that monadic probabilities should necessarily be independent of time.

Monadic probabilities can be identical with collectivistic probabilities. For instance, if we toss the same dice repeatedly we get a random sequence of the six possible states; and if we toss a large number of identical dice, each once only, we get a random sequence of the same six states, each state belonging to different dice. The two random sequences are statistically equivalent. In the former case we deal with the monadic probabilities of the states in the represental tossing of the given dice and in the latter case with the collectivistic probabilities of the same states in a tossing of the represental dice. Except for this subtle difference in the mode of reference, the two situations are identical.

On the other hand, corresponding to some collectivistic probabilities, there may not exist monadic probabilities at all. For instance, the states of parenthood of men in a population can lead to a random sequence of the possible states C_n 's; but the state of parenthood of any particular man, like Mr. Jones, taken over his lifetime, leads to the regular sequence C_0, C_1, C_2, \dots with only a few terms. So the probabilities associated with the states of parenthood can only be collectivistic and never be monadic.

Failure to recognize the distinction between the collectivistic probabilities and the monadic probabilities has been the cause of much confusion in interpreting quantum mechanics over the last eight decades.

4. Probability Theory and Statistics

The laws of probability theory are relevant *only to* random sequences and *not to* disordered sequences. First we consider the relation between the theory of probability and statistics which deals with random sequences.

The relation between probability theory and statistics is similar in *some* respects to the relation between geometry (a branch of mathematics) and mensuration (an empirical science of measurement). In geometry we treat the radius of a sphere as an unspecified constant and express all the properties of the sphere in terms of the radius; all these relations are mathematically exact. In the case of a (real) solid spherical object existing in nature, the measured values of the diameter (and not the radius *per se* as it cannot readily be measured) would invariably have a spread (irrespective of the accuracy of measurement); similarly, the measured values of the surface area, as well as those of volume, would also have a spread. The mean value of the area as well as that of the volume, would be related only *approximately* to the mean value of the diameter in accordance with the laws of geometry. Approximateness persists irrespective of how the mean value of the diameter is selected. Important point is that the principles of geometry *per se* do not offer any one specific criterion for selecting a numerical value for the radius of a (real) solid spherical object (for applications of the laws of geometry to the real spherical object).

Similarly, in the case of statistical data all the relative frequencies $\omega(n)$'s, $\omega(mn)$'s, $\omega(lmn)$'s, ..., would have a spread of values, irrespective of from which long segment of any of the random sequences they are estimated. The best values of $\omega(mn)$'s, $\omega(lmn)$'s, ..., would only be *approximately* related to the best values of $\omega(n)$'s in accordance with the laws of probability. Approximateness persists irrespective of how these best values of $\omega(n)$'s are selected. Important point is that the principles of probability theory *per se* do not offer any one specific criterion for selecting the numerical values of the primary probabilities (for applications of the laws of probability theory to a random sequence).

Mathematical exactness is an inherent feature of both geometry and probability theory whereas approximateness is an inherent feature of both mensuration and statistics.

In one *crucial* aspect the relation between probability theory and statistics is different from the relation between geometry and mensuration. In both geometry and mensuration we deal with the same quantities: point, line, distance, angle, area, volume, etc. But in the cases of probability theory and statistics, *only* the possible states are common to both. The particulars, mutual exclusion of the possible states in a particular, the sequences of states, randomness, and the relative frequencies belong to the province of statistics, whereas the represental, coexistence of the states in the represental, and the probabilities belong to the province of probability theory. *Thus the concept of randomness is foreign to the probability theory and the concept of probability is foreign to the random sequences* (and hence to statistics). Thus statistics and probability theory are two distinct disciplines.

Their *only* meeting point is that the numerical values of the relative frequencies of the states in a random sequence are *approximately* equal to the numerical values of the corresponding probabilities. Much confusion (and concomitant controversy) can be avoided if we realize these points.

One basic feature of the probability theory is that from the given set of numerical values of the primary probabilities, we cannot construct a *unique* random sequence. This is consistent with the basic feature of statistics that, belonging to the same population, there can be many statistically equivalent random sequences, none unique. That even an unspecified primary probability distribution admits of many statistically equivalent random sequences, plays a vital role in determining the “the state of dynamic equilibrium” of the classical, as well as the quantum, statistical assemblies (as we shall see in Sec. 5).

In the theory of probability *per se* we deal with a set of possible states $S_1, S_2, \dots, S_n, \dots$ and their probabilities $P(1), P(2), \dots, P(n), \dots$, in an abstract manner (the represental being implicit). Neither the identity of the particulars nor their states need be specified; (the numerical values of) the probabilities are treated as unspecified constants. So in applying the laws of probability to any statistical data we should first: 1) identify the reference class (denoted by a universal) to which the large number of particulars under consideration belong, 2) identify all the possible states of these particulars, 3) make sure that the states of the particulars recorded in any spatial or temporal order lead to an actual or a conceptually possible random sequence. 4) And then make use of the numerical values of the primary probabilities evaluated from a random sequence (or from symmetry considerations in simple cases).

It is important to note that statistical analysis can be made of any *disordered* sequence of states also. In such a case we deal with the relative frequencies of the possible states, and those of the groups of ordered states, and their interrelations; these are not expected to conform to the laws of probability even approximately. Because of this, in statistics “mathematical treatment is often applied to statistical data, without any probability consideration.” [22]. Thus statistics, as an independent empirical science, has a much wider range of applications than the theory of probability *per se* warrants.

Evaluation of Primary Probabilities from a Random Sequence

There is no unique method of determining the numerical values of the primary probabilities from a given random sequence of states. Here we just mention a reasonably good method. Let us consider the random sequence (S2.2) relevant to the states of parenthood of particular men in a population. The approximate relations between the primary probability $P(0)$ and the various relative frequencies are given by :

$$\omega(0) = N_0/N \approx P(0), \tag{4.1}$$

$$\omega(00) = N_{00}/(N/2) \approx P(00) = P^2(0), \tag{4.2}$$

$$\omega(000) = N_{000}/(N/3) \approx P(000) = P^3(0), \tag{4.3}$$

$$\omega(0000) = N_{0000}/(N/4) \approx P(0000) = P^4(0), \text{ etc.} \tag{4.4}$$

Now consider the quantity

$$y_N(0) = 1 + (N_0/4) + (2N_{00}/N2!) + (3N_{000}/N3!) + \dots \tag{4.5}$$

$$\approx 1 + P(0) + (P^2(0)/2!) + (P^3(0)/3!) + \dots \tag{4.6}$$

$$\approx e^{P(0)}. \tag{4.7}$$

By counting the numbers $N_0, N_{00}, N_{000}, \dots$ in a segment of S2.2 having N number of elements we can find the individual relative frequencies $\omega(0), \omega(00), \omega(000), \dots$ and evaluate $y_N(0)$ which would be approximately equal to $e^{P(0)}$. As we choose successively larger number of elements in the segment, we may expect $\omega(0), \omega(00), \omega(000), \dots$ to assume more stabilized values and the value of $\ln y_N(0)$ also to assume a more stabilized value which may be accepted as the empirically evaluated value of the probability $P(0)$. This method is sensitive to the presence, as well as to the absence, of the relative frequencies of the groups of ordered states $C_{00}, C_{000}, C_{0000}, \dots$; so, this is better than plotting N_0/N as a function of N . Similarly we can obtain well stabilized values of $P(1), P(2), \dots$. From these we may get a set of values of $P(n)$'s, which satisfy the basic condition that their sum is *ex-*

actly unity.

The plot of $\ln y_N(0)$ versus N does not, of course, lead to a limit in the sense of a mathematical limit because of the approximateness inherent in the statistical data. For the same reason there is no point in trying to obtain a set of more accurate (?) values of $P(n)$'s than is necessary for the purposes of statistical analysis. The question whether probabilities $P(n)$'s have, with reference to a population of particular men, inherently a unique set of values is equivalent to the question whether any real object (like the Earth, a volley ball, a billiards ball, a steel ball in a ball-bearing, a hydrogen atom, a helium nucleus, etc.) which we like to identify as a sphere has inherently a unique radius. No measurable quantity which can have a continuous range of values can have a unique value.

We now show how certain factors can influence primary probabilities without affecting the possible states. Consider a non-regular polyhedron whose unequal faces, $F_1, F_2, \dots, F_n, \dots$ are cut in such a way that the polyhedron can rest on any of these surfaces on a horizontal plane. Let us define as "state of rest" of the polyhedron, the face on which it comes to rest after it has been shaken well in a vessel and tossed on a plane. In any particular tossing the state of rest can only be one of the states $F_1, F_2, \dots, F_n, \dots$. In repeated tossings we get a random sequence of these possible states with respective relative frequencies.

The state of rest of the polyhedron in the represental tossing may be put as

$$F = \sum P(n) F_n, \quad (4.8)$$

where $P(n)$ is the probability of the state F_n . All the states of rest F_n 's coexist with respective probabilities.

Let us tilt the plane by a small angle θ with respect to the horizontal. In repeated tossings we get a random sequence of the same possible states but with a new set of relative frequencies. This random sequence is not statistically equivalent to the previous one (generated with θ equal to 0). The state of rest in the represental tossing may now be put as

$$F' = \sum P'(n) F_n \quad (4.9)$$

where $P'(n)$'s are new probabilities of the (same old) states F_n 's. If θ is sufficiently large, probabilities of some states may become zero.

The way the polyhedron is cut, which determines the faces F_n 's, may be referred to as *the internal physical condition* of the polyhedron. The angle of tilt θ , which determines the relative frequencies, may be referred to as *the external physical condition* specifying the statistical data. Each tossing may be identified as the process of subjecting the polyhedron each time to the external physical condition for generating a random sequence. It is important to note that the external physical condition does not affect the possible states *per se* but only affects the numerical values of the probabilities associated with these states. We shall see presently that this is so in quantum mechanics also.

5. Probability Theory and Statistical Physics

We deal with probability in both probability theory and in statistical physics (which consists of quantum mechanics and classical statistical mechanics). The definition of probability, as well as the laws of probability theory, are common to both. But there are many fundamental differences between probability theory and statistical physics regarding probability. First we consider the differences regarding possible states and the numerical values of the primary probabilities.

In statistics we deal with macroscopic objects whose possible "states" are easily identifiable (like the states of height, eye color, parenthood, etc.) and treat the primary probabilities of these states as unspecified constants. Their numerical values are estimated only by making use of some statistical data. In physics, the laws of physics determine uniquely the possible states of physical systems but there is no unique method of determining the primary probability distributions. 1) In the case of an atomic system (like a *single* hydrogen atom) the monadic probability distributions are determined uniquely by the laws of quantum mechanics (without any reference to probability theory) [23]. 2) In the case of some collectivistic quantum phenomena, the collectivistic probability distributions are determined uniquely by the laws of quantum mechanics (without any reference to probability theory) [23]. 3) In the case of an assembly of classical systems in statistical equilibrium, the collectivistic probability distribution is determined by a combination of the laws of classical mechanics and the laws of probability theory [24]; this is so in the case of atomic systems also [25]. As all these three types have already been dealt

with in detail from the point of view of physics, we now consider them from the point of view of probability theory.

5.1. Monadic Probability Distributions Determined by the Laws Quantum Mechanics

Consider an atomic system consisting of a particle of mass m_0 bound to a time-independent potential $V(\mathbf{r})$, existing over a duration of time in the stationary state n , specified by a complex function $\psi_n(\mathbf{r}, t)$ which is a solution of the Schrödinger equation, a law of physics (which involves m_0 and $V(\mathbf{r})$); here \mathbf{r} is the space coordinate. In this case, the real *time-independent* quantity $\psi_n^*(\mathbf{r})\psi_n(\mathbf{r})d\mathbf{q}$ is the probability $P_n(\mathbf{r})d\mathbf{q}$ that the particle exists within an element of volume $d\mathbf{q}$ at \mathbf{r} at the represental instant of time. As it corresponds to the particle of *the same* system, $P_n(\mathbf{r})$ is a monadic probability distribution. That $P_n(\mathbf{r})$ is independent of time shows that there is consistency between the laws of quantum mechanics and the theory of probability. The importance of this result cannot be overemphasized.

Similarly, using the laws of quantum mechanics, one can derive from $\psi_n(\mathbf{r})$ another complex function b_{np} where the time-independent quantity $b_{np}^*b_{np}d\mathbf{p}$ is the probability $P_n(\mathbf{p})d\mathbf{p}$ that the particle momentum \mathbf{p} lies between \mathbf{p} and $\mathbf{p} + d\mathbf{p}$ at the represental instant of time; this is also a monadic probability distribution relevant to the same system in the same state $\psi_n(\mathbf{r})$.

For proper understanding of *the theory of probability*, it is important to recognize the following subtle points. 1) The laws of physics determine the complex function $\psi_n(\mathbf{r}, t)$ as such, and do not determine the primary probability distribution $P_n(\mathbf{r})$ as such. This is so in the case of $P_n(\mathbf{p})$ also. 2) Existence of the probability distribution $P_n(\mathbf{r})$ only means that there are many successive instants of time t_1, t_2, t_3, \dots at which the particle exists at the positions $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots$, respectively, leading to a random sequence of \mathbf{r} values characterized by the relative frequency distribution $\omega_n(\mathbf{r})$ which is approximately equal to $P_n(\mathbf{r})$. Similarly, the existence of $P_n(\mathbf{p})$ only means that there are many successive instants of time t'_1, t'_2, t'_3, \dots at which the particle momentum is $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \dots$, respectively leading to a random sequence of \mathbf{p} values with $\omega_n(\mathbf{p})$ approximately equal to $P_n(\mathbf{p})$. 3) Both the probability distributions $P_n(\mathbf{r})$ and $P_n(\mathbf{p})$ are derived from the same complex function ψ_n using the laws of quantum mechanics and both are primary probability distributions; but they are relevant to two *distinct* random sequences, one of \mathbf{r} values and the other of \mathbf{p} values of the same particle in the same state $\psi_n(\mathbf{r})$. 4) According to the laws of quantum mechanics, a joint probability distribution for \mathbf{r} and \mathbf{p} does not exist. This means that the set of instants of time t_1, t_2, t_3, \dots and the set t'_1, t'_2, t'_3, \dots do not have a common instant of time. Why this is so and what this signifies are problems of physics. From the point of the theory of probability we have to recognize that the existence of individual probability distributions of two quantities (like \mathbf{r} and \mathbf{p}) belonging to the same entity does not mean that a joint probability distribution for the two quantities should necessarily exist.

5.2. Collectivistic Probability Distributions Determined by the Laws of Quantum Mechanics

When a large number of (particular) identical systems, each initially in the same stationary state ψ_i are individually subjected to an extra potential V_e one after the other, each system would come to exist (after being subjected to V_e) in one of the possible stationary states ψ_n 's, leading to a random sequence of the states ψ_n 's, in which the states ψ_n 's occur with relative frequencies $\omega(n)$'s. These systems (after being subjected to V_e) are said to form a statistical assembly.

With reference to the assembly of systems, we can think of a represental system in the state ψ which is related, according to the principle of superposition of quantum mechanics, to the states ψ_n 's by the relation

$$\psi = \sum a_n \psi_n, \tag{5.1}$$

where the complex quantities a_n 's are uniquely determined by the Schrödinger equation which now involves m_0 , $V(\mathbf{r})$ and V_e . Here m_0 and V determine the possible states ψ_n 's of the represental system; and V_e determines a_n 's; if V_e is a function of time t , so are a_n 's. The potential V_e does not affect the stationary states ψ_n 's but determines only a_n 's associated with these states. We may identify m_0 and $V(\mathbf{r})$ as the internal physical conditions and V_e as the external physical condition. Evidently, the situation here is similar to tossing of a polyhedron, discussed in Sec.4. The real numbers $a_n^*a_n$'s are the primary probabilities $P(n)$'s of the states ψ_n 's, uniquely determined by the laws of physics. Needless to state that $\omega(n)$'s are approximately equal to $P(n)$'s. These are, of course, collectivistic primary probabilities.

The above reasoning is applicable to the following statistical quantum phenomena: 1) Scattering of particles

by a localized short-range potential $V_e(r)$ (or by a pair of such potentials, or a lattice of such potentials); 2) Transition of atomic systems from one possible stationary state ψ_i to another state ψ_n under the influence of a time-dependent potential $V_e(r,t)$; 3) Leakage of particles through a barrier formed by a potential $V_e(r)$; 4) Transmission of neutrons through a neutron interferometer; (v) Spin correlations in the Bohm experiment, etc.

5.3. Probability Distributions Determined by the Laws of Classical Mechanics and the Laws of Probability Theory

For the sake of definiteness let us consider (a gas consisting of) a large number N_0 of identical and independent (particular) particles confined within a large enclosure [24]. Evidently the particles collide with one another exchanging energy and momentum according to the laws of classical mechanics. It is envisaged that as a result of repeated collisions between the particles “the state of dynamic equilibrium” is reached in the sense that fractions of the total number of particles having different energies would remain almost the same at all subsequent instants of time. We refer to it as an assembly of particles in statistical equilibrium. Our object is to find the primary probability distribution $P(E)dE$ that the represental particle exists in the states having energy between E and $E + dE$. This is achieved by making use of the laws of classical mechanics *and* the laws of probability theory; here the possible states are determined by the laws of physics and the probability distribution is determined by the laws of probability.

According to classical physics the state of any particular particle at any instant of time t is specified by its position \mathbf{r} and momentum \mathbf{p} . We have shown that statistical properties are independent of the regions of space within which the individual members of the assembly exist. So, so far as the statistical properties of the assembly of particles are concerned, the possible states of the particles can be specified by momentum \mathbf{p} only (or equivalently by energy E only). This means that statistical properties of a gas are independent of which particular particle exist at which point (within the enclosure of the assembly) at any instant of time.

The laws of classical physics being universal, the exchange of energy and momentum between any two colliding particles in a gas is independent of the space-time region over which the collision takes place. So, there would be no relation between the energy E and the space-time region of each particle after the collision. As a result of this, the E values of the particles in the gas considered at an instant of time in any spatial order, lead to a random sequence; for the same reason, the E values of any single particular particle considered at a large number of well separated successive instants of time also lead to a statistically equivalent random sequence. Thus the origin of randomness in classical statistical mechanics (a branch of physics) is the same as in statistics (a branch of applied mathematics). This justifies the application of probability theory to the assembly of particles.

If, at any instant of time, $N(E)$ is the number of particles having energy between E and $E + dE$, the relative frequency $\omega(E) = N(E)/N_0$ is approximately equal to the probability $P(E)dE$ that the represental particle exists in the states having energy between E and $E + dE$. To begin with we treat $P(E)$ as an unspecified function as in probability theory. Applying the central limit theorem of probability theory to the random sequences of the states of the particles at different instants of time, it can be shown that the number $N(E')$ of particles having energy between E' and $E' + dE$ has a Gaussian distribution where the most probable number $\langle N(E') \rangle$ is *exactly* equal to $P(E')dEN_0$; $\langle N(E') \rangle$ is also the time-averaged time-independent mean number of particles. Our object is to find the distribution of $\langle N(E) \rangle$ as a function of E ; for the sake of simplicity we denote $\langle N(E) \rangle$ by $N(E)$. Imposing the conditions that the total number of particles N_0 is constant and the total energy is conserved, it can be shown that

$$N(E) = 1/\left[\exp(\alpha + \beta E)\right] = N_0 P(E) \quad (5.2)$$

Thus we get a well defined mathematical expression for the probability distribution $P(E)$, where α and β depend on the total number of particles and the total amount of energy in the gas. This is the well known Maxwell-Boltzman distribution. In the case of a (real) gas of atoms or molecules in thermal equilibrium, β is identified as $1/kT$ where k is the Boltzmann constant and T the temperature of the gas.

The above reasoning can be extended to quantum statistical thermodynamics also leading to Bose-Einstein and Fermi-Dirac distributions [25]. In these cases also the possible states are determined by the laws of physics and the probability distributions are determined by the laws of probability theory.

We now mention the other fundamental differences between probability theory and statistical physics regard-

ing probability. 1) In probability theory the possible states are denoted by symbols like C_n 's; they may be just symbols as in the case of the states of the eye color or may denote also numbers as in the case of the states of parenthood. In quantum mechanics, ψ and ψ_n 's are well defined complex functions of r and t , determined by the laws of physics. Thus, in spite of the formal similarity between (4.1) of probability theory and (5.1) of quantum mechanics, the two theories differ fundamentally. 2) The laws of quantum mechanics determine the *complex* functions like ψ and ψ_n 's, and the *complex* numbers like a_n 's; and the squares of their moduli are identified as the primary probabilities. There is no analogous situation in probability theory. 3) All the results of probability theory are essentially probabilities. Though many results of quantum mechanics are probabilities, in yielding well defined values of energy, the square and one Cartesian component of angular momentum, electric and magnetic multipole moments, selection rules, the Bragg angles, etc., quantum mechanics is deterministic and not probabilistic. 4) The probability theory *per se* invariably treats the primary probabilities as unspecified constants, whereas quantum mechanics provides, in many cases, the laws (of nature) which determine uniquely the numerical values of the primary probabilities. Thus, so far as primary probabilities are concerned, quantum mechanics deals with probability *at a more basic level* than the probability theory does.

For the sake of completeness and the benefit of physicists, we may state the following. All the reasoning in Secs.5.1 and 5.2 is based on Einstein's statistical interpretation of quantum mechanics [13] [23] in which a clear distinction is made between the stationary states ψ_n 's of particular systems (akin to particular men) and the superposed state ψ of the represental system (akin to the represental man) relevant to a statistical assembly of particular systems (akin to a population of particular men). In contrast to this, in the conventional interpretation of quantum mechanics based on Bohr's views [6], no such distinction is made between ψ and ψ_n 's. Similarly, the reasoning given in Sec. 5.3 differs fundamentally from that of Gibbs [26] [27] whose interpretation of probability is based on the views prevalent at the beginning of the last century; but the final results are the same.

6. Subjective Application of Probability Theory to a Particular

We have developed the theory of probability with reference to the represental (an object of thought). We now investigate the scope and limits of applying this theory to a particular (an object of perception), both in statistics and in statistical physics.

When we are ignorant of the identity of the man who is coming round a corner, we may refer to him by the universal, a man, by making such a statement as "a man is coming round the corner", though we know definitely that he can only be a particular man (as he is an object of perception). Ignorance of the identity of the man being on our part, such a reference to a particular man by the universal, a man, is subjective. Once we come to know of his identity that he is Mr. Jones, our reference to him by the universal (a man) ceases.

Similarly, when we are ignorant of the state of parenthood of the man who comes round the corner, can we refer to him by the represental and ask ourselves the question: "what is the probability that the state of parenthood of the man coming round the corner is C_1 ?", though we definitely know that he can only be a particular (as he is an object of perception) with only one definite state of parenthood. Again, ignorance being on our part, such a reference to a particular by the represental is subjective.

In such a subjective application of probability theory to a particular man (of whose state we are ignorant), we may regard the numerical values of the probabilities $P(0), P(1), \dots, P(5)$ as our subjective degrees of expectation (DE) of the state of parenthood of the man being C_0, C_1, \dots, C_5 respectively; *all these 6 DE coexist (subjectively) in our mind*. The moment we come to know that the man (is Mr. Jones who) has 2 children, our (subjective) reference to him by the represental ceases and so do all our (subjective) DE of different states. Here *the numerical values* (of DE) are objective (because they have been evaluated) and *the expectations (per se* of the different states) are subjective.

Above reasoning looks innocent but is deceptive. Because of the *misleading* role it has played (purely for historical reasons) in the conceptual development of probability theory, we consider this problem at some length.

First let us consider the simplest case of tossing of a (much harassed) fair coin. Here each tossing is a *particular* event, leading to only one of the two possible states of rest denoted by H and T . Evidently, in repeated tossings we get a monadic random sequence of H 's and T 's. In the *represental* tossing both the states coexist with their respective probabilities, $P(H)$ and $P(T)$, each equal to $1/2$. In the subjective application to any particular tossing, the two DE coexist in our mind from before the tossing of the coin to the instant we come to know of the result. In this case there would be no difficulty in the subjective application of probability theory to a particular (tossing). This is so in the case of a fair dice also. In the case of the tossing of an irregular polyhedron with

many (ten or more) faces, it is doubtful whether the human mind can comprehend simultaneously as many (ten or more) distinct degrees of expectation (even when their numerical values are given).

What is the probability that Mr. Jones who is 70 years of age at present, would survive up to 75? From the old statistical data (relevant to the population to which Jones belongs) we can find instances of a large number of particular 70 year old men who survived (S) and did not survive (D) up to 75. This leads to a random sequence of S 's and D 's. From the sequence we can estimate the values of $P(S)$ and $P(D)$ of "the represental 70-year old man". Can this numerical value of $P(S)$ be accepted as our subjective DE of Mr. Jones surviving up to 75?

In evaluating $P(S)$ we have not taken into account the factors such as the present state of health of Mr. Jones, his ability to bear future medical expenses, the hereditary traits of his family, willingness of his close relatives to take good care of him in old age, his personal habits of smoking and drinking, etc., which crucially affect in a rather complicated manner his longevity. In this case there is no unique reference class to which Mr. Jones may be accepted as belonging and hence no meaningful numerical value of $P(S)$. The subjective application of probability theory to a particular is not valid in this case (and in similar cases).

Let us consider a case in which there is no ambiguity regarding the selection of the reference class and the primary probabilities are uniquely determined by the laws of nature. Radioactive decay of atomic nuclei is a case in point. All Radon nuclei (${}_{86}\text{Rn}^{210}$) belong to a kind referred to as a Radon nucleus (a universal). These nuclei decay through emission of alpha particles to become Polonium nuclei (${}_{84}\text{Po}^{206}$). A particular nucleus in the state Rn at some instant of time t_1 would suddenly decay at a later instant of time t_1' into the state Po. Another particular nucleus in the state Rn at some other instant of time t_2 would suddenly decay at a later instant of time t_2' into the state Po. Thus if we consider a large number of Rn nuclei one after the other, we get a collective random sequence of $t'-t$ values. Radioactive decay is a statistical phenomenon.

According to the law of radioactive decay of physics, if a large number $N(0)$ of nuclei are in the state Rn at time $t = 0$, the number of nuclei $N(t)$ in the same state Rn at time t is approximately $N(t) = N(0)\exp(-0.26t)$ and that in the state Po is, of course, $N(0) - N(t)$, where t is measured in hours. The number 0.26, called the decay constant, is characteristic of the decay. Any particular nucleus is, at time t , either in the state Rn or in the state Po. The represental nucleus exists in both the states Rn and Po with probabilities $P(t) = \exp(-0.26t)$ and $1 - P(t)$ respectively. Typical values of $P(t)$ are: 1 for $t = 0$; 0.77 for $t = 1$; 0.60 for $t = 2$; 0.46 for $t = 3$; 0.35 for $t = 4$; and so on. Here t is measured from the instant at which the represental nucleus is in the state Rn; this is a fundamental feature of nuclear decay.

With this background we may consider the subjective application of probability theory to a particular. Let two physicists, A and B place at 12 noon one particular Radon nucleus in a detector system which records the time t_0 at which the Radon nucleus decays, and go away from the detector system. At 12 noon both of them have the same DE of the nucleus being in the state Rn at different times in future as given by the function $P(t)$; their typical DE are as given above.

Let A alone come back at 1 pm, look at the detector system, and learn that the nucleus has not decayed. Because of this new (objective) knowledge that the nucleus is in the state Rn at 1 pm (*i.e.*, P (at 1 pm) = 1), he chooses $t = 0$ at 1 pm in the equation $P(t) = \exp(-0.26t)$ and reevaluates his new DE. His typical DE now become: 0.77 for 2 pm; 0.6 for 3 pm; 0.46 for 4 pm; and so on. Being unaware of the state of the nucleus at 1 pm, B would continue to hold the same DE as she did at 12 noon; her typical DE would be 0.77 for 1 pm; 0.60 for 2 pm; 0.46 for 3 pm; and so on.

Again, let A alone look at the detector system at 2 pm and learn that the nucleus has decayed into Po at 1:37 pm. With this new (objective) knowledge, he can no longer treat the particular nucleus as the represental nucleus; all his DE become irrelevant and cease to exist in his mind. Again, being unaware of the decay of the nucleus at 1:37 pm, B would continue to hold the same DE as she did at 12 noon. Thus different observers hold at the same time different degrees of expectation depending on their (objective) knowledge regarding the state of the nucleus at different times, though there is absolutely no ambiguity in this case either in choosing the represental nucleus (*i.e.*, the reference class) or in the numerical values of the probabilities. Again, in this case (and in similar cases) the subjective application of probability theory to a particular is invalid. Needless to state that the decay of the particular Po nucleus at 1:37 pm is a natural and spontaneous process, and is unaffected by the presence of the detector system and the two physicists A and B.

The theory of probability does lend itself to *subjective application* to a particular "event" under some special conditions which are satisfied in popular betting games. Such betting games provide too limited a perspective for developing a general theory of probability of universal appeal. It is high time that the theory of probability, a

lofty branch of *pure* mathematics, is rescued from gamblers and Dutch book makers.

7. Represental Approach and von Mises' Approach

Von Mises has developed a mathematically rigorous theory of probability [21] [22]. Von Mises *defines* the probability of a possible state as the *limiting* relative frequency of the state in an infinitely long random sequence of the possible states so that the relative frequencies have well defined numerical values which obey the mathematically rigorous laws of the probability theory. Here what is (at best) a *numerical equality* between the limiting relative frequency and the probability is (erroneously) accepted as the definition of *the concept* of probability as well; the concept and the numerical value are not distinguished. Evidently, the concept of density is not the same as the numerical value of density.

In contrast, in the represental approach we *define* probability of a possible state as the relative extent to which the state *coexists* with all the possible states (in the represental). The coexistence is the essence of the *concept* of probability and the *relative extents* are *merely the numerical values* of the probabilities. Further, for the development of the (mathematical) theory of probability *per se* it is not at all necessary to associate any numerical values with the primary probabilities; they are treated as unspecified constants (just as the radius of a circle in geometry). It is shown that the numerical values of primary probabilities relevant to many atomic phenomena are uniquely determined by the laws of nature without any reference to random sequences. Thus it is not only wrong but even unnecessary *to define* probability with reference to a relative frequency (either in a finite or an infinite random sequence).

Von Mises defines randomness in terms of insensitivity to place selection. Place selection involves deriving from a given infinite sequence of the states A_i 's, each state A_i having a well defined limiting frequency $\omega(i)$, a subsequence by retaining or rejecting, according to a set of rules, an entry at the serial number n depending on the states entered at the preceding serial numbers up to $n-1$. If the states A_i 's retain, in the derived sequence, the same limiting relative frequencies as in the original sequence, then the original sequence is said to be a random sequence. The set of rules is not unique in the sense that there can be many sets of rules.

This procedure for place selection involves retaining or rejecting the *individual* states in a sequence. So, this procedure is insensitive to the presence, as well as to the absence, of the *groups of ordered* states in the sequence; and hence it cannot distinguish between the disordered sequences of the type (S2.1) (to which the laws of probability are not applicable) and the random sequences of the type (S2.2) (to which the laws of probability are applicable). Thus von Mises' criterion for randomness in a sequence is rather unsatisfactory.

In contrast to this, in the represental approach the criteria for randomness are well specified (Sec. 2). Apart from being much simpler than von Mises' criterion of place selection, these criteria distinguish between the disordered sequences and the random sequences.

As he *defines* probability as the relative frequency in a random sequence belonging to a collective, von Mises observes "First the collective and then probability. 'We shall not speak of probability until a collective has been defined' " [22]. In our approach to the theory of probability *per se*, we deal with only a set of possible states and their probabilities in an abstract manner, without any reference to a population (of men) or a statistical assembly (of atomic systems) existing in nature (Sec.3); this is in agreement with the distinctive feature of pure mathematics which invariably does not make references to human knowledge.

Von Mises' definition of probability does not admit of subjective application of probability to a particular at all, whereas the represental approach does under special conditions (Sec. 6).

8. Concluding Remarks

We have recognized that human knowledge is essentially based on three levels of cognition which involve three epistemological concepts: particulars (objects of perception on which alone human experience and knowledge depend), universals (objects of thought which are necessary for forming general conclusions, for communication, and for expressing the laws of nature), and representals (objects of thought which *have now become necessary* for dealing with statistical data). Thus the present approach to probability is based on a firm epistemological foundation.

In this approach, we: 1) distinguish between the disordered sequences and the random sequences; 2) distinguish between the particulars, the universal, and the represental; 3) define probability unambiguously; 4) distinguish between the concept of probability and the numerical value of probability; 5) distinguish between the primary probabilities and the derived probabilities; 6) treat the primary probabilities as unspecified constants 7)

show that the primary probabilities relevant to some atomic phenomena are uniquely determined by the laws of nature, and only in the absence of such laws are they evaluated from the given statistical data; 8) show that the subjective application of the probability theory to a particular is admissible under some special conditions; 9) explicitly recognize that statistics and probability theory are two distinct disciplines; and 10) make explicit that the concept of randomness is foreign to the theory of probability and the concept of probability is foreign to random sequences (and hence to statistics).

Atoms are the simplest type of systems existing in nature. Atomic phenomena offer a much better insight for understanding probability than statistics. So it is desirable that all students of probability should have a basic course in quantum mechanics and all students of physics should have a basic course in probability theory, with the clear understanding that quantum mechanics is not an extension of probability theory, nor probability theory of quantum mechanics.

As early as 1863, Quetelet introduced the concept of *average man* into population statistics. If the epistemological significance of this concept that it involved an intermediate level of cognition (between those of the particulars and the universal) was given due recognition, much of the later controversy in interpreting probability and quantum mechanics would have been avoided. The present approach is essentially due to the “average man”, to whom I am most grateful.

References

- [1] Nagel, E. (1939) *Principles of the Theory of Probability*. University of Chicago Press, Illinois.
- [2] Black, M. (1967) Probability. In: Edwards, P., Ed., *The Encyclopedia of Philosophy*, Vol. 6, Macmillan and Free Press, New York.
- [3] Fine, T.L. (1973) *Theories of Probability*. Academic Press, New York.
- [4] Howson, C. (1995) *British Journal for Philosophy of Science*, **46**, 1-32. <http://dx.doi.org/10.1093/bjps/46.1.1>
- [5] Jaynes, E.T. (2003) *Probability Theory: The Logic of Science*. Cambridge University Press, Cambridge. <http://dx.doi.org/10.1017/CBO9780511790423>
- [6] Bohr, N. (1949) *Discussions with Einstein on Epistemological Problems in Atomic Physics*. In: Schilpp, P.A., Ed., *Albert Einstein: Philosopher Scientist*, Library of Living Philosophers, Evanston. (Reprinted, Harper and Row, 1959)
- [7] Hartle, J.B. (1968) *American Journal of Physics*, **36**, 704. <http://dx.doi.org/10.1119/1.1975096>
- [8] Einstein, A., Podolsky, B. and Rosen, N. (1935) *Physical Review*, **47**, 777.
- [9] Einstein, A. *Autobiographical Notes*.
- [10] Ballentine, L.E. (1970) *Reviews of Modern Physics*, **42**, 358. <http://dx.doi.org/10.1103/RevModPhys.42.358>
- [11] Newton, R.G. (1980) *American Journal of Physics*, **48**, 1029. <http://dx.doi.org/10.1119/1.12299>
- [12] Home, D. and Whitaker, M.A.B. (1992) *Physics Reports*, **210**, 224. [http://dx.doi.org/10.1016/0370-1573\(92\)90088-H](http://dx.doi.org/10.1016/0370-1573(92)90088-H)
- [13] Einstein, A. (1936) *Journal of Franklin Institute*, **221**, 313. [http://dx.doi.org/10.1016/S0016-0032\(36\)91045-1](http://dx.doi.org/10.1016/S0016-0032(36)91045-1)
- [14] Woozley, A.D. (1967) *Universals*. In: Edwards, P., Ed., *The Encyclopedia of Philosophy*, Vol. 8, Macmillan and Free Press, New York.
- [15] Cramer, H. (1946) *Mathematical Methods of Statistics*. Princeton University Press, Princeton.
- [16] Savage, J. (1954) *The Foundations of Statistics*. John Wiley, New York.
- [17] Fisher, R.A. (1973) *Statistical Methods and Statistical Inference*. Hafner Press, New York.
- [18] Quetelet, A. (1869) *Physique Sociale ou sur le Development des Faculties de L’homme*. Brussels.
- [19] Kyberg Jr., H.E. and Smokler, H.E., Eds. (1964) *Studies in Subjective Probability*. John Wiley, New York.
- [20] de Finetti, B. (1974-1975) *Theory of Probability* (Vol. I, 1974; Vol. II, 1975). John Wiley, New York.
- [21] von Mises, R. (1964) *Mathematical Theory of Probability and Statistics*. Academic Press, New York.
- [22] von Mises, R. (1957) *Probability, Statistics and Truth*. Allen and Unwin, London.
- [23] Umakantha, N. (2009) Einstein’s Statistical Quantum Mechanics. *Il Nuovo Cimento*, **124B**, 135-154.
- [24] Umakantha, N. (2011) A New Approach to Classical Statistical Mechanics. *Journal of Modern Physics*, **2**, 1235-1241. <http://dx.doi.org/10.4236/jmp.2011.211153>
- [25] Umakantha, N. (2010) A New Approach to Quantum Statistical Thermodynamics. *Il Nuovo Cimento*, **125B**, 396-406.
- [26] Gibbs, J.W. (1960) *Elementary Principles of Statistical Mechanics*. Dover, New York. (Reprint of 1902 Edition)
- [27] Tolman, R.C. (1938) *The Principles of Statistical Mechanics*. Oxford University Press, Oxford.