Securing the Foundations of Probability Theory

Randolph L. Gerl

Abstract: Several traditional problems in probability theory are discussed and a resolution to them is proposed. The use of probability theory in the study of physical reality is contrasted with its use in pure mathematics and the latter is found to be problematic. The proposed resolution is postulated to work for all physical reality but is inclusive enough to cover many situations in pure mathematics.

Introduction

Probability theory must be self-consistent in all situations since it is so crucial to the functioning of the modern world. As scientists have shown, physical existence is fundamentally quantum in nature and, for quantum theory to be well founded, probability theory must be well founded as well. Current probability theory suffers from some foundational problems, for example the existence of non-measurable sets which do not allow probabilities to be defined and, for another example, issues over finite additivity vs countable additivity. Ideally, it should be possible to define probability measures for all events. Sigma algebras have been put forward to solve some of the problems with infinite sample spaces and this approach has been partially successful. Also, the conceptual basis for probability theory is still debated although in practice existing theory works well; references [1] through [11] touch on some of these issues as do many online and offline publications. As will be discussed, standard probability theory also introduces contradictions. Tragically, nonsensical propositions are common throughout much of the literature on probability theory. For example, false propositions like zero probability events are not impossible or zero probability events happen all the time. These false propositions even appear in some textbooks.

Discussion

The aim of this paper is to provide an axiomatic basis for probability theory that is totally free from these problems. It is useful to start with what is standard probability theory without sigma algebras. This minimalist approach contains some of the basics of standard probability theory:

A probability space is a set {*S, F, P*} where *S* is the set of outcomes, *F* is a subset of the power set of *S*, and *P* is a function, *P*:*F*-->[0, 1], which defines the probability measure on *F*. The function *P* is subject to:

$$(1)\ P(A) \geq P(\textit{the empty set}) = 0, \qquad \textit{For all } A \in F$$

$$(2)\ P\left(\bigcup_{K=1}^{T} A_K\right) = \sum_{K=1}^{T} P(A_K), \qquad \textit{Where all } A_K \textit{ are disjoint } \in F$$

$$(3)\ P(S) = 1$$

Where, $T \in \mathbb{N}$ or $T=\omega$ (the smallest transfinite ordinal). Countable additivity is provided in condition (2) and the issue of defining probability measures on uncountable sets is not yet dealt with. With this minimalist definition of a probability space, it can be shown that it is impossible to randomly choose an integer since there is no uniform probability measure on the integers. The well-known proof for this is a relatively simple proof by contradiction. Proof: assume that there exists a uniform probability measure on the integers and *P* is the probability of selecting any integer. Then, for any *P*>0 condition (2) requires that *P(S)* is infinite and that contradicts condition (3). If *P*=0, then condition (2) requires that *P(S)* is zero and that also contradicts condition (3). Since *P*=0 and *P*>0 are the only options, the original assumption must be false because it results in contradictions. So randomly picking an integer is proven impossible using probability theory because no uniform probability measure exists for the integers. It is important to note that countable additivity was used in the proof.

Similar proofs can be constructed to show that there is no uniform probability measure on the set of natural numbers, rational numbers, prime numbers, odd numbers, etc., and so it is impossible to randomly pick any of them. A proof by contradiction can also be given that shows it is impossible to randomly choose a real number. Proof: use the integers to break up the real number line into half-open unit segments [*k*, *k*+1) for every integer *k*. Assume it is possible to randomly choose some real number *r*. Now, *r* must reside inside one of the half-open unit segments [*k, k*+1) and call the segment it resides in [*j*, *j*+1). But that means the integer that defines that half-open unit segment, *j*, was randomly chosen and that contradicts the fact that it is impossible to randomly choose an integer. So, the original assumption that it is possible to randomly choose a real number must be false. It is impossible to have a uniform probability measure on the real numbers. (*One could imagine defining a measure on the rational interval [0, 1) that assigned probabilities based on ratios of lengths but that would contradict countable additivity.*)

With the minimalist approach used so far in this paper, it can be shown that it is impossible to randomly pick a real number from the half-open unit interval [0, 1). A proof by contradiction requires the axiom of choice. Proof: Partition the half-open unit interval [0, 1) into an uncountable number of partitions, *P*, where *x* and *y* (both members of [0, 1)) are in the same partition if and only if $x - y \in \mathbb{Q}$ (the set of rational numbers). By the axiom of choice, it is possible to pick one member from each partition and let the set P_0 be the set containing those picks. A countable partition of [0, 1), call it *CP*, is defined such that each partition is indexed by each rational number *q* in $\mathbb{Q}^{[0,1)}$ so that [0, 1) is partitioned into the disjoint sets CP_q and let $CP_0=P_0$. CP_q is the set of numbers in [0, 1) which are found by adding *q*, counterclockwise, to all the members in CP_0 where the interval [0, 1) now has the topology of a circle. The sets CP_q are a Vitali partition of [0, 1). Since the countable collection of sets CP_q are indexed by a countable collection of rational numbers, a bijection between those rational numbers and the integers can be established. The sets CP_q can now be indexed by the integers and call that indexing CP_i where *i* is an integer. Now, assume that it is possible to randomly choose a real number in [0, 1) and call that number *r*. Now, *r* must be in one of the sets CP_i and suppose it is in CP_k. But that means the integer, *k*, that indexes the set CP_k was randomly chosen and that contradicts the fact that it is impossible to randomly choose an integer. So, the original assumption that it is possible to randomly choose a real number from the half-open interval [0, 1) must be false. It is impossible to have a uniform probability measure on the interval [0, 1). (*This is intuitively sensible for another reason: there can be no uniform probability measure on the set of rational numbers in the interval [0, 1) since they have a bijection to the integers*).

The traditional approach to probability introduces a contradiction here by then claiming it is possible to randomly select a real number in the interval [0, 1), for example, and it then uses Lebesgue measure theory to calculate the probability that the randomly chosen real number is inside a subinterval of the unit interval by using ratios of the lengths of the two intervals. But in doing this a contradiction has been introduced into the theory: the traditional approach now assumes it is possible to do something that was proven above, with a minimalist probability space, to be impossible. By supplementing the axiomatic conditions (1), (2), and (3) above with a Lebesgue measure option for uncountable sets with various constraints on the sigma algebra, *F*, the theory now becomes self-contradictory. The theory now says it is both possible and impossible to randomly select a real number from, for example, the half-open unit interval [0, 1). What has happened here is that countable additivity cannot be made consistent with the fact that some events are not Lebesgue measurable. Eliminating events that are not Lebesgue measurable and making *F* a strict proper subset of 2^*S* does not get rid of the contradiction. Furthermore, excluding non-Lebesgue measurable sets from *F* excludes valid events. Ideally, probability theory should be able to calculate their probabilities and say how to add up $\aleph_1$ or more probabilities.

So, self-consistency and a solid foundation for probability theory suggests that Lebesgue measure theory should not be used to calculate probabilities for uncountable situations. This means that the traditional approach to probability density functions and continuous random variables must be revised to have a secure foundation for probability theory. In such a secure foundation, probability density functions and continuous random variables must be thought of as continuous approximations for discrete but extremely numerous finite cases. The use of the Lebesgue measure in probability theory was motivated by the desire to exclude troublesome events. But doing probability theory with uncountable sets is not something that applies to the physical world. In the physical world it cannot be claimed that a dart board, for example, contains an infinite number of locations. Physical reality is governed by quantum gravitational physics and the uncertainty principle suggests that the number of locations on a dart board, for example, will be an eigenvalue of some yet to be discovered quantum gravitational operator (and that eigenvalue is large but finite). The fact that a dart hits a location on a dart board is proof that the probability of it hitting there is not zero. The fact that a dart hits at a specific location is proof that physical space cannot be the idealized spaces of pure mathematics. The degrees of freedom for locations on a physical dart board must be large but finite. The same things must be true for all other "continuous" situations in the physical world. This doesn't mean that continuous approximations are of no use when dealing with dart boards, for instance. Those approximations are going to be easier to use than the true discrete quantum nature of such objects where the number of locations on a dart board is some huge quantum number and the probability of selecting any location is some very tiny but non-zero number. Nor can problems like these be solved by introducing infinitesimals into the theory, see [12] for example.

Probability theory should not be formulated for the idealized abstract Platonic world of pure mathematics because it is ultimately founded on statistical observation in the physical world. Historically, probability theory is defined as mathematical modeling of the outcomes of experiments, for example, tossing coins, collecting statistics from observations or games of chance. So, the mathematical models for probability theory should relate to the physical world and not the abstract world of purely mathematical spaces. Throwing darts at the real number line is not anything that can take place in the physical world and so there is no way to do experiments and get data about it. Reformulating probability theory to solve the problems mentioned here, and other problems not mentioned here, should include the following: (a) make the theory self-consistent by removing the Lebesgue measure, (b) expand countable additivity to include uncountable additivity to take over the job that sigma algebras, with their Lebesgue measures, used to do, (c) ensure that it is possible to define probability measures for all events, and (d) ensure that the theory doesn't permit false propositions like probability zero events can happen. Making these changes to the minimal theory given above results in an improved theory and definition for probability spaces:

A probability space is a set $\{S, F, P\}$ where S is the set of outcomes, F is the power set of S, and P is a function, P:F-->[0, 1], which defines the probability measure on F. The function P is subject to:

$$(1)\ P(A) \geq P(\textit{the empty set}) = 0, \qquad \textit{For all } A \in F$$

$$(2)\ P\left(\bigcup_{K=1}^{O_T} A_K\right) = \sum_{K=1}^{O_T} P\left(A_K\right), \qquad \textit{Where all } A_K \textit{ are disjoint } \in F$$

$$(3)\ P(S) = 1$$

All four changes are included here; for (a) and (b) the sets A_K have their indexes well-ordered up to the ordinal O_T (and no longer limited to natural numbers and ω), with $O_T \leq O_S$ and O_S is the unique ordinal isomorphic to S. The ordinal O_S has the same cardinality as S. A probability measure cannot be defined on a proper class because, by definition, the powerset operation is done on sets. So (2) is now uncountable additivity and the sum should be absolutely convergent. For (c) and (d) the event set, F, is now the powerset of S and no longer just a subset of it. A consequence of this improved definition of a probability space is that uncountably infinite sample spaces must have sparse probability measures. That is, in those cases most events have zero probability and those that don't are few and far between.

Conclusion

In this paper, a new definition of a probability space solves various problems that arise in the traditional approach. The removal of Lebesgue measures eliminates the problem of more than a trillion trillion events occurring daily on Earth that always have probability zero. For example, locations on Earth that are struck by photons or the locations of the center of masses of various objects when they come to rest. In those cases, traditional probability theory says that probability zero events are all that ever happen and non-zero probabilities for them never happen. In those cases zero probability is all that there is and that makes probability theory useless for them. This aspect of traditional probability theory is totally contrary to observations and common sense. With this improved approach, events with probability zero cannot happen. An event has probability zero if and only if it is impossible. Events with 100 percent probability must and always do occur. The almost surely and almost never philosophy never made observational sense and is removed from the theory. This formulation of probability spaces should be able to accommodate the yet to be discovered correct quantization of General Relativity and the yet to be discovered unified theory of nature.

REFERENCE:

[1] Finite additivity versus countable additivity: De Finetti and Savage N. H. Bingham (2010)

[2] Statistical Implications of Finitely additive Probability, Joseph B. Kadane, Mark J. Schervish, Teddy Seidenfeld

[3] On Qualitative Probability / sigma-Algebras, C. Villegas

[4] Stanford Encyclopedia of Philosophy, Interpretations of Probability

[5] Interpreting Probability: Controversies and Developments in the Early Twentieth Century: Cambridge University Press

[6] Stanford Encyclopedia of Philosophy: Quantum Logic and Probability Theory

[7] Countable Additivity and Subjective Probability, Jon Williamson

[8] Countable Additivity, Idealization, and Conceptual Realism, Yang Liu

[9] Convergence to the Truth Without Countable Additivity, Michael Nielsen, journal of Philosophical Logic

[10] The Origins of de Finetti's Critique of Countable Additivity, Eugenio Regazzini, University of Pavia

[11] E.T. Jaynes's Solution to the Problem of Countable Additivity, Elliot, Colin Vol. 87, Issue. 1

[12] MIT Open courseware, https://ocw.mit.edu/courses/24-118-paradox-and-infinity-spring-2019/resources/mit24_118s19_lecnote12/

[13] MIT open courseware