

Neglibility in Mathematics and Physics

Cristian S. Calude¹, Solomon Marcus² and Karl Svozil^{1,3}

¹Department of Computer Science The University of Auckland
Private Bag 92019, Auckland, New Zealand

Romania Academy, Calea Victoriei 125, Bucharest, Romania

³Institute for Theoretical Physics, University of Technology Vienna
Wiedner Hauptstraße 8-10/136, A-1040 Vienna, Austria

March 26, 2014

1 Introduction

2 A probabilistic infirmity

Let us start with the imaginary experiment in which we choose a real number between zero and one at random (uniformly), say $1/3$, and then ask some person to guess it. What is the probability that the person guesses it correctly? The answer is zero because the probability for picking any particular number, in particular $1/3$, is equal to zero. It is very unlikely, but not impossible, to guess the correct answer, so “the probability has to be more than zero”; otherwise we have a problem!

Here is an argument showing that the probability p that one correctly guesses any number is zero. We prove this result by showing that p is smaller than any positive real number r , hence it is zero! Call x the number randomly picked. Imagine that the interval $[0,1]$ is painted white. Pick any positive real number $r \leq 1$. Then there is a sub-interval of length r of the interval $[0,1]$ containing x . Imagine that this sub-interval is painted black, so now we have a black strip of length r onto the original white strip, and the chosen number x is in the black strip. What is the probability that the guessed number lands on the black strip? The probability of an event measured on the interval $[0,1]$ is equal to its length, hence the answer is r : indeed, $r = \frac{x}{1}$ is the proportion of the white strip that is covered in black. But in order for the guessed number to equal x , it has to land in the black strip, so the probability p of correctly guessing x cannot be larger than the probability of guessing a number on the black strip, hence $p \leq r$. As r has been arbitrarily chosen, it follows that $p = 0$. The probability of correctly guessing a number in $[0,1]$ is zero as a single point in the interval $[0,1]$ has zero length, yet it is still a non-empty subset of the interval!

Does this result indicate an infirmity of the way we calculate probabilities?

To understand better the phenomenon we ask the following question: what does it mean to pick a real number at random? Being allowed to pick between infinitely many possibilities is an *approximation*. It does not actually occur in nature, but it is a very useful way to think (model) certain physical phenomena.¹ For start, as in the argument above, we compute the probability of a number falling into an interval. For instance, one can ask what the chance is that a random real number is between 0.3332 and 0.3333, which means that it is approximately equal to 0.333. This is clearly not problematic. Next, we compute the probability of hitting a specific number. This probability will be zero for every number, as shown above. However, there is a crucial difference between the two probabilities: the first probability is of “practical” importance because it can be subject to experiment, while the second probability is just “imaginary”, it has no practical meaning.

Still this “explanation” does not alleviate the intuitive feeling that something is not right.

3 Negligible sets

4 From measure to probability

Intuitively, probability is a measure which quantifies the likeliness that an event will occur. The theory of probability based on Kolmogorov’s axioms [5, 6] is a mathematical discipline of great foundational simplicity and broad applicability.

Like all mathematical theories, probability theory is formally presented in terms of abstract concepts that can be considered separately from their meaning. In Kolmogorov’s formulation, sets are interpreted as events and probability is a measure on a class of sets. The terms, manipulated by specific mathematical and logical rules, lead to formal results.

Being based on measure theory, probability theory inherits some properties of the “mother theory”, in particular the unintuitive possibility that an event of probability zero may happen.

A simple such example is the event of flipping—in a fair coin-tossing experiment—only infinitely many heads after some moment in time; that is, the outcome of the experiment is a sequence which ends in infinitely many heads. This event has probability zero, but it is not impossible: in fact, the event contains infinitely many sequences.

From a mathematical point of view, *impossibility is a stronger statement which is not captured by [classical] probability theory*. This obvious deficiency can, in part, be mitigated by observing that *if the probability that an event is greater than zero, then that event exists*. Some mathematical arguments successfully use this method of proof.

If an event is sure, then it will always happen, and no outcome not in this event can possibly occur. If an event is almost sure, i.e. the probability of the event is one, then

¹This is a model, not a representation of what nature “actually is” or “might be”.

outcomes not in this event cannot be excluded, they are theoretically possible. The probability of such an outcome occurring is *smaller than any fixed positive real*, and therefore must be zero. Hence, one cannot definitively say that these outcomes will never occur, but one can assume that *for all practical purposes*, abbreviated, *fapp*—an expression used by physicists—this to be true.

5 Probability, randomness and applicability

For applications, the formal results are interpreted or translated back into the problems of the domain of interest. This seems simple and unequivocal, but in fact it is puzzling and difficult. Actually, this is a serious philosophical problem called the “applicability problem” [10], which according to Kolmogorov [7] is

[t]he problem [...] to describe the reasons why mathematical probability theory can be applied at all to phenomena of the real world.

More precisely, what domain are we describing when we take the formal apparatus of events and probability and interpret it in terms of a physical theory, say statistical mechanics or quantum mechanics? A major difficulty comes from the variety of possible probabilities one can use in a given domain: how can we decide which mathematically possible probability is the “right probability”, i.e. the probability which actually describes what the domain accepts as the “intuitive probability”?

There exist various interpretations of probability that can be grouped into two classes: subjective and objective. Subjective probability is a measure of a degree of belief. The Bayesian probability, arguably the most popular version of subjective probability, uses expert knowledge and experimental data to deduce/compute probabilities. Objectivists use probability to describe some objective or physical state of affairs. The main objective probability is the frequentist probability, according to which the probability of a random event denotes the relative frequency “in the long run” of an experiment’s outcomes, when the experiment is repeated “under the same conditions”. All these approaches—which can be presented in several different forms—have merits as well as drawbacks [12].

The simple identification of probability with frequency proportions in a finite repeat of an experiment is too simple/naive theory for a number of reasons: one is that although frequencies should in some way cluster around probabilities they cannot be equal to them. For example tossing N times a fair coin seldom produces a frequency exactly equal to one-half and it will never lead to this result if N is odd. The standard “solution” of this problem is to consider infinite trials: identifying probability with the frequency in the limit (“in the long run”) is then justified by the laws of large numbers from probability theory. Unfortunately, the relative frequency in the limit is itself problematic in several ways, one being related to the order. Indeed, the limit of a series depends on order of the terms of the series. An even more disturbing fact is related to the distribution of outcomes in

the series. An alternating head/tail sequence of outcomes in the coin-tossing experiment gives exact one-half limit frequency, but would we say in this case that the probability of head outcome in the trial was one-half? Clearly, for a credible application we need some assurance that the outcomes in the trial occurred in some “randomly” way. This requires some understanding and knowledge of what randomness is, a difficult problem. Significant progress in this direction has been made by Richard von Mises, A. Church, A. Wald, P. Martin-Löf, A. N. Kolmogorov and others using a new type of complexity developed in algorithmic information theory, see [2, 3].

Martin-Löf [8] theory of randomness is rooted in the way statisticians reject hypotheses. A statistical hypothesis is rejected with a certain confidence if the corresponding outcome is sufficiently improbable given that the hypothesis were true. This intuition can be rigorously translated into a mathematical algorithmic notion of “Martin-Löf test of randomness” which gives Martin-Löf’s notion of random sequence: a sequence which passes all Martin-Löf tests of randomness. Technically, Martin-Löf’s approach starts by defining a class of negligible sets of sequences—more precisely, sets of sequences which are “constructively” null sets according to Lebesgue probability²—and then defines randomness by taking the complement of the union of all these classes. In this way Martin-Löf random sequences form a class of constructive measure one: in a constructive way, almost every sequence is Martin-Löf random.

Two important facts have been proved, see [2, 3]. First, mathematically the notion of “perfect/true random sequence”, which intuitively would require the absence of any pattern or correlations between successive outcomes is *vacuous*. Secondly, there exist degrees of randomness to which there is no upper limit.

Frequently a set is negligible in one sense, but large in a different sense, for example the set of Martin-Löf random sequence: constructive measure one and constructive meagre.

5.1 Dependence on dimensionality

Negligibility may also be dependent on the notion of dimensionality. Recall, as an example, that the standard Lebesgue measure of the Cantor set is zero – and thus the Cantor set is negligible from that point of view.

Nevertheless, the Cantor set is isomorphic to the unit interval on the binary continuum. If we take the proper Hausdorff measure, however, this artifact is alleviated.

6 Negligibility in physics

Motivated by problems of statistical physics, ergodic theory, see [13, 11], studies dynamical systems with an invariant measure and related problems. Negligible sets abound in this

²That is, they can be covered in a constructive way with rational intervals of arbitrarily small length. In contrast with the classical Lebesgue theory, the union of all constructively null sets is itself a constructively null set.

area. Here are just two examples: a) in an ergodic system the phase average of phase functions equals their infinite time averages, *except*, possibly, for a set of measure zero of initial phase points, b) the set of all ergodic (with respect to Lebesgue measure) automorphisms of the forms a co-meager set in the space of all automorphisms of the unit interval which preserve the Lebesgue measure, [9].

By what right can we assume that negligible sets (sets of measure/probability zero or meagre sets) can be physically ignored? After all, they may be non-empty, even worse, infinite.

One possibility is to use special properties of the negligibility of the set one feels should be ignored, e.g. measure zero sets have no interior points.³ Intuitively, a measure zero set has an intrinsic instability as all its points can be “driven out” from the set at the smallest interference from outside. **Prof. Marcus: can you give other stronger, similar properties of measure zero sets or meagre sets?**

An argument against ignoring—from a physical point of view—negligible sets is the following by preparing an experiment in a state belonging to a negligible set. For example, the quantum experiment in [1]. In fact, in general fixing the value for an experiment will make the system into space of a smaller dimension, hence a measure-zero set.

With pro/against examples for physically ignoring negligible sets, we need a new way to treat the infirmity, the goal of this section.

6.1 Negligible observables

Suppose one is interested in *negligible observables* in the sense that some outcomes or events occur “almost never.” In order to give any formal meaning to this statement one needs to assume that any such observables are embedded in continua.

For the sake of an example, consider, for instance, the following proposition: if one chooses a random element from the continuum interval $[0, 1]$, then chances that this element is rational, or, alternatively, computable, non-random, or algebraic, are negligible.

Claims of whether or not negligible observables are relevant for physics [14] are relative to the set theoretic assumptions entering physical models: with respect to classical continua (such as the real numbers), these considerations may appear meaningful. If, however, physical continua are just modelled by dense sets, say the rationals or vectors formed by Pythagorean triples intersection with the three dimensional unit sphere [4], then negligibility may no longer hold.

As sophisticated as these considerations may appear, they are very relevant for large areas of physics; in particular also for deterministic chaos, where the choice of the initial state, subjected to a deterministic evolution, determines the future type of (non-)random behaviour.

³That is, for each point x in a measure zero set and every open set O containing x , there exist points in the set not contained in O .

7 Non-standard and other approaches

References

- [1] Alastair A. Abbott, Cristian S. Calude, Jonathan Conder, and Karl Svozil. Strong Kochen-Specker theorem and incomputability of quantum randomness. *Physical Review A*, 86:062109, Dec 2012.
- [2] Cristian Calude. *Information and Randomness—An Algorithmic Perspective*. Springer, Berlin, second edition, 2002.
- [3] R. Downey and D. Hirschfeldt. *Algorithmic Randomness and Complexity*. Springer, Berlin, 2007. to appear.
- [4] C. D. Godsil and J. Zaks. Coloring the sphere. University of Waterloo research report CORR 88-12, 1988.
- [5] A. N. Kolmogorov. *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Springer, Berlin, 1933.
- [6] A. N. Kolmogorov. *Foundations of the Theory of Probability*. Chelsea, New York, 1956.
- [7] A. N. Kolmogorov. On logical foundations of probability theory. In Jurii Prokhorov and Kiyosi Ito, editors, *Probability Theory and Mathematical Statistics*, volume 1021 of *Lecture Notes in Mathematics*, pages 1–5. Springer Berlin / Heidelberg, 1983.
- [8] P. Martin-Löf. The definition of random sequences. *Information and Control*, 9:602–619, 1966.
- [9] J. C. Oxtoby and S. M. Ulam. On the existence of a measure invariant under a transformation. *Annals of Mathematics*, 40(3):560–566, 1939.
- [10] Christopher P. Porter. Kolmogorov on the role of randomness in probability theory. preprint, 2014.
- [11] Céar Ernest Silva. *Invitation to Ergodic Theory*. American Mathematical Society, Providence, R.I., 2008.
- [12] Lawrence Sklar. *Physics and Chance*. Cambridge University Press, Cambridge, 1993.
- [13] Peter Walters. *An Introduction to Ergodic Theory*. Springer-Verlag, Heidelberg, 1982.
- [14] Eugene P. Wigner. The unreasonable effectiveness of mathematics in the natural sciences. Richard Courant Lecture delivered at New York University, May 11, 1959. *Communications on Pure and Applied Mathematics*, 13:1–14, 1960.