What We Talk About When We Talk About Probability

Zhiyang Xun

July 3, 2020

> If people do not believe that mathematics is simple, it is only because they do not realize how complicated life is. — *John von Neumann*

Contents

1 Introduction

We have taken the course of probability theory for a whole semester and have seen a lot of beautiful results based on the axiomatic probability theory. For a long time, I have been enchanted by the power of axiomatic systems in mathematics. It is such a miracle to construct a wonderland out of several concise axioms.

However, a question has always been haunting me. What is the relationship between the beautiful axiomatic systems and the real world? I mean, I have never seen a σ-algebra in my life, but I shout loudly to my friends when we play Monopoly: "My probability of winning is greater than yours!" It is weird that we have used words like "probability" since our childhood, but we have never realized what we are referring to exactly.

This paper analyzes the intuition behind probability and tries to anwer the question: "What are we talking about when we talk about probability?" The analysis will be both mathematical and philosphical. We try to show that there is a correspondence between the axiomatic probability theory and the real world. The correspondence is surprisingly nontrivial.

Main Contribution of the Paper: This paper raises three brand new questions to the philosophy of probability and tries to answer them. The questions associate the philosophy of probability with the traditional philosophy of epistemology, ontology, and free will. This association has never been established as far as I have read.

2 Three Philosophical Questions

Here we list three philosphical questions for you to think about.

1: Three questions

1. Does probability exist before human being is evolved?
2. If you say that the probability of getting "heads" for a coin tossing is 0.5, but I insist that it should be 0.6, can we both be right?
3. If there is no determinism in the real world, there is no logical inferrence, thus there is no probability in the real world. However, if determinism is true, then everything is determined, where is probability?

These questions may seem tricial or even ludicrous at first sight. Actually they are not. We will talk about them more detailedly in Section 7.

3 Intuition behind "Probability"

To figure out what "probability" accurately implies, the quickest way might be to look it up in the dictionary. List 2 shows the result. It is not that concise due to the fact that when we use the word "probability", we do not always refer to the same thing.

2: Definition of "probability"

1. the quality or state of being probable
2. something (such as an event or circumstance) that is probable
3. (a) i. the ratio of the number of outcomes in an exhaustive set of equally likely outcomes that produce a given event to the total number of possible outcomes
      ii. the chance that a given event will occur

   (b) a branch of mathematics concerned with the study of probabilities
4. a logical relation between statements such that evidence confirming one confirms the other to some degree

— *Dictionary by Merriam-Webster*

Definition 1 seems to be quite straightforward, but it cannot satisfy us, since it uses the word "probable" to explain "probability". If we dig deeper to find out what "probable" implies, we will get List 3.

3: Definition of "probable"

1. supported by evidence strong enough to establish presumption but not proof
2. establishing a probability
3. likely to be or become true or real

— *Dictionary by Merriam-Webster*

The definition starts to get not that straightforward. In the three explanations of "probable", the "establishing a probability" one cannot be taken by us since we must avoid circular definitions. The other two explanations are counterintuitive. When we say "quality", we mean something that can be measured and be represented by a number. However, can we really measure how strong the support is or how likely something to become real? (Remember we cannot simply say "probability" now!)

Although the quantity is impossible to measure. We do have an intuition in our mind that how likely I will win this game. This intuition comes from my mind but I cannot tell what it really is. If you win, you win; if I win, I win. How can I prove to you that I am more "likely" to win? Now, probability indicates an agent's degree of confidence. It is not an objective quantity, but purely comes from someone's own belief. When I say "There's large possibility that he won't come today.", what I really mean is how confident I am to say he won't come today.

Besides the quantity, let us look at other definitions. Definition 2 is the meaning when we use the word "probability" to represent an event. Definition 3 gives several definitions but the vague words like "chance" or "equally likely" cannot satisfy us. 3.(a).i is an interesting one. It follows from the interpretation given by Laplace's classical definition of probability theory. We will talk about that in Section 5 later.

Actually, both definition 3 and 4 are not that intuitive, those explanations actually come from important philosphical ideas.

4 Where Do Axioms Come From?

Before going any further, we must keep in mind that all the axiomatic systems come from the real world. We have already seen the intuition behind "probability", and mathematicians come up with some axioms based on those intuition to help us analyze the world more easily.

Let us review how Kolmogorov has formalized the world.

4: The Kolmogorov axioms

A σ-algebra F on Ω is called an event space. Let P be a function from F to the real numbers obeying:

1. $P(E) \in \mathbb{R}, P(E) \geq 0$ for every $E \in F$.
2. $P(\Omega) = 1$.
3. Any countable sequence of disjoint sets $E_1, E_2, \ldots$ satisfies

$$P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i).$$

Apparently, no one can associate those formulas with the real world instantly. Thus we need to interpret the world to the probability space. We want to find a way to understand what

$$P(\text{It will rain tomorrow})$$

actually means.

How can we judge whether an interpretation is good for formalization? George Salmon proposed three criteria:

> **Admissibility**. We say that an interpretation of a formal system is admissible if the meanings assigned to the primitive terms in the interpretation transform the formal axioms, and consequently all the theorems, into true statements. A fundamental requirement for probability concepts is to satisfy the mathematical relations specified by the calculus of probability...
> **Ascertainability**. This criterion requires that there be some method by which, in principle at least, we can ascertain values of probabilities. It merely expresses the fact that a concept of probability will be useless if it is impossible in principle to find out what the probabilities are...
> **Applicability**. The force of this criterion is best expressed in Bishop Butler's famous aphorism, "Probability is the very guide of life."...
>
> — George Salmon, *The Foundations of Scientific Inference*

Among them, admissibility and ascertainability seem to be two very clear criteria. They both seem pretty natural to accept. However, let us consider one case: if there is an inadmissible interpretation that has perfect ascertainability and applicability, should we abandon it? Even though they are not compatible with the Kolmogorov system, we should still accept it, because it is the axioms that serve the real world, but not vice versa.

The applicability can be viewed in many different ways. For example, we expect the interpretation can be applicable to making rational decisions. One thing that worths noticing is that all the discussion above did not mention frequency. However, in the real world, almost every time we want to check if the probability is correct, we compare it with the frequency. There is a connection between probability and frequency in the real world, that is why we want our interpretation also applicable to frequency.

Actually, all of these criteria can be generalized to judge interpretation to other axiomatic systems, but discussing interpretation of probability seems to be the most interesting one. Let us see two interpretations now.

5 The Classical Probability

The most natural interpretation for probability is that we can take probability as an objective value. The most famous example for probability is tossing coins. When you toss a fair coin, you feel that there is no difference between heads and tails for a toss. Thus you may claim that

$$P(\text{HEADS}) = P(\text{TAILS}) = \frac{1}{2}.$$

Indeed, in such circumstances our intuition tells us that probability is shared equally among all the possible outcomes, so the probability of an event is simply the fraction of the total number of possibilities in which the event occurs. Based on that idea, Laplace gives a famous definition.

Blaise Pascal (1623 - 1662) was a French mathematician, physicist, inventor, writer and Catholic theologian. He wrote that, "The theory of chances consists in reducing all events of the same kind to a certain number of equally possible cases, that is to say, to cases whose existence we are equally uncertain of, and in determining the number of cases favourable to the event whose probability is sought. The ratio of this number to that of all possible cases is the measure of this probability, which is thus only a fraction whose numerator is the number of favourable cases, and whose denominator is the number of all possible cases."

This definition is not rigorous at all because it did not tell us what "events of the same type" specifically refers to. Intuitively, 'heads' and 'tails' are equally likely outcomes of tossing a fair coin; but if their kind is 'ways the coin could land', then 'edge' should presumably be counted alongside them. What's more, we have no idea how to examine whether two things atr "equally uncertain".

As a result, although it seems that there is an objective "probability" in this interpretation, the probability relies on a rational agent to give an assignment and we cannot persuasively explain why the agent assigned it such a value. In a sense, it is not totally objective.

One example to illustrate that the probability is not well-defined is Bertrand Paradox. Here we give a similar version. A factory produces cubes with side-length between 0 and 1 randomly; what is the probability that a randomly chosen cube has side-length between 0 and 1/2? Alice might answer that the probability is 1/2 because "side-length between 0 and 1/2" is exactly half of "side-length between 0 and 1". However, Bob might say that the probability is 1/4 since the side-length between 0 and 1/2 corresponds to face-area between 0 and 1/4, which is a quarter of "face-area between 0 and 1". They are both right. If I say the probability is 1/8, it also makes sense. However, there should not be so many values for the same probability. Now we have so many values because it is judged by different agents and words like "randomly" or "equally uncertain" highly rely on an agent to interpret it.

Now let us check whether this classical interpretation meets the criteria. The classical probabilities are nonnegative and the probability for universe is 1 , but they are only finitely additive. Thus it is only admissible to a weak version of the Kolmogorov axioms, where σ-algebra turns into a normal algebra.

It is also ascertainable. Classical probability is interesting because it assigns probability to events just based on the judgement of a rational agent. All we need to do is to record the assignment. Thus we have the ascertainability.

From our real-life experience, classical probability has great applicability. However, it is actually hard to connect this definition with the frequency. If we find that the frequency of getting a HEADS is 0.51, how can we use classical probability to explain this? Finding 100 states for tossing a coin is both clumsy and unrealistic. We must say its applicability is limited.

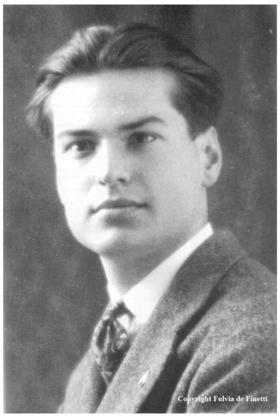
The classical probability is limited in many ways. For example, probability is impossible to be irrational here. A lot of work has been done to perfect the classical interpretation, but the main idea is still the principle of indifference. We subdivide the event into small cases that we are ignorant about the difference. Critics accuse that we cannot extract information from ignorance – if we are ignorant about the cases, how can we assign a probability to it confidently?

6 Subjective Interpretations

The idea of regarding probability as subjective is actually quite natural. If it is hard for us to find objective in the real world, and we can only expect a rational agent to assign probability to events. Why don't we just admit that probability is a purely subjective thing?

> By degree of probability, we really mean, or ought to mean, degree of belief.
>
> — De Mogan

As we have discussed in Section 2, when we say "probability", what we what to express is the degree of belief. Based on the belief, we make decisions that benefit ourselves most. So why don't we use the decision we make to define probability?



Bruno de Finetti (1906-1985) was an Italian probabilist statistician and actuary. He wrote that, " Let us suppose that an individual is obliged to evaluate the rate p at which he would be ready to exchange the possession of an arbitrary sum S (positive or negative) dependent on the occurrence of a given event E, for the possession of the sum pS; we will say by definition that this number p is the measure of the degree of probability attributed by the individual considered to the event E, or, more simply, that p is the probability of E (according to the individual considered; this specification can be implicit if there is no ambiguity)."

Does this new way of interpreting probability meet the criteria? Especially, is it really admissible? Axiom 1 and 2 are easy to check, but for two mutually exclusive event A and B, does $P(A \cup B) = P(A) + P(B)$ always hold?

Suppose in your belief $P(A \cup B) < P(A) + P(B)$. Then I can buy from you a bet on $A \cup B$ for $P(A \cup B)$ units, and sell you back bets on A and B individually. You would be happy to pay $P(A)$ and $P(B)$ respectively. Then you have lost $P(A) + P(B) - P(A \cup B)$ no matter what the result is. In gambling, we say that you are in a Dutch book now.

A **Dutch book** is a set of odds and bets which guarantees a profit, regardless of the outcome of the gamble. It is associated with probabilities implied by the odds not being coherent, namely are being skewed.

It is fascinating that we can prove that a subjective probability is admissible if and only if it will not bring you to a Dutch book.

Although the result is elegant, some critics point out that the utility of a bet may not equal to the expected amount of income. Suppose we have 2 packages. In package A you can get 1 billion for sure; in package B you have a chance of 0.1 to get 20 billion. I believe most of us will choose package A even though the expectation of package B is much higher. To solve this problem, mathematicians try to replace the "quantity" with "utility". However, we need to admit that people are complicated. Perfectly depicting people's choice seems to be impossible.

7 Assocation Between Probability and Philosophy

Let us get back to the three questions we raised in Section 2.

5: Question 1

Does probability exist before human being is evolved?

For the first question, it is similar to the question that whether probability is objective or subjective. If you think that probability is subjective, probability is just people's belief. Then there shouldn't be probability when there is no human.

Even if you believe that probability is objective, the answer might still be no. We can compare probability to other quantities like "mass". Is there mass when there is no human? Most people will say yes, since mass seems always exist in the real worldvwaiting for people to measure; while some will say no, because in a way we did not discover the mass, we invented it. In the real world there are only "objects", human attached the attribute "mass" to them.

George Berkeley(1685-1753) was an Irish philosopher. He asked that "If a tree falls in a forest and no one is around to hear it, does it make a sound?" "But, say you, surely there is nothing easier than for me to imagine trees, for instance, in a park [...] and nobody by to perceive them. [...] The objects of sense exist only when they are perceived; the trees therefore are in the garden [...] no longer than while there is somebody by to perceive them."

Probability is a translation of uncertainty, but "uncertain" also comes from people's mind. If we think that probability's existence relies on human, what about uncertainty? If uncentainty also does not exist, it sounds even weirder.

6: Question 2

If you say that the probability of getting "heads" for a coin tossing is 0.5, but I insist that it should be 0.6, can we both be right?

All the interpretations of probability requires a rational agent to turn the read-world into relations and numbers. Is it possible that we disagree with each other but we are both rational? When we were talking about ascertainability, we wanted to find a way to determinine probability of an event. However, in real life, how can I prove I am the right one and you are the wrong one? Normally, we would use frequency to check probability, but there is no strict independent and identical repetitions in the real world. What if there is a coin that is unfair in the first toss but using some technique to become fair starting from the second toss? We can never find that from observing frequency.

If we view probability as a degree of belief, how can you accuse me because my belief is too strong or too weak? If we can both be right even if we have different assignment, does that mean ascertainability can never be satified?

7: Question 3

If there is no determinism in the real world, there is no logical inferrence, thus there is no probability in the real world. However, if determinism is true, then everything is determined, where is probability?

Determinism is the philosophical view that all events are determined completely by previously existing causes. If you believe whether things will happen are already determined, there should be another zero-one law in probability – if something is doomed to happen, then its probability is 1; if something will not happen, the probability is 0.

Fortunately, if we choose the subjective interpretation, there could still be belief even if things are determined. Nevertheless, if things are determined, do you still believe a rational agent can be rational? From the perspective of an oracle, when a rational agent bets on something that will not happen, it seems totally irrational.

Even if we try to abandon determinism, things are not getting easier, because abandoning determinism is abandoning the universal law of cause and effect. The law of cause and effect is actually a basis for any logical analysis.

David Hume(1711-1776) is a Scottish Enlightenment philosopher, historian, economist, and essayist. He thinks liberty requires determinism. If our actions were not necessitated in the above sense, they would ”have so little in connexion with motives, inclinations and circumstances, that one does not follow with a certain degree of uniformity from the other.”

As Hume said, if we still want to analyze people's actions or use the word "rational", we must admit determinism. This led us to a paradox: in either case of determinism, there is hardly any space for probability.

After discussing the three questions, we got no answer but more questions. It only proves that the philosophical ideas behind probability have so many things to think about. People might say that they completely understand a mathematical theorem, but they will never say they can completely understand a philosophical idea. As Von Neumann said, "If people do not believe that mathematics is simple, it is only because they do not realize how complicated life is."

Reference

- George Salmon, *The Foundations of Scientific Inference*
- *Interpretations of Probability*, https://plato.stanford.edu/entries/probability-interpret/
- *Dictionary by Merriam-Webster*, https://www.merriam-webster.com/
- *Blaise Pascal - Wikipedia*, https://en.wikipedia.org/wiki/Blaise_Pascal
- *Bruno de Finetti - Wikipedia*, https://en.wikipedia.org/wiki/Bruno_de_Finetti
- *George Berkeley - Wikipedia*, https://en.wikipedia.org/wiki/George_Berkeley
- *David Hume - Wikipedia*, https://en.wikipedia.org/wiki/David_Hume