Daniel Michaeli
5. March 2020

Different Interpretations of Probability

The concept of probability is a fundamental part of our universe and its random events. It becomes very apparent when discussing topics such as gambling, weather forecasting and elections, but even events like predicting how well your favorite NBA team will fare during the playoffs or estimating the safety of a particular mode of transportation are all based on probabilites on some level. Everyone understands that by the term "fitfy-fifty" we mean two equally likely possibilites, but what exactly does that mean and how did we come up with that number in the first place? These questions spark a very interesting discussion about how we should interpret the numbers of probability theory in the most intuitive way possible. It turns out that there is no one-size-fits-all solution, and instead we've come up with various ways to understand the numbers that dictate the randomness of the universe. The math itself does work out, but the interpretation debate is more of a philosophical one in hopes of understanding our reality.

Please note that this is not a rigorous or completely accurate description of the interpretations. The intent is not to delve deep into the different interpretations and the math defining them, but rather to introduce different approaches and ways of thinking about the same phenomenon, and how well these different ideas explain specific scenarios.

Classical probability is the study of events where the different outcomes are **equiprobable**, meaning they all have the same likelihood of occuring. Examples of such events are flipping a coin or rolling a dice, with the geometrical structure of the coin and dice providing the equiprobability of the event. Classical probability defines the probability of a specific outcome as the ratio between the amount of different **favorable outcomes** and the total amount of different outcomes. For example, the probability of rolling a six is 1/6, where the 1 represents the only favorable outcome and the 6 represents all possible outcomes. In a coin flip the probability of getting a heads (or tails) is 1/2, or "fifty-fifty", where the 1 represents the favorable outcome and the 2 represents both the heads and tails, i.e. all possible outcomes. This idea makes intuitive sense: increasing the amount of favorable outcomes will increase the likelihood of randomly picking one of them, and therefore increase the probability. But how come this exact fraction will predict the frequency of successful trials for repeated trials with great accuracy? Why is it that, if rolling a dice 10 or 20 times in a row, you will on average roll the number six just about one time for every six rolls? What is the connection between the number of favorable outcomes and the ratio of successful trials?

For these simple events the **frequentist interpretation** comes in handy. This interpretation treats the ratio as a fundamental probability of that particular event. Probabilities for different events can kind of be thought of as numbers deeply rooted in reality itself, in much the same way as constants of nature, like the speed of light in a vacuum or Newton's gravitational constant. It is essentially impossible to determine the specific probability of an event occuring with only one trial: assuming you have never seen a dice roll before or know anything about how gravity works you would not really have any idea what the probability of rolling a six is. By repeating the event and tracking the successful and total outcomes we can define a ratio, a **success frequency** between them **(successful outcomes)/(total outcomes)**. What we find is that as more trials are performed this success frequency gets closer to the probability fraction we mentioned earlier, i.e. the ratio between favorable outcomes and total outcomes. If you rolled a dice 2 times and calculated the **(successful outcomes)/(total outcomes)** ratio it probably would differ quite a bit from the mathematically calculated 1/6 probability, but if you instead rolled it fifty, hundred or a thousand times your ratio would converge to the defined probability with more and more accuracy. The frequentist would argue that the probability of the dice roll, 1/6, is exactly the ratio of **(successful outcomes)/(total outcomes)** given an infinite amount of trials. The mathematically calculated probability of an event is essentially the frequency of success out of an infinitely large amount of repeated trials. This is practically impossible but the theory itself is simple and objectively defined, and models simple events like dice rolls, coin flips and card games really well. The frequentist interpretation shines particularly in gambling where the amount of repeated trials with equiprobable outcomes can be rather large, resulting in a close convergence to the mathematically calculated probability.

That is not to say, however, that this interpretation is completely unchallengeable. First of all, the definition itself lacks mathematical rigor that is paramount in mathematics itself. Unlike a limit of a function, these probability limits can never truly be proven or mathematically calculated, only estimated from the data of a finite amount of trials beforehand. There is no way to ever perform an infinite amount of coin flips, so how can we be 100% sure that the success frequency of getting a heads, given an infinite amount of coin flips, actually will converge to exactly 1/2? What if the coin suddenly started to behave differently after the first 300 flips, and only landed heads up after that? There is no mathematically proven way for us to know that the success frequency will converge to the exact mathematically calculated probability, only well founded assumptions that work well in practice but fail to create a rigorous definition.

Coin flips and blackjack rounds are easy to repeat, but what about one-off events? There are many events with some element of randomness to them that due to circumstances just don't repeat themselves in the same way like a simple coin flip. Despite that we still like to assess probabilities of single-case events like weather forecasts, elections and other events in similar fashion. When saying that "Donald Trump has a 60% chance of winning the 2020 election" or "there is a 10% chance of rain" the frequentist interpretation suddenly breaks down. An election is a very complicated event, where the likelihood of Trump winning depends on a multitude of factors, such as the other candidates in the race, the particular time and the relevant political and economic situation of the country as well as the rest of the world. It can never truly be repeated with the same probabilities, and it certainly isn't an equiprobable event - that would literally be like choosing the winner by spinning a lottery wheel. For events like these we often refer to the **subjective interpretation / Bayesian interpretation** of probability. This interpretation defines probability as a subjective assessment of ones own metric of certainty that an event will occur. After you've flipped a coin and covered it with your hand there exists an uncertainty of whether it landed heads or tails. However, this uncertainty is a "problem" of your own mind, and not a fact about the coin itself. The coin flip already has a determined outcome, regardless of you knowing the answer or not. A frequentist would explain the 1/2 probability with the infinite trials limit, whereas the subjectivist would argue that based on his/her knowledge and information about the world they would assess the coin flip a confidence value of 50% of landing heads. Another example: with 0% being defined as impossible and 100% as guaranteed, by stating that "there is a 10% chance of rain tomorrow" you basically describe your low confidence that it will rain tomorrow. As more information becomes available these confidence levels will update to reflect a new value. If a satellite image suddenly revealed a gigantic dark cloud sweeping across the U.S. the subjectivist's probability of 10% chance of rain would increase to a higher confidence value.

Like I already mentioned this approach of intuitive understanding works really well in single-case events where repititions are impossible. For example, a subjectivist could argue that there is a 50% chance of the human race surviving a nuclear warfare based on subjective confidence and information, while the frequentist wouldn't even be able to assess a probability due to the lack of ever performing a number of trials. The cons of this approach is that subjectivity is often biased and inferior to objective definitions. Also, the frequentist interpretation seems more intuitive and easy to grasp when it comes to describing the more simple, equiprobable events like dice rolls and coin flips.

To summarize: probability can be thought of in multiple ways, that differ from each other quite a bit. None of them are superior or perfect, but they're all very useful for describing different situations in a hopefully somewhat intuitive way in hopes of understanding the meaning behind the numbers that govern stochastic processes.

These interpretations go a lot more in depth, and anyone who wants to learn more about them as well as other interpretations can find great resources on the internet. The main goal of writing this text was to give you some interesting food for thought as well as emphasize the importance of interpretating and trying to grasp what actually lies behind the math itself. This applies to all fields of mathematics, as well as other sciences where calculations and formulas are used. By learning about the underlying principles and derivations you will not only understand what you're doing, but also have an easier time applying the knowledge to different problems in a creative manner.