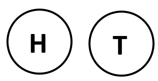
Axioms of Probability Lecture 1:

What is Probability?

Let's begin with the question that you, no doubt have already heard too many times. Toss an ordinary coin. What is the probability that it comes up heads? What does this question even mean?

One might reason as follows: When I toss a coin, one of two potential universes comes into being.



Since there are only 2 possibilities, I can express by 1/2 the idea that one out of two possibilities will happen. This type of thinking is useful but incomplete. For example, you will either live to be 1000 years old or you will not. These two possibilities do not have equal weight and one isn't as likely to happen as the other. Instead, we can think of probability as a measure of nature's preference of a particular outcome. When an experiment is performed a great many times, the proportion of trials that resulted in success divided by the total number of trials could be taken as a measure of this preference. The number 1/2 indicates indifference. 1/2 means that the outcome will be allowed to happen as frequently as it is prevented. Nature simply doesn't care! (Incidentally, when you ask a girl on a date, 1/2 is the number that you should fear! If the probability that she accepts your invitation is very close to 0, then she hates you and hate means that she is not indifferent to you).

Let's run some experiments with the coin! For n tosses, define $n(H)$ to be the number of outcomes that resulted in Heads. The sequence $\frac{n(H)}{n}$ is a non-deterministic sequence. Before running the experiments, the terms of this sequence are unknown. The table below illustrates what has happened when a dime was tossed 20 times on two different occasions, Run 1 and Run 2 respectively.

Trial	Run 1	n(H)/n	Run 2	n(H)/n
1	H	1/1	H	1/1
2	T	1/2	H	2/2
3	T	1/3	T	2/3
4	T	1/4	H	3/4
5	T	1/5	H	4/5
6	T	1/6	H	5/6
7	T	1/7	T	5/7
8	T	1/8	H	6/8
9	T	1/9	H	7/9

10	T	1/10	T	7/10
11	H	2/11	H	8/11
12	H	3/12	T	8/12
13	T	3/13	T	8/13
14	H	4/14	T	8/14
15	H	5/15	T	8/15
16	H	6/16	H	9/16
17	H	7/17	T	9/17
18	T	7/18	H	10/18
19	T	7/19	T	10/19
20	H	8/20	T	10/20

Notice that finitely many trials cannot communicate the whims of the universe precisely. The estimates at trial n do not have to be the same! However, notice that the diversity and unpredictability of the terms in the sequence appears to go down. The rows of the infinite table start to look more and more alike.

Trial	n(H)/n Run 1	n(H)/n Run 2	n(H)/n Run 3	n(H)/n Run 4	n(H)/n Run 5
10	0.5	0.6	0.8	0.7	0.6
100	0.43	0.5	0.58	0.51	0.54
1000	0.49	0.504	0.474	0.513	0.495
10000	0.4985	0.4952	0.5003	0.5028	0.4968

These observations lead us to be more firmly convinced in our initial philosophy: Give nature enough opportunity and she will show you what she prefers.

Informal Definition of Probability:

Let A be an event (simply an outcome that may happen). Then the probability that A occurs is a number $0 \leq P(A) \leq 1$ given by

$$P(A) = \lim_{n \to \infty} \frac{n(A)}{n}$$

where $n(A)$ is the number of trials out of n that resulted in outcome A.

Remark: This means that every time we run the number of trials of this sequence to infinity, the sequence converges to the same number $P(A)$. On the small scale, the fate of the sequence may not be known, but its ultimate fate is determined.

It is hard to wrap one's mind around the concept of probability, partly because in most cases it is impossible to verify the convergence empirically. Even if we could run the trials of the sequence to infinity infinitely many times, it still wouldn't prove that the sequence converges all the time.

Sometimes we can presume to know this limit, because nature's intent appears to be already known:

Ex. You are betting on Boxer A to win against Boxer B, because Boxer A is twice as good a fighter as Boxer B. What is the probability that you will win the bet?

Solution: The statement "Boxer A is twice as good a fighter as Boxer B" means that given many matches $n = n(A) + n(B)$ between the two boxers, $n(A)$ of the matches will be won by Boxer A and $n(B)$ of the matches will be won by Boxer B (assume that there are no ties) with $n(A) \approx 2n(B)$. Thus,

$$P(A) = \lim_{n\to\infty} \frac{n(A)}{n} = \lim_{n\to\infty} \frac{n(A)}{n(A) + n(B)} = \lim_{n\to\infty} \frac{2n(B)}{2n(B) + n(B)} = \frac{2}{3}$$

Ex. Who among you hasn't heard Abba's The Winner Takes it All? There is a line in this song: "The Gods may throw the dice. Their minds as cold as ice. And someone way down here, loses someone dear." What does this mean about the probability of any one outcome?

Solution: The Gods are indifferent. They don't prefer one outcome to the other. If the die is rolled many times $n = n(1) + n(2) + n(3) + n(4) + n(5) + n(6)$ and $n(k)$ is the number of times event k occurs, we would expect $n(1) \approx n(2) \approx n(3) \approx n(4) \approx n(5) \approx n(6)$. Thus

$$P(1) = \lim_{n\to\infty} \frac{n(1)}{n} = \lim_{n\to\infty} \frac{n(1)}{n(1) + n(2) + n(3) + n(4) + n(5) + n(6)}$$

$$= \lim_{n\to\infty} \frac{n(1)}{6n(1)} = \frac{1}{6}$$

Although we will soon describe a mathematically precise definition of probability, the spirit of the subject lives in the running frequency interpretation. It is this interpretation that will guide us in defining such important concepts as conditional probability and expectation. The frequency limit interpretation will also help us make sense of the calculations that we obtain.

Question: If the probability of an event is 0, does this mean the event cannot happen?

Solution: Recalling that probability is a limit, we can easily imagine that the answer is no. Success may occur only finitely many times, making the numerator of the limit of the ratio $\frac{n(E)}{n}$ stop at a finite value while the denominator *n* drags continue to drag this ratio to 0.

For example, suppose you order a package online and the delivery is equally likely to arrive at any time *t* between 12:00 and 13:00. What is the probability that this package will get delivered at time $t = t_0$?

Note that if we divide the time interval $[0, 1]$ into n subintervals, the package is equally likely to arrive in any one of the subintervals $\left[\frac{k}{n}, \frac{k+1}{n}\right]$ for $0 \le k \le n - 1$. It is easy to see that $P\left[\frac{k}{n}, \frac{k+1}{n}\right] = \frac{1}{n}$ and that t_0 belongs in one of these intervals. In particular, $P(t_0) \le \frac{1}{n}$ for any n. Hence $P(t_0) = 0$. In other words, the probability that your package arrives at any particular time is 0. Do you see what is happening?

Question: If the probability of an event is 1, does this mean the event always happens?

Solution: Take the previous delivery problem and select any countable collection of moments in time. The probability that the package will not be delivered on any of the moments of time which you have selected is 1, even though the actual moment of delivery is in some countable collection.

To understand this phenomenon, imagine running the package delivery experiment once and observing that the package was delivered at time t_0 . If you could run this experiment again and again, you will not ever observe the same exact outcome. Events that occur without repetition have 0 probability.

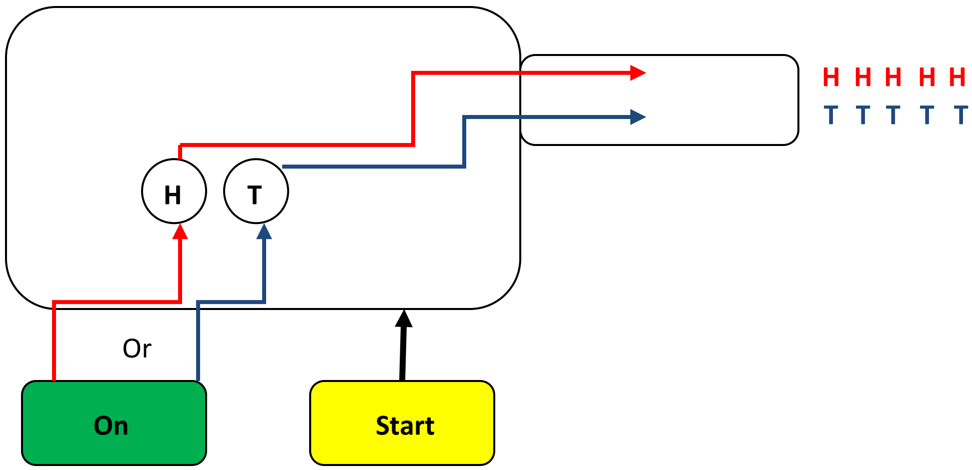
Question: Do there exist events that do not have a probability?

Solution: When we think of probability as a limit of a sequence, we are quickly reminded of divergent limits. Could there be experiments, whose outcomes aren't behaving the same, after we restart the setup?

Imagine the following machine:

As soon as you turn it on, the machine flips a coin and enters either state H or state T. We perform the experiment by clicking the button Start, at which point the machine communicates its current state.

When we try to calculate $\lim_{n\to\infty}\frac{n(H)}{n}$ we sometimes get 1 and sometimes 0. Thus $P(H)$ does not exist.



One may argue that we do not really restart the experiment unless we turn off the machine before pressing Start again, but we also do not restart the universe when we flip a coin. If we have a record of the previous tosses, the conditions in which the new toss is performed are not the same as they were for the first trial. There is no reason why the preference settings would have to remain fixed.

Moreover, when you study measure theory, you become aware of the existence of non-measurable sets. These are the sets that do not satisfy the 3 axioms of probability that we are about to discuss and are therefore qualify as events that do not have a likelihood of occuring.