

Probability – A Review of some basic concepts

Probability is a subject in mathematics that gets covered in high school, usually without much context. In quantum physics it provides an essential perspective on how the microcosm behaves. We will not use much of the mathematics of probability in the course, except in a qualitative way. So the following is just intended to refresh your memories on some basic concepts and to introduce the notion of probability distributions.

The Probability of a single event (often called the *a priori* probability) is the relative frequency for that event to occur in a large number of trials (or the number of occurrences in large set of trials divided by the number of trials).

Examples:

Suppose there is a population in which 501,000 boys are born out of 1 million births -> $501,000/1,000,000 = 0.501$ or 50.1% That will be the probability for a boy to be born in any single birth or set of births in that population.

Other examples of *a priori* probabilities (P):

$P(\text{heads on tossing a coin}) = 1/2$

$P(\text{getting 6 on one roll of die}) = 1/6$

$P(\text{picking 4 of clubs in full deck of cards}) = 1/52$

$P(\text{picking winning number in a 4 digit lottery}) = 1/9999$

We can often assign P without trials, assuming randomness and the equality of all possibilities.

Most instances of probability involve more complicated processes than one single event. The question is, how do we handle multiple events, like two rolls of dice or three picks of cards?

Combining probabilities involves two different procedures:

* If two events (A and B) are **mutually exclusive** (i.e. the occurrence of one event excludes the other, and vice versa) their probabilities add

$P(\text{A or B}) = P(A) + P(B)$

e.g. What is the probability that one roll of a die can produce 1 OR 2? $1/6 + 1/6 = 2/6 = 1/3$.

* If two events are **independent** their probabilities multiply

$P(\text{A and B}) = P(A) \times P(B)$

e.g. What is the probability that one roll of a pair of dice will produce two 6's? $1/6 \times 1/6 = 1/36$.

Consider **coin tossing** to see how these rules of probability work.

The probability of one head in **one toss** = $1/2$. This is the *a priori* probability, based on the assumption that there is an equal probability for heads or tails.

For **two tosses** of a coin there are more possible outcomes.

The probability of two heads in two tosses = $1/2 \times 1/2 = 1/4$. These are independent events.

What is the probability of 1 head & 1 tail in two tosses?

Consider the two (mutually exclusive) possibilities for this, HT or TH.

$P(\text{HT}) = P(H) \times P(T) = 1/2 \times 1/2 = 1/4$. These are two independent events.

$P(\text{TH}) = P(T) \times P(H) = 1/4$. Again these are independent events.

But $P(\text{H\&T in any order}) = 1/4 + 1/4 = 1/2$, because the two orderings are mutually exclusive.

Note that the probability of no heads in two tosses = $1/2 \times 1/2 = 1/4$, which is the probability of two

tails. So the total of all combinations is 1.

Three tosses gets into more complex analyses. There are 8 possible outcomes,

HHH
 HHT HTH THH
 HTT THT TTH
 TTT

First notice the pattern 1 3 3 1. Does this look familiar? Now if each H or T was different from every other H or T we would have for 3 distinct H(orT)'s there would be 3 ways to choose positions for 1st H-T \times 2 ways for 2nd \times 1 way for 3rd or 3! (the number of permutations of 3 distinct objects). But we assume that all H's or T's are the same (indistinguishable) so for 3H's or 3T's we have $3!/3!=1$ (the number of combinations – we divide out the indistinguishable rearrangements).

For 2H's (or 2T's) the pair can be interchanged so $3!/2!=3$ combinations. Hence the 1331 pattern.

Now what is the probability to get three heads in 3 tosses? There is only one combination, so $P(\text{HHH}) = 1 \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = 1/8$ which is the same as $P(\text{TTT})$.

What is the probability to get two and only two heads in three tosses? Now there are 3 combinations, so

$$P(\text{two H}) = P(\text{HHT or HTH or THH}) = 3 \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = 3/8.$$

You can see that this is the same as the probability for only one head also, since that is the probability for two tails. If the coins were weighted somehow, this would not be the case. Then the *a priori* probability for a head would be different than for a tail.

Four tosses can be dealt with in the same way. There are 16 possible outcomes,

HHHH
 HHHT HHTH HTHH HHHT
 HHTT HTHT HTTH THHT THTH TTHH
 TTTH TTHT THTT TTTT
 TTTT

Now we have 1 4 6 4 1 combinations. For distinguishable coins there would be 4! combinations. But for four heads we have 4! indistinguishable arrangements giving $4!/4! = 1$. For three heads there are 3! arrangements that are indistinguishable, so $4!/3! = 4$ is the number of combinations that are left. When we have two heads and two tails, there are 2! ways to rearrange the heads and 2! ways to rearrange the tails, resulting in $4!/(2!2!) = 3! = 6$ outcomes for two and two. The probability for four heads is

$$P(\text{HHHH}) = 1 \times (\frac{1}{2})^4 = 1/16.$$

And for three heads

$$P(\text{three H}) = P(\text{HHHT or HHTH or ...}) = 4 \times (\frac{1}{2})^3 \times (\frac{1}{2})^1 \text{ where the third power of } \frac{1}{2} \text{ is for the three heads and the first power of } \frac{1}{2} \text{ is for the tail.}$$

For two heads only

$$P(\text{two H}) = P(\text{HHTT or HTTH or ...}) = 6 \times (\frac{1}{2})^2 \times (\frac{1}{2})^2 \text{ where the powers of } \frac{1}{2} \text{ are for the two heads and for the two tails.}$$

Suppose the coin is weighted so that the *a priori* probability for a head is, for example, 2/3. Then the probability for two heads would involve the powers $(2/3)^2$ for the two heads and $(1/3)^2$ for the two tails.

A pattern should be emerging now. Mathematicians and physicists like general formulas so that we don't have to go through this whole procedure every time we want to know some particular probability. All of the above coin tossing examples are summarized in a rule. Consider n tosses of

an unweighted coin or one toss of n coins. For k heads and $n-k$ tails there are

$$\frac{n!}{(n-k)!k!}$$

combinations, so the probability is

$P(k \text{ heads and } n-k \text{ tails}) =$

$$\frac{n!}{(n-k)!k!} \left(\frac{1}{2}\right)^{(n-k)} \left(\frac{1}{2}\right)^k$$

This is called the **binomial distribution**. When the coins are weighted the formula changes in a simple way. Suppose p is the *a priori* probability for a head. Then the *a priori* probability for a tail will have to be $q = 1 - p$.

So the formula becomes

$P(k \text{ heads and } n-k \text{ tails}) =$

$$\frac{n!}{(n-k)!k!} (p)^k (1-p)^{(n-k)}$$

This is useful when you consider the probabilities in throwing dice. Suppose you ask questions like what is the probability for getting a pair of sixes out of 7 rolls of a single die? That is equivalent to asking for the probability of two heads out of 7 tosses of a **weighted** coin with the weight $p = 1/6$ for the head and $q = 5/6$ for the tail (or all the faces that are not six).

For very large n this formula leads to the **normal distribution**. And when p is very small compared to 1 the form becomes the **Poisson distribution**. These are special formulas that have wide applicability. But the binomial distribution will be sufficient for our purposes.