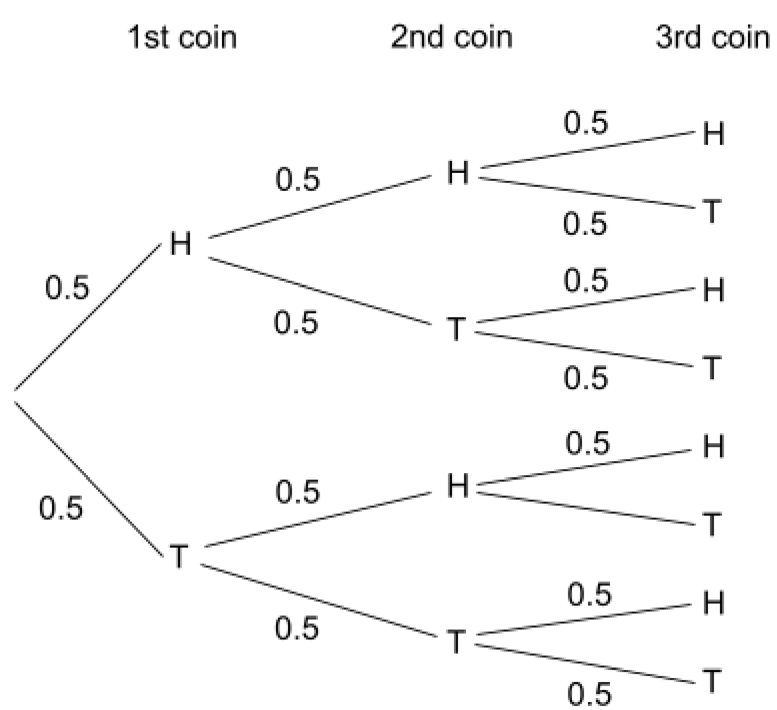
Probabilities with Two Outcomes

There are many situations we can think of that can end up in two ways. A flipped coin can land on either heads or tails, a light can be on or off, and a door can be open or closed. If we ask someone whether they've been to London before, they can answer in two ways: yes or no. An event which has two possible outcomes like this is known as a binomial event, the prefix 'bi' of course meaning two. We can even turn events that generally have more than two outcomes into a binomial event - for example, if we roll a dice, we can take the outcome to be whether it landed on an odd or even number.

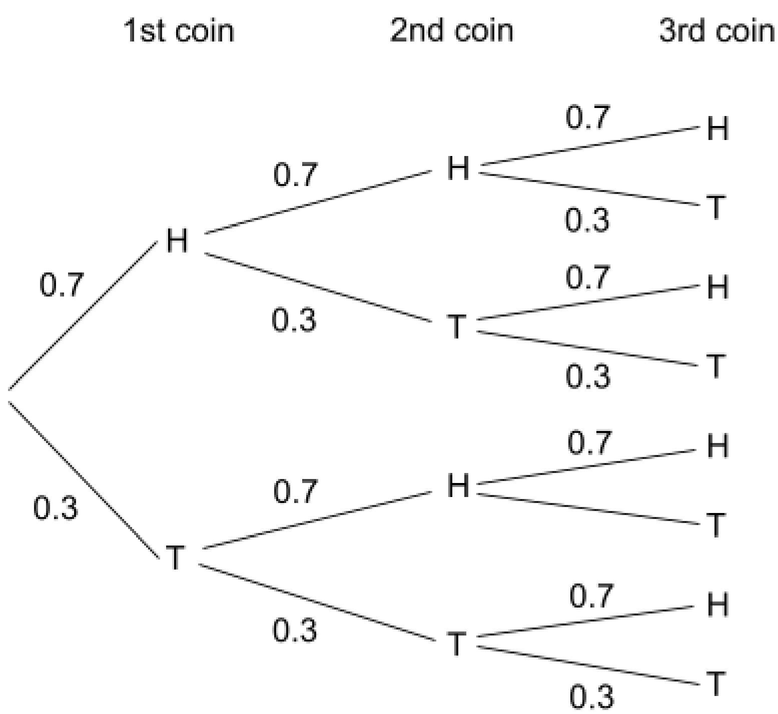
Naturally mathematicians want a way to calculate the probability of an outcome occurring a certain number of times. If we flip a fair coin three times, what is the probability that exactly one of these will result in a head? We call this a binomial experiment. A binomial experiment is essentially repeating a binomial event several times - we must also be careful that each trial (event) is independent of the others. That is, the outcome of one event won't affect the outcome of any others. If we do not adhere to this rule, we may encounter famous probability problems such as the Monty Hall problem and the Three Prisoners problem.

Let's go back to the problem of finding the probability of exactly one head from three flips of a coin. We can represent this problem quite nicely with a tree diagram.



Each of the numbers represents the probability of the event it is pointing towards happening (in this case, H is heads and T is tails). As it is a fair coin, the chance of a head is 0.5, and the chance of a tail is 0.5. In order to find the probability of exactly one head from the three flips, we need to find all the paths in the tree diagram in which there is exactly one head. Should the first flip be a head, the other two necessarily are tails. This leaves the combination HTT. If the first flip results in a tail, then the next one may be either a head or a tail and still hold the requirement. If the second flip is a head, the third must be a tail, and vice versa. This gives us the other two combinations of THT and TTH. We must now find the probability for each of these combinations and add them together to find the overall probability. HTT has a probability of $0.5 \times 0.5 \times 0.5 = 0.125$ (recall that the probability of two independent events is the separate probabilities multiplied). The other two combinations result in 0.125 also. When added together, the probability of exactly one head is $0.125 + 0.125 + 0.125 = 0.375$.

Now imagine we have a biased coin where the probability of a head is 0.7 (and so the probability of a tail is $1 - 0.7 = 0.3$). We can still represent this with a tree diagram.



And we can still work out the probability of exactly one head (try it out yourself - the answer should be 0.189).

Tree diagrams can certainly be helpful but they become tedious to use when we want to work with a higher number of trials. Imagine drawing out a tree diagram for 10 trials - the final column would have 1024 outcomes! Luckily we have more efficient ways of calculating binomial probabilities, thanks to several mathematicians including Blaise Pascal and Isaac Newton. We can use something called binomial expansion, which is the expansion of powers of a binomial. So, in general:

$$(p+q)^n$$

Note that p and q can be any letter we wish, although these are the accepted ones to use. When applying binomial expansion to probabilities, we can think of p as the probability of a success, and q as the probability of a fail. In the case of the weighted coin above, we can define success as a head. So $p = 0.7$, the probability of a success (head), and $q = 0.3$, the probability of a fail (tail). n is the number of trials that we are performing in the binomial experiment, for example if we flip three coins n will be 3. If we calculate the binomial expansion $(p+q)^3$, we will get the following:

$$\begin{aligned}
&(p+q)^3 \\
&= (p+q)(p+q)(p+q) \\
&= p^3 + 3p^2q + 3pq^2 + q^3
\end{aligned}$$

Now let's suppose we are calculating the probability of exactly one head from the three *biased* coins as above (heads = 0.7, tails = 0.3). We will let p be a head and q be a tail. The powers in the binomial expansion represent the number of that outcome occurring. For example, $3p^2q$ tells us that this particular part of the binomial expansion represents two heads (p^2) and one tail (q is the same as q^1). The three at the start means there are three combinations where this holds true (the HTT, THT, and TTH we discussed earlier). In order to calculate the probability of exactly one head, we will need the term $3pq^2$. From there it is a simple matter of plugging in the probability of a head into p and of a tail into q. So:

$$\begin{aligned}
&3pq^3 \\
&= 3(0.7)(0.3)^3 \\
&= 0.189
\end{aligned}$$

It can still be tedious to calculate binomial expansions (although certainly not as much as tree diagrams). There are two main methods that can be used to do them more efficiently and quickly, namely Pascal's triangle, and the binomial theorem.

Pascal's triangle was created by mathematician, physician and philosopher Blaise Pascal. It can give us the coefficients of each term in the binomial expansion, and we can work out the powers as p goes down by one each time and q goes up by one each time. Pascal's triangle can be made by writing ones diagonally, then for each entry, adding the number above and to the left with the number above and to the right:

1	$n = 0$
1 1	$n = 1$
1 2 1	$n = 2$
1 3 3 1	$n = 3$
1 4 6 4 1	$n = 4$
1 5 10 10 5 1	$n = 5$
1 6 15 20 15 6 1	$n = 6$

If we want to calculate $(p+q)^5$, we can find the coefficients in the 6th row (the first row is $n = 0$ as anything to the power of zero is one):

$$
\begin{aligned}
&(p+q)^5 \\
&= 1p^5 + 5p^4q^1 + 10p^3q^2 + 10p^2q^3 + 5p^1q^4 + 1q^5 \\
&= p^5 + 5p^4 + 10p^3q^2 + 10p^2q^3 + 5pq^4 + q^5
\end{aligned}
$$

From here we can calculate probabilities of binomial experiments just as before, e.g. three successes and two fails is given by $10p^3q^2$.

This method can be very useful if we need to manually do some binomial expansion - it is easy to draw out Pascal's triangle to a fairly high power. The other method, the binomial theorem, is a purely formulaic way to do it and as such can be used by computers and calculators extremely efficiently. On the other hand, it can get just as tedious as a tree diagram if we try to do it manually.

The binomial theorem was invented by Isaac Newton and gives a formula that can be used in full to calculate binomial expansions. It is as follows:

$$(p+q)^n = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k}$$
$$where \ \binom{n}{k} = \frac{n!}{k!(n-k)!} \ (this\ is\ pronounced\ 'n\ choose\ k')$$

It is essentially telling us to apply the 'n $choose$ k' function for every integer 0 up to n. This gives us the coefficient for each term (just know that this gives the coefficients in Pascal's triangle). As k is increasing by one for each term, $p^k q^{n-k}$ is a mathematical way for applying the 'increase the p power by one and decrease the q power by one' rule we made earlier when using Pascal's triangle. The sigma (Σ) notation then tells us to add all of these terms together. The binomial theorem is just the method we used before, but in a purely formulaic form. This allows it to be used on computers, and also provides a rigorous way for us to do these calculations. If we want to calculate the binomial expansion $(p+q)^4$, we can do it using the binomial theorem as follows:

$$\begin{aligned}
&(p+q)^4 \\
&= \sum_{k=0}^{4} \binom{4}{k} p^k q^{4-k} \\
&= \binom{4}{0} p^0 q^{4-0} + \binom{4}{1} p^1 q^{4-1} + \binom{4}{2} p^2 q^{4-2} + \binom{4}{3} p^3 q^{4-3} + \binom{4}{4} p^4 q^{4-4} \\
&= 1p^0q^4 + 4p^1q^3 + 6p^2q^2 + 4p^3q^1 + 1p^4q^0 \\
&= q^4 + 4q^3p + 6q^2p^2 + 4qp^3 + p^4
\end{aligned}$$

You can check this holds with the coefficients shown in Pascal's triangle, and the powers do increase/decrease according to how we expect.

The binomial theorem does boast one advantage over Pascal's theorem that helps massively with calculating binomial probabilities. Instead of producing the entire binomial expansion, we can use the binomial theorem to produce just the term we need to calculate the probability. We do this by setting k to the number of successes we want. The theorem is also useful for extremely large values of n.

Imagine we have a coin which has a 0.6 probability of flipping a head (and so a 0.4 probability of flipping a tail). We perform a binomial experiment on this coin in which we flip it 20 times, and we want to know the probability that exactly eight of them will flip heads. As you will know by now, this number of trials is extremely hard to calculate with a tree diagram, and it would be tedious to calculate the whole binomial expansion. Instead, we can plug the value of eight into k in the binomial theorem. We will get:

$$\begin{aligned} &\binom{n}{k} p^{k} q^{n-k} \\ &= \binom{20}{8} p^{8} q^{20-8} \\ &= 125970 p^{8} q^{12} \end{aligned}$$

We now have just the term we need, instead of the whole expansion. So to calculate the probability of exactly eight heads, we must substitute 0.6 into p and 0.4 into q.

$$\begin{aligned} &= (125970)(0.6^{8})(0.4^{12}) \\ &\approx 0.0354974 \end{aligned}$$

One of the things I like about binomial probability is that it can all be boiled down into this relatively simple formula. What seems like a tedious process - drawing out tree diagrams, calculating binomial expansions - are actually just the processes we go through to find out that it can be done simply by plugging in the number of successes to a variable k.