r) Probability

Step 1) The Probability Scale

The word probability is linked to the word probably. If someone says I will probably be there tonight, it means they are more likely to be there than not be there tonight, it is likely they'll be there. If someone says they will probably not heave cereal for breakfast, it means they are less likely to have cereal than something else, they are unlikely to have cereal.

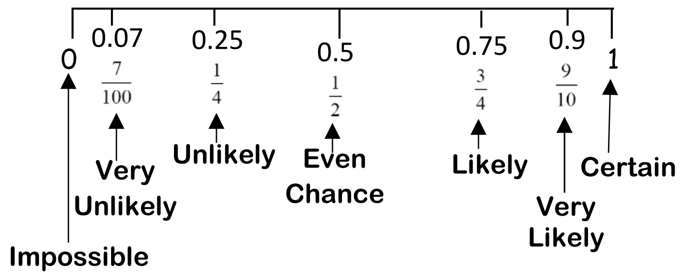
This language of talking about how likely something is to happen is the language of probability. But the words likely, unlikely, probably and probably not are all a bit vague. So we mathematicians have developed a scale from 0 to 1 to describe how likely something is in a very precise way.

If something is impossible to happen, that is if it is certain not to happen, we call it probability of 0. If something is certain to happen, we say the probability is 1. Everything else falls somewhere in between, the larger the probability, the closer to 1, the more likely it is that it will happen. The closer to 0, the less likely. A probability of 0.5 (which is the same as ½) means that the thing is equally likely to happen as not happen.

For short hand we write the probability of an event as an equation with a capital P on the left, the event we are talking about in brackets and the number between 0 and 1 representing the likelihood of it happening, on the right hand side.

P(me being there tonight) = 0.7 means the probability of me being there tonight is 0.7 (which is likely)

P(red token from hat) = 0.3 means the probability of picking a red token from a hat is 0.3 (which is unlikely).



There is no exact place where something moves from being likely to very likely (is it 0.85, or 0.9, or 0.99?) but the numbers gives us a very precise way of measuring and comparing the probabilities of different things.

Step 2) The Probability of an Event – P(Event)

We have a very exact way of finding a probability in a theoretical situation. Let's say we have a coin.

$$P(Event\ A) = \frac{Number\ of\ possible\ event\ As}{Total\ number\ of\ possible\ events}$$

For example let's look at a coin. When flipping a coin, there are two possible outcomes, flipping a head or a tail. So the probability of flipping a head would be written as…

$$P(Flipping\ a\ head) = \frac{1}{2}$$

If you roll a dice, there are 2 multiples of 3 (3 and 6) and 6 possible outcomes in total (1, 2, 3, 4, 5 & 6) so…

$$P(Roling\ a\ multiple\ of\ 3\ on\ dice) = \frac{2}{6} = \frac{1}{3}$$

Let's say we had a bag with the letters of the word ELEPHANT each on a card. If we picked a card at random we could say there are 2 ways of getting an E, and 8 possible cards to pick from in total, so.

$$P(pick\ an\ E\ from\ ELEPHANT) = \frac{2}{8} = \frac{1}{4}$$

All probabilities are built up in this same way, and they will always give an outcome between 0 and 1.

Step 3) Experimental Probability

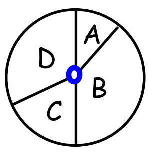
Often in real life we work with types of outcomes that aren't so predictable or aren't as equally likely to occur as one another. Tokens being picked from a hat might act in predictable ways, but arrows being fired at a target, or people answering general knowledge questions will not give equally likely outcomes. For this we can experiment, either firing lots of arrows at the target, or asking lots of people our questions, and base our probabilities on the results of these experiments.

Experimental probability is found by dividing the number of successful experiments by the total number of experiments.

$$P(Event\ As) = \frac{No\ of\ event\ A\ outcomes}{Total\ number\ of\ tests}$$

A spinner has 4 sections but is unevenly weighted. A tally was taken to show how many times it landed on each section in a number of test spins.

Section	Tally	Frequency
A	𝍸 𝍸 ||	12
B	𝍸 𝍸 𝍸 𝍸 𝍸 𝍸 𝍸	35
C	𝍸 𝍸 𝍸 𝍸 𝍸	25
D	𝍸 𝍸 𝍸 𝍸 𝍸 |||	28
Total		100



We simply calculate the probabilities by dividing the number of times the spinner landed on that section (A, B, C or D) by 100 (the total number of tests)

For example

$$P(Section\ A) = \frac{No.of\ tests\ landing\ on\ section\ A}{Total\ tests} = \frac{12}{100}$$

We can then simplify the fraction or turn it to a decimal $\frac{12}{100} = \frac{3}{25} = 0.12$

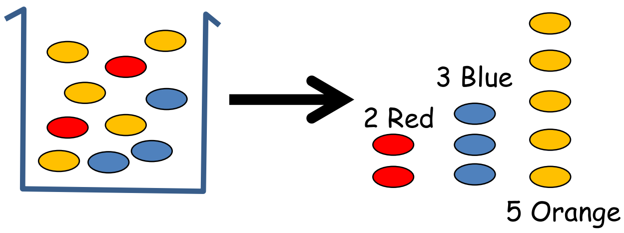
Section	Frequency	Probability (from formula)	Probability (simplified fraction)	Probability (decimal)
A	12	$\frac{12}{100}$	$\frac{3}{25}$	0.12
B	35	$\frac{35}{100}$	$\frac{7}{20}$	0.35
C	25	$\frac{25}{100}$	$\frac{1}{4}$	0.25
D	28	$\frac{28}{100}$	$\frac{7}{25}$	0.28

We can also call this the **relative frequency** because it says how many times a particular outcome happens, relative to all the tests you did.

In fact with experimental probability, we say that the relative frequency is the probability.

Step 4) Sum of Probabilities of all Possible Outcomes is 1

A jar contains 2 red, 3 blue and 5 orange sweets.



P(blue) =$\frac{3}{10}$, P(orange) = $\frac{5}{10}$& P(red) = $\frac{2}{10}$

Sum of all the probabilities is

$$\frac{3}{10} + \frac{5}{10} + \frac{2}{10}$$

$$= \frac{10}{10}$$

$$= 1$$

In fact, the sum will always be 1. This is because the bottom of each fraction is the total frequency of all the possible outcomes. The top of each fraction is the frequency for each possible outcome, but one we add them together we have the frequency of all the particular outcomes, added together. In other words…

$$\frac{Sum\ of\ freq\ of\ each\ outcome}{Total\ frequency\ of\ all\ possible\ outcomes}$$

$$= \frac{Total\ frequency\ of\ all\ possible\ outcomes}{Total\ frequency\ of\ all\ possible\ outcomes} = 1$$

We can use this to find out missing probabilities.

There are green, purple and yellow counters in a cup. One is picked at random. The probability of getting either a green or a purple is $\frac{7}{10}$. What is the probability of getting yellow one?

$$P(red\ or\ purple) + P(yellow) = 1$$

$$\frac{7}{10} + P(yellow) = 1$$

so $P(yellow) = 1 - \frac{7}{10} = \frac{3}{10}$

Another interesting fact (corollary) from this is that the probability of something happening or not happening is one.

On a dice $P(rolling\ a\ 6) = \frac{1}{6}$

$$P(not\ rolling\ a\ 6) = P(rolling\ a\ 1or2or3or4or5) = \frac{5}{6}$$

$$\frac{1}{6} + \frac{5}{6} = \frac{6}{6} = 1$$

This will always be true, as an event happening or not happening covers all possible outcomes for that event (it must either happen or not happen).

Step 5) Sample Space Diagrams

So now we move on to probabilities involving more than one event.

A sample space diagram, is a type of 2-way table to represent all the possible combined outcomes for two independent events, where all the outcomes for each event are equally likely.

Independent events are ones where the outcome of one cannot affect the other.

For example, if you flip a coin, and roll a dice, whether you flip a head or a tail it won't affect the role of the dice. Similarly, if you role a 1, 2, 3, 4, 5 or 6 this won't affect how likely it is you flip a head or a tail on the coin. So these two events are independent.
Now there are two possible outcomes on the coin flip and they are both equally likely, with probability $\frac{1}{2}$

Similarly, all the 6 outcomes of rolling the dice are equally likely with probability $\frac{1}{6}$

This is a perfect example for a sample space diagram, each event has equally likely outcomes, and the two events are independent (don't affect each other).

Here is how the diagram looks.

		Dice			
		1	2	3	4
Coin	Heads	H,1	H,2	H,3	H,4
	Tails	T,1	T,2	T,3	T,4

You can see that there are 8 equally likely outcomes, with the probability of each outcome is $\frac{1}{8}$. This diagram gives you a visual sense of how all the possible outcome pairs fit together.

We can use these diagrams to answer quite complex questions. For example, I role two separate dice, one blue, one green, and multiply the answer. What is the probability that this product is an odd number.

Now the roles of the two dice are independent, and the outcomes for each dice are equally likely.

This time, instead of writing Blue 4, Green 3 in the sample space diagram, we will write the product of the two dice.

Note: The product of two numbers is what you get when you multiply them.

		Blue Dice					
	X	1	2	3	4	5	6
Green Dice	1	1	2	3	4	5	6
	2	2	4	6	8	10	12
	3	3	6	9	12	15	18
	4	4	8	12	16	20	24
	5	5	10	15	20	25	30
	6	6	12	18	24	30	36

There are 9 odd products.

There are 36 outcomes in total.

$$P(Odd\ Product) = \frac{9}{36} = \frac{1}{4}$$

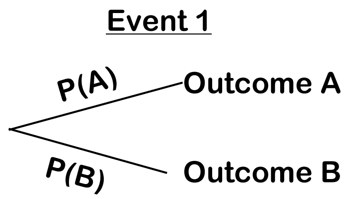
However if we had a weighted blue dice that landed more frequently on the 6, the outcomes in the column representing 6 would have higher probabilities than the other columns, and so we couldn't make calculations as if each cell in the outcome grid was equally likely.

When the outcomes are not equally likely we need a tree diagram.

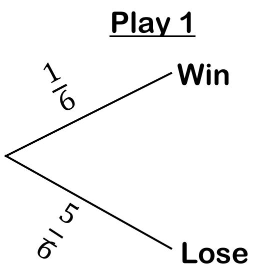
6) Tree Diagrams of Independent Events

A tree diagram is a way of working with multiple events, where the possible outcomes for each event aren't equally likely.

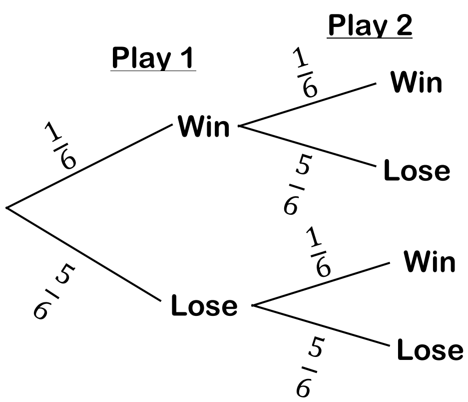
They are set up with the event title at the top, with a branch for each possible outcome. The outcome name is written at the end of the branch, and the probability of that outcome written along the branch as follows…



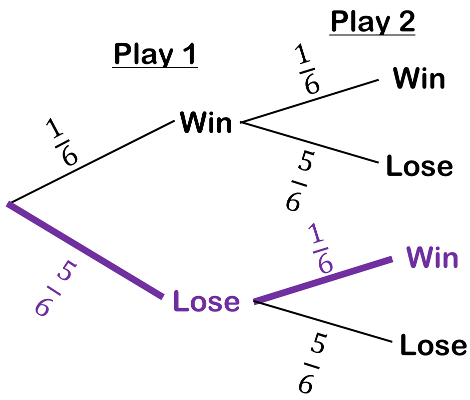
Let's take a game where you win if you role a 6 on a dice, and lose if you role any other number.



Now we can add a second play of the game, one for if we won the first game, and another for if we lost it.



Let's look at a particular line along the branches of the tree.

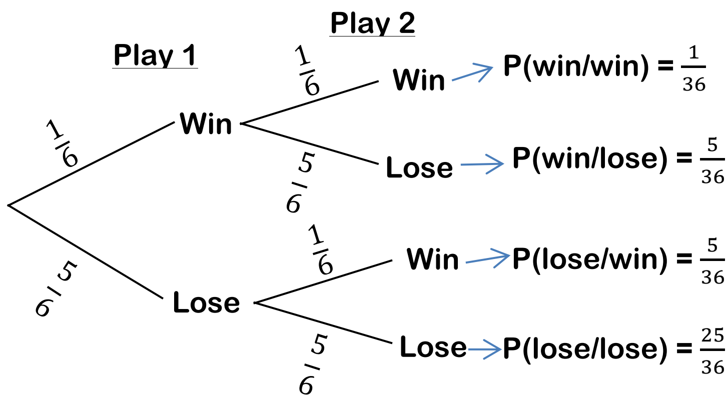


The line we have made purple represents playing the first game and losing, then playing the second game and winning.

When you combine probabilities of independent events you have to multiply them.

$$P(lose\ 1st, win\ 2nd) = \frac{5}{6} \times \frac{1}{6} = \frac{5}{36}$$

We can calculate the probability of all the outcome pairs in the same way…



Finally we can add probabilities along the right hand side.

For example…

P(win exactly 1) = P (win/lose) + P(lose/win) = $\frac{10}{36}$

P(same result twice) = P (win/win) + P(lose/lose) = $\frac{26}{36}$

Why for we multiply probabilities for combining outcomes (and is times) and divide when finding the probability of one or another outcome occurring (or is add).

Firstly let's think about the probability of all the possible outcomes. Here there are four possibilities for he four separate paths along the tree. We could have

1) Win first, win second

2) Win first, lose second

3) Lose first, win second

4) Lose first, lose second

These are all the possible outcomes, so their probabilities should add up to 1.

Let's find out if they do...

1) P(win/win) = $\frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$

2) P(win/lose) = $\frac{1}{6} \times \frac{5}{6} = \frac{5}{36}$

3) P(lose/win) = $\frac{5}{6} \times \frac{1}{6} = \frac{5}{36}$

4) P(lose/lose) = $\frac{5}{6} \times \frac{1}{6} = \frac{25}{36}$

If we add them all up we get

$$\frac{1}{36} + \frac{5}{36} + \frac{5}{36} + \frac{5}{36} = \frac{36}{36} = 1$$

So this is a consistent way of making sure all the probabilities add up to 1 (see step 4).

Another way of thinking about this is that the possibility of one or another outcome, must be higher than just one of the outcomes (unless one of the probabilities is 0). The only way to be certain increasing a number between 0 and 1, using another number between 0 and 1, is by adding.

So or is add!

And why do we multiply when looking at possible consecutive outcomes, in other words looking at event A and then event B. Why is and times.

Well the chance of flipping a tail on a coin is ½ . It must be less likely to do this twice in a row than to do it once, so the probability must get smaller. The only consistant wat to make a number between 0 and 1 smaller, using another number between 0 and 1 is by multiplying.

In the coin context, exactly how much less likely is it that we will get a second tail in a row, it is half as likely. So we want ½ of a ½ which we get by multiplying.

So P(Tail, tail) = ½ of a ½ = ½ x ½ = ¼

So or is times

There's a couple of good ways of memorising this.

One is using the "adorable andrex puppy." This was a very cute Labrador retriever puppy used in a very successful advertising campaign for andrex loo roll.

"adorable andrex puppy."

Add-or able gives add is or

and-re-x gives and is times

Or a brilliant song to the tune of winter classic "O Christmas Tree" goes...

Probability tree, probability tree,
We multiply across thee.
And when we're done, we can sum (+) down,
All the products (x), that we have found.
Probability tree, probability tree,
We multiply across thee.

An independent event, is one where separate goes don't effect each other. For example, rolling a dice, won't change the probabilities for a second roll of the dice.

Now we will look at dependent events.

7) Tree Diagrams of Dependent Events

Dependant events are ones where the outcome of the first event has an effect on the second event.

Let's say there are 3 green counters and 7 blue counters in a hat.

If we pick a counter at random (no peeking!) we get the following probabilities.

P(Green) = $\frac{3}{10}$

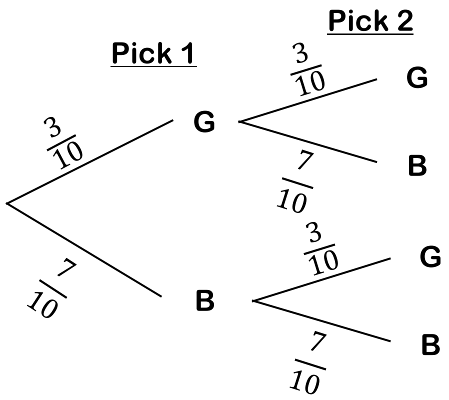
P(Blue) = $\frac{7}{10}$

If we put the counters back and do a second pick we will get the same number of counters of each colour and so the same probabilities. So these two picks would be independent events.

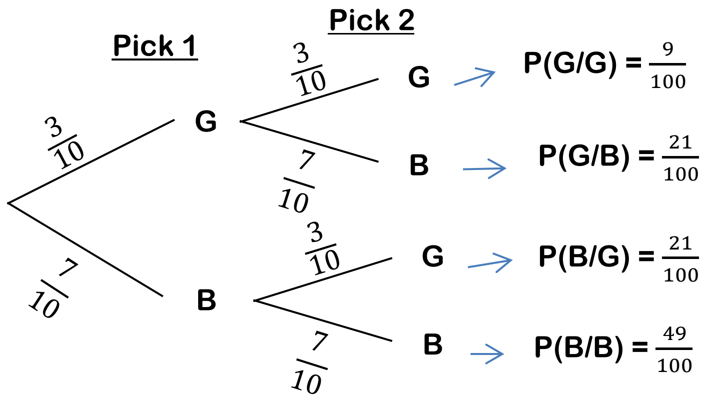
However, if we do not put the counter back for the second pick, then the probabilities for the second pick depend on whether we picked a green or a blue counter on our first pick. This is because if we picked a blue first there will be one less blue counter for the second pick (as well as one less counter overall) meaning a lower chance of getting a blue. If we picked a green counter on the first pick then there will be one less green counter (as well as one less counter overall) and so a lower chance of getting a green on the second pick.

This type of event pairing is called a dependent event. The probabilities for the second event are dependent on what happens in the first event.

So let's look again at our hat with 3 green and 7 blue counters, and construct an independent event tree diagram for two picks from the hat when we do put the first counter back.

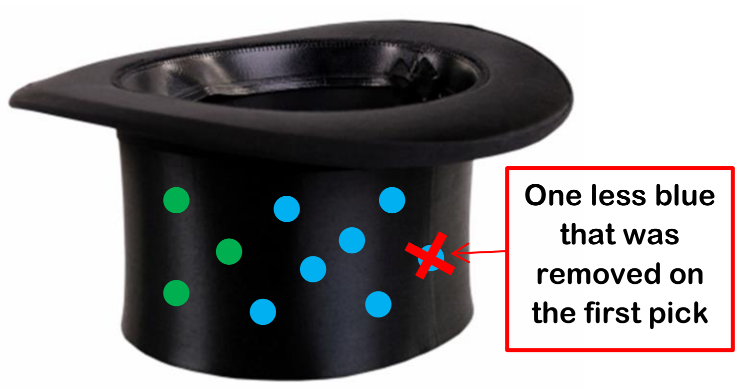


Let's just work out the probabilities of each of the four possible pairs of outcomes: Green/Green; Green/Blue; Blue/Green; and Blue/Blue; then add them up to be sure they make 1, and that this is a consistent group of probabilities.



$$\frac{9}{100}+\frac{21}{100}+\frac{21}{100}+\frac{49}{100}=\frac{100}{100}=1$$

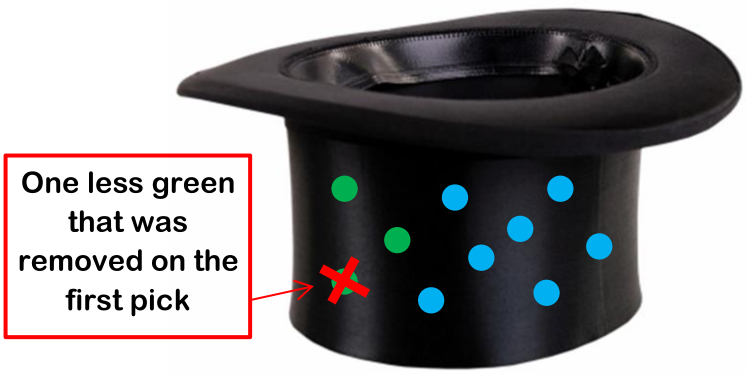
But what happens if we don't put the counter back between picks. The outcome of the first pick, will effect the outcome of the second pick. Let's say we are interested in the probability of the second pick's outcome being blue. If the first pick was blue then the hat looks like this…



So after first picking a blue, we have 3 greens and 6 blues left.

So P(Blue) = $\frac{6}{9}=\frac{2}{3}$

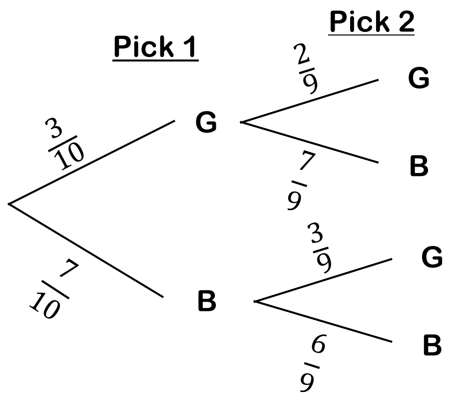
If however we had picked a green on the first pick the hat would like this.



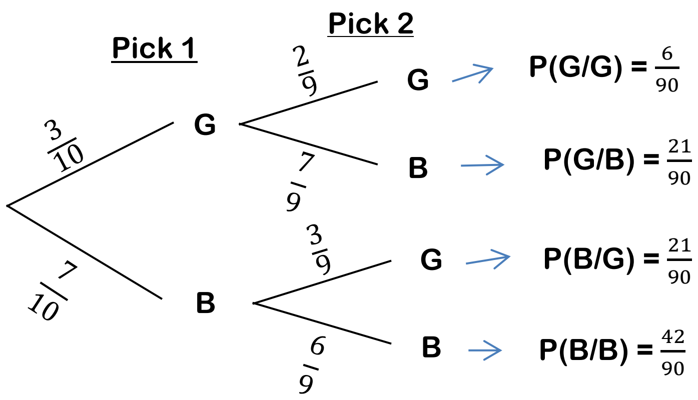
So after picking a green, we have only two greens left, and seven blues.

In the case P(Blue) = $\frac{7}{9}$

In this way, adjusting the second picks based on how the outcomes of the first pick affect the numbers of greens and blues we can build a tree diagram for these two dependent events.



Again let's find the probabilities of each of the four possible pairs of outcomes: Green/Green; Green/Blue; Blue/Green; and Blue/Blue; then add them up



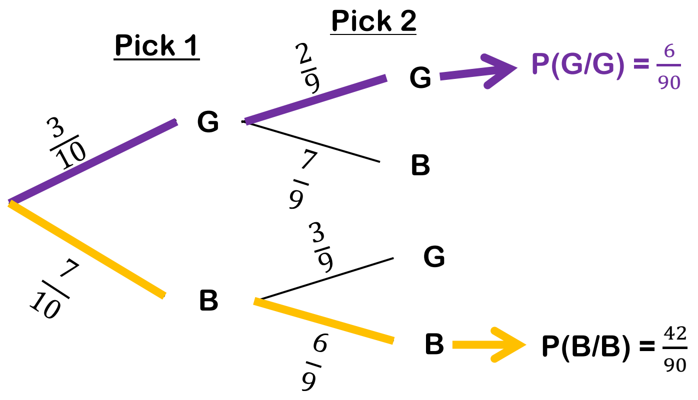
$$\frac{6}{90}+\frac{21}{90}+\frac{21}{90}+\frac{42}{90}=\frac{90}{90}=1$$

It is worth noting that some of these "end" probabilities can be simplified, but it is soften easier not to simplify them with tree diagrams, so that if you end up adding the "end" probabilities because of an "or" question then you already have fractions with common denominators.

Finally let's use this tree diagram to find the probability that both picks are the same colour (when we don't put the first counter back after the first pick).

So the probability of a green then a green (represented on the tree diagram below as the purple pathway) is found by

P(Green/Green) = $\frac{3}{10}\times\frac{2}{9}=\frac{6}{90}$



And the probability of picking two blues (represented on the tree diagram below as the orange pathway) is

P(B/B) = $\frac{7}{10}\times\frac{6}{9}=\frac{42}{90}$

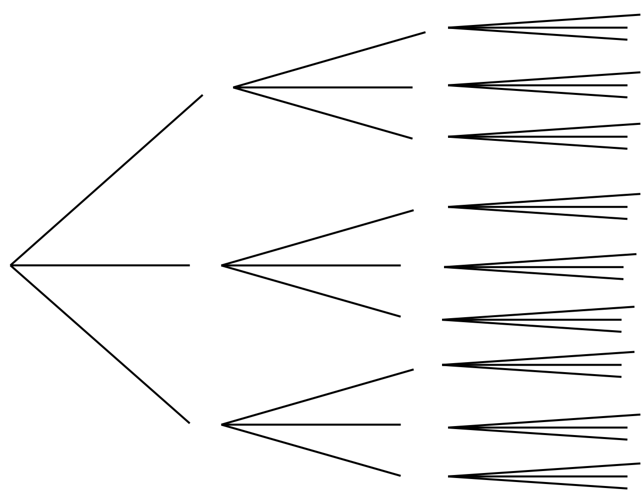
So P(both same colour) = $\frac{6}{90}+\frac{42}{90}=\frac{48}{90}=\frac{24}{45}=\frac{8}{15}$

Finally it is worth noting that you don't always need to draw a tree diagram.

For example there are 2 orange, 3 red, and 1 purple counter in a hat. What is the probability of picking 3 reds in a row?

Firstly if we do this with putting the counters back as independent events.

To draw the whole three stage tree diagram with 3 branches for each event would be very complex!



This is mega complex even without the outcomes and probabilities written in.

But we can work out many things without drawing the diagram.

When putting the counters back each time, the probability of picking a red each time is $\frac{1}{6}$

P(3 consecutive reds) = $\frac{3}{6} \times \frac{3}{6} \times \frac{3}{6} = \frac{27}{216} = \frac{9}{72} = \frac{1}{8}$

You may be screaming that it is way simpler to work simplify P(Red) = $\frac{3}{6} = \frac{1}{2}$ and then do

P(3 consecutive reds) = $\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$

And you'd be right, in this example it is much easier to simplify the fraction first!

But finally let's look at the case where you do not put the counter back each time. Surely here we have to draw the big horrible 27-pointed probability tree because the probabilities change each time.

Well no we don't.

P(3 consecutive reds without replacement)

$$= \frac{3}{6} \times \frac{2}{5} \times \frac{1}{4} = \frac{6}{120} = \frac{1}{20}$$

On each consecutive pick we simply have one less red (-1 from top), and one less counter over all (-1 from bottom).

P(3 Reds) = $\frac{1}{20}$