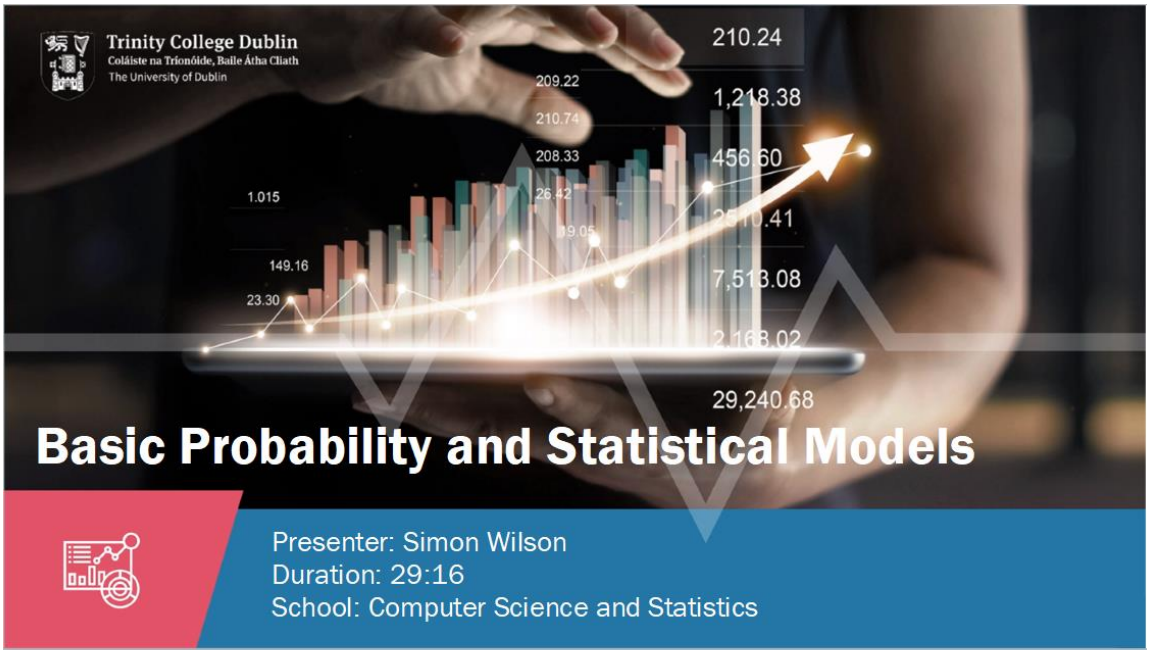
Basic Probability and Statistical Models

Slide 1: **Introduction**

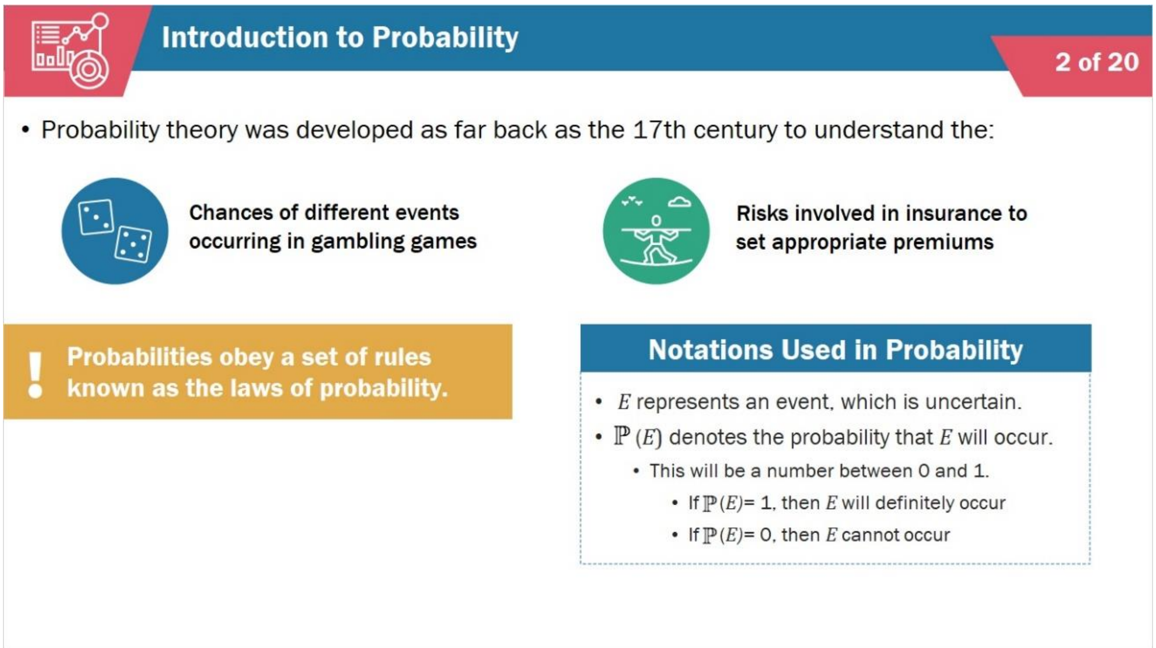


Welcome to this presentation on basic probability and statistical models. My name is Simon Wilson and I will lead this presentation, where we explore some of the important tools that we use to develop and justify statistical methods.

The first of these is probability; the way that, mathematically at least, we can quantify and manipulate uncertain situations. As we saw in the last session, the presence of random variation is the principal motivation behind using statistical methods and probability is the area of mathematics that deals with such variation. Understanding the concepts behind statistical methods, as well as their properties, requires some knowledge of probability.

Within probability, statistical models are how we define the nature of the random variation in a population. We have already looked at the normal curve or distribution as one sort of random variation, and here we will look at that in more detail, as well as some other sorts of variation that we commonly see.

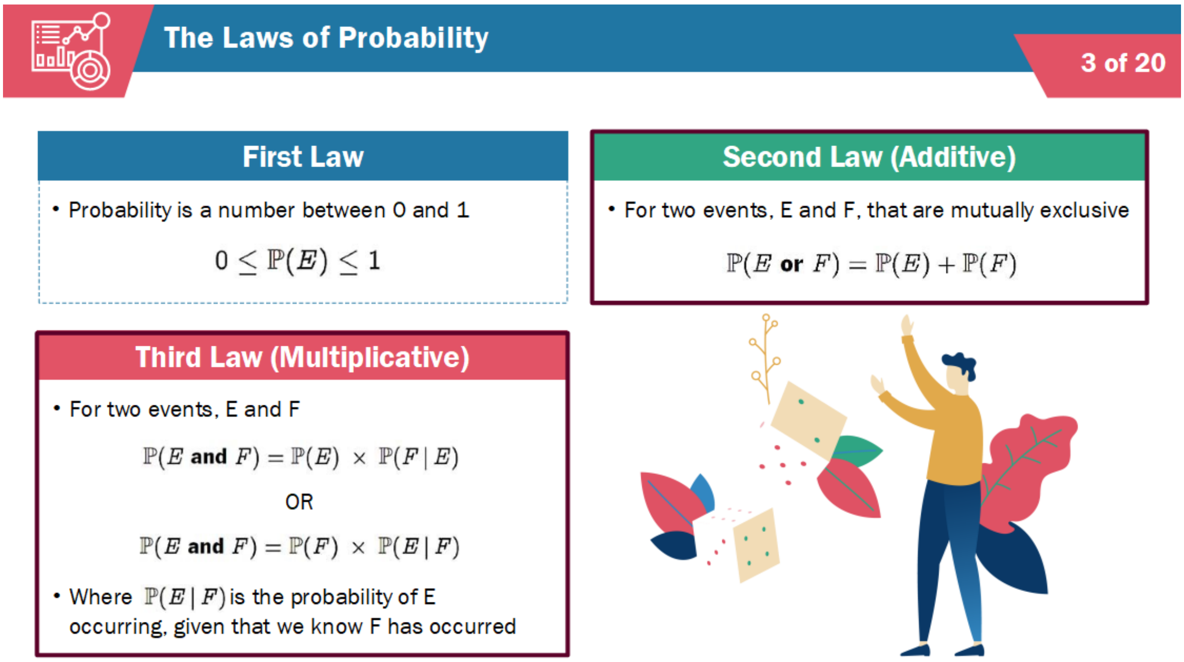
Slide 2: Introduction to Probability



Probability theory has a long history, going back to the 18th and even 17th centuries. It was developed largely by mathematicians with an interest in games of chance, and in particular, that success in gambling with such games often relied on understanding the chances of different events occurring. Another motivation was insurance, as the availability of fire, life and other forms of insurance became more common, and insurers needed a good understanding of the risks that were being insured in order to set premiums appropriately.

Mathematicians quickly realised that probabilities always seemed to obey a set of rules, that these days are called the laws of probability. To discuss these, let's start off with thinking about the probability of some event that we call 'E'. We do not know if 'E' will occur or not; it is uncertain. 'E' could be throwing a head when tossing a coin, or that it rains tomorrow, or that next month's winning lottery numbers contain the number 37. We use the notation 'P(E)', as written on the slide, to denote the probability that E will occur. By convention, this will be a number between 0 and 1, with a probability of 1 denoting certainty that E will occur, and 0 denoting certainty that E will NOT occur.

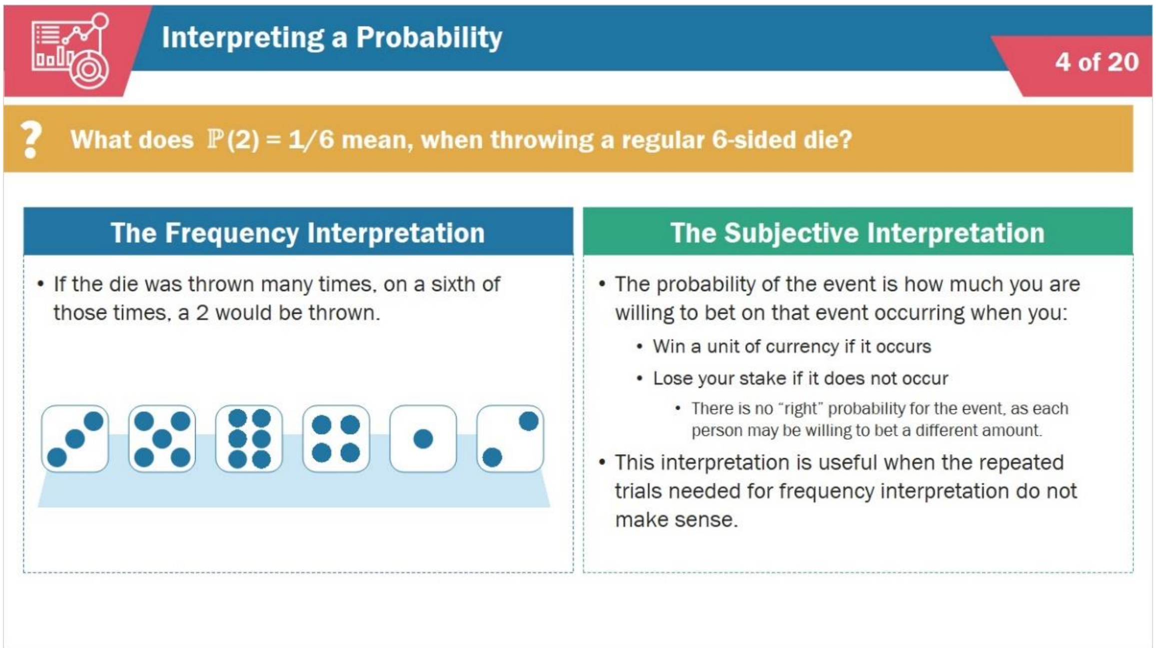
Slide 3: The Laws of Probability



That a probability is a number between 0 and 1 is the first law (or property) of probability. There are two others, as given here, that tell you what are the probabilities of two combinations of two events, namely the probability that both of these events occur, as well as the probability that one or both of the events occur. Taking two events, let's say E and F, the additive law says that the probability that E *or* F occurs is the sum of their individual probabilities, but that is under the condition that these events cannot both occur at the same time (what we call mutually exclusive events). The third, or multiplicative law, tells you that the probability of E and F both occurring is the product of the probability of one of the events and the probability of the other one conditional on the first one having occurred. P(F|E) means the probability that F occurs, given that we know E has occurred; This law also can be written with E and F swapped.

In the Apply/Reflect section of this session, you'll get a chance to look at some examples of these laws in more detail.

Slide 4: ## Interpreting a Probability

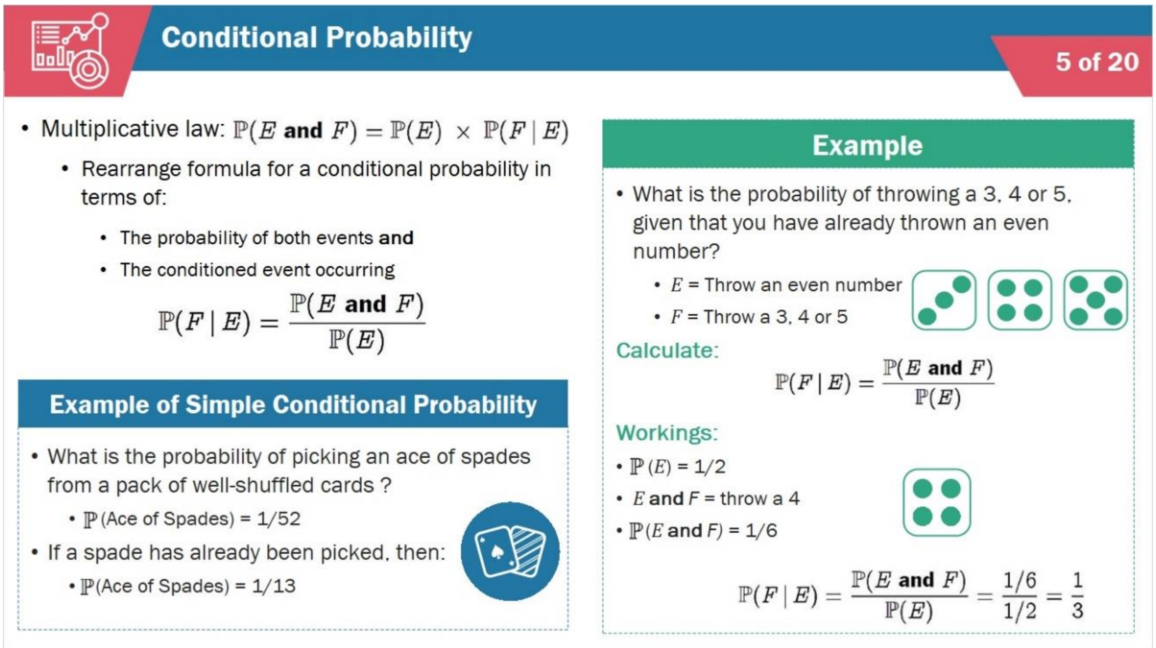


When we say that the probability of getting a 2 when throwing a regular 6-sided die is 1/6, what does that actually mean? It turns out that there is not a simple answer to that question, and these days there are in fact two ways that we usually think about that meaning.

The first, and the one that is useful for most purposes and in particular for the motivation behind some of the statistical methods that we look at later on in this module, is that it refers to the proportion of times that the event occurs over a very long sequence of trials where it could have occurred – the frequency interpretation So we interpret the probability of a 2 being 1/6 as that if we threw the die very many times then on $1/6^{th}$ of those throws would we see a 2.

The second is the subjective interpretation, which associates a probability more closely with gambling. The probability of an event is how much you're willing to bet on that event occurring, where you win one unit of currency if it does occur and lose your stake if not. So, in this case, we are willing to bet $1/6^{th}$ of a euro, let's say, on rolling a die and getting a 2, where we win a €1 if that happens. In this interpretation, there is no 'right' probability for an event as each of us may be willing to bet a different amount. It is a useful interpretation where the idea of repeated trials, that we need for the frequency interpretation, does not make much sense. For example, in assessing whether Shakespeare wrote all of the plays attributed to him, we might quantify our uncertainty in this by a probability, but the frequency interpretation, where there would have to be very many Shakespeares and a certain proportion of them wrote all of the attributed works and the rest did not, may not make much sense.

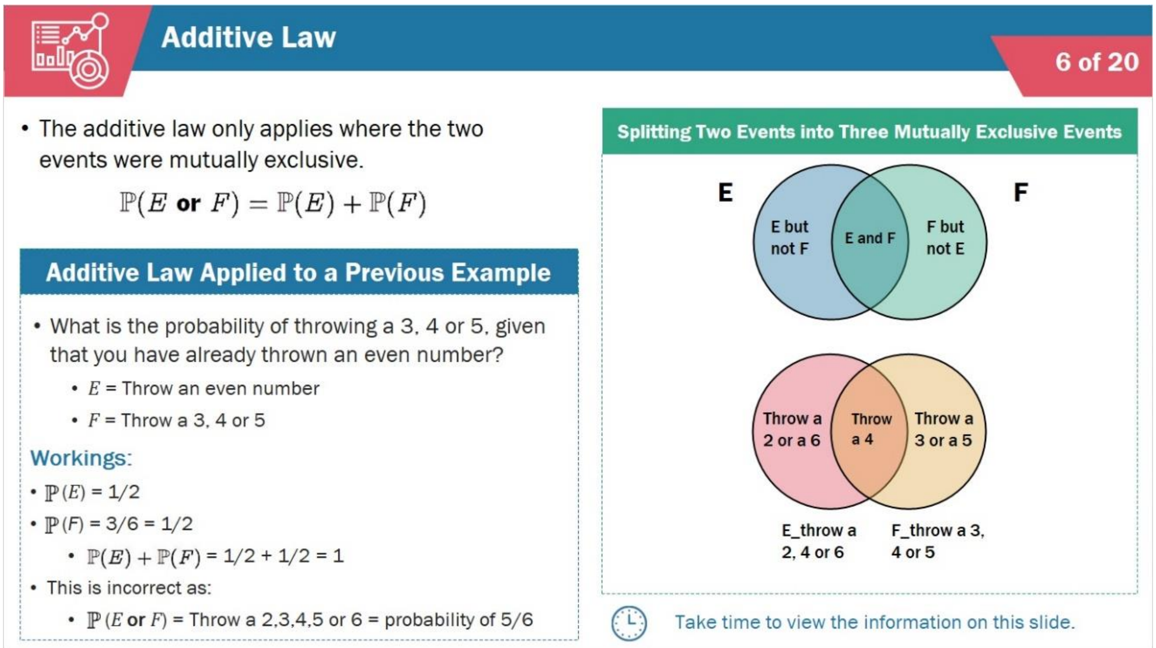
Slide 5: Conditional Probability



Recall the multiplicative law, that uses the idea of the probability of one event conditional on knowing another has occurred. Clearly this knowledge can change a probability. For example, the probability of picking an ace of spades from a pack of well-shuffled cards is 1/52 (it is one out of the fifty two possible cards). If you know that a spade has been picked then, conditional on this having occurred, the chance is now 1/13 (it is one out of the thirteen spades). Just re-arranging the multiplicative law gives us a formula for a conditional probability in terms of the probability of both events and the conditioned event occurring.

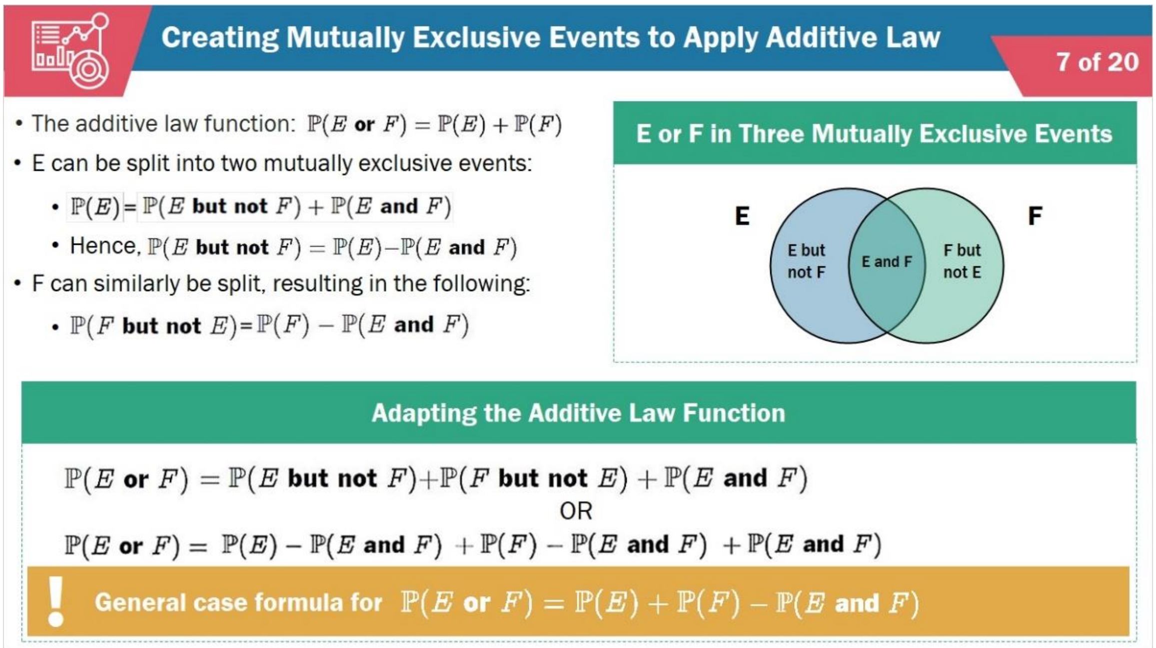
Let’s take a look at another example here to illustrate the formula and show that it does produce the sensible answer that we might have got directly. We want to look at the probability of throwing a 3, 4 or 5, given that we know that we’ve already thrown an even number. So, the event E is to throw an even number and the second event F is to throw a 3, 4 or 5. If we apply our formula: the probability of throwing an even number is 1 /2 as we can either throw an odd or an even number. If we look at the probability of throwing a 4 next, then that is 1/6 – there are six possibilities on the die. Hence the answer is 1/3 which makes sense since only one of the three outcomes in F is even.

Slide 6: Additive Law



Remember that the additive law only applied to the case where the two events were mutually exclusive or could not occur at the same time. In the next few slides, we'll use the laws of probability to derive the general formula for the probability of either or both events E and F occurring. This is a useful example of how we can use these laws to derive other properties of probabilities. The two events from the example in the last slide are not mutually exclusive, since throwing a four would result in both events occurring, and indeed we see that the additive law does not give us the right answer for the probability of E or F. The first step is to note that we can always take two events and split them up into three mutually exclusive ones, as illustrated here with a Venn diagram, done for both the general case and for our example.

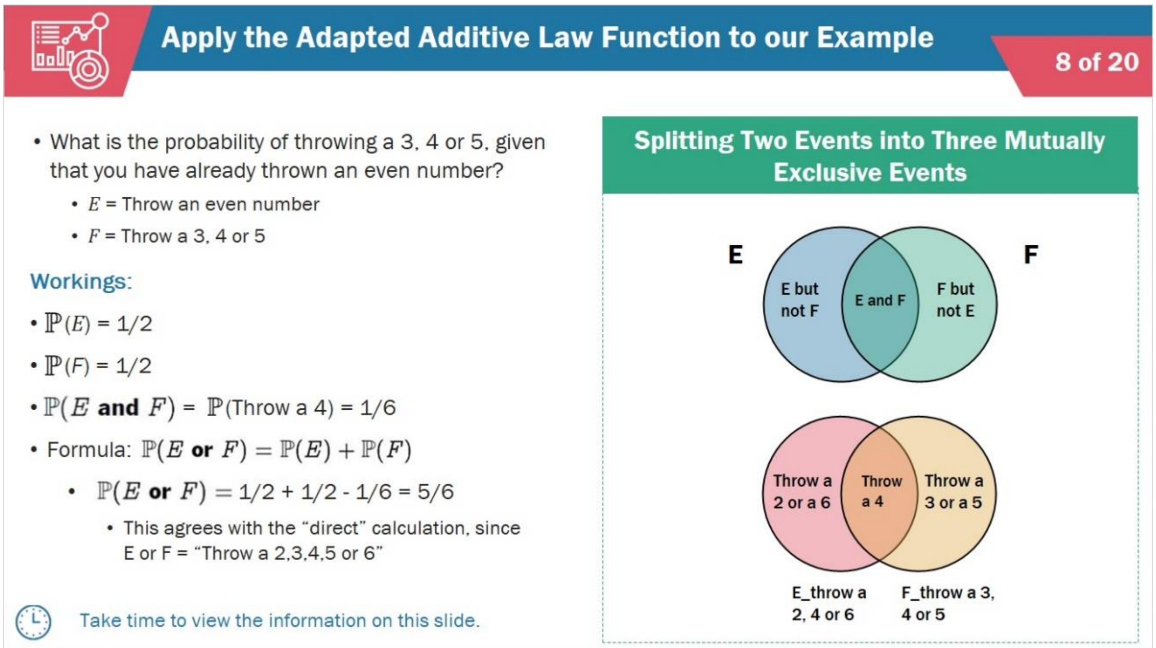
Slide 7: Creating Mutually Exclusive Events to Apply Additive Law



If we look at the event E by itself, we see that it can be split into two mutually exclusive ones: “E but not F” or “E and F”, which allows us to use the additive law to get an expression for the probability of E but not F. Repeating this for F gives us a similar expression for F but not E. As mentioned in the last slide, we can split up the event “E or F” into the three mutually exclusive events “E but not F”, “F but not E” and “E and F”. One and only one of these can occur at the same time.

By our additive law for mutually exclusive events, we can write the probability of E or F as the sum of the probabilities of these three events. Then we can substitute in our expressions for the probability of E but not F, and for F but not E, and this gives us the general case formula, that the probability of E or F is the sum of the probabilities of E and of F, minus the probability that both occur.

Slide 8: **Apply the Adapted Additive Law Function to our Example**

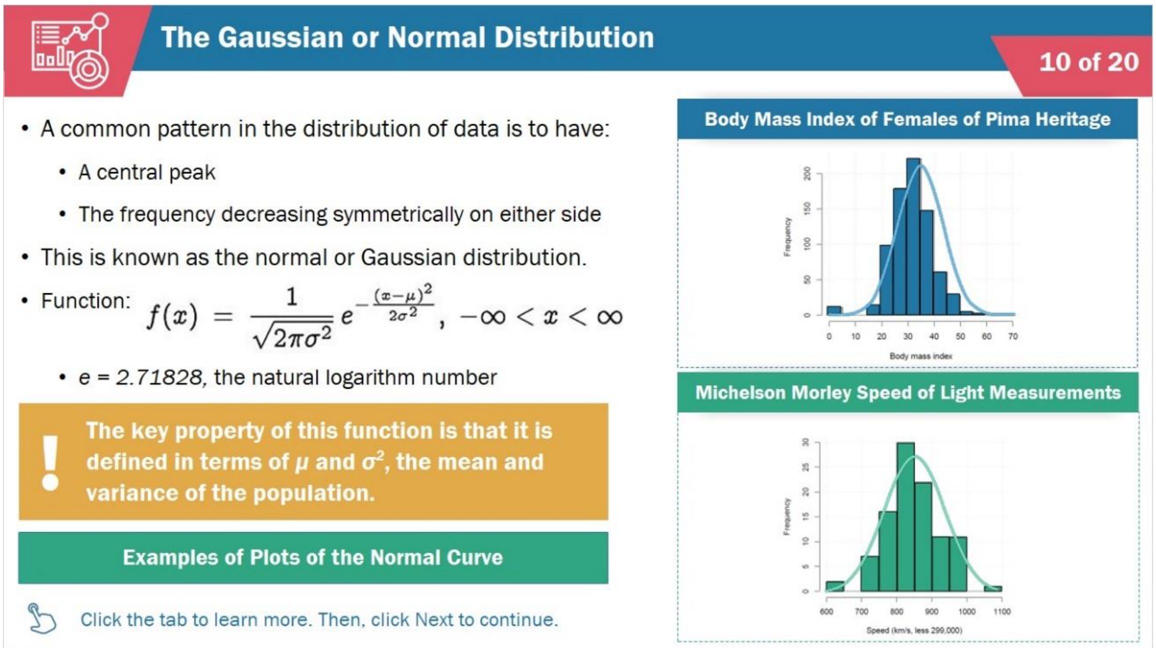


This formula gives us the sensible answer for our example, as seen here.

Slide 9: **Section: Normal Curve or Distribution**

Now we move on to the next part of the presentation, to revisit the normal curve or distribution as a description of the random variation in a population.

Slide 10: The Gaussian or Normal Distribution

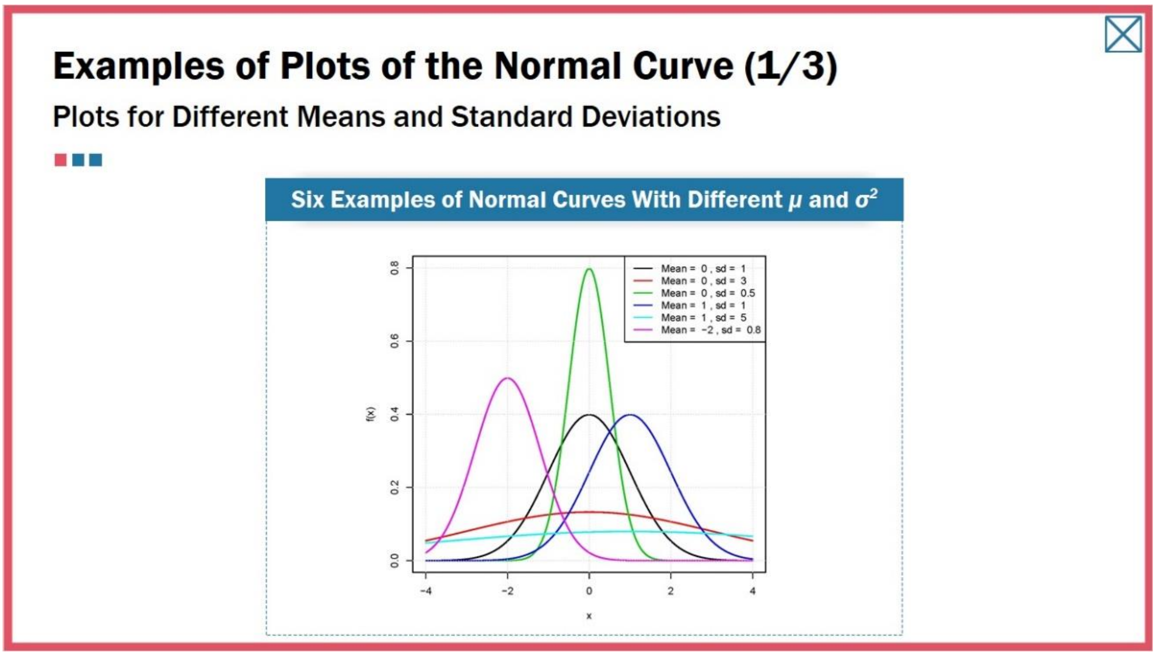


Recall that a very common pattern in the distribution of data is to have a central peak with the frequency decreasing symmetrically on either side of this peak. Two examples are shown here. This pattern is known as the normal or Gaussian distribution. Also note that we will use the terms ‘curve’ and ‘distribution’ interchangeably; both refer to the mathematical description of the random variation in a population.

In mathematics we can write down a function that represents an idealised version of the distribution of such data. If we calculated the value of this function for different values of x, and plotted them, with x on the horizontal axis and the function value on the vertical axis, we would see that it takes on the normal curve shape. The expression that we use is here, where e is the natural logarithm number, approximately 2.71828. The key property of this function for our purposes is that it is defined in terms of two numbers, usually denoted by the Greek letters mu and sigma squared. These are the mean and variance of the population.

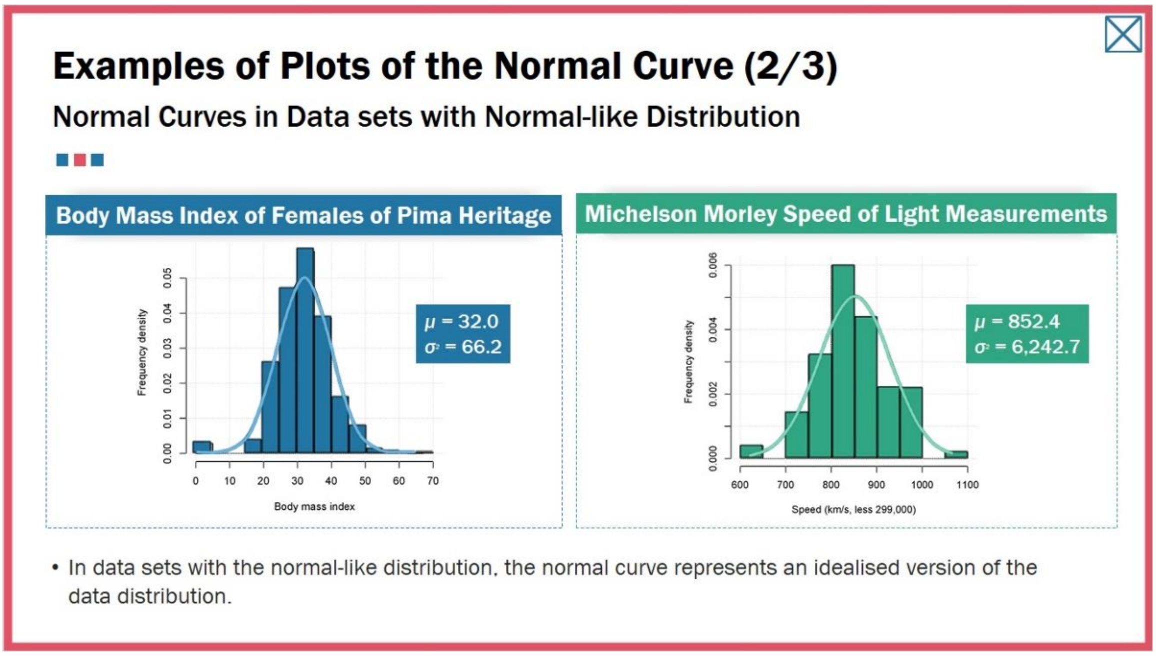
Click the tab to see examples of plots of the normal curve. When you are ready, click Next to continue.

Tab 1: **Examples of Plots of the Normal Curve**



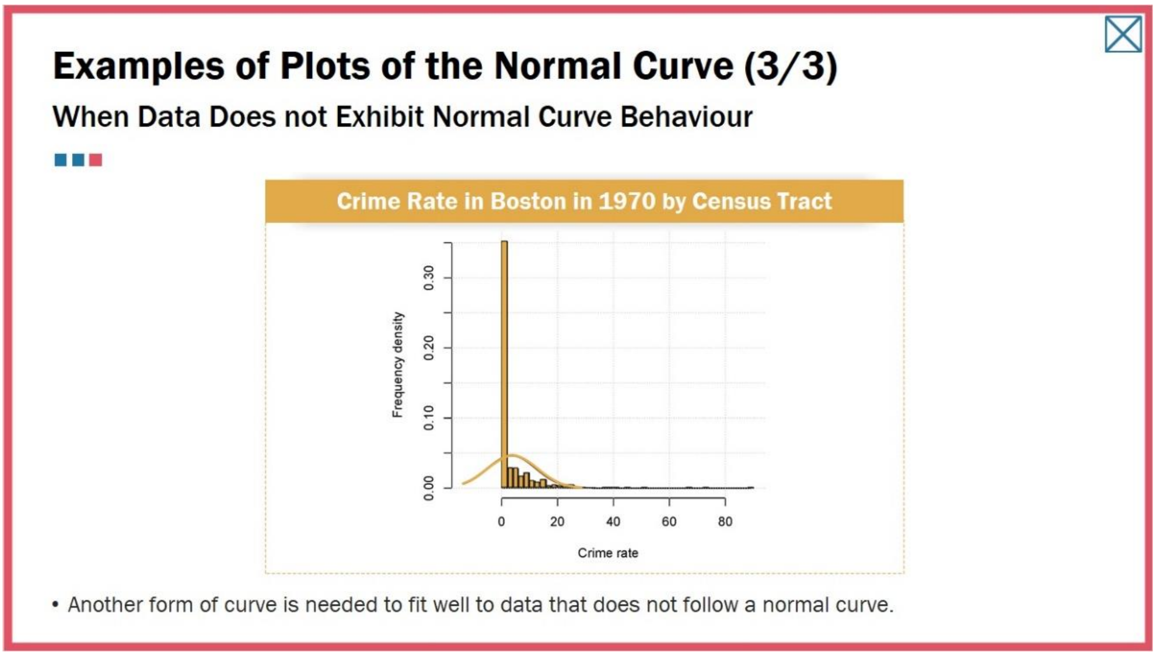
Here are six examples of plots of the normal curve for different means and standard deviations (the square root of the variance). One sees that in all cases the curve has the distinctive 'bell-shape' to it. The mean controls the location of the central peak, and the variance controls how peaked or flat the curve is about that peak. A larger variance means a population with more variation about the mean, hence a flatter curve.

Tab 1.1: **Normal Curves in Data sets with Normal-like Distribution**



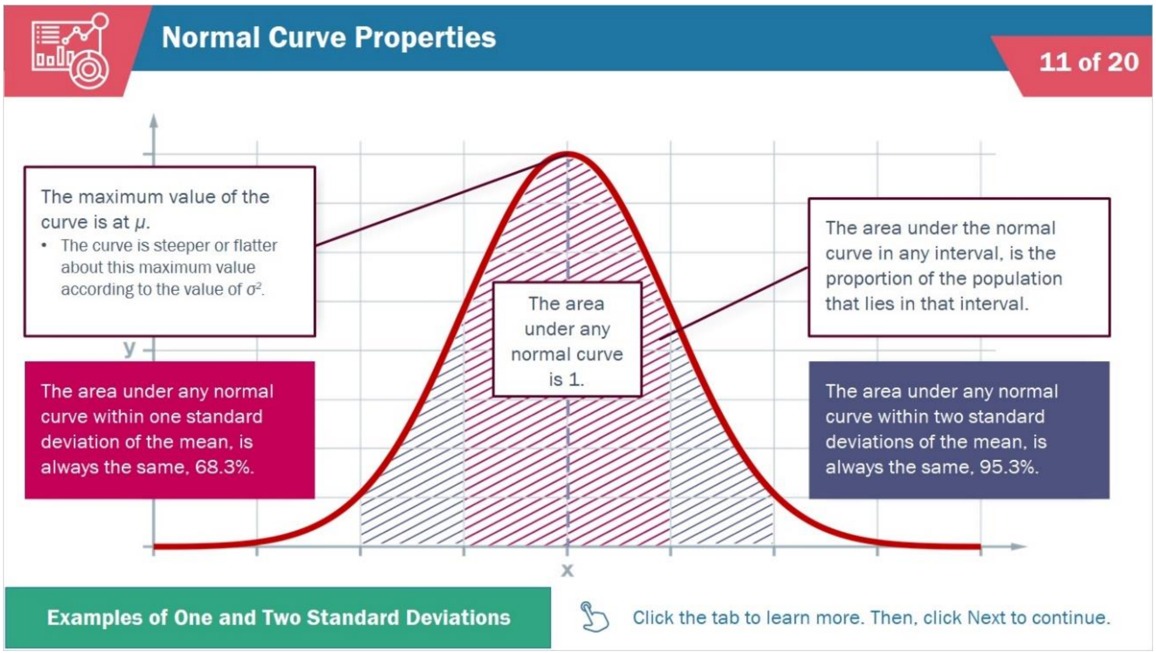
If we take the two example data sets with the normal-like distribution of observations, we can overlay the normal curve with the same mean and variance as the data set with the frequency density histogram (recall that this is where we scale the height of the histogram bars so that the total area of the histogram is 1) and indeed we see that we get a good match between the two. The normal curve does represent a smoothed-out or idealised version of the data distribution.

Tab 1.2: When Data Does not Exhibit Normal Curve Behaviour



Not all data follow a normal curve and that is illustrated here with the data from Boston on crime rates by district of the city. A normal curve with the same mean and variance as this data set is clearly not a good fit. We would need some other form of curve to fit well to these data.

Slide 11: Normal Curve Properties

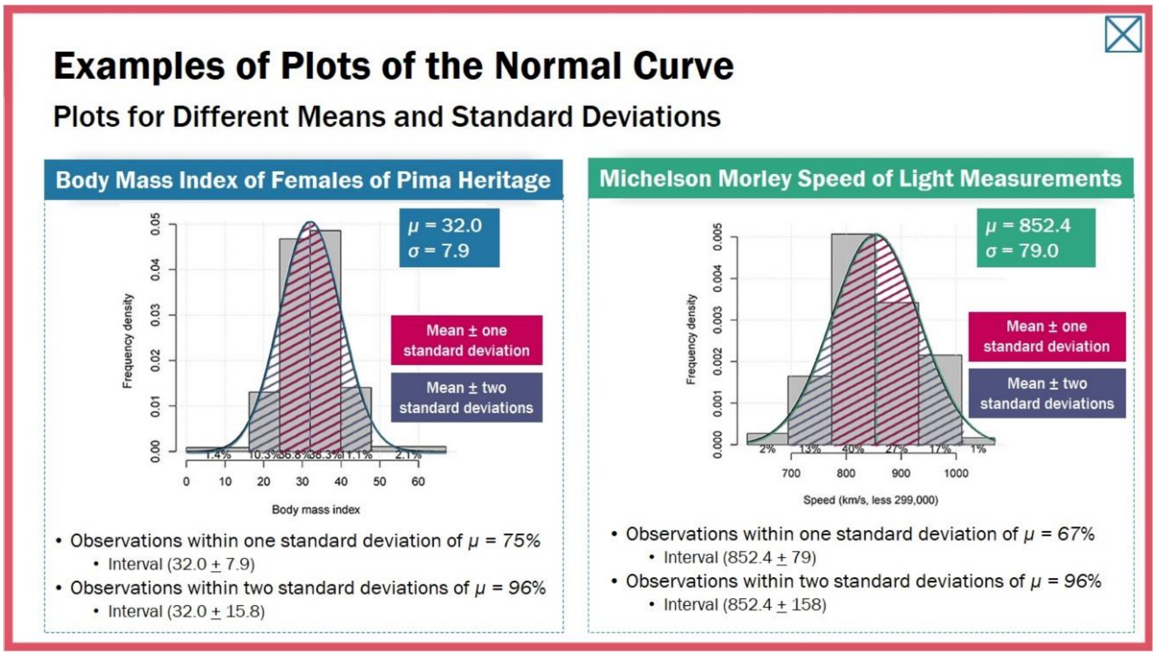


We have seen how the mean and variance values affect the shape of the normal curve. The maximum value of the curve is at mu and the curve is steeper or flatter about this maximum according to the value of sigma squared. There are some other properties of this curve that are useful. The first is that the area under any normal curve, like a frequency density histogram, is 1, as it represents every possible value that an observation from the population could take, from minus infinity to plus infinity. Further, just as the area of a single bar in the frequency density histogram is the proportion of observations in that interval, the area under the normal curve in any interval is the proportion of the population that lies in that interval. In other words, thinking about our frequency interpretation of a probability, it's the probability that an observation randomly sampled from the population takes a value in that interval.

Further, the area under any normal curve within one standard deviation of the mean, that is to say between the mean minus the standard deviation and the mean plus the standard deviation, is always the same, regardless of the mean and variance values. The value is about 68.3%, representing the probability of an observation lying in that interval. Similarly the area under the curve or probability of an observation within two standard deviations of the mean is always the same, about 95.3%.

Click the tab to see how these last two properties apply to our two data sets. When you are ready, click Next to continue.

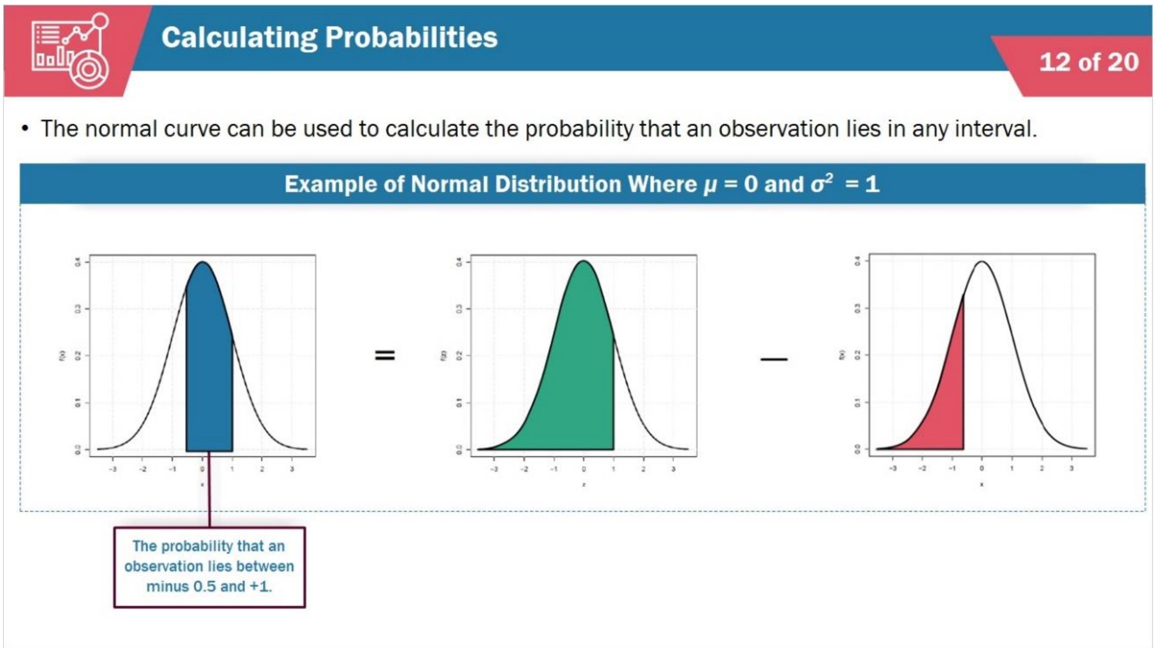
Tab 1: Examples of Plots of the Normal Curve



These last twp properties actually hold quite well for our two data sets that appear to follow the normal distribution, as can be seen here. In these plots, the histogram bins for each data set are set to be from the mean plus or minus one and two standard deviations, and the percentage of observations in each bin is written down. For our Pima heritage data set, 75% of observations are within one standard deviation of the mean and 96% of the observations are within two standard deviations of the mean.

For the light speed data, 67% of the observations are within one standard deviation of the mean and 96% of the observations are within two standard deviations of the mean.

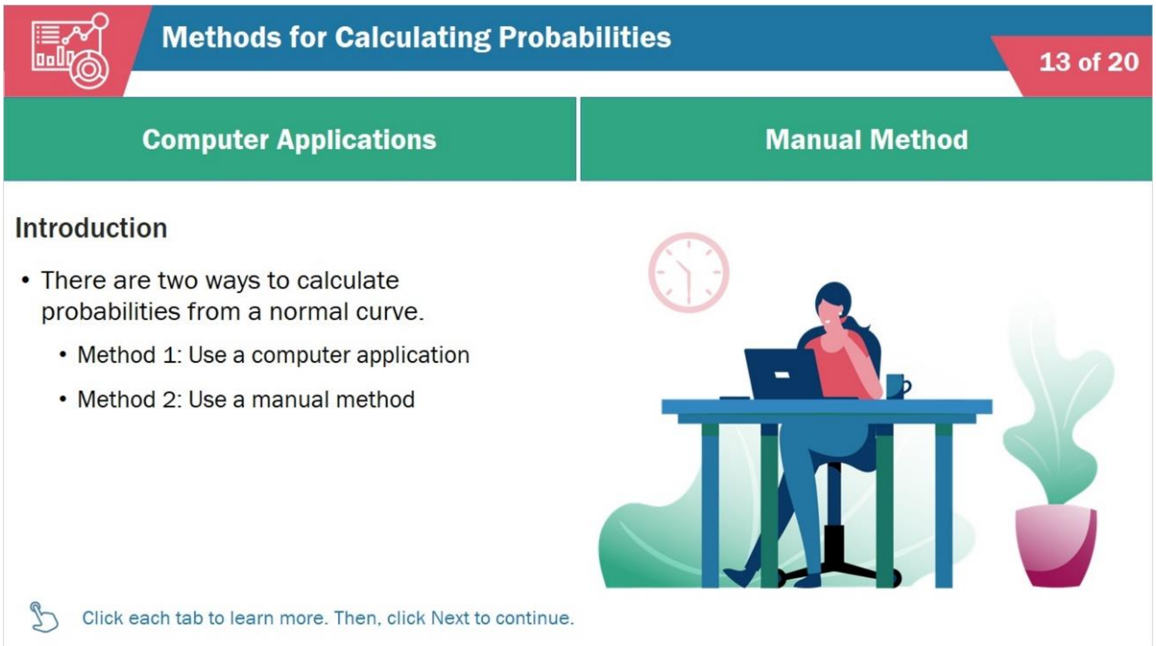
Slide 12: Calculating Probabilities



We can use the normal curve to calculate the probability that an observation lies in any interval. As we noted, this is just the area under the normal curve over that interval. Here we see an example where we have taken the normal distribution with mean 0 and variance 1. The shaded area is the probability that an observation lies between minus 0.5 and 1.

Another thing to note is that this area is the difference in the area to the left of 1 and the area to the left of minus 0.5 under the curve.

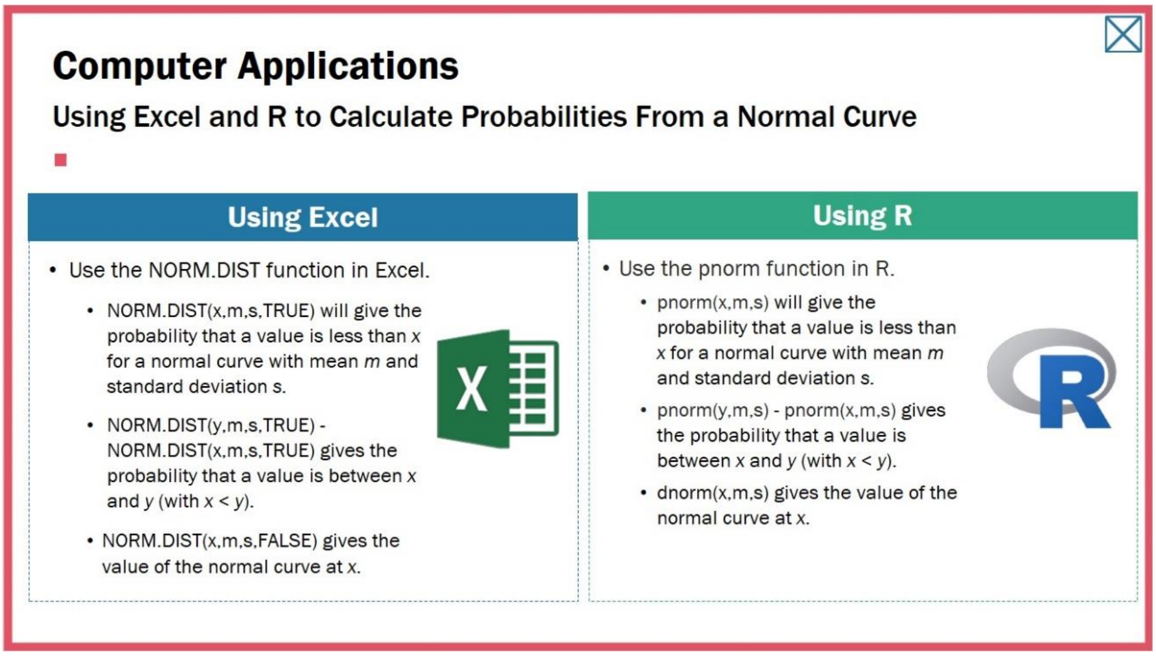
Slide 13: Methods for Calculating Probabilities



There are two ways to calculate these probabilities from a normal curve. The first, and easiest, is if you have access to a computer that runs software that calculates these areas for you. Both Excel and R have commands to calculate the probability, or area under a curve, of an interval. The second way is a manual method that has been used for much longer and pre-dates the availability of computers.

Click the tabs to learn about each method. When you are ready, click Next to continue.

Tab 1: Computer Applications



When using Excel to calculate the probability on an interval, , the NORM.DIST function gives you the area to the left of a value under a normal curve with the mean and standard deviation that you specify. In Excel, NORM.DIST(x,m,s,TRUE) will give the probability that a value is less than x for a normal curve with mean m and standard deviation s

Therefore NORM.DIST(y,m,s,TRUE) minus NORM.DIST(x,m,s,TRUE) gives the probability that a value is between x and y (with x < y).

Note that NORM.DIST(x,m,s,FALSE) gives the value of the normal curve itself at x.

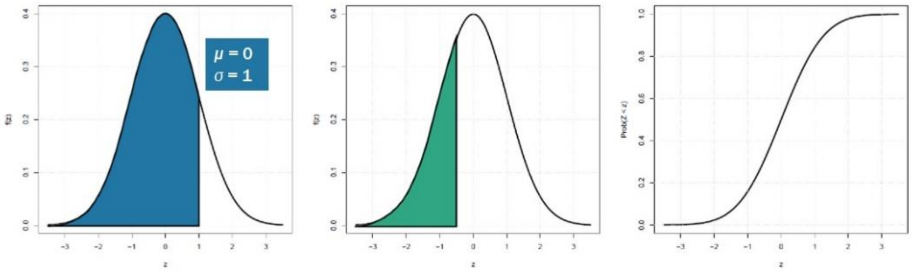
In R, the equivalent function is pnorm. By taking the difference of this function at two values, you get the area between them and hence the interval probability.

Tab 2: Manual Method

Manual Method (1/4)

Calculating the Probabilities From a Normal Curve Manually

- A normal distribution with mean = 0 and variance = 1 is called the standard normal distribution.
  - *Z* denotes a quantity that follows the standard normal distribution.



Example Using Standard Normal Distribution

The example uses the normal distribution with mean 0 and standard deviation 1, and this is called the standard normal distribution. We usually denote a quantity that follows such a distribution with Z.

Tab 2.1: Tabulating Values

Manual Method (2/4)

Tabulating Values

- Mathematicians evaluated the area to the left of a value for a standard normal distribution over many different values and tabulated them.

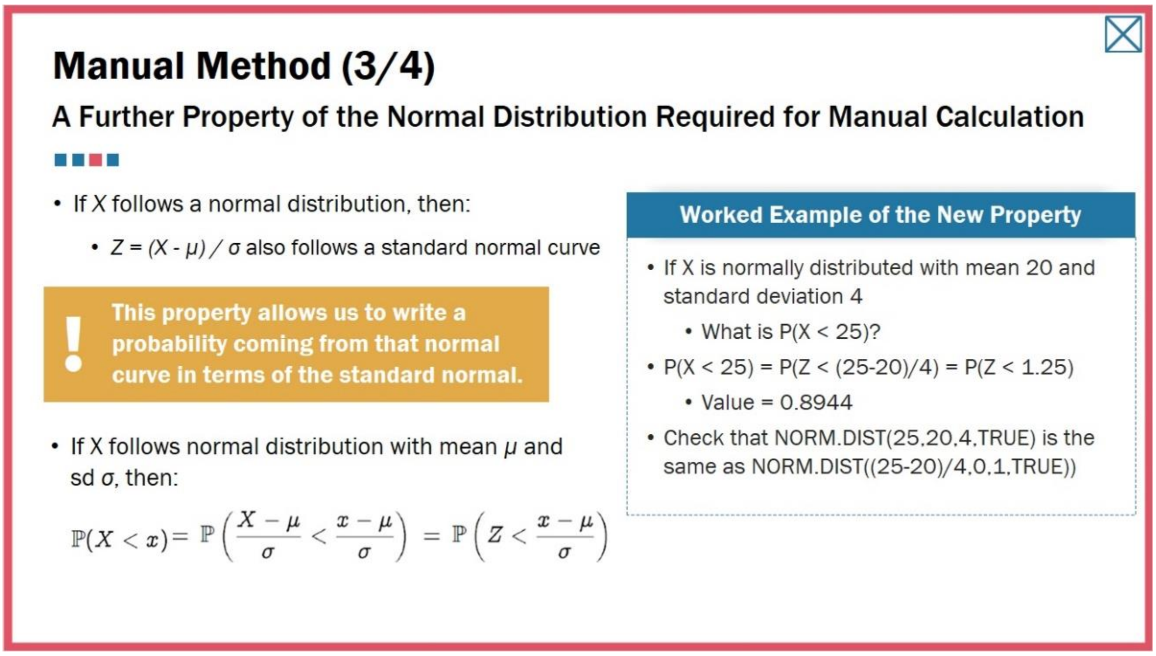
The table gives the value of P(Z < z), where Z is standard normally distributed. For example, P(Z < 1.23) =

z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817

Values of P(Z < z), Where Z is Standard Normally Distributed

Mathematicians evaluated the area to the left of a value for a standard normal distribution over many different values and put those in a table; a laborious task in the pre-computer age. This can be done in a variety of ways, such as plotting the curve on finely-graded graph paper so that the area can be counted up, or using approximate formula for the area that could be calculated by hand. Here we see such a table, actually generated by the NORM.DIST function in Excel in this case but in the past published in a book along with various other useful tables of values for statistics and probability. To determine the value of say P(Z<1.23), go to 1.2 in the z column and then move across the row, until you come to the corresponding number for column 0.03. The value is 0.8907.

Tab 2.2: A Further Property of the Normal Distribution Required for Manual Calculation



This manual method of calculating probabilities from the normal curve, requires some knowledge of a further property of the normal distribution. if some quantity, let's call it X, follows a normal distribution then the quantity that we get by subtracting off its mean and dividing by its standard deviation also follows a normal curve, and in particular it follows the standard normal (with mean 0, variance 1). This property allows us to write a probability coming from that normal curve in terms of the standard normal.

Suppose X follows a normal distribution with mean m and standard deviation s. Look at the probability of the event that the quantity X is less than some value x.

So, if we want P (X<x), then we can write:

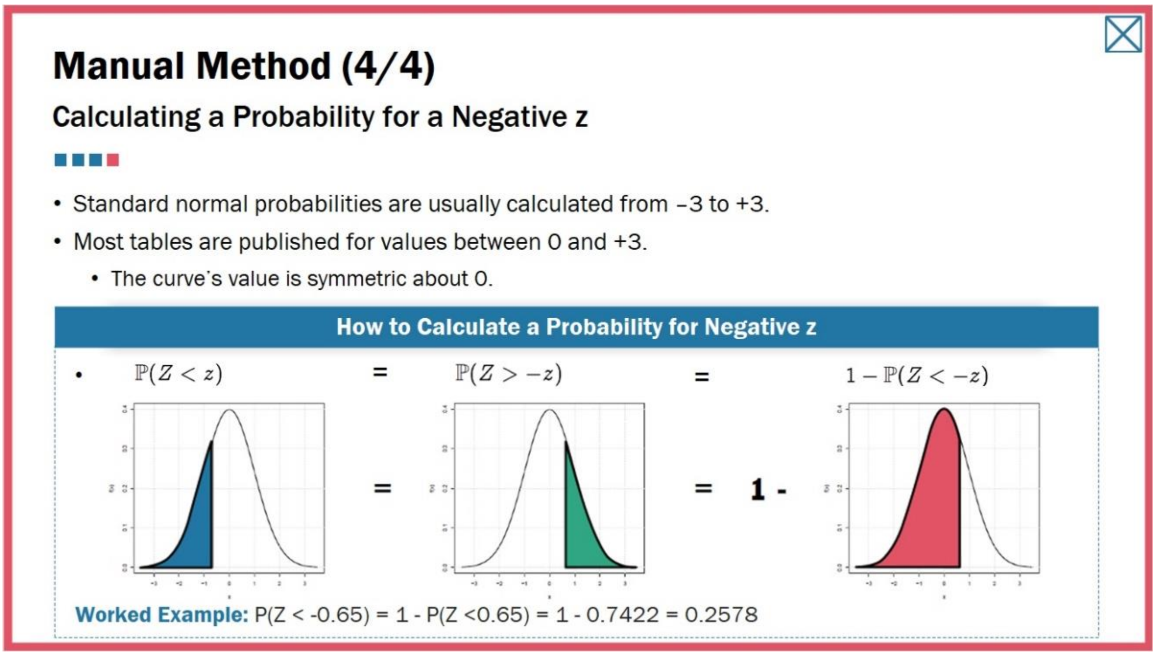
$$P(X < x) = P\left(\frac{X-\mu}{\sigma} < \frac{x-\mu}{\sigma}\right) = P(Z < \frac{x-\mu}{\sigma})$$

A rather roundabout and equivalent way of stating the same event is to say that (X – mu)/sigma is less than (x – mu)/sigma. But the property of the normal distribution means that (X – mu)/sigma follows a standard normal curve. Hence the probability that X is less than some value x is the same as the probability that a standard normal distributed value is less than (x – mu)/sigma that we could look up from our table.

Let's work through an example. If X is normally distributed with mean 20 and standard deviation 4, what is P(X<25)?

Let's slot the values into the function. To determine the value of Z<1.25, we go back to our table and look up the value for 1.25 which is 0.8944. You can check this yourself using the NORM.DIST function in Excel, as illustrated on the slide.

Tab 2.3: Calculating a Probability for a Negative z

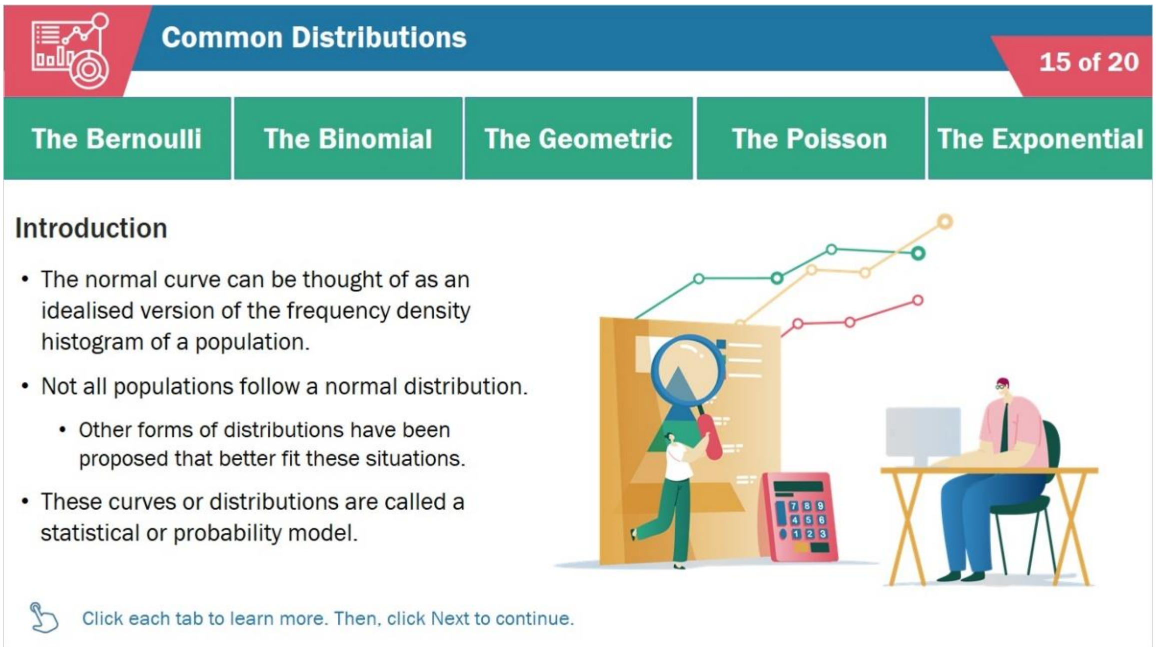


Standard normal probabilities are usually calculated from -3 to +3, because the area to the left of -3 is close to 0 and the area to the left of +3 is almost 1. Most published tables, however, just list values from 0 to +3. What happens if we want to calculate a probability for a negative z, for example something like P(Z < -0.65)? Fortunately, another property of the standard normal curve comes to help us; the curve's value is symmetric about 0 or, in other words, the value of the curve at a positive value is the same as that at the negative of that value. This means that the area to the left of a value is the same as the area to the right of its negative. Since the total area under a normal curve is 1, this is 1 minus the area to the left of its negative. The example illustrates this for the case of P(Z < -0.65).

Slide 14: **Section: Statistical Models**

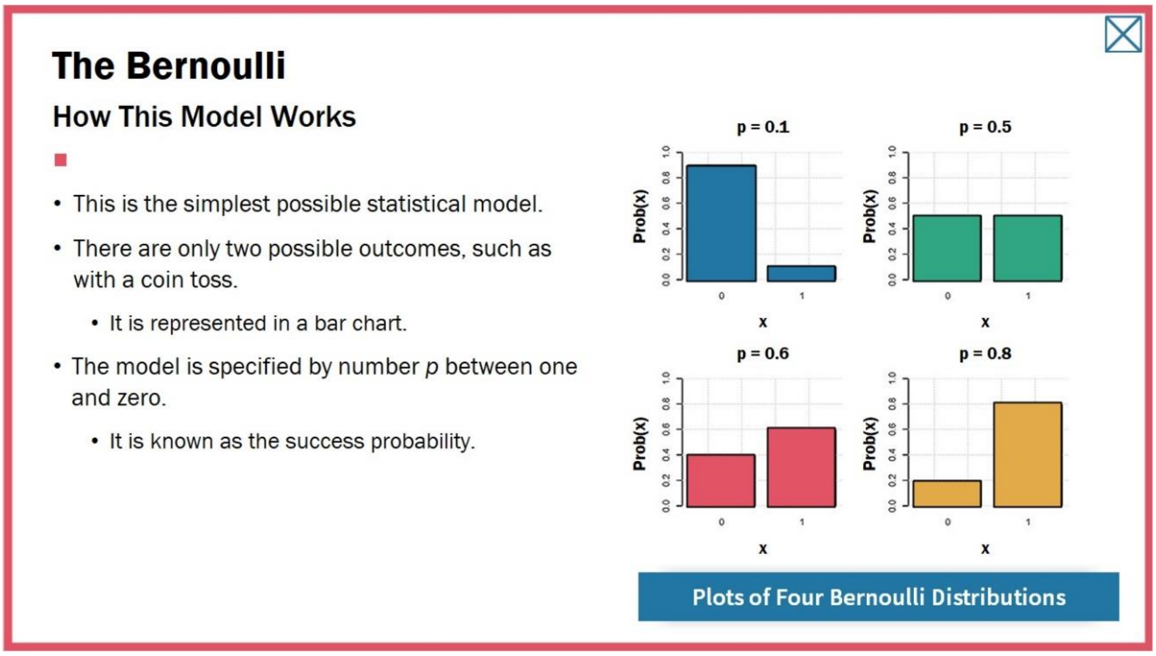
We are now going to look at statistical models. They are one of the most used and useful tools in statistics as they allow us to both understand random variation better and develop statistical tools for data analysis that have a solid mathematical basis.

Slide 15: **Common Distributions**



As we have noted, the normal curve can be thought of as an idealised or smoothed-out version of the frequency density histogram of a population. As we have also seen, not all populations follow a normal distribution and so many other forms for distributions have been proposed that better fit these other situations. We call these curves or distributions a statistical or probability model Click the tabs to learn about some of the other most common distributions that we encounter. When you are ready, click Next to continue.

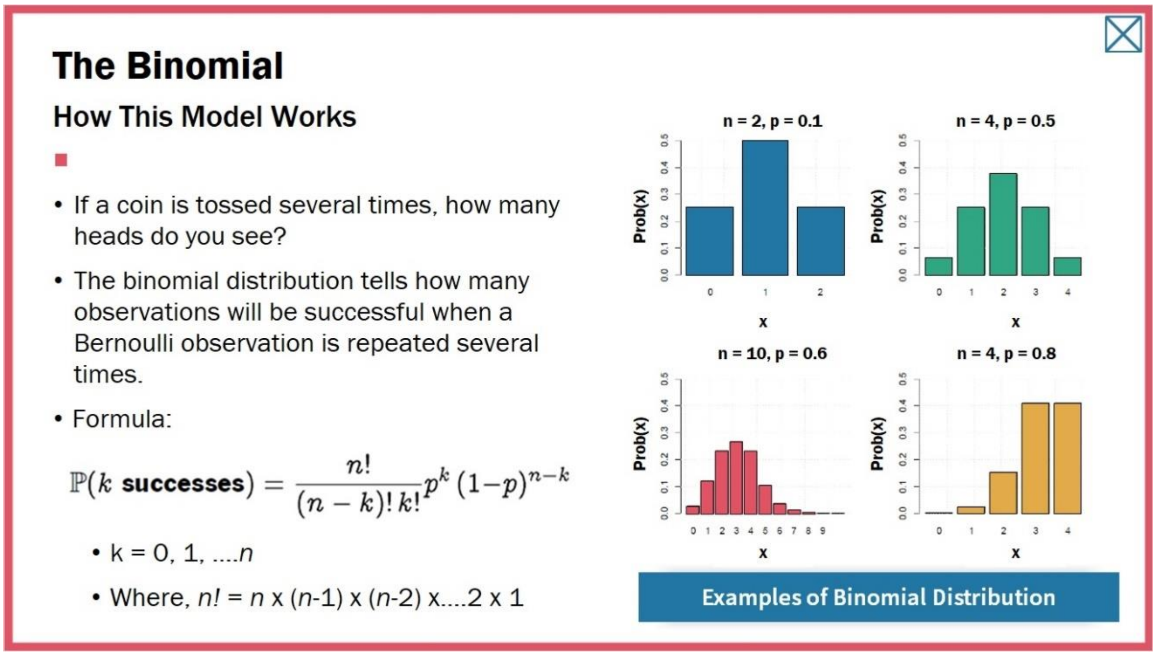
Tab 1: The Bernoulli



What is the simplest possible random situation? It’s the coin toss: there are only two possible outcomes and we don’t know which will occur. This is the Bernoulli distribution. Note that it makes no sense to plot a curve associated with this type random variation, as there is not a continuous set of possible values but only two of them, usually given the values 0 and 1. We represent this distribution as a bar chart, with the probability for each of the two values represented by the height of the bars.

This distribution or model is specified by some number p, between 0 and 1, that is the probability of getting a 1. This is often called the success probability, with the outcome ‘1’ usually associated with a successful or more desirable outcome. The probability of the other value, or 0, is 1 minus this since there are no other possible values. Here we see plots of four Bernoulli distributions with different values of p.

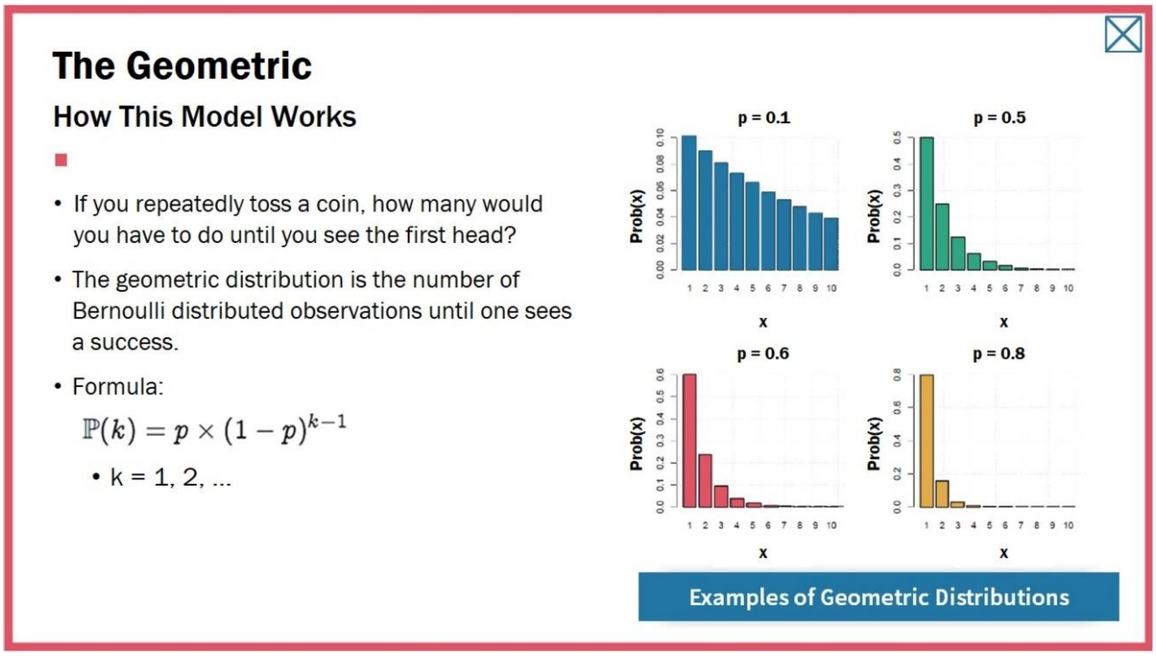
Tab 2: The Binomial



Suppose one tosses a coin several times. How many heads do we see? More generally, repeat a Bernoulli observation several times, and how many of those observations are 1 (or are a 'success')? The binomial distribution tells you this. It is another discrete distribution where one cannot draw a continuous curve like the normal; if there are n Bernoulli observations then the possible values of the binomial are the integers from 0 to n.

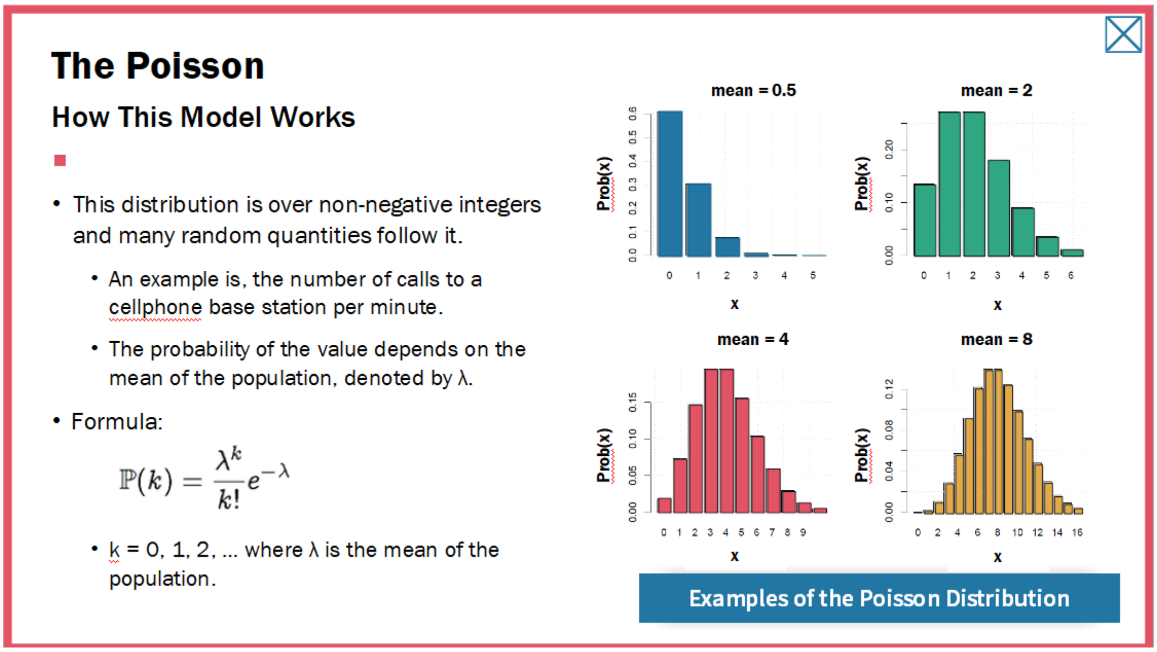
The formula for the probability of k successes is displayed on the slide - $\frac{n!}{(n-k)!k!}\, p^k(1-p)^{n-k}$. The notation n-exclamation mark, called n factorial, is the product of the first n integers. Some examples of the binomial distribution are given here as well for different numbers of observations n and success probability p.

Tab 3: **The Geometric**



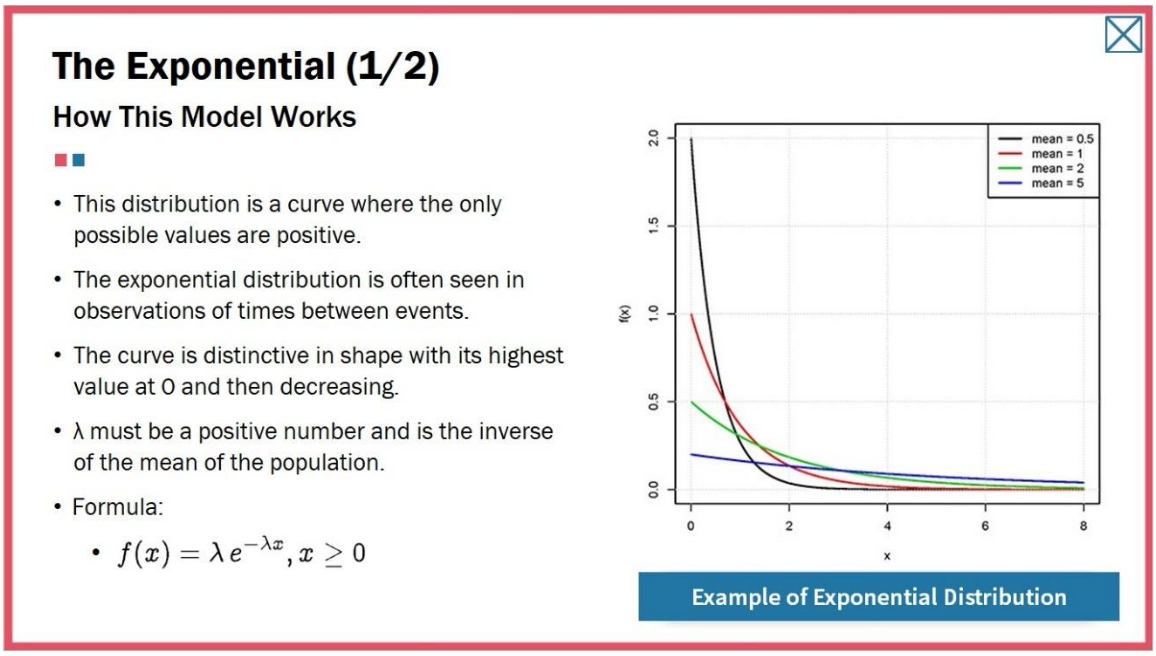
Now repeatedly toss a coin until you see the first head. How many tosses did that take? The geometric distribution describes that number of tosses. More generally, one can say that the geometric is the number of Bernoulli distributed observations until one sees a success. The formula for the probability that it takes k observations for this to happen is displayed on the slide, $\mathbb{P}(k) = p \times (1-p)^{k-1}$, along with examples for different p. For the geometric, the possible values are the positive integers.

Tab 4: **The Poisson**



The Poisson distribution is the distribution over the non-negative integers (so 0, 1, 2 and so on) and we find that many random quantities follow it. Examples are the number of calls to a cell phone base station per minute, number of transactions made with a major internet retailer per second, or the number of particles emitted by a radioactive source. The probability of an observation taking the value k is displayed on the slide. The probability of each value is given in the slide, $\mathbb{p}(k) = \frac{\lambda^k}{k!}\,e^{-\lambda}$, and depends on the mean of the population, usually denoted by the Greek letter lambda. Four examples of this distribution for different means are given.

Tab 5: The Exponential Curve or Distribution



The final distribution that we look at is also a curve, like the normal distribution, but one where the only possible values are positive. The exponential distribution is often seen in observations of times between events. The time between server failures in a data centre, or the time between vehicles on a highway, are often exponentially distributed.

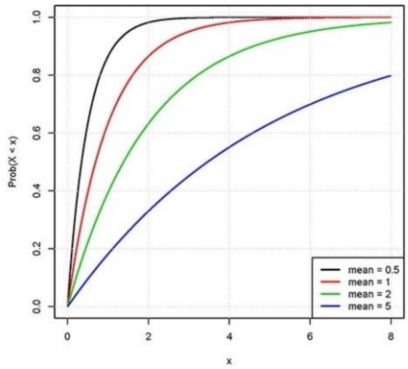
The curve has a distinctive shape, with its highest value at 0 and then decreasing. The rate at which it decreases is governed by the value of a parameter that is also usually denoted lambda. Lambda must be a positive number and in fact lambda is the inverse of the mean of the population. The smaller the mean (and so the larger is lambda), so the more steeply the curve declines to 0. Like the normal curve, there is an area of 1 under all of these curves, regardless of the value of lambda. The formula for the curve is displayed on the slide, $f(x) = \lambda e^{-\lambda x}, x \geq 0$.

Tab 5.1: Formula for the Area Under the Curve to the Left of any Value

The Exponential (2/2)

Formula for the Area Under the Curve to the Left of any Value

- The area under the curve to the left of any value has the following formula:
  - $\mathbb{P}(X \leq x) = 1 - e^{-\lambda x}, x \geq 0$



Example of Four Exponential Curves

Once nice property of the exponential curve is that there is a nice formula for the area under it to the left of any value. The probability that a value is less than x is given by the expression on the slide. No such simple formula exists for the normal curve.

Slide 16: Calculating Probabilities in Excel

Calculating Probabilities in Excel

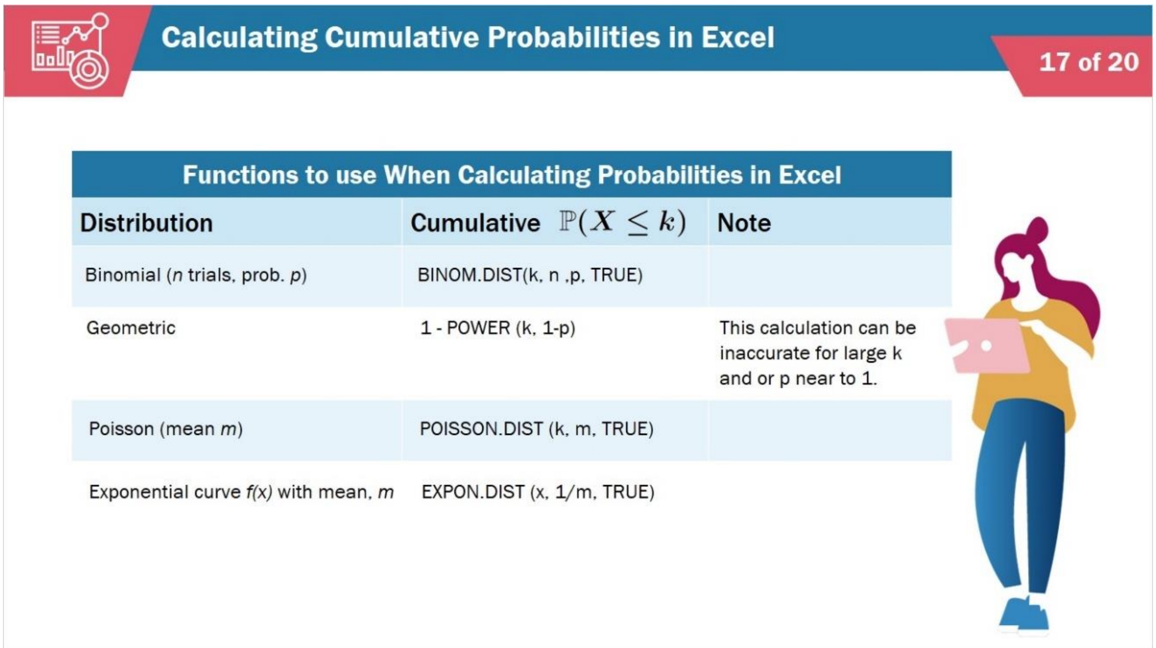
16 of 20

Functions to use When Calculating Probabilities in Excel

Distribution	P(k)	Note
Binomial (*n* trials, prob. *p*)	BINOM.DIST(k, n ,p, FALSE)	
Poisson (mean *m*)	POISSON.DIST (k, m, FALSE)	
Geometric	p * POWER (k-1, 1-p)	This calculation is not always accurate.
Exponential curve *f(x)* with mean, *m*	EXPON.DIST (x, 1/m, FALSE)	

Excel has explicit functions for calculating the binomial and Poisson probabilities, given the values of the parameters that are needed to specify the distribution, and we can use other functions to calculate the geometric probability with the warning that this latter calculation may not be accurate in certain circumstances. Excel also has a function to calculate the value of the exponential curve with a given mean.

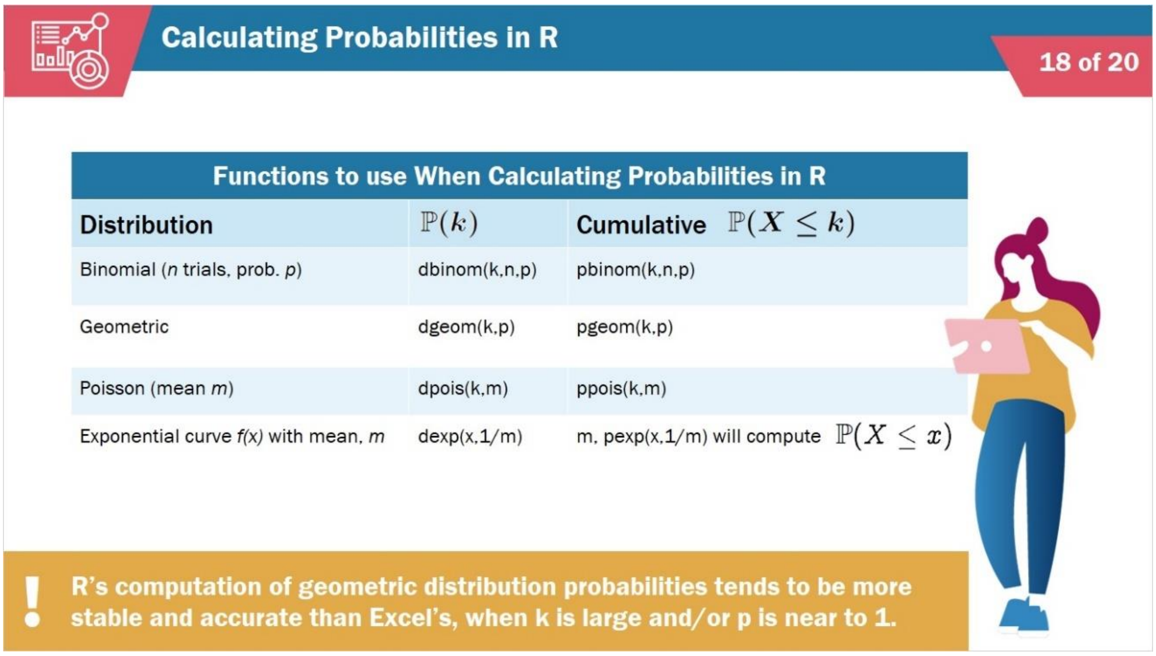
Slide 17: Calculating Cumulative Probabilities in Excel



Distribution	Cumulative $\mathbb{P}(X \leq k)$	Note
Binomial (*n* trials, prob. *p*)	BINOM.DIST(k, n ,p, TRUE)	
Geometric	1 - POWER (k, 1-p)	This calculation can be inaccurate for large k and or p near to 1.
Poisson (mean *m*)	POISSON.DIST (k, m, TRUE)	
Exponential curve *f(x)* with mean, *m*	EXPON.DIST (x, 1/m, TRUE)	

Excel can also calculate the so-called cumulative probabilities for the binomial, geometric and Poisson distributions, that is the probability that value is less than or equal to any integer k. It can also calculate the area under the exponential curve that gives one the probability of being less than or equal to any value x.

Slide 18: Calculating Probabilities in R

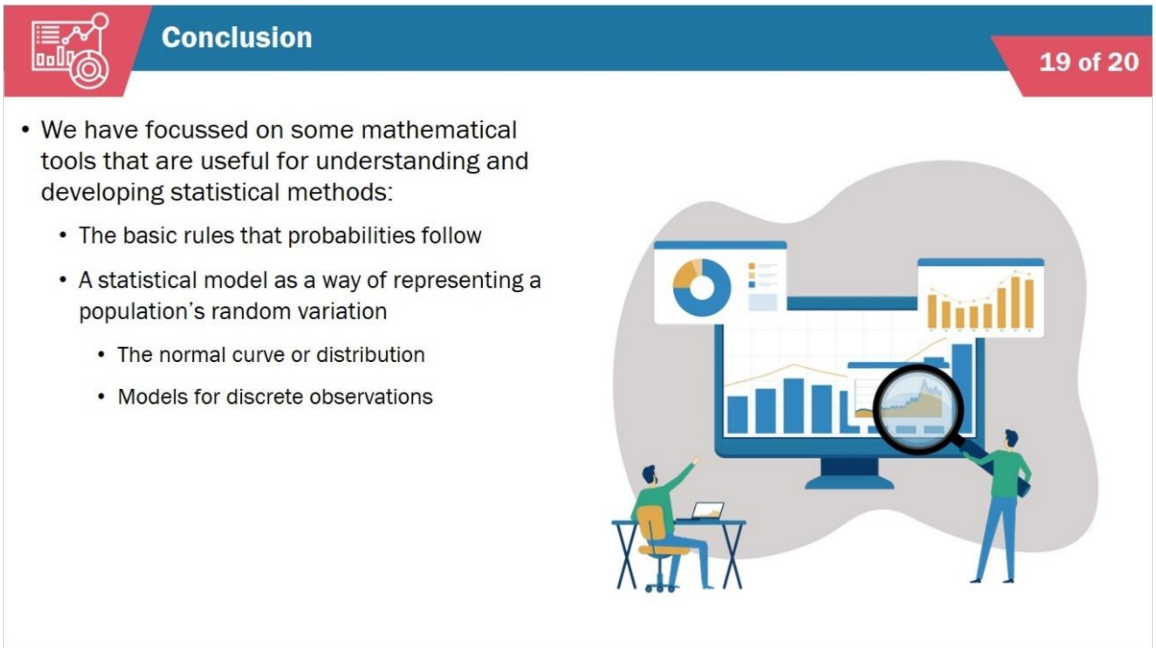


Distribution	$\mathbb{P}(k)$	Cumulative $\mathbb{P}(X \leq k)$
Binomial (*n* trials, prob. *p*)	dbinom(k,n,p)	pbinom(k,n,p)
Geometric	dgeom(k,p)	pgeom(k,p)
Poisson (mean *m*)	dpois(k,m)	ppois(k,m)
Exponential curve *f(x)* with mean, *m*	dexp(x,1/m)	m, pexp(x,1/m) will compute $\mathbb{P}(X \leq x)$

! R's computation of geometric distribution probabilities tends to be more stable and accurate than Excel's, when k is large and/or p is near to 1.

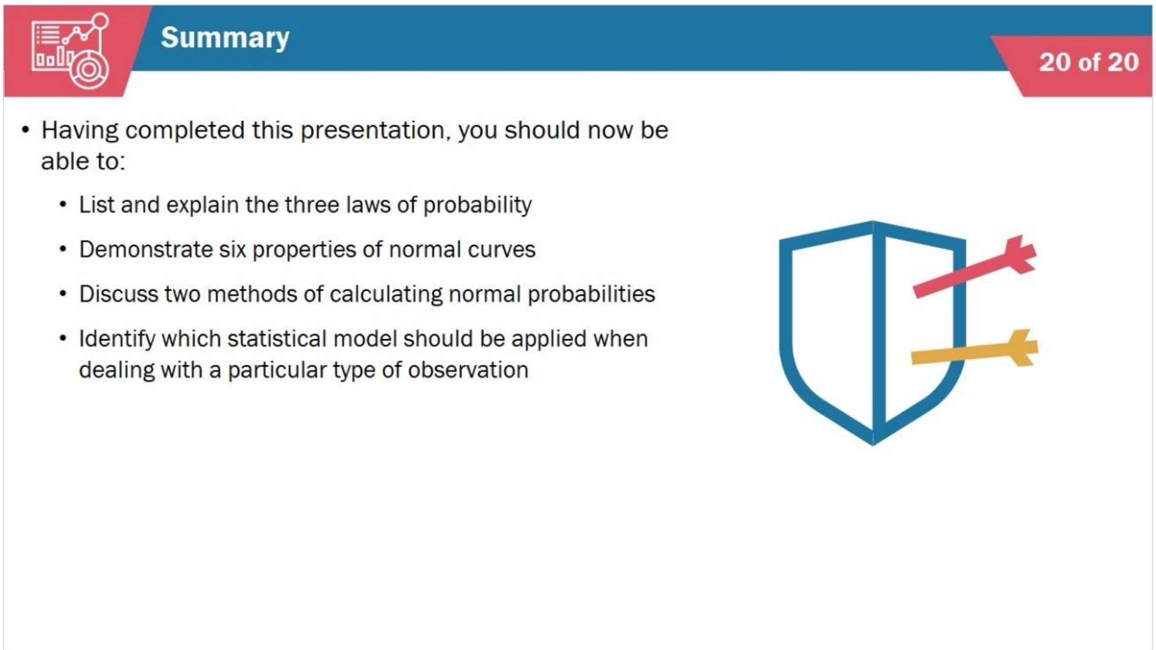
Here, we give the equivalent functions used in R. R's computation of geometric distribution probabilities tends to be more stable and accurate than Excel's when k is large and p near to 1.

Slide 19: **Conclusion**



In this presentation, we have focussed on some of the mathematical tools that are useful for understanding and developing statistical methods. We looked at the basic rules that probabilities follow, and then looked at the idea of a statistical model as a way of representing a population's random variation. Most time was spent with the normal curve or distribution, as this is the most important one, but we also looked at some of the others, in particular some models for discrete observations such as the Bernoulli and Binomial.

Slide 20: **Summary**



Having completed this presentation, you should now be able to:

- List and explain the three laws of probability
- Demonstrate six properties of normal curves
- Discuss two methods of calculating normal probabilities
- Identify which statistical model should be applied when dealing with a particular type of observation