AUDITING SIMPLE RISK ASSESSMENTS

This chapter introduces the most basic ideas of probability and risk and shows how they can help us audit simple risk assessments.

These are the sort of casual risk assessments that pop up in conversation and on risk registers. Even at this simple level you will find a lot of surprises and helpful insights.

To start with, in the world of business, 'risk' has a high profile and 'probability' is a word a lot of people try to avoid. In the world of mathematics the situation is reversed, with 'probability' the undisputed king and 'risk' an afterthought, sneaking in from theories about investment portfolios.

As you read on, remember how this book is designed. It's a series of concepts and terms, each of which will help you in your work. Tackle them in order, patiently and carefully. Your objective is to learn as much as you can, not to finish the book as quickly as possible.

1 PROBABILITIES

A lot of ideas about **probabilities** are controversial among theorists or take a while to understand, but what we know for certain is that probabilities *work*. There are people who talk about and benefit from using **probabilities** and this has been true for hundreds of years.

One of the great pioneers of the mathematics of **probability** was Frenchman Pierre-Simon Laplace (1749–1827). In the introduction to his book, *Théorie Analytique des Probabilités*, he wrote that 'que la théorie des probabilités n'est, au fond, que le bon sens réduit au calcul,' which means 'the theory of probability is just common sense reduced to calculation.'

> **' what we know for certain is that probabilities work '**

Probabilities are stated about things that might happen or, more broadly, about things that might be true. For example, consider the statement 'the probability that Happy Boy wins the 3.15 p.m. race at Kempton Park is 0.12.' The thing that might happen is Happy Boy winning. The statement that might be true is that 'Happy Boy will win'.

It is also generally agreed that **probabilities** are numbers between 0 and 1 inclusive and that a probability of 0 means something is considered certainly not true or not going to happen, while a probability of 1 means it certainly is true or certainly will happen.

Sometimes **probabilities** are expressed as percentages between 0 and 100%. Sometimes they are given as odds, as in '3:1 against', which translates to a probability of 0.25, or 25% if you prefer. Sometimes they are given as proportions as in 'one in four', which is also a probability of 0.25.

Take care when translating between different styles. In the song 'Five to One' by the Doors, Jim Morrison equates 'five to one' with 'one in five', but of course that should be one in six.

2 PROBABILISTIC FORECASTER

It is also clear that **probabilities** come from many sources, which I'll call **probabilistic forecasters**. Mostly they come from people (e.g. weather forecasters, tipsters, research companies, managers in companies), from mathematical formulae, and from computer systems. Some of these **probabilistic forecasters** restrict themselves to a very narrow topic, while others are prepared to give **probabilities** for a wider range of propositions or outcomes.

One question of great interest to auditors and many others is how good the **probabilities** from a particular **probabilistic forecaster** are.

3 CALIBRATION (ALSO KNOWN AS RELIABILITY)

How can you assess the **probabilities** provided by a **probabilistic forecaster**? There are two ways:

1 Look at how the **probabilities** are worked out (which includes looking at any data used).
2 Compare the **probabilities** to reality and see how well they match up.

The second method is the easiest to understand and is easy to do if you have enough data. You can't make any assessment from just one example unless the **probabilistic forecaster** says something is certain and turns out to be wrong.

However, if you have lots of **probabilities** from the same source and you know what actually happened or what the truth was then you can calculate various scores that show how good the source is.

There are two main qualities that good **probabilities** must possess, and one of them is **calibration**.

If a **probabilistic forecaster** of some kind is well **calibrated** then, over time, the frequencies of actual results will agree with the **probabilities** given. For example, suppose for a year a forecaster gives a **probability** of rain tomorrow and we record whether or not there was rain. The forecaster is perfectly **calibrated** if it rained on 10% of the days when the forecaster gave a **probability** of 0.1 of rain, rained on 20% of the days when the forecaster said the **probability** of rain was 0.2, and so on. The extent to which the proportions of days with rain agree with the **probabilities** given for those days is **calibration**.

There are a number of formulae for calculating overall **calibration** across a range of forecasts, but it is a good idea to look at **calibration** at each level of **probability**. A good *average* **calibration** score may hide problems, most likely with poor **calibration** for extreme events.

4 RESOLUTION

Furthermore, **calibration** is not a complete measure of good **probabilities**.

Imagine that, over a typical year, it rains on half the days over a particular town. Every day the forecaster says the **probability** of rain is 0.5, regardless of the season or recent weather, thus demonstrating high **calibration**. We expect more don't we?

The extra thing we expect is that the forecast is responsive to conditions and when the opportunity arises to give **probabilities** for rain that are higher or lower than average the forecaster does so, and in the right direction. These more informative **probabilities** are said to have higher **resolution**. Again, there are alternative formulae for calculating **resolution**.

Higher **resolution** is usually achieved by taking more circumstances into consideration. The weather forecaster could consider not only the identity of the town, but also the season and recent weather. If the forecaster is clever enough to reach the limit of what can be predicted from these circumstances it might be time to gather additional data, perhaps from rainfall radar, weather stations out to sea, and from satellites.

However, there is a limit to how far this can be taken. The more circumstances the forecaster chooses to use, the harder it is to adjust for them all accurately because there are fewer directly comparable past experiences to use as a guide.

A key point to understand is that there is no such thing as *the* **probability** of something happening or being true. We must always think about the **probability** given what knowledge of circumstances we choose to take into consideration, and there are always options to choose from.

> ***there is no such thing as the probability of something happening or being true***

The perfect **probabilistic forecaster** would give **probabilities** of rain of 1 or 0, and would always be right. These **probabilities** would have maximum possible **resolution** and **calibration**.

Incidentally, published examples illustrating **calibration** and **resolution** are nearly always in terms of weather forecasting because that is the field of study where these ideas have been developed, but they apply to any **probabilities**.

5 PROPER SCORE FUNCTION

If you want to motivate a forecaster to give you well **calibrated**, high **resolution probabilities** and want to give some kind of bonus as encouragement then you need to use a **proper score function**.

This is a formula that calculates an overall measure of forecasting skill that gives the forecaster no incentive to lie and every incentive to give the best **probabilities** possible. The Brier Score and the Ignorance function (a logarithmic score) are both **proper score functions**.

Ignorance is a function based in information theory and shows the amount of information, in bits, that learning the outcome provides. For example, if you are certain that an outcome will happen and it does then you receive no information, i.e. you learn nothing you don't already know. However, if your **probability** for that outcome is less than 1 then you will learn something from seeing it happen. If you are convinced that something is impossible and yet it happens then your Ignorance is infinite, an interesting comment on closed mindedness.

probabilities are … particularly good for thinking about what policies we should adopt

Ignorance can also be interpreted as the time to double your money by betting on outcomes where all outcomes carry equal payouts. Even more interesting is that if you are betting against someone else then your Ignorance needs to be lower than theirs to expect to gain money! Clearly, the quality of **probabilities** has practical importance.

As I mentioned, any approach to assessing **probabilities** needs lots of examples of **probabilities** to work with. Even an idiot can guess right occasionally, so **probabilistic forecasters** need to be judged over a longer term. I often think that **probabilities** are more helpful as a guide to what we should expect over a long series of outcomes, so they are particularly good for thinking about what policies we should adopt.

Probability assessments also need to be made across a defined set of forecasting tasks. For example, it would be grossly unfair to assess a weather forecaster's **calibration** using **probability** judgements for the outcomes of financial investments.

6 AUDIT POINT: JUDGING PROBABILITIES

When people get to practice giving **probabilities** and receive feedback they usually get better at it.

The ideas of **calibration** and **resolution** show that we can judge a person's ability to provide **probabilities**, even if they are just based on gut feelings.

However, to do this we need a reasonable amount of data about **probabilities** they have given and what actually happened. It is also inappropriate for forecasts about things that people will try to change in response to the forecasts.

Some organizations would find that they do have these data and could work out **calibration** and **resolution** numbers, as well as plot graphs showing how **probabilities** given compared to reality.

If that's possible and it hasn't been done, shouldn't it be considered? **Probabilities** might turn out to be surprisingly well **calibrated,** perhaps even to the extent that people feel they can be used in cost-justifying investments in controls. Alternatively, it may be that feedback would be useful for improving the quality of **probabilities** people work with.

7 PROBABILITY INTERPRETATIONS

Not everyone who uses **probabilities** interprets them in the same way and misunderstandings can occur with practical and painful consequences.

‘misunderstandings can occur with practical and painful consequences’

The explanations below focus on what most people actually think and do today, rather than going through all the many proposals made by philosophers, scientists, lawyers, and others down the centuries.

Unless you've studied the meaning of **probabilities** in great depth do not assume you know this already!

8 DEGREE OF BELIEF

In everyday conversations we often say 'probably'. For example, last weekend I was introduced to a man at a party whose name was 'Charles', though I'm not entirely sure now, it was *probably* 'Charles'.

This is **probability** interpreted as a **degree of belief**. Specifically, it is a measure of how much I believe a statement to be true. In my example, the statement was 'The name of the guy was Charles.'

If I think this statement is certainly true then my **probability** of it being true is 1. If I think this is certainly not true then my **probability** of it being true is 0. If I'm not sure either way then my **probability** will be somewhere between these extremes.

This interpretation of **probability** makes it something personal. To Charles (or whoever it was) the name is quite certain. Indeed I was quite a bit more confident when we were first introduced. **Probability** interpreted this way depends on information (and memory!).

However, that doesn't mean it is purely subjective, because these **probability** numbers can still be tested and different people with the same information and instructions should come up with similar numbers.

Interpreting **probabilities** as **degrees of belief** is much more common, more important, and more scientifically respectable than many people think.

In 1946, physicist, mathematician, and electric eel expert Richard Threlkeld Cox (1898–1991) showed how some very simple, common-sense requirements for logical reasoning about uncertain statements led to the laws of mathematical **probability**. He improved on his thinking in 1961 and others have also refined it, notably Edwin Thompson Jaynes (1922–1998), another physicist, writing shortly before his death.

When stating **probabilities** it is good practice to make clear what information about the circumstances of your prediction you are using. As mentioned earlier, different choices will give different **probabilities**.

For example, if you are referring to your **degree of belief** that it will rain on your garden tomorrow you might decide to take into account nothing about the seasons or the weather, or you could say 'given that it is a day in August', or 'given that the weather forecast on TV said it would rain all day', and so on.

You can't say 'given all the facts' because tomorrow's weather would be one of them, and saying 'given all the facts I know now' is likely to lead to confusion as your knowledge continues to increase over time. Ideally you should make a clear, sensible choice, and communicate it.

9 SITUATION (ALSO KNOWN AS AN EXPERIMENT)

Other common interpretations of **probability** focus on the narrower topic of outcomes. This is the explanation most likely to be shown in a textbook on **probability**.

The outcomes in question are those of a **situation,** or **experiment** (e.g. tossing a coin, drawing a card from a shuffled deck, driving a car for a year and recording the cost of damage to it, paying the total claims on an insurance policy).

The word **experiment** is rather misleading because it doesn't have to be an experiment in the usual sense. It is really just any situation where the outcome has yet to be discovered by the person who is doing the thinking. This includes things that have yet to happen and also things that have happened but whose outcome is not yet known to the thinker.

In this book I've used the word **situation** instead of **experiment** to help you keep an open mind.

Situations are things we have to define, and to some extent they are an arbitrary choice. They define a collection of, perhaps many, episodes in real life that have happened and/or may happen in the future. Each of these episodes will be unique in some way, but if they meet our definition of the **situation** then they are examples of it. For example, 'drawing a card from a shuffled deck' could be our choice of **situation**, but it might have been 'drawing a card from a shuffled deck of Happy Families cards' or 'drawing the top card from a deck of ordinary playing cards shuffled by a professional conjuror.'

clear definition is important

In effect our choice of **situation** is the same as our choice of which circumstances to take into consideration.

Our choice of **situation** makes a difference, and clear definition is important.

10 LONG RUN RELATIVE FREQUENCY

Another common interpretation of **probabilities** focuses on the outcomes from **situations** we see as inherently difficult or even impossible to predict.

Suppose I vigorously flip a fair coin in the traditional way. What is the **probability** of getting heads? Most people will answer confidently that it is 50% or ½ or 0.5, or perhaps 50:50, or evens, depending on their preferred language. This is a **probability** we feel we know.

An idea that captures a lot of our intuitions about **probability** is that it has something to do with what would happen if we could study the outcomes of many examples of a **situation**.

If we could toss that fair coin billions of times and record the proportion of heads we would expect it to be very close to 50% (see Figure 1). So, when we say the **probability** of heads next time is 0.5 that is consistent with the idea that if we did the same thing lots of times then half the time the outcome would be heads.

In this interpretation, **probability** is a property of the real world independent of our knowledge of it.

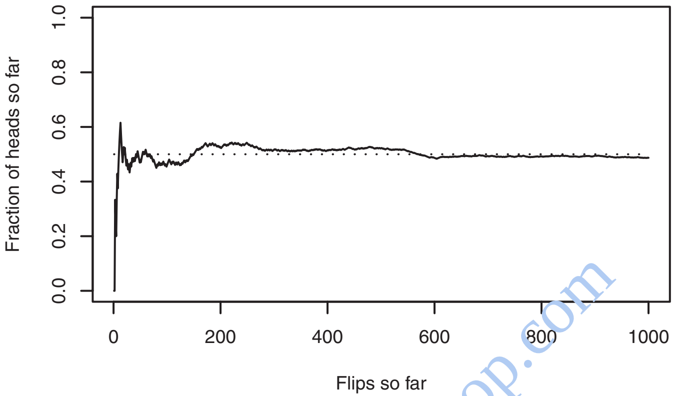

Figure 1 Gradually converging on the long run relative frequency of heads from flipping a fair coin

For this book I've called this the **long run relative frequency (LRRF)**.

As ever, our choice of which **situation** we imagine repeating is crucial and any given occasion could be an example of many different **situations**.

Probability numbers, as always, must lie between zero and one. Zero means that heads never turns up, while one means it always does.

11 DEGREE OF BELIEF ABOUT LONG RUN RELATIVE FREQUENCY

Unfortunately, **probability** based purely on **long run relative frequency** doesn't always score well on Ignorance and other forecasting skill scores. The problem is that it ignores an additional source of uncertainty that is often present.

Imagine I vigorously flip a coin that is clearly bent. What would you say is the **probability** of getting heads this time?

It feels different doesn't it? Our basic schooling on **probability** tends to focus on **situations** like card games, dice rolling, and coin tossing where we usually assume we know what the **long run relative frequencies** should be because we assume that all outcomes are equally likely. Problems set in school textbooks usually say that the coin is 'fair', meaning that you can assume the usual **probabilities**.

In real life things are rarely so convenient. Individual outcomes aren't always equally likely. We don't know exactly what the **long run relative frequencies** would be. We're uncertain about them. Coins look a bit bent and are tossed by conjurors we don't trust. We can't repeat things billions of times. We have to make estimates.

We can't repeat things billions of times. We have to make estimates

So, here's another interpretation. **Probability** can also mean our **degree of belief** about **long run relative frequencies**.

In this interpretation **probability** depends on our knowledge. The real world has a **long run relative frequency** and statements about that are what the **degrees of belief** apply to.

Mathematicians sometimes use **probability** to mean just **long run relative frequency**. On other occasions they use **probability** to mean their **degree of belief about long run relative frequency**. They may even use both ideas in the same problem, calling them both **probability**. Using the same name both times is confusing so separating the two ideas can be very helpful.

12 DEGREE OF BELIEF ABOUT AN OUTCOME

Mathematicians sometimes use **probability** to mean **degree of belief about an outcome**. For example, the statement 'heads next time' could be true or false. This interpretation of **probability** applies the **degree of belief** idea to a statement like that.

The **degree of belief** about an outcome can be calculated from the **long run relative frequency** and the **degree of belief about long run relative frequency**. Alternatively, you can just jump to a **degree of belief about an outcome** by some other means, such as intuition.

I know the different **probability interpretations** take a while to sink in but stick at it because this is where some huge practical mistakes have been made. Here's an example that might just do it for you.

Picture that bent coin I flipped a page ago and imagine you had the opportunity to flip it and learn more about its tendency to come up heads. From a handful of flips you couldn't know the true **long run relative frequency** of that coin. That means you don't know the **probability** of heads in the **long run relative frequency** sense.

However, you could start to form a view about what are more likely values for the **probability** (**LRRF** interpretation) and you could express this in terms of **probabilities** (**degree of belief about LRRF** interpretation) that each **LRRF** is the true one.

So what is the **probability** of getting heads next time that you would use for gambling purposes? (This is the one that represents your **degree of belief** that the outcome will be 'heads' and is your **probability** in the **degree of belief about an outcome** interpretation.)

That **probability** you could get by intuition (the traditional way) or by combining the **probability** (**LRRF** interpretation) with the other **probabilities** (**degree of belief about LRRF** interpretation) in a particular way to get to a **probability** (**degree of belief about an outcome** interpretation).

The theory of **probability** works fine whichever interpretation you use, but the problems come when different interpretations get confused or inappropriately left out.

13 AUDIT POINT: MISMATCHED INTERPRETATIONS OF PROBABILITY

Obviously things can go wrong if one person means **probability** in one sense and another thinks they mean something else. We can even confuse ourselves.

Ask someone for the **probability** of heads from tossing an unfamiliar bent coin and many will answer 'I don't know', revealing that they are thinking in **long run relative frequency** terms. They are seeing **probabilities** as characteristics of the real world, independent of their knowledge of them.

Take that same person to their favourite sporting event and ask them what they think of the odds on a famous competitor and they will happily take a view. This is true even though they still don't know what the **long run relative frequency** of that competitor winning that event is, and could never know it. In this context they activate the **degree of belief about an outcome** interpretation, without realizing they have done so.

The most dangerous version of this confusion is where one person is thinking in terms of **long run relative frequencies** and offers **probability** information to someone else who thinks they are getting **degree of belief about an outcome** information.

The speaker is giving what could be wild guesses about the real world, without mentioning that they are guesses. The listener does not realize this crucial information is being left out. In this simple misunderstanding uncertainty is ignored and the listener comes away thinking things are much better understood than they really are. This happens so often that I will be returning to it repeatedly in this book.

14 AUDIT POINT: IGNORING UNCERTAINTY ABOUT PROBABILITIES

Focusing on **long run relative frequencies** and forgetting that we aren't certain of them is a mistake. It may happen due to mismatched interpretations of **probability**, or it may be that the uncertainty is ignored for some other reason, such as convenience or a desire to seem authoritative.

Whatever the reason, the consequence is that risks are generally underestimated and too little is done to use available data to help get a better view.

15 AUDIT POINT: NOT USING DATA TO ILLUMINATE PROBABILITIES

People often fail to use available data to firm up **probabilities**. This may be because they think of **probabilities** as nothing more than subjective guesses about outcomes.

More often it is because they focus on the **long run relative frequency** idea and think any data used must be from past occurrences of the *identical* circumstances to those now expected. Unable to find data that are from identical circumstances in the past, they give up on data altogether.

Identical circumstances never happen; that would require repeating the history of the universe. What does happen is recurrence of circumstances that match the definition of one or more **situations** that we have chosen. It is also possible to generate quite good **probabilities** by taking into account the degree of similarity between **situations**.

The trick is to think of definitions for **situations** that include the occasion for which we want a **probability**, seem to capture a lot of what is important, and for which we have data.

In doing so we must accept that the more narrowly we define the **situation**, the more relevant past data will be, but the fewer data we will have to work with. Put another way, narrow **situation** definitions give us high uncertainty about highly informative **long run relative frequencies**.

For example, a construction company that builds houses, flats, and some large buildings like schools might have years of data on estimated and actual costs and times to complete its projects. It would be a mistake to think that, because every project is unique in some way, past experience is no guide to future cost estimates. It might be that using data from its house construction in the last two years gives a helpful distribution of estimates that, at the very least, enables baseless optimism to be challenged.

The key points are that we don't need to repeat identical circumstances and may have more relevant data than we realize.

16 OUTCOME SPACE (ALSO KNOWN AS SAMPLE SPACE, OR POSSIBILITY SPACE)

Having covered the main **interpretations** of **probability** it's time to go back to the idea of a **situation** and explain some more of the thinking and terminology behind the most common textbook version of **probability** theory.

In this approach, the next thing to define for a **situation** is its **outcome space**, otherwise known as its **sample space** or **possibility space**. This is the set of all possible elementary outcomes from the **situation**. We also need a way to name or otherwise refer to the outcomes.

For example, tossing a coin once is usually said to have the **outcome space** {heads, tails} but if you let it drop onto a muddy sports field it might be more accurate to say {heads, tails, edge}. If you prefer shorter names then that could instead be {h, t, e}. It's another option.

What is an elementary outcome? That's something else to be decided and written down. There are options and to some extent the decision is an arbitrary one. However, some choices are easier to work with than others. For example, if you can define your outcomes in such a way that they are equally likely, then that makes life a lot easier.

Sometimes the outcomes are more like combinations of outcomes. For example, the outcomes from tossing two coins one after the other could be defined as {(H,H), (H,T), (T,H), (T,T)} with the first letter representing one coin and the second representing the other. Another example would be measurements of newborn babies, where each outcome could be represented by a bundle of facts such as its weight, length, sex, and skin colour.

In real life **situations** we usually have a number of different ways to characterize what could happen. For example, we might be interested in health and safety, or money, or time. Each possibility, if chosen, will give us a different **outcome space**.

The phrase **sample space** is what the mathematicians most often use, for historical reasons, but it is misleading (again) because sampling in the usual sense isn't normally involved. In this book I've used the less common term **outcome space** so you don't have to keep reminding yourself to forget about sampling.

17 AUDIT POINT: UNSPECIFIED SITUATIONS

Many so-called 'risks' for which people are asked to give a 'probability' do not describe adequately the **situation** they apply to. For example, there may be a 'risk of theft' but over what time period, involving which assets, and measured in what way? Unless this vagueness is cleared up it's hard to say anything meaningful about how big the 'risk' is, even broadly and without numbers.

Consider reviewing a sample of risk descriptions and recommending some kind of quality improvement work.

Different styles of risk analysis require clarity on different points, so you are looking for any statement that seems vague and should also consider whether important qualifications have been left out altogether. It is very common to forget to state the time period for a 'risk'. For example, 'Risk of theft in the next year' is much less likely than 'Risk of theft at any time in the future.'

18 OUTCOMES REPRESENTED WITHOUT NUMBERS

The outcomes in an **outcome space** can be represented in a variety of ways. One way is without numbers. For example, if beads of different colours are put into a bag and shaken, and then one is drawn out, the outcomes might be represented by colours, e.g. {Red, Blue, Green }.

'if the risk is 'Loss of market share' then surely it matters how much market share is lost'

This is important because some concepts in risk mathematics do not apply if the outcomes are not represented by numbers.

A lot of the things we call risks and put on risk registers are worded so that there are just two outcomes and they're not represented by numbers. Those two outcomes are {'The risk happens', 'The risk does not happen'}.

This is simple, but usually much *too* simple and tends to mean we cannot think about important nuances. For example, if the risk is 'Loss of market share' then surely it matters how much market share is lost. The problem is not lack of numbers but failure to capture the richness of potential outcomes. Most mathematical risk analysis is much more informative.

19 OUTCOMES REPRESENTED WITH NUMBERS

In other **outcome spaces** the outcomes are represented by numbers.

20 RANDOM VARIABLE

Often what people are interested in is not the outcome but, instead, a number that depends on the outcome. For example, if you roll two dice when playing Monopoly it is the total of the dice you care about.

And when people enter a lottery they are interested in how many of the balls selected at random in the draw match the balls they bet on. They are not really interested in exactly which balls are drawn. A lot of risk management in businesses focuses on money.

A **random variable** is, strictly speaking, neither random nor a variable, but is a rule that links each outcome to a unique number. Given an outcome it returns the appropriate number. People often talk about **random variables** as if they represent the actual outcome (which is not yet known). In other words, they treat them as if they are the numbers returned rather than the rule, but this usually doesn't lead to mistakes.

Random variables, by convention, always return what mathematicians call 'real' numbers, which for our purposes just means they don't have to be whole numbers, but can be anywhere on the continuous number line.

Sometimes the way outcomes are linked to numbers can seem a bit arbitrary. For example, when the outcome space is {success, failure} these outcomes are often mapped to one and zero respectively.

Traditionally, **random variables** are usually given names that are capital letters from the English alphabet and the runaway favourite choices are X and Y.

In practice the definition of a **random variable** is a matter of choice and needs to be clear.

21 EVENT

An **event**, in mathematics, means a subset of the **outcome space**. For example, if you've chosen the **situation** of tossing a coin and letting it fall on a muddy sports field and the **outcome space** {heads, tails, edge} then you could define a number of possible **events** having one or more outcomes in them, such as an **event** you could call 'valid outcome' defined as the set {heads, tails}.

What **events** are defined is yet another free choice. The **event** 'valid outcome' is likely to be useful when talking about coin tossing on a muddy field, but of course you could look at it in other ways.

Events involving discrete outcomes can be defined by listing all the outcomes included or by stating some rule for membership.

Events involving outcomes that could be anywhere on a continuum of numbers are often defined by giving the top and bottom of the range of numbers to be included in the event. Another common technique is to give one number, defining the **event** as all outcomes with numbers less than or equal to that number.

Random variables can be used to succinctly define **events**. For example, if the name X is given to a **random variable** returning the total of two fair dice thrown together then:

1 $\{X = 4\}$ is the event that contains all the outcomes that add up to 4, i.e. {(1,3), (2,2), (3,1)}; and
2 $\{X < 3\}$ is the event that contains all the outcomes that add up to less than 3, i.e. {(1,1)}.

This is the traditional notation and I hope it is clear what is intended. If not then it may be that you've noticed the mistake, which is to write as if X is the value returned by the **random variable**, not the **random variable** itself. Perhaps a clearer notation would be something like

$\{X(w) = 4\}$ where w represents the outcome from the **situation**, and $X(w)$ is the usual way to show the value returned when a function (e.g. X) is applied to an input value (e.g. w).

An **event** is not necessarily something sudden, dramatic, or memorable. This idea is very different to our ordinary idea of an 'event' and this causes some confusion. Procedures for risk management tend to be written as if 'events' are dramatic things with all or nothing results, like explosions. But in reality most situations where 'risk' needs to be managed are not like this. There are a few explosions but far more slightly surprising outcomes of undramatic **situations**. It is better to use the mathematical idea of an **event** and this is more consistent with the vast majority of 'risks' that people think of.

> ***An event is not necessarily something sudden, dramatic, or memorable***

22 AUDIT POINT: EVENTS WITH UNSPECIFIED BOUNDARIES

Many 'risks' on risk registers have a form like 'inadequate human resources'. We imagine a scale of human resources and a zone towards the bottom that is 'inadequate'. Unfortunately, the level below which human resources are inadequate is unspecified (and probably unknown) making the 'risk' unspecified too.

23 AUDIT POINT: MISSING RANGES

Another problem with 'risks' like 'inadequate human resources' is that the choice of the word 'inadequate' is rarely the result of careful thought. It could have been replaced by 'less than expected' or 'zero' with little comment by most people. Choosing 'inadequate' as the definition for the **event** removes from consideration other ranges that might be surprising and require planning for. I call these *missing ranges*. They are very easy to check for and point out.

24 AUDIT POINT: TOP 10 RISK REPORTING

Many people in senior positions have been encouraged to believe that they need to focus on the 'top 10 risks'. I wonder how they would feel if they understood that **events** are defined by people and can be redefined to suit their purposes.

Imagine you are a manager in a risk workshop and somebody has just suggested a risk for inclusion in a risk register that (1) you would obviously be responsible for, (2) will probably be in the top 10, and (3) you can't do much about. You don't want the risk to be in the top 10 and to get beaten up by the Board every quarter so you say, 'That's a really interesting risk, but I think to understand it fully we need to analyse it into its key elements.'

You then start to hack the big 'risk' into smaller 'risks', keeping on until every component is small enough to stay out of the top 10.

The point is that the size of a 'risk' is heavily influenced by how widely it is defined. Most of the time the level of aggregation of risks is something we set without much thought, so whether something gets into the top 10 or not is partly luck.

Auditors should highlight this issue when found and suggest either the level of aggregation of 'risks' be controlled in some way or top 10 reporting be abandoned and replaced by a better way of focusing attention.

25 PROBABILITY OF AN OUTCOME

In researching for this book I consulted several different sources and got several different explanations of **probability** theory, with slightly different terminology and slightly different notation.

The reason for this is historical and understanding it may help to make sense of it all.

In the beginning, **probability** theory was focused on winning in games of chance. It concentrated on situations where there was just a finite number of outcomes, such as the roll of a die, or a hand in a card game.

It made perfect sense to talk about the **probability of an outcome** and to calculate the **probability** of an **event** by adding up the **probabilities** of the outcomes they included. (Remember that an **event** is a subset of the **outcome space**, so it's a set of outcomes.)

The sum of the **probabilities** of all the outcomes from a **situation** is one, because it is certain that one of those outcomes will result, by definition.

Later, people moved on to think about **situations** where the outcomes could be any point on a continuum, such as the life of an electric light bulb. In this example the life could be, theoretically, any amount of time. Even between a lifetime of 10 minutes and a lifetime of 11 minutes there is an infinite number of possible lifetimes. (In practice we can't measure accurately enough to recognize that many but in principle it is true.)

This revealed an awkward problem. The **probability of the outcome** being *exactly* equal to any particular point on the continuum seemed to drop to zero, and yet the outcome had to be somewhere on the continuum. How can adding up lots of zeroes give the result one?

To get around this problem, **probability** was defined in a different way specifically for these continuum situations, but still starting with outcomes and building up from there.

26 PROBABILITY OF AN EVENT

Then in 1933 the Russian mathematician Andrey Nikolaevich Kolmogorov (1903–1987) did some very fancy maths and showed how both problems (with and without outcomes on a continuum) could be dealt with using one approach.

Although Kolmogorov's approach has been accepted for decades it still hasn't reached every textbook and website.

Kolmogorov's thinking is a mass of mind-boggling terminology and notation (which I'm not going to go into) and was mostly concerned with applying the fashionable ideas of measure theory to **probability**. Yet one of the key ideas behind it is simple: since starting with **probabilities for outcomes** hasn't worked neatly for us, let's start with **probabilities for events** instead.

27 PROBABILITY MEASURE (ALSO KNOWN AS PROBABILITY DISTRIBUTION, PROBABILITY FUNCTION, OR EVEN PROBABILITY DISTRIBUTION FUNCTION)

The result of Kolmogorov's hard work was the notion of a magical thing called a **probability measure** that tells you what **probability** number is associated with each **event**. (The word 'measure' here indicates Kolmogorov was using measure theory, but you don't have to in order to associate **probability** numbers with **events**.)

The alternative name **probability function** (which lacks the link to measure theory) is a good one because, in mathematics, a function is simply a rule, table, or formula that takes one or more inputs and consistently replies with a particular output. For example, a function called something like 'square' might return the square of each number given to it. (A **random variable** is also a function.)

In the case of **probability**, you tell the **probability function** which **event** you are interested in and it returns the **probability** that goes with it.

The alternative names are used in different ways by different authors, which can be confusing, particularly when **probability distribution** is used to refer to something that does not give **probabilities**.

The way the **probability measure** is designed depends on what type of outcome is involved and what is a convenient way to identify the **event**.

For example, if the **outcome space** for coloured balls pulled out of a bag is {Red, Blue, White, Black} then a **probability function** called Pr (one of the common name choices) might be used, as in these examples:

1 Pr({White}) = 0.3 means that the **probability** of pulling out a white ball is 0.3.
2 Pr({Black}) = 0.2, means that the **probability** of pulling out a black ball is 0.2.
3 Pr(Monochrome) = 0.5, where Monochrome = {White, Black}, means that the **probability** of pulling out a black or white ball is 0.5.

In writing these examples for you I've been quite strict and made sure that the thing inside the ( ) parentheses is a set of outcomes. However, the notation used is not always so careful.

Remember that **events** can also be specified using **random variables**.

The impact of Kolmogorov's work may have been huge for the theoretical foundations of **probability**, but it has made little impact otherwise so most of us don't need to know any more about it.

28 CONDITIONAL PROBABILITIES

Mathematicians have a habit of leaving out information to keep their formulae looking simple, expecting readers to guess the rest from the context.

Formulae about **probabilities** give countless examples of this. The usual way to write 'the **probability** of event *A* occurring' is:

$$P(A)$$

"Mathematicians have a habit of leaving out information to keep their formulae looking simple"

But what **situation** are we talking about? What is the **outcome space**? Or, put it another way, what parts of our knowledge about the circumstances surrounding the **event** of interest are we choosing to use for the purposes of this **probability** number? For example, if we are interested in the outcome of tossing a coin, do we say this is an example of coin tossing, of tossing this particular coin, or of coin tossing on a muddy field? If the coin is to be flipped by a conjuror do we take into account the fact that he has just bet us £100 it will be heads?

Usually, little or even none of this is stated in the formula, with the obvious risk of confusion or mistakes. For good reason, people sometimes point out that *all* **probabilities** are **conditional probabilities**.

However, there is a standard notation for showing information that defines the **situation** or otherwise shows what parts of our knowledge of circumstances are being used. This is the notation for **conditional probabilities**. For example, a way to write 'the **probability** of event *A* occurring given this is an instance of a **situation** with outcome space *S*' is this:

$$P(A \mid S)$$

You say this to yourself as 'the **probability** of *A* given *S*.'

When new information arrives we are not obliged to use it in every **probability** we state. However, for **probabilities** where we do use the new information this effectively redefines the **situation**.

For example, suppose our initial **situation** was 'drawing a playing card from a shuffled deck' but later we learn that the deck has been shuffled and the card drawn by a conjuror. This new information redefines the **situation** quite dramatically.

In symbols, if we want to show 'the **probability** of event *A* occurring given this is an instance of a **situation** with outcome space *S*, and given the outcome is already known to be within event *B*', we write:

$P(A \mid S, B)$

In this particular example this makes the new outcome space, in effect, *B*, because *B* is entirely within *S*.

For some, perhaps all, occasions where we want to use **probabilities** the addition of more and more information might eventually allow us to predict the outcome with complete certainty, in theory.

29 DISCRETE RANDOM VARIABLES

At this point **probability** theory starts to focus on **events** defined using **random variables**.

Random variables are functions that give 'real' values, i.e. numbers that could, in principle, lie anywhere on a continuous number line from zero all the way up to infinity (∞), and indeed from zero all the way down through minus numbers to minus infinity (–∞). In symbols, they are in the range (–∞,∞).

However, when a **random variable** is defined for the **outcome space** of a **situation**, it may well be limited to returning just certain values within that huge range. For example, if the **random variable** represents the total of two dice then it can only take the specific values 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12, even though it is a real number.

Random variables are classified into three types according to the values they can return once hooked up to an **outcome space**. The simplest type is the **discrete random variable**.

Discrete random variables can return only a finite number of values, or an infinite but countable number of values.

To illustrate the meaning of 'countable', the set of numbers {1, 2, 3 … and so on forever} has infinitely many elements but they are countable, whereas the number of numbers on the real number line between 0 and 1 is infinite and not countable. Countable infinity is much smaller!

30 CONTINUOUS RANDOM VARIABLES

The other type of **random variable** that gets a lot of attention is the **continuous random variable**. This type (1) can return an uncountably infinite number of values but (2) the **probability** of returning any particular value is always zero.

That usually means that the value is somewhere on a continuum of numbers and no particular value is special.

If your brain is still functioning at this point you may be wondering how the **probability** can always be zero. Surely an outcome of some kind is inevitable, by definition, so the sum of the **probabilities** for all the individual outcomes must be one. How can the sum of lots of zeros be anything other than zero?

Good question, and perhaps it makes more sense to think of those zeroes actually being infinitesimally small 'nearly zeroes' so that what is really happening is that infinitely many infinitesimally small things are being added together. Only by cunning mathematical reasoning can the value of such a sum be worked out.

A huge proportion of the applied risk analysis done by mathematicians in business and elsewhere involves **continuous random variables** (though it is not necessary to go through the reasoning about infinity each time).

Incidentally, the Ignorance function mentioned in connection with **proper scoring rules** can only be applied to **discrete random variables**, but applying it to **continuous random variables** simply involves slicing the continuous case into lots of little pieces. This is just a reminder that in most cases where we **model** the world with continuous variables the reality is that we cannot and do not measure to infinite accuracy. Money, for example, is usually tracked to two decimal places, not to infinite precision, which would involve quoting some numbers to infinitely many decimal places!

the reality is that we cannot and do not measure to infinite accuracy

31 MIXED RANDOM VARIABLES (ALSO KNOWN AS MIXED DISCRETE-CONTINUOUS RANDOM VARIABLES)

Discrete and **continuous random variables** get so much attention it is easy to get the impression that they are the only types that exist. Not so, and in fact **random variables** of the third type are applicable to most of the 'risks' people put on risk registers.

These forgotten **random variables** are so unloved that it took me a while to find their proper name: **mixed random variables**.

Like the **continuous random variables** they can take an uncountably infinite number of values, but these hybrids can give special values whose **probability** of occurrence is more than zero.

For example, suppose that the **random variable** is for the useful life of a light bulb. Some light bulbs don't work at all, while others go on for a period we don't know in advance.

This means that the **probability** of lasting exactly zero seconds is more than zero, but the **probability** of any particular lifespan beyond this is zero.

32 AUDIT POINT: IGNORING MIXED RANDOM VARIABLES

Perhaps because they don't get much attention **mixed random variables** tend to get left out.

People don't think of using them in their risk analysis and instead behave as if everything is either discrete or continuous.

This is important because such a high proportion of 'risks' on risk registers are best described by a **mixed random variable**.

It is true that there are very few well known distribution types that are mixed and software does not support them directly, in most cases. However, a mixed type can easily be built from a combination of **discrete** and **continuous random variables**.

For example, to express the lifespan of a light bulb you can use a **discrete random variable** to say if it fails immediately or not, and then a **continuous random variable** to show the **probability distribution** of its lifespan assuming it at least gets started.

Be alert for this mistake when reviewing risk management procedures, templates, and models.

33 CUMULATIVE PROBABILITY DISTRIBUTION FUNCTION

There is one type of **probability distribution function** that can capture, and graph, the nuances of **random variables** of any type.

This kind of function is called a **cumulative probability distribution function**. It gives the **probability** that the value returned by a **random variable** will be *less than or equal to* any particular value.

The graph of a **cumulative probability distribution function** always rises from left to right, as in Figure 2.

Take a moment to think this through a few times because we are not used to seeing this kind of graph.

Cumulative probability distribution functions are extremely useful in risk analysis because they can be used in many different situations, even when other types of function are too fussy to be applied.

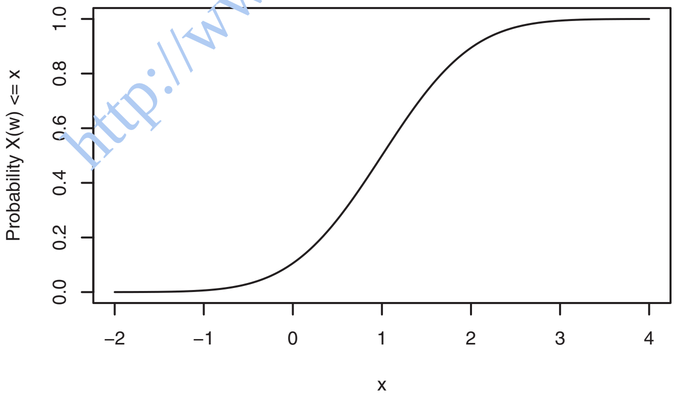

Figure 2 A cumulative probability distribution function shows the probability that the random variable returns a value less than or equal to x

For example, suppose a 'risk' has been written that says 'The cost of fire damage to our warehouses during this year.' Imagine that there's a good chance that this will be zero, because fires are rare. However, if a fire starts then the cost could be anywhere from tiny (a slight scorching) to catastrophic, with a large building burned to the ground.

A **cumulative probability distribution function** can capture all this. For cost values less than zero (we gain money) the cumulative **probability** is zero. That's not going to happen. At a cost of exactly zero the **probability** will be the chance of no fire damage during the year. For higher and higher values of cost the cumulative **probability** will gradually increase, ultimately getting closer and closer to one (see Figure 3).

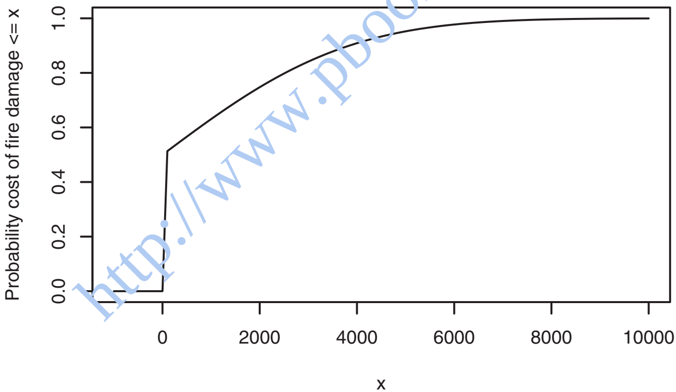

Figure 3 Cumulative probability distribution function for cost of fire damage, x

34 AUDIT POINT: IGNORING IMPACT SPREAD

The usual treatment of items on a risk register is to ask people for the probability of 'it' happening and the impact if 'it' does.

But what if the impact could be anything over a wide range? For example, how do you estimate a single impact level for a risk item like 'Loss of market share'? Surely it depends on how much market share is lost, among other things. I call this 'impact spread' and my study of published risk registers shows that virtually all risk register items have impact spread for at least one reason and often for several.

The question on the risk register requires an answer that is a single number or category, and there are several ways people could choose one. They could pick the first level of impact within the range that comes to mind. They could pick the level that seems most representative, or most likely, or is the **probability** weighted average of the whole range, or a halfway point, take something at random, or pick something that combines with the **probability** number to give the priority rating they think the 'risk' should have.

If we want the impact ratings to mean something then we need to define how people should reduce the range to a single point, or change our technique.

The two recommendations auditors should consider first are these:

- Define the required impact rating as the **probability** weighted average impact over the whole range of possibilities. This means that when it is combined with the **probability** it gives something meaningful.
- Change the rating system so that it asks for a **probability** of at least some impact, and then the **probability** of impact greater than one or more other thresholds. This technique elicits a simplified variant of the **cumulative probability distribution function** and is easier to explain.

35 AUDIT POINT: CONFUSING MONEY AND UTILITY

When we talk about 'impact' another possible confusion is between a measure such as money and how much we value the money. The word 'utility' is often used to mean the real value we perceive in something.

For example, a financial loss like losing £1 million is surely more important if this amount would destroy your company.

When we talk casually about 'impact' there is always the danger of overlooking this point and flipping from thinking in money terms to acting as if it is really utility we are talking about.

The two ways of thinking give different answers. Suppose we have two 'risks', one of which can lead to losses in a narrow range, with the average being £100,000. The other also has an average of £100,000 but the range of possibilities is much larger with a possibility of losses that ruin the company.

Is it fair to treat these two losses as having the same impact? In financial terms their average is the same but if we translate to utility and *then* take the average the second risk is considerably worse.

Some organizations try to express a 'risk appetite', which is supposed to help employees respond consistently and appropriately to risks, especially the bigger ones. If averages (or other midpoints) from money impact distributions are being used then the risk appetite initiative is seriously undermined.

36 PROBABILITY MASS FUNCTION

A fussier probability distribution is the **probability mass function**, which only applies to **discrete random variables** (see Figure 4).

A **probability mass function** gives the **probability** that the **random variable** will return any particular value.

The importance of **probability mass functions** perhaps goes back to the early focus on the **probabilities** of outcomes as opposed to **events**.

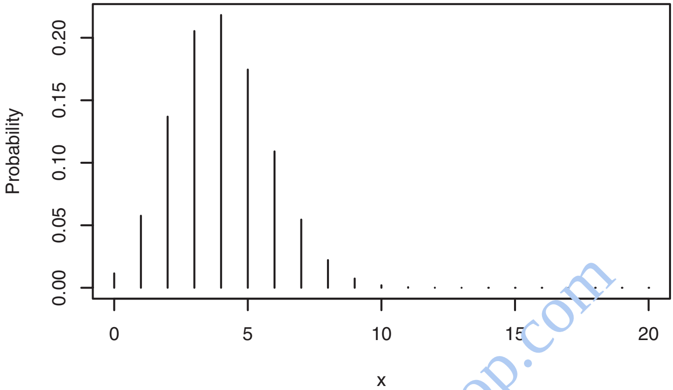

Figure 4 A probability mass function

Also, if you have the **probability mass function** then you can calculate the **probability of any event**.

37 PROBABILITY DENSITY FUNCTION

Obviously a **probability mass function** won't work with a **continuous random variable** because the **probability** of any particular value being returned is always zero, and that's a problem with the **mixed type** too.

For **continuous random variables** only it is possible to create a function called a **probability density function** that returns not **probability**, but something called **probability** *density*.

Graphs like the one in Figure 5 have **probability density** on the vertical axis, not **probability**, so in that sense they are not **probability distributions** at all.

The *area* under one of these **probability density function** graphs is what represents the **probability**. If you want to know the **probability** that the **random variable** will return a value somewhere between two

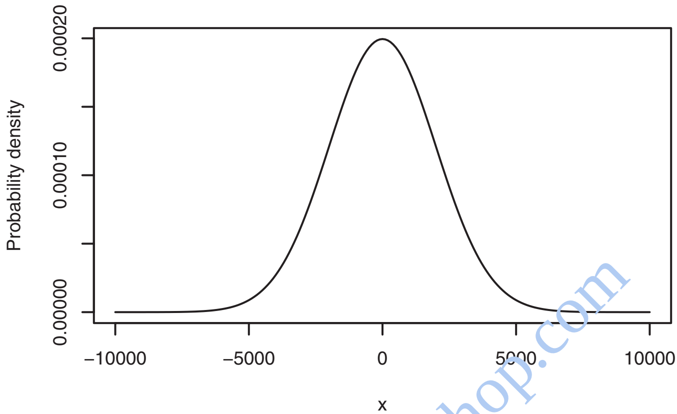

Figure 5 A probability density function

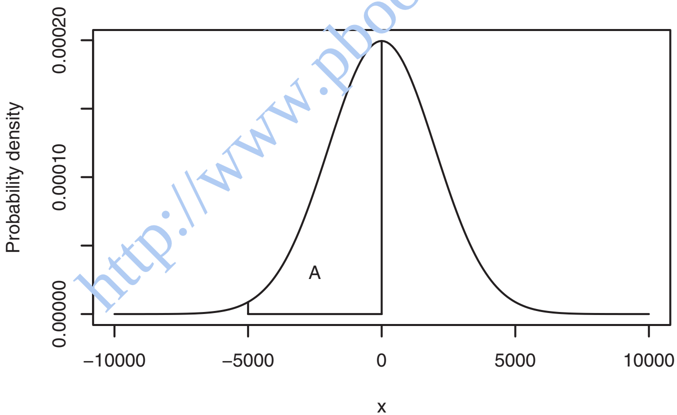

Figure 6 The area, A, under the curve is the probability of x being between –5000 and 0

numbers then you need the area under the **probability density function** curve that lies between those two values. The total area under the curve is always one (see Figure 6).

Again, the importance of **probability density functions** perhaps goes back to the days when **probability** theory was focused on outcomes. They are an attempt to give a number for each possible outcome, which is sort of like **probability** even though it isn't **probability**. If you have the **probability density function** then you can calculate the **probability of any event**.

38 SHARPNESS

One quality of **probabilities** that tends to contribute to high **resolution** is **sharpness**. **Sharpness** is simply use of **probabilities** that are near to zero or one, and it does not imply that those **probabilities** are also well **calibrated**.

The choice of the word **sharpness** is now easy to understand in terms of **probability density functions**.

Imagine Figures 7(a) and 7(b) represent forecasts for the change in value of a portfolio of investments over two periods of 24 hours. In Figure 7(a), which is for the first period of 24 hours, one forecast-

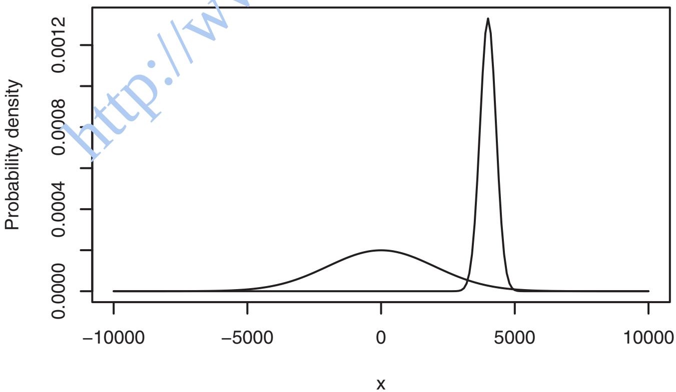

Figure 7 (a) Forecasts for the first day with a wide distribution and a much sharper distribution, equally well calibrated

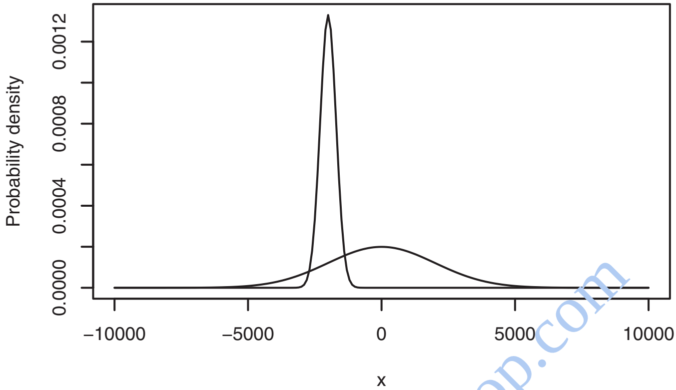

Figure 7 (b) Forecasts for the second day with the wide distribution unchanged but the sharp distribution more responsive to circumstances

The more we try to take into consideration, the less directly relevant past experience we can draw on

ing approach gives a well **calibrated** but widely spread **probability distribution** while the other, equally well **calibrated**, distribution is much **sharper**. Figure 7(b) shows the forecasts for the second period of 24 hours and the widely spread distribution is unchanged while the **sharp** forecast has taken more circumstances into account and is different from the previous day.

The more we try to take into consideration, the less directly relevant past experience we can draw on. We have the chance to achieve high **resolution**, but without much history as a guide we risk poor **calibration**. It's a balancing act and understanding it is a hot research topic.

39 RISK

Finally, we have arrived at **risk**. The reason for your long wait is that mathematicians don't really have much use for the word in either of its main senses.

In everyday conversations we often talk about 'risks', meaning nasty possibilities that can be listed and counted. Mathematicians have **events** and **random variables** instead, and they are much better defined ideas, free from the associations with danger and losses that tend to make 'risk' an entirely negative idea.

In everyday conversations we also talk about how much 'risk' we face, meaning a *quantity* of some nasty possibility. The concept of **probability** was invented centuries ago and when combined with values of outcomes it does everything that 'risk' does and so much more.

the idea of … a number that represents some notion of 'risk' has caught on

However, there is a mathematically oriented concept of **risk**. Its development may owe something to influential work on portfolio theory by American economist Harry Markowitz (1927–), in which he used a number to represent the spread of possible returns from investments and called it 'risk'. This was done to make some of his mathematics more convenient and is justified only by some rather specific assumptions about how investors value investment returns and about how those returns are distributed. However, these finer points have long been ignored and the idea of applying a formula to a **probability distribution** to produce a number that represents some notion of 'risk' has caught on.

In this approach, **risk** is a number calculated using a function that takes as its input the **probability distribution** of a **random variable**.

There is no agreed function for calculating **risk**. There are already several to choose from and more will probably be invented in future. Under close scrutiny all of these have shortcomings.

Before I explain some of these alternative **risk** functions it will be helpful to explain something that is often used as part of them and is generally very useful.

40 MEAN VALUE OF A PROBABILITY DISTRIBUTION (ALSO KNOWN AS THE EXPECTED VALUE)

Another function that takes a **probability distribution** as input and returns a number is the **mean**, otherwise known by the highly misleading name of **expected value**. This is the **probability** weighted average of all outcomes, and only works when the outcomes are represented as numbers.

For example, if we think of the **probability mass function** for a fair die rolled properly, then the outcomes and their **probabilities** are:

$$P(1) = \frac{1}{6},\ P(2) = \frac{1}{6},\ P(3) = \frac{1}{6},$$

$$P(4) = \frac{1}{6},\ P(5) = \frac{1}{6},\ \text{and}\ P(6) = \frac{1}{6}$$

The **probability** weighted average of these is:

$$1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 3 \times \frac{1}{6} + 4 \times \frac{1}{6} + 5 \times \frac{1}{6} + 6 \times \frac{1}{6} = 3\tfrac{1}{2}$$

No, I haven't made a mistake; the **expected value** from rolling a die is 3½ – which is an impossible outcome. In this case the **expected value** is also an impossible value.

In ordinary language, if we 'expect' something that means we either think it jolly well ought to happen or that it is more likely than not. In mathematics, an **expected value** does not need to be more likely than not and might not even be possible.

41 AUDIT POINT: EXCESSIVE FOCUS ON EXPECTED VALUES

When **expected values** come into a conversation (e.g. about forecasts) other outcomes tend to be forgotten. An **expected value** supported by pages of spreadsheeting gets a credibility it rarely deserves.

Auditors should check for this in a variety of situations and recommend taking a broader view and considering more possible futures.

42 AUDIT POINT: MISUNDERSTANDING 'EXPECTED'

The word 'expected' has two ordinary meanings as well as its mathematical meaning and this can lead to confusion.

First, people might think that 'expected' means 'more likely than not', i.e. a fairly highly level of confidence in a prediction. If the business case for a project says its value is 'expected' to be £2.5 million then non-mathematical readers might think that means a very confident prediction of a value of exactly £2.5 million (give or take a few thousand perhaps). It could really mean that the project's proposers have almost no idea what the true value is but the **probability** weighted average of their wild guesses is £2.5 million.

If there is a risk of this misunderstanding taking place then the auditor should point it out. Since giving only **expected values** is poor practice, the obvious recommendation is to provide more information about other possible results.

Second, people might think that 'expected' means 'ought to happen'. Let's imagine the spreadsheet says the **expected** cost of a project is £6.3 million. That means the **probability** weighted average of the guesstimates is £6.3 million. It does not mean that the cost of the project *ought* to be £6.3 million and therefore that's what the budget should be.

Turning **expected values** into budgets or other types of target is a mistake. It is much better to look at the whole **probability distribution** and take a view based on that fuller information.

43 AUDIT POINT: AVOIDING IMPOSSIBLE PROVISIONS

In putting together an initial budget for the 2012 Olympic Games the UK government faced a difficult choice. How much should it include for VAT?

This VAT payment would be a tax paid by the UK government to the UK government, but its inclusion in the budget was still important because funding was not just coming from the general public purse.

Either the games would be declared VAT exempt or they wouldn't. What would you have put in the budget? One perfectly sensible option would have been to budget for the **expected value** of the VAT, i.e. the total VAT bill multiplied by the **probability** of having to pay it at all. How good this is depends on how you value differences between budget and actual, but using the mathematician's favourite, the **expected value** of the budget errors squared, it turns out that the **expected value** for VAT is a great choice.

However, you can imagine that for many people this must have seemed a bizarre choice. It was a budget guaranteed to be wrong. In fact they decided to put nothing in the budget at all and were surprised to find, a year or so later, that VAT would be charged.

44 AUDIT POINT: PROBABILITY IMPACT MATRIX NUMBERS

Here's one that could embarrass a lot of people. Another potential problem with risk register impact and probability ratings comes from the way people sometimes combine them for ranking and selection.

Imagine that the method for combining probability and impact ratings into one rating is defined by the usual grid. Let's say it's a 5 by 5 grid for the sake of argument, looking like Figure 8:

Probability	VL 1	L 2	M 3	H 4	VH 5
5 VH	5	10	15	20	25
4 H	4	8	12	16	20
3 M	3	6	9	12	15
2 L	2	4	6	8	10
1 VL	1	2	3	4	5
	Impact				

Figure 8 Probability impact matrix with 25 cells

There are 5 levels of probability ranging from very low (VL) to very high (VH), and the same for impact. The levels have also been given index numbers from 1 to 5. The combined score is found by multiplying the two indices together and is shown in the cells of the matrix.

Oh dear. What people imagine they are doing is taking the **expected value** of the impact, or something like it, but the numbers being used are not probability and impact but the index numbers of the rows and columns.

When you look at the ranges of impact and probability that define each level they are usually of unequal sizes. For example, 'very low' impact might be '£1–1,000', 'low' impact might be '£1,001–10,000', and so on. Typically the levels get much wider each time.

This means that, often, the index numbers are more like the logarithms of the impact and probability so multiplying them gives you something more like 'the logarithm of the impact raised to the power of the logarithm of the probability'! However you look at it, this is a mistake.

What it means is that 'risks' get ranked in the wrong order and if you have a habit of reporting on only 'risks' over a certain rating then the set of 'risks' selected for reporting will usually be the wrong set.

45 VARIANCE

This is a function whose result is often used as **risk**. It is the **expected value** of the square of differences between possible outcome values and the **mean** outcome. That means it gets bigger the more spread out the possible values are.

The way it is calculated depends on what sort of **probability distribution** is involved.

As with other **risk** numbers it is calculated from the **probability distribution** of a **random variable**. For example, if the **random variable** represents the result of rolling a 6-sided die then the **probability** of each of its six discrete outcomes is 1/6 and its **mean** is 3.5 as we have already seen. Its **variance** is:

$$\frac{1}{6}\times(1-3.5)^2+\frac{1}{6}\times(2-3.5)^2+\ldots+\frac{1}{6}\times(6-3.5)^2$$
$$=2\tfrac{11}{12}$$

Variance can also be calculated for actual data about past events, but this is not **risk**, though it is sometimes taken as an estimate of **risk**, and may be calculated with a slight adjustment in order to be a better estimate.

46 STANDARD DEVIATION

This is just the square root of the **variance**, i.e. multiply the **standard deviation** by itself and you get the **variance**.

As with the **variance**, it gets bigger with more dispersed outcomes.

Also like **variance**, **standard deviation** can be calculated for actual data about past events, but this too is not **risk**, though it is sometimes taken as an estimate of **risk**.

47 SEMI-VARIANCE

A problem with the **variance** and **standard deviation** is that they increase with the *spread* of the **probability distribution**. That means that the possibility of something extremely good happening makes the **risk** number larger. This does not agree with our intuitive idea that **risk** is a bad thing.

> ***the possibility of something extremely good happening makes the risk number larger***

Alternative **risk** functions have been invented to try to focus more on the bad outcomes, such as lost money, and one of these is the **semi-variance**.

This is the **expected value** of the squared difference between outcomes below the **mean** and the **mean** itself. In other words, it is the **variance** but ignoring outcomes above the **mean**.

48 DOWNSIDE PROBABILITY

This is another **risk** number that focuses on possible disappointment. It is the **probability** of not getting an outcome better than some target outcome.

What is taken as the target is a free choice and needs to be defined, but could be a target rate of return for an investment, for example. Outcomes better than the target are ignored. The **downside probability** is a function of the target chosen, and will be higher for more ambitious targets.

49 LOWER PARTIAL MOMENT

This combines ideas from the **semi-variance** and the **downside probability**. It is the **expected value** of the squared difference between outcomes below some target or threshold and the target itself.

50 VALUE AT RISK (VAR)

Another **risk** function that focuses on the downside is **value at risk**, and it has become the most famous.

This is calculated as the loss such that the **probability** of things turning out worse is less than or equal to a given **probability** threshold. The **probability** threshold is something that has to be chosen and is usually small.

For example, a bank might model the value change over the next 24 hours of a collection of investments. The loss such that a loss at that level or worse is only 5% likely is their 5%, 1 day **VaR** for that particular portfolio. Put another way, they are 95% confident they won't lose more than the **VaR** over the next 24 hours (see Figure 9).

Like some other **risk** functions, **value at risk** is sensitive to the extremes of a **probability distribution**, which are very difficult to know accurately, and it says nothing about the *very* extreme possibilities. For these and other reasons it has come in for some severe criticism and been cited as contributing to the credit crunch of 2007–2009.

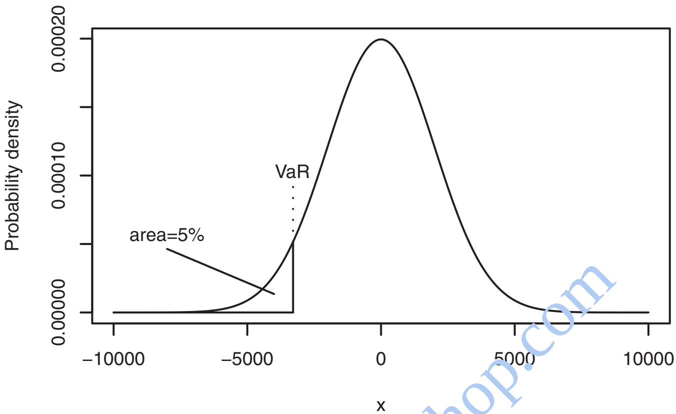

Figure 9 Value at risk based on 5% confidence shown on a probability density function

The name itself is somewhat misleading. It sounds like it represents how much money we currently have invested. A bank might have billions invested but say its **VaR** is just millions. The rest is safe? Hardly.

Value at risk is a common **risk** measure for market risk, i.e. risk related to the value of a portfolio of assets traded on a market (usually a financial market).

VaR is usually calculated on the basis of the market value of the portfolio, not the returns from it (which also include payments such as dividends).

It is usual to assume that the composition of the portfolio is not changed during the period. In reality, trading may well happen during the period so the value of the portfolio will also change for that reason.

It is also common to assume that the **expected** value change of the portfolio during the period is zero. Consequently the only thing that needs to be understood is the variability of market values. **VaR** calculated on this basis is called absolute **VaR**. For periods of just one day this simplifying assumption is not unreasonable in most cases, but for longer periods it may be better to calculate the relative **VaR**, which involves calculating the **expected** market value as well.

Finally, there are alternative bases for calculating market values.

In the chapter on finance we'll look in more detail at how the **probability distributions** used to calculate **VaR** are derived.

51 AUDIT POINT: PROBABILITY TIMES IMPACT

Some years ago I asked a large audience of business continuity managers how they would define 'risk'. The most popular answer was to define it as 'probability times impact'. It is hard to think of a less appropriate definition.

'Probability times impact' is shorthand for the **expected value** of the probability distribution of impact. It is a good candidate for a best guess. It is the number most people would use as an estimate for the impact if forced to give just one number. A risk is something *unexpected*, so 'probability times impact' is the opposite of risk!

More successful ways to define 'risk' in terms of **probability** and impact have held on to the whole distribution rather than reducing it to one number.

The practical problem that 'probability times impact' causes is that outcomes other than the **expected value** get forgotten and uncertainty about those outcomes is ignored, leading to a massive, systematic understatement of **risk**. Business continuity managers should be particularly upset by this because it means that the extreme outcomes they focus on drop out of sight!

Auditors should identify when 'probability times impact' is being used, highlight the problem, and recommend something better.

Table 1 Top ways to go wrong with a risk register

Isolation from the business model
Incoherent mix of mental models
Impossible impact estimates
Undefined situations or events
Focusing on the 'top 10'
Taking risk as probability times impact
Index number multiplication
Confusing money with utility
Ignoring impact spread
Narrow perceptions due to poor calibration and lack of links between events