Gerald's Column
by Gerald Fitton

During the last few months I hope that I have got you thinking about the meaning of probability. This month I want to introduce you to a related concept, Risk Assessment.

Decisions and Risks

Mathematicians like definitions because that way they can give a special meaning to a commonly used word. By using a word you think you know they can fool you into believing that the subject matter is much simpler than it really is.

A Decision is something over which you have control. For example you can choose to take an umbrella with you when you go out. You make the Decision. On the other hand there are many Events over which you have no control. For example, you go out (your Decision) you take an umbrella (your Decision) and it may or may not rain (an Event over which you have no control). At best you will have an estimate of the Risk that it might rain. The extent of the Risk can be quantified as a Probability; you don't control its value.

Estimating Probabilities

In my earlier articles I suggested that there were two ways of estimating probabilities; then I introduced you to yet another. These three methods of estimating probabilities are:

1 We can use Symmetry to build up a mathematical model of the Probability Distribution Function (PDF). A good example is the unbiased coin. We use this useful concept of symmetry to give equal probability to Heads or Tails. By using the symmetry of such simple events we can build up quite sophisticated Probability Distribution Functions such as the Binomial, Poisson, Normal, Chi square and 't' PDFs.

   We have seen that it is possible to choose a PDF which is not suitable or one which is totally irrelevant to the event in which we have greatest interest. You will remember that the model of an unbiased coin was totally irrelevant to the pre 1920 sterling silver double headed half crown which I inherited from my grandfather. If you based a Decision to bet on Tails on the assumption that the unbiased coin PDF was valid then you would lose with monotonous regularity.

2 We can make or procure measurements of the statistic for a set of events which we believe are representative of the future event in which we have an interest. For example we could have made lots of measurements about the promptness of buses and used that data to forecast the probable length of our wait for a bus—which will never come.

   This method of estimating probabilities relies on the assumption that the results of past measurements can be applied to the event about which we have to make a Decision. You will remember that I was betting my library book fine on the arrival time of a bus.

3 We can use what I might call a personal estimate. People making Decisions often make them on the basis of their experience, knowledge or they simply have a 'hunch'. In my controversial opinion—and speaking as a mathematician—this is the 'best' method!

> Making a successful Decision in the case of the double headed coin and the diverted bus scenario rely on making such a personal decision, even though that Decision is to give up. Indeed, in my personal experience, those people who can make Decisions only when supported by a well respected PDF or incontrovertible experimental data make the worst kind of decision maker. Certainly I would never trust their judgement!
>
> Successful entrepreneurs often play their 'hunches' and might be hard put to rationalise their reasons for making the often excellent Decisions which they do. On the other hand addictive gamblers often convince themselves that the probability of winning is greater than other methods indicate. Some do 'hunches' well—some don't!

So, for my money, method 3 (faith in your own personal judgement) is the most important protocol for making Decisions. Your ability to quantify Risk will be improved by using methods 1 & 2, or employing staff to do the sums, but these two methods are no substitute for the greater inherent reliability of the much maligned and denigrated method 3.

Decision Making

It is my guess that those of you who are asked to provide a mathematical justification for a Decision for which the outcome is uncertain will choose to use by default a method based on calculating Expected Values (EV). It is a standard and well documented method which if I were not to include it would lay me open to well deserved criticism—a sin of omission.

It is not the only analytical technique. Indeed, in my view, Expected Value is not the best way to arrive at a set of numbers which are those most relevant to making a good Decision. I shall discuss these other techniques—but not this month.

Expected Values (EV)

Spreadsheets are a wonderful tool for all but the most simple EV calculations. I have chosen a somewhat contrived example to demonstrate this method using spreadsheet matrix multiplication. Although the example is simple the method can be expanded to much more complicated and complex scenarios.

I wonder if you have ever considered becoming a Market Trader. Let's pretend that's what we do. Next Saturday we shall go along to the market and we shall take with us either umbrellas or parasols. We have to pick them up from our supplier early in the week and, in this contrived example, we are able to pick up either all umbrellas (which will sell well if it is Wet) or all parasols (which will sell well if it is Dry). We are not allowed a mixture.

To keep the example simple, umbrellas and parasols are completely interchangeable in their contribution to our profits, etc, What we have to do is to select either umbrellas or parasols and sell as many as we can.

In this simplified example I have two pieces of information (knowledge) which are regarded as indisputable, the 'Outcome Matrix' and the 'Probability Matrix'. Of course, neither is indisputable—they are both estimates! The answer returned by the matrix calculation is the 'Expected Value Pay off Matrix'. It is these Expected Values which are used to make the Decision whether to take umbrellas or parasols to the market.

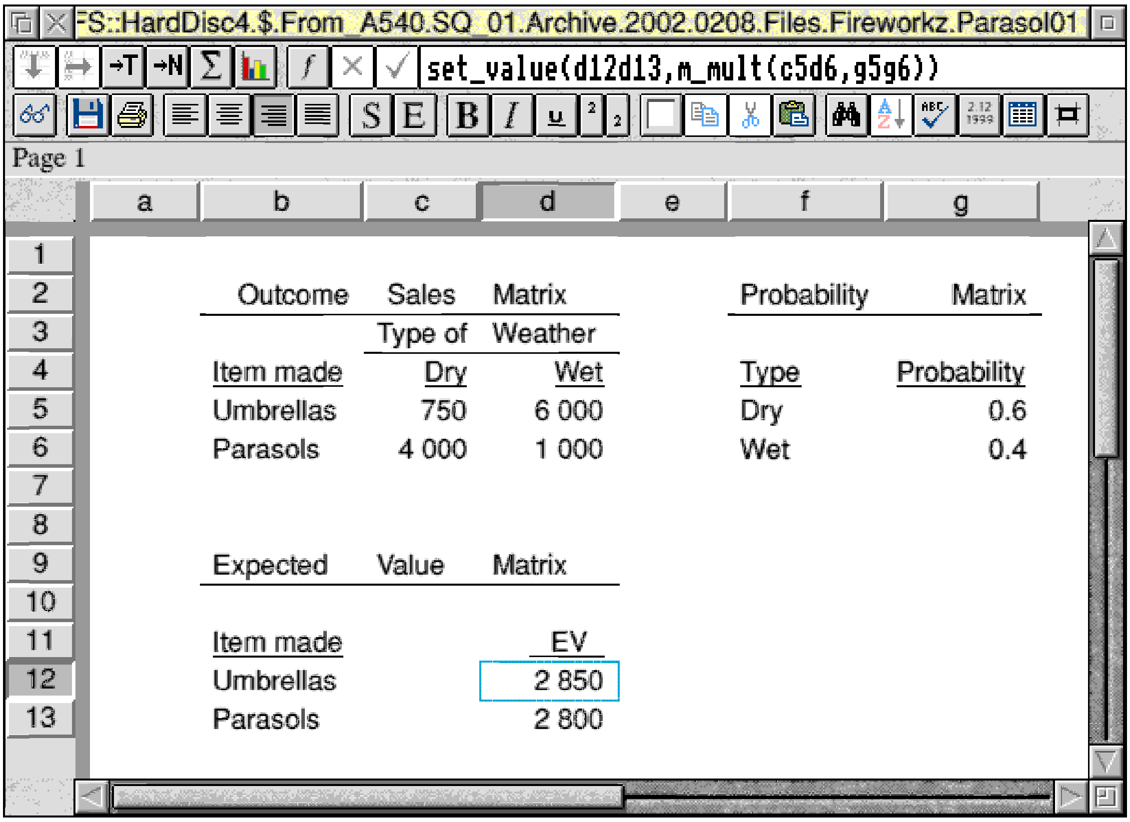
The Outcome Matrix

In this simple example Saturday's weather is either Dry or Wet. If we take umbrellas on a Dry day then we'll sell only 750 units; however, if it is Wet we'll do a lot better, we'll sell 6000 units. If our Decision is to take parasols rather than umbrellas then the corresponding Dry and Wet sales are 4000 and 1000 respectively. Of course these sales forecasts are not indisputable even though we assume so; they are estimates with a statistical distribution!

I am sure that you will see that if we make the correct Decision and if we're lucky with the weather then we'll sell either 4000 parasols (lovely) or 6000 umbrellas (even better). If we make the wrong decision our sales are down to either 1000 parasols if it's Wet. Worse, if it is Dry then we sell a mere 750 umbrellas.

The Meteorological Data

The Probability Matrix contains information about the weather. The interpretation of the data is that 60% of Saturdays are Dry and 40% are Wet. Although not at all relevant to the EV calculation I'd like you to ponder on the origin of these probabilities. For now let's say only that it is a method 2 estimate based on measurements taken over weeks or even years.



	a	b	c	d	e	f	g
1							
2		Outcome	Sales	Matrix		Probability	Matrix
3			Type of	Weather			
4		Item made	Dry	Wet		Type	Probability
5		Umbrellas	750	6 000		Dry	0.6
6		Parasols	4 000	1 000		Wet	0.4
7							
8							
9		Expected	Value	Matrix			
10							
11		Item made		EV			
12		Umbrellas		2 850			
13		Parasols		2 800			

The EV Matrix

There are different conventions for laying out the Outcome and Probability Matrices. Naturally I think the one I use is best! It is the one shown in the screenshot above.

In my version of the Outcome Matrix I write the choices over which I have control (and about which I must make a Decision) down the side and I write the Outcomes of the ‘chancy’ Event (over which I have no control) across the top. I write the Probability Matrix as a vertical column to the right of the pay off Matrix.

If I have enough space then I write the EV matrix to the right of both of these in the manner of a horizontal multiplication sum such as a * b = c. In the screenshot the EV matrix is below the two ‘knowledge’ matrices.

First I shall describe a manual version of the calculation and then the matrix version.

Let us start by assuming that our Decision is to sell umbrellas. Indeed, suppose we decide to take umbrellas every Saturday forever! On the Dry Saturdays we’ll sell 750 units and on the Wet ones we’ll sell 6000 units. It is Dry 60% of the time so that in six market days out of every ten (approximately) we’ll sell 750 units. It is Wet 40% of the time so that on four market days out of every ten we’ll sell 6000 units. On average we’ll sell 2850 units per market day. The calculation is: (750 * 0.6) + (6000 * 0.4) = 2850.

Similarly if our Decision is to take parasols then we’ll be a little worse off. We’ll sell (4000 * 0.6) + (1000 * 0.4) = 2800 units per market day.

If we have no idea whether the weather will be Dry or Wet then, over a long period of time, we’ll be 50 units per market day better off choosing umbrellas. We must decide to sell umbrellas because, over a long period of time, we’ll sell 50 more units per market day than if we always took parasols.

There is no such thing as absolute probability. We might be willing to pay for a meteorological forecast; it depends on what it costs and its accuracy. I shall discuss quantifying the value of more accurate information in a future article.

Matrix Multiplication

Laying out the matrices in my a * b = c manner facilitates the use of matrix multiplication (which is available in all but the simplest of spreadsheets) to find the EV Matrix.

The formula in cell d12 of the spreadsheet consists of a set_value(destination,source) and a m_mult(matrixA,matrixB) function.

Let’s deal with the multiplication first. The matrix I have called matrixA is the block c5d6, the Outcome Matrix. The matrix I have called matrixB is the block g5g6, the Probability Matrix. The matrix multiplication works across the rows of matrixA and, simultaneously, down the column(s) of matrixB. Each element of matrixA is multiplied by the corresponding element of matrixB and the results are added.

The sum of the products of every combination of rows of matrixA and column(s) of matrixB is calculated. The result is another matrix which is stored in most spreadsheets as a ‘hidden’ array. All that the set_value(destination,source) function does is to expand this ‘hidden’ array into the destination block, in our case the block d12d13.

Extension

I am sure that you will be able to visualise an expansion of this calculation.

For example, instead of the weather being Dry or Wet we could have included a 'Showery' category. The matrix multiplication formula I have included in cell d12 will still return a pair of answers which will be the expected number of units sold.

Similarly the categories of items made (umbrellas and parasols) could be expanded to include a greater range of options such as a waterproof parasol. Matrix multiplication will still return the expected number of units sold.

If umbrellas are more profitable than parasols (for example the profit on two pretty parasols equals the profit on three black umbrellas) then this can be taken into account with a further matrix or array multiplication.

Through the life of a project many Decisions have to be made at different times.

The complete set of Decisions and Outcomes can be developed into a set of interdependent matrix multiplications to return a set of expected values. These expected values can be used as the basis for making all the many Decisions within what is called a Decision Tree.

The Value of Information

The method of Expected Values is useful when we have probabilities available to us derived either from a reliable PDF, method 1, or from past experience, method 2, or even from a hunch, method 3, to estimate the probabilities of various outcomes.

Often it is possible to obtain more information and use it to reduce the extent of the uncertainty. Such information has a value. Usually it has a cost in money or time (or something more valuable than either). I shall explain how a relatively simple extension of this month's spreadsheet can be used to quantify the value of that extra information.

The Percentage Player

Using the Method of Expected Values to analyse Risk and thus to take Decisions is useful if the probability values are reliable and if similar projects are undertaken regularly. Over the long term the outcomes will 'average out' so that the benefits of 'good luck' will eventually compensate for the inevitable 'bad luck'. I have had to cover Expected Values because it is 'expected' of me and I can't let you down! I am aware that if I left it out I would receive so much email that my provider would cut me off!

Taking the Big Risk

Personally I don't place a lot of weight on this method of decision making. Life can be so much much more exciting by resolving not be a 'percentage player' playing the game the EV way. I remember being told by a wise businessman: "There is no Profit without Risk!"

There are times when the decision to be taken is unique. There may be no opportunity to ‘average out’ the good and bad outcomes. For example, if you ‘bet’ the whole of your wealth on a unique business venture, and lose—then you may not get a second opportunity. You may risk your life for a good cause; and you may lose. But was it a ‘good bet’?

In such cases probabilities are less relevant to the Decision than are the consequences that the outcome might have on future opportunities or on even more important things.

Classically there are three main criteria for considering the consequences of unique or rare but risky decisions. They are called MaxiMax, MaxiMin and MiniMax. You will have to be patient; the quantification of these criteria and the use of a spreadsheet to execute the relevant calculations must await another day.

Feedback

In the meantime if you wish to comment, criticise or contribute then please email or write to me. Or you may prefer to submit a note to Paul for inclusion in his Comment Column.