

1 Introduction

Complex decisions

Management decisions are often complex because they involve:

1. Risk and uncertainty: you may have to choose between various options, with various levels of risk. Stay in your current job, move to a new job with a different company, set up your own business.
2. Multiple objectives: you may want to maximise your immediate or long term earning potential, increase the amount of control you have over your work, increase the amount of time you can spend with your family.
3. A complex structure: if you decide to set up your own business, there are many steps to take and it can be done in a number of ways. It would be hard to summarise everything simply and clearly.
4. Multiple stakeholders: what is best for you may or may not be best for your partner and/ or your children, for example.

People have problems coping with this complexity. The human mind has limited information processing capacity and memory, and it is difficult to take all factors into account. To cope with this complexity we naturally tend to simplify problems, so that we can find a solution. This can lead to inconsistency and biases if not done in a scientific way.

The Role of Decision Analysis

How can Decision Analysis help us? The key lies in the word “analysis”, which refers to breaking something down into constituent parts. By doing this, we can solve simpler problems as components of a more complex whole.

When analysing a problem, you will need to make assumptions about the real problem which might be greatly simplifying. This can raise consciousness about key issues, and also potential dangers due to uncertainty.

The process of analysing a problem will also generate a clear rationale for a decision. Even if some of the assumptions you have to make can be questioned, the reasoning for the decision is clear.

This process may also make clear that there is important information that you lack to make an effective decision. This helps set priorities on what information needs to be obtained (and also what is not needed).

Note that the main purpose of decision analysis is to yield insights and understanding about the decision problem rather than to impose an optimal solution.

Rationality

The basic assumption underlying Decision Analysis is that of rationality. In any analysis we will be able to quantify the outcomes in a way to say some are more and less preferable, and it is assumed that we should always choose the “best”.

Our assumptions may lead us to a single optimal solution. If the decision maker fully accepts the assumptions of the model as (effectively) true, then the decision indicated by the analysis is the one that should be implemented.

Often, assumptions will be more questionable. What if there is a conflict between the results of the analysis and the intuition of a decision maker?

It may be that the analysis failed to capture some important aspect of problem, for example inadequately accounting for risk or potential reputational damage.

Alternatively, perhaps intuitive preferences were only partly formed or were inconsistent, or were based on past experience which no longer applies to changed circumstances.

Exploring this conflict can lead to deeper insights and understanding about the decision problem. This may lead to a new and more sophisticated analysis.

Applications of Decision Analysis

Decision Analysis has a wide range of applications. Anywhere where there are complex decisions to be made, a logical analysis using such methods is an important component in coming to a sound judgement.

An interesting set of examples are given in Goodwin and Wright. These include:

- 1) Strategic decision making at the Du Pont chemical company
- 2) Structural decision problems in response to the Chernobyl disaster
- 3) Selecting research and development projects in pharmaceutical companies
- 4) Systems acquisition for the US military
- 5) Supporting top level political decisions in Finland
- 6) Automating advice giving in building societies
- 7) Optimal fund allocation in a shampoo company

2 Decision-making under uncertainty

The maximin criterion

Suppose that a food manufacturer must choose the number of batches of a perishable product to make each day.

Each batch produced costs \$800 to make, and yields a revenue of \$1000 if it is sold. If it is not sold, it has to be thrown away.

Assume that the demand on any given day may be one or two batches, and so either one or two should be produced. This is summarised in the following table, including the profit made under each outcome.

Action/Demand (no. of batches)	1	2
Produce one batch	\$200	\$200
Produce two batches	-\$600	\$400

The Maximin criterion says to find the worst outcome for each choice you might make, and then pick the largest profit in this case. It assumes the worst case scenario.

The small number in row 1 is 200 which is larger than the smallest in row 2, -600. Thus we choose the option associated with row 1, i.e. to produce one batch.

The Expected Monetary Value (EMV) criterion

In the previous example, we had no measure of how likely the two possible levels of demand were. Suppose we have a good idea of how likely each is, through past experience.

Suppose that the probability of the demand being one batch is 0.3, and so the probabilities of it being two batches is 0.7. This is summarised as follows:

Action/Demand (no. of batches)	1 (prob 0.3)	2 (prob 0.7)
Produce 1 batch	\$200	\$200
Produce 2 batches	-\$600	\$400

We can calculating expected profits as follows:

Produce one batch:

$$\text{Expected daily profit} = (0.3 \times \$200) + (0.7 \times \$200) = \$200.$$

Produce two batches:

$$\text{Expected daily profit} = (0.3 \times -\$600) + (0.7 \times \$400) = \$100.$$

Thus the optimal choice is to produce one batch.

Sensitivity analysis: the above decision relies on the probabilities being as we assumed. What if they are different?

We shall draw a simple figure (see associated slides from Goodwin and Wright Chapter 1) to illustrate this.

Sensitivity analysis and Limitations of the EMV criterion

Often a problem is not simply to maximise expected income. For instance you may not take an option that has maximum expected return if it had a small change of catastrophe.

For example, should you insure your house against structural damage, fire, flood etc. The insurance company that you deal with will be hoping to make a profit from a policy it sells you. If it is on average making a profit, you are making a loss.

It makes sense for you to take out this type of insurance (but consider the above logic for other types of insurance), because if you do not and your house burns down you will be ruined. The damage done by a big loss is more than the benefit from an equivalent large gain.

The amount of money involved is still small for the insurance company, so they can disregard this, and so think in terms of expected reward and sell you the policy.

The key thing to consider is your attitude to risk.

Utility functions

The attitude to risk of a decision maker can be incorporated into a utility function. For each outcome there is a single value allocated to it.

The simplest case is if our outcomes can be expressed in terms of monetary values. We will use utility functions if we are not simply interested in expected value, for instance if a large loss would be disastrous.

We consider the following conference organising problem.

A conference organiser can choose between two venues, the Luxuria Hotel, and the Maxima Centre.

If there is high attendance then the expected profit at the Luxuria Hotel is \$30000, if low attendance \$11000. The equivalent figures for the Maxima Centre are \$60000 for high attendance, and an \$10000 loss for low attendance.

If the Luxuria Hotel is chosen, the probability of high attendance is 0.6, if the Maxima Centre, then this is 0.5.

We can calculating expected profits as follows:

Choose Luxuria Hotel:

$$\text{Expected profit} = (0.6 \times \$30000) + (0.4 \times \$11000) = \$22400.$$

Choose Maxima Centre:

$$\text{Expected profit} = (0.5 \times \$60000) + (0.5 \times -\$10000) = \$25000.$$

Thus the Maxima Centre is best if using the Expected Monetary Value criterion.

A utility function typically allocates a value 1 to the largest possible outcome, and 0 to the smallest. In this case the largest is \$60000 and the smallest $-\$10000$. What is the value of \$30000?

We ask our organiser, whether she prefers a guaranteed \$30000 or \$60000 with a probability p , and $-\$10000$ with probability $1 - p$. She will have her preferences for various p (e.g. $p = 0.6$ would give a slightly higher expected value for the lottery, but the consequences of the loss might mean she still prefers the safe \$30000).

When we find a p which she is indifferent between the choices, that is the utility of \$30000.

Suppose that she chooses the following utilities:

Money	Utility
\$60000	1.0
\$30000	0.85
\$11000	0.60
-\$10000	0.0

We can find the expected utilities as follows:

Choose Luxuria Hotel:

$$\text{Expected profit} = (0.6 \times 0.85) + (0.4 \times 0.6) = 0.75.$$

Choose Maxima Centre:

$$\text{Expected profit} = (0.5 \times 1.0) + (0.5 \times 0.0) = 0.5.$$

Thus the Luxuria Hotel is best if using this utility function.

Interpreting utility functions

The utility function described in the previous example is “risk-averse”. If you plot the utilities on the y -axis of a graph against monetary value on the x -axis, you will get a slope that is decreasing as the monetary value increases.

Our organiser’s utility function meant that she was neutral between a guaranteed \$30000 or \$60000 with a probability 0.85, and \$ – 10000 with probability 0.15.

Similarly she was neutral between a guaranteed \$11000 or \$60000 with a probability 0.6, and –\$10000 with probability 0.4.

Most people are “risk-averse” like this.

A “risk-neutral” utility function would just be a straight line. Someone using the Expected Monetary Value criterion is effectively using a risk-neutral utility function. Any “risk-averse” function will lie above this straight line.

A “risk-prone” utility function has a continuously increasing slope, and will lie below the risk-neutral line.

Utility functions for non-monetary attributes

Sometimes we face a problem where the outcomes are not described in monetary terms. We can still use the idea of utility to decide between options in this case, as long as it is clear what is a good outcome and what is a bad outcome.

For instance if we wish to develop a new product, and the outcome is the time it takes to develop it, then a short time is good and a long time is bad.

We can use the same lottery argument as before. When given a number of options (e.g. 1 year, 2 years or 8 years) we allocate utility 1.0 to the best option (1 year) and 0.0 to the worst option (8 years).

To find the utility of values between these two (e.g. 2 years), we consider obtaining that outcome with certainty against a lottery of the best outcome with probability p and the worst one with probability $1 - p$.

The value of p where we would be indifferent between the two options is the utility of the intermediate value that we are considering.

For example if this value is 0.95 in the above case, we are saying that $u(2 \text{ years}) = 0.95 u(1 \text{ year}) + 0.05 u(8 \text{ years}) = 0.95(1.0) + 0.05(0.0) = 0.95$.

Allais's paradox

We shall consider the following example, which illustrates that people do not always operate according to a logical idea of utility.

Consider the following decision. You can choose one of the following:

Option A: \$1000000 with probability 1.

Option B: \$1000000 with probability 0.89, \$5000000 with probability 0.1 and \$0 with probability 0.01.

Most people choose option A.

Now consider the alternative decision. You can choose one of the following:

Option X: \$5000000 with probability 0.1, \$0 with probability 0.9.

Option Y: \$1000000 with probability 0.11, \$0 with probability 0.89.

Most people choose option X.

Are these choices logically consistent?

Setting this up in terms of utility theory, we have three ordered sums of money. \$5000000 is the largest and so must have utility 1.0, \$0 is the smallest and so has utility 0.0. What could the utility of \$1000000, which we shall denote by $u(1M)$, be?

Option A has value $u(1M)$.

Option B has value $0.89u(1M) + 0.1 \times 1.0 + 0.01 \times 0.0$
 $= 0.1 + 0.89u(1M)$.

Option A was preferred to option B, so that

$u(1M) > 0.1 + 0.89u(1M) \Rightarrow u(1M) > 0.1/0.11 = 10/11$.

Option X has value $0.1+0=0.1$.

Option Y has value $0.11 u(1M)+0$.

Option X was preferred to option Y, so that

$0.1 > 0.11u(1M) \Rightarrow u(1M) < 10/11$.

Thus we have a contradiction.

This does not invalidate utility theory. Rather it shows that sometimes the choices that people make are perhaps not strictly logical. Nevertheless, these choices are consistently made, and this issue should be taken into account.

Multi-attribute Utility

Often we have more than one important attribute to try to optimise, e.g. cost and time.

Often making improvements in one comes at a cost to the other. For example we can speed up work by paying overtime to the workforce or hiring extra workers at some financial cost.

We thus need to decide in some way the relative importance of the different attributes.

Suppose that a project manager has a project to complete with time and cost to consider. Ideally he wants the project completed in 12 weeks, which was the timeframe specified by the customer.

Completing the project on time gains customer goodwill. He can improve the chance of doing this by hiring extra workers and running 24 hour shifts.

Assume that there are two options only:

- 1) Work normally, in which case with probability 0.1 the project is completed on time at a cost of \$50000, with probability 0.6 it is 3 weeks overdue at cost \$60000 and with probability 0.3 it is 6 weeks overdue at cost \$80000.
- 2) Hire extra labour, in which case with probability 0.8 the project is completed on time at a cost of \$120000, and with probability 0.2 it is 1 week late with a cost of \$140000.

How do we choose the best option?

Utility independence

Attribute A is utility independent of attribute B, if the decision makers preferences for gambles involving different levels of A, but the same level of B, do not depend on the level of attribute B.

This means that when comparing utilities we will be able to add separate utility functions for the different attributes; otherwise it is more complicated.

Utility functions for overrun time and project cost

We shall consider the following multi-attribute utility function

$$u(x_1, x_2) = k_1u(x_1) + k_2u(x_2) + k_3u(x_1)u(x_2),$$

where $k_3 = 1 - k_1 - k_2$.

First we find a single attribute utility function for each measure in turn, assuming independence.

For the overrun time, no overrun is best with utility of 1.0, and a 6 week overrun worst, with utility 0.0. If the project manager is indifferent between a certain overrun of 3 weeks, and a gamble offering no overrun with probability 0.6 and a 6 week overrun of 0.4, we get $u(3 \text{ weeks})=0.6$. Similarly, we assume that the manager finds $u(1 \text{ week})=0.9$.

We repeat the process with the cost. Suppose that we arrive at the following table for the project managers utilities for overrun and cost

Overrun (weeks)	Utility	Cost of project (\$)	Utility
0	1.0	50 000	1.00
1	0.9	60 000	0.96
3	0.6	80 000	0.90
6	0.0	120 000	0.55
		140 000	0.00

We need now to find the values of the k_i s.

Suppose that we consider either an outcome with the best overrun time (0 weeks) and the worst cost (\$140000) against a lottery, between the best cost and overrun times (0 weeks and \$ 50000) with probability k_1 and the worst of both (6 weeks and \$140000) with probability $1 - k_1$.

The bigger the value of k_1 , the more important that we consider the overrun time. Suppose that $k_1 = 0.8$.

We can find k_2 similarly, by changing the certain outcome to the worst overrun time but the best cost. Again, the larger k_2 , the more important we think the cost is. Suppose that $k_2 = 0.6$.

This means that relatively speaking, the project manager believes that the overrun time is the more important factor.

This finally gives us that $k_3 = 1 - k_1 - k_2 = -0.4$.

We can now find the utilities for any combination of overrun time and cost. For example

$$u(3 \text{ weeks, } \$ 60000) = 0.8 \times 0.6 + 0.6 \times 0.96 - 0.4 \times 0.6 \times 0.96 = 0.8256.$$

3 Decision trees

Constructing a decision tree

The initial node in a decision tree is the first decision of the decision-maker. It is represented by a square box. All nodes which represent a decision point are represented by such a square.

Leading off from this node are several branches, each representing a possible decision.

Typically, before the decision-maker has the chance to make a further decision, chance will determine which of various different scenarios that they face. Thus at the end of the first branches there will be a vertex which represents a chance occurrence (due to actions of the environment, the market etc.). Such a vertex is represented by a circle. All nodes representing a chance event are represented by such a circle.

Leading off from such a node are several branches, each represented a possible occurrence, accompanied by the probability that it occurs.

At the end of each of these branches, there might be an outcome (which specifies the reward to the decision-maker) or another square box, representing another decision to be made.

It is possible that there might be more than one square box or more than one circle in succession, but generally these can be consolidated, so we only consider alternating paths of squares and circles.

Finding the optimal policy

A policy is a a plan of action which specifies the decision at all square nodes that can be reached under the policy.

To find the optimal policy we work back from the end nodes of the tree. These end nodes, or terminal nodes, will have a value which may be a monetary value.

If the previous node on a branch is a circle, then its value is found by summing the products of the values of subsequent nodes and the probabilities of their occurrence.

If the previous node is a square, the value here is just the largest of the values of all of the subsequent nodes, since the decision-maker will choose the one giving the largest value.

We shall consider an example to illustrate this (see the food processor example in the Decision trees slides).

If the values in our problem are not monetary ones, but utilities, there is absolutely no difference to how the process is carried out.

We shall consider the same example where monetary values are replaced by utilities.

Decision trees for continuous probability distributions

In the decision tree problems that we have seen so far, there have only been a limited number of outcomes for any course of action, e.g. two outcomes, high or low sales.

However, for some problems there may be an effectively infinite number of possibilities. Actual sales could be any value up to many thousands or more. Thus we effectively have a continuous probability distribution.

One possibility is to approximate this continuous distribution with a discrete distribution, including only a small number of values. These could be a high, medium and low value.

The high value could be that which has only a 5% chance of being exceeded, the medium one that which has a 50% chance of being exceeded, and the low one that which has a 95% chance of being exceeded. This is the extended *Pearson-Tukey approximation*.

Probabilities must be allocated to the high, medium and low values. A natural choice would be 0.1, 0.8 and 0.1 respectively. However this significantly underestimates the variance. The actual probabilities often used are 0.185, 0.63 and 0.185 respectively.

Influence diagrams

Influence diagrams are a simple way of representing the interactions between decisions and events due to outside factors/ chance.

Just as in decision trees, there are boxes and branches between the boxes, where a square box represents a decision and a circle represents an event.

There are four possible sequences of two boxes. The first and the second can each be a circle or a square.

Two squares means a sequence of two decisions, where the first decision is known before the second is made.

Two circles mean two chance events that follow each other, e.g. the weather can influence the number of customers at a shop.

The other two orderings have either a decision made following a chance event,

or chance affecting the outcome of a decision, after the decision has been made.

Sometimes a decision-maker will construct an influence diagram to understand how key aspects interact in a conceptual way, before then using it to construct a decision tree.

4 Revising judgements in the light of new information

Bayesian methods

Suppose that we believe that there is a probability of 0.3 that a virus is present in the soil in a farm, and so a probability of 0.7 that it is absent. This is our *prior probability*.

A laboratory has a test for the presence of the virus.

If the soil is infected, the test has a probability of 0.9 of indicating that the virus is present.

If the soil is not infected, the test has a probability of 0.2 of indicating that the virus is present.

This provides *new information* to update our probability estimates.

There are thus four possibilities with the following probabilities:

Virus is present and the test says it is present: $0.3 \times 0.9 = 0.27$

Virus is present and the test says it is absent: $0.3 \times 0.1 = 0.03$

Virus is absent and the test says it is present: $0.7 \times 0.2 = 0.14$

Virus is absent and the test says it is absent: $0.7 \times 0.8 = 0.56$

Bayes' theorem says that the conditional probability of A given B, $P(A|B)$, satisfies

$$P(A|B) = \frac{P(A \text{ and } B)}{P(B)}.$$

Let A denote the event that the virus is present, and B denote the event that the test indicates that the virus is present.

We further denote A^c as the complement of A , i.e. the event that the virus is not present. Similarly we denote B^c as the event when the test indicates that the virus is not present.

Suppose that the test indicates that the virus is present. The updated probability of the virus being present is then $P(A|B)$, where $P(A \text{ and } B)=0.27$ and $P(B)=P(A \text{ and } B)+P(A^c \text{ and } B)=0.27+0.14=0.41$.

Thus $P(A|B)=0.27/(0.27+0.14)=0.66$. This is the *posterior probability*.

Similarly, if the test said that the virus was not present, we would need to find $P(A|B^c)$. This leads to the posterior probability of the virus being present being $0.03/(0.03+0.56)=0.05$.

Thus for a real problem we might be faced with the possibility of acquiring more information at some cost (e.g. buying the test in the above). This would lead to the figure in the accompanying slides.

The value of new information

In the above we have considered the possibility of buying new information that might help us make a decision. How much is that information potentially worth?

Suppose that the profit from the crop would be \$90000 if the virus was not present, but there would be a loss of \$20000 if it was present.

Further suppose that if the farmer decided not to plant the crop, then a profit of \$30000 could be made.

If there was no test available, the choice would be between

A - plant crops with expected return $0.7 \times \$90000 + 0.3 \times -\$20000 = \$57000$.

B - plant alternative crop with return \$30000.

Here option A is clearly the best.

You want to plant the crop if and only if there is no virus present. If there was a perfectly accurate test available then the return (without the cost of the test) would be

C - $0.7 \times \$90000 + 0.3 \times \$30000 = \$72000$

Buying this test would increase profits by $\$72000 - \$57000 = \$15000$. Thus you would be willing to pay up to this amount to have this test available.

Suppose we only have the imperfect test previously described. Assuming that this test is accurate enough that if it indicated the virus was present you would not plant the crop, and if it indicated it was absent you would plant it (and this assumption should be properly checked using the decision tree, but the test would be useless if this was not the case) then the following gives us the value of the test.

Recall that the probability that the test indicates the virus is present is 0.41, and so it is 0.59 that it indicates it is absent.

Recall also that if the test says the virus is absent, then it is actually present with probability ≈ 0.05 (actually 0.0508 to 4dp).

If the test says that the virus is present, then we obtain $0.9482 \times \$90000 + 0.0508 \times -\$20000 = \$84411$.

This means that the expected profit with the imperfect information from the test is

$$D - 0.41 \times \$30000 + 0.59 \times \$84411 = \$62102 \approx \$62100.$$

The expected value of the imperfect information (EVII) is thus $\$62100 - \$57000 = \$5100$.