# Decisions

What needs to be in a model for decisions under uncertainty?
- Options
- Outcomes and their values (£, time etc)
- Probabilities of the outcomes

Decision maker controls options, but not the outcomes

# Decision Ingredients

The basic components of a decision analysis are:

- A space of possible decisions, $D$.
- A set of possible outcomes, $\mathcal{X}$.

By choosing an element of $D$ you exert some influence over which of the outcomes occurs.

## Definition (Loss Function)

A loss function, $L : D \times \mathcal{X} \to \mathbb{R}$ relates decisions and outcomes. $L(d, x)$ quantifies the amount of loss incurred if decision $d$ is made and outcome $x$ then occurs.

An algorithm for choosing $d$ is a *decision rule*.

## Example (Insurance)

- You must decide whether to pay $c$ to insure your possessions of value $v$ against theft for the next year:

$$d = \{\text{Buy Insurance}, \text{Don't Buy Insurance}\}$$

- Three events are considered possible over that period:

$$x_1 = \{\text{No thefts.}\} \qquad x_2 = \{\text{Small theft, loss } 0.1v\}$$
$$x_3 = \{\text{Serious burglary, loss } v\}$$

- Our loss function may be tabulated:

| $L(d, x)$ | $x_1$ | $x_2$ | $x_3$ |
|---|---|---|---|
| Buy | $c$ | $c$ | $c$ |
| Don't Buy | 0 | $0.1v$ | $v$ |

# Uncertainty in Simple Decision Problems

- As well as knowing how desirable action/outcome pairs are, we need to know how probable the various possible outcomes are.
- We will assume that the underlying system is independent of our decision.
- Work with a probability space $\Omega = \mathcal{X}$ and the algebra generated by the collection of single elements of $\mathcal{X}$.
- It suffices to specify a probability mass function for the elements of $\mathcal{X}$.

## Example (Insurance Continued)

- There are 25 million occupied homes in the UK (2001 Census).
- Approximately 280,000 domestic burglaries are carried out each year (2007/08 Crime Report)
- Approximately 1.07 million acts of "theft from the house" were carried out.
- We might naïvely assess our pmf as:

$$
\begin{aligned}
&\text{“no theft”} && p(x_1) = \frac{25 - 1.07 - 0.28}{25} && = 0.946 \\
&\text{“small theft”} && p(x_2) = \frac{1.07}{25} && = 0.043 \\
&\text{“serious burglary”} && p(x_3) = \frac{0.28}{25} && = 0.011
\end{aligned}
$$

# The EMV Decision Rule

- If we calculate the expected loss for each decision, we obtain a function of our decision:

$$\bar{L}(d) = \mathbb{E}\left[L(d, X)\right] = \sum_{x \in \mathcal{X}} L(d, x) \times p(x)$$

- The *expected monetary value* strategy is to choose $d^\star$, the decision which minimises this expected loss:

$$d^\star = \arg\min_{d \in D} \bar{L}(d)$$

- This is sometimes known as a *Bayesian decision.*
- A justification: If you make a lot of decisions in this way the you might expect an averaging effect...

## Example (Still insurance)

- Here, we had a loss function:

| $L(d,x)$ | $x_1$ | $x_2$ | $x_3$ |
|---|---|---|---|
| Buy | $c$ | $c$ | $c$ |
| Don't Buy | $0$ | $0.1v$ | $v$ |

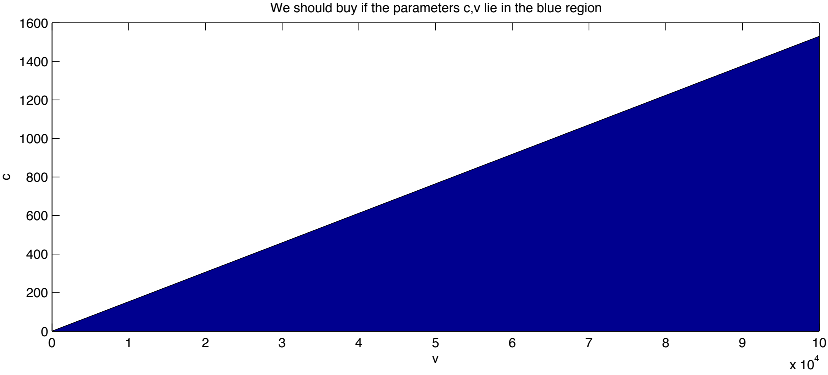
- And a pmf:

$$p(x_1) = 0.946 \qquad p(x_2) = 0.043 \qquad p(x_3) = 0.011$$

- Which give us an expected loss of:

$$\bar{L}(\text{Buy}) = 0.946c + 0.043c + 0.011c \qquad = c$$
$$\bar{L}(\text{Don't Buy}) = 0.946 \times 0 + 0.0043v + 0.011v \qquad = 0.0153v$$

- Our decision should, of course, depend upon $c$ and $v$.
- If $c < 0.0153v$ then the EMV decision is to buy insurance:



# Alternative Formulation

- Rather than defining a *loss function*, we could work with a *reward function*:

$$R(d, x) = -L(d, x)$$

- Leading to an expected reward:

$$\bar{R}(d) = \mathbb{E}\left[R(d, \cdot)\right] = -\mathbb{E}\left[L(d, \cdot)\right] = -\bar{L}(d)$$

- And the EMV rule becomes choose

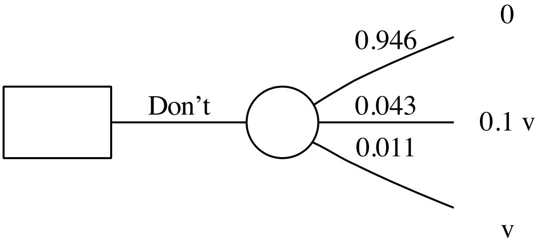
$$d^{\star} = \arg\max_{d \in D} \bar{R}(d)$$

# Graphical Representation: Decision Trees

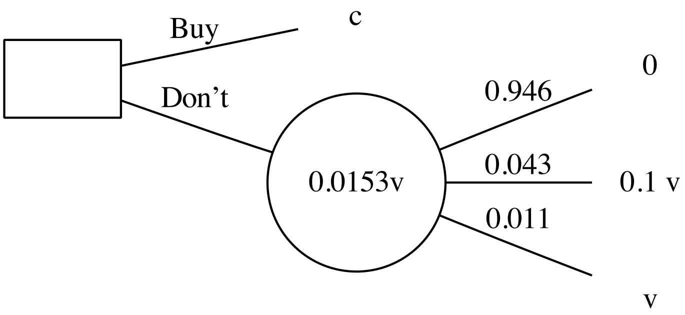
Drawing a decision tree:

1. Find a large piece of paper.
2. Starting at the left side of the page and working chronologically to the right...
    2.1 Indicate decisions with a □.
    2.2 Draw forks from decision *nodes* labelled with the decisions.
    2.3 Indicate sets of random outcomes with a ◯.
    2.4 Draw edges from random event *nodes* labelled with their (conditional) probabilities.
    2.5 Continue iteratively until all decisions and random variables are shown.
    2.6 At the right hand end of each path indicate the *loss/reward*.

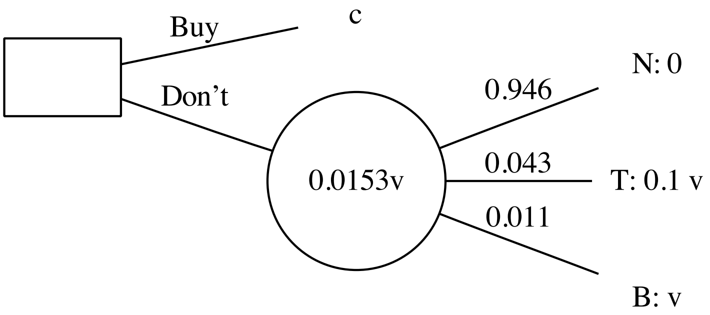
In the case of the insurance example, start with the first possible decision and we obtain:



only one outcome if we buy insurance:



In more complex examples, we should label the random events (say $N$ for no robbery, $T$ for small theft and $B$ for burglary...



# Calculation and Decision Trees

First, we fill in the expected loss associated with decisions:

- starting at the RHS of the graph, trace paths back to ◯ nodes.
- Fill in the rightmost ◯ nodes with the (conditional[4]) expected losses (the probabilities and losses are indicated at the edges and ends of the edges).
- For each decision node which now has values at the end of each branch, find the branch with the largest value.
- Eliminate all of the others.
- This produces a reduced decision tree.
- Iterate.
- When left with one path, this is the EMV decision!

[4] On all earlier events – i.e. ones to the left

# Do Not Laugh at Notations[5]

- At this point you may be thinking that this is a silly picture and that you'd rather just calculate things.
- That's all very well...
- but it gets harder and harder as decisions become more complicated.
- This graphical representation provides an easy to implement recursive algorithm and a convenient representation.
- This lends itself to automatic implementation as well as manual calculation.

[5]Invent them, for they are powerful. *RP Feynman.*

# More Complicated Cases

Consider this case:

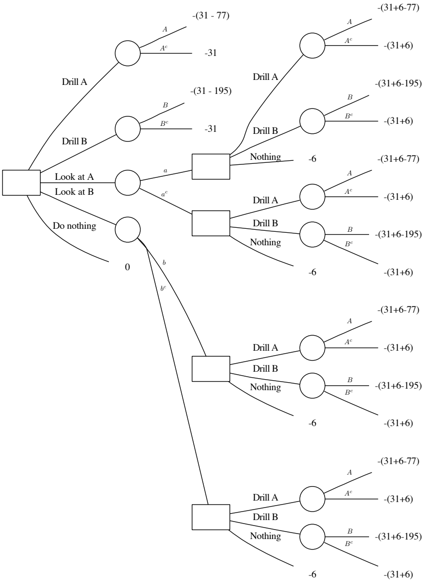
- You may drill (at a cost of £31M) in one of two sites: field A and field B.
  - If there is oil in site $A$ it will be worth £77M.
  - If there is oil in site $B$ it will be worth £195M.
- Or you may conduct preliminary trials in either field at a cost of £6M.
- Or you can do nothing. This is free.

This gives a set of 5 decisions to make immediately. If you investigate site $A$ or $B$ you must then, further, decide whether to drill there, in the other site or not at all (we'll make things simpler by neglecting the possibility of investigating both).

## Decision Trees — Example

Drilling: £31M
Investigation: £6M
Oil in A worth £77M
Oil in B worth £195M

We begin by constructing the tree without probabilities.



# Your Knowledge

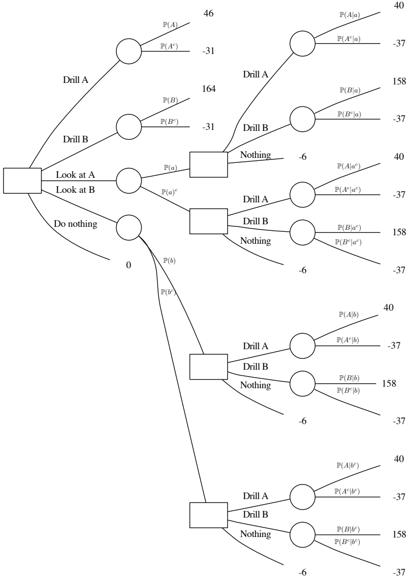
- The probability that there is oil in field $A$ is 0.4.
- The probability that there is oil in field $B$ is 0.2.
- If oil is present in a field, investigation will advise drilling with probability 0.8.
- If oil is not present, investigation will advise drilling with probability 0.2.
- The presence of oil and investigation results in one field provides no information about the other field.

# What do you know – formally?

Let $A$ be the event that there is oil in site $A$ and let $B$ be the event that there is oil in site $B$. Let $a$ be the event that investigation suggests there is oil in site $a$ and let $b$ be the event that investigation suggests that there is oil in site $b$.
The information on the previous page becomes:

- $\mathbb{P}(A) = 0.4$
- $\mathbb{P}(B) = 0.2$
- $\mathbb{P}(a|A) = \mathbb{P}(b|B) = 0.8$
- $\mathbb{P}(a|A^c) = \mathbb{P}(b|B^c) = 0.2$

## Decision Trees — Example




Then work out what each probability should be.

# Bayes Rule is Needed

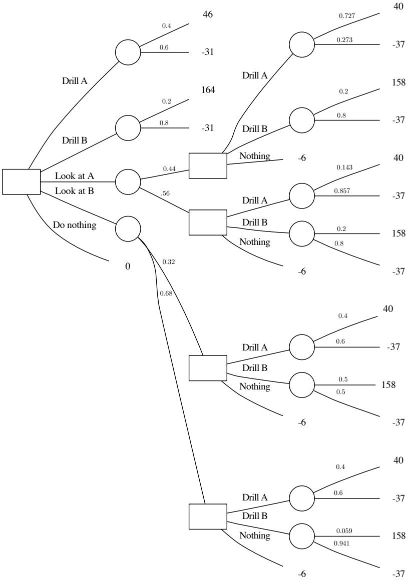
We really need to know the probability that oil is present in a field given that investigation indicates that there is (we know the converse).

$$\begin{aligned}\mathbb{P}(A|a) =& \frac{\mathbb{P}(a|A)\mathbb{P}(A)}{\mathbb{P}(a|A)\mathbb{P}(A) + \mathbb{P}(a|A^c)\mathbb{P}(A^c)} \\ =& \frac{0.8 \times 0.4}{0.8 \times 0.4 + 0.2 \times 0.6} = 0.727\end{aligned}$$

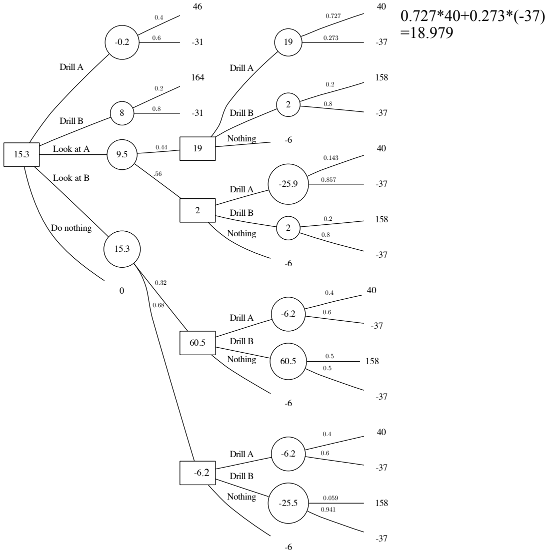
$$\begin{aligned}\mathbb{P}(B|b) =& \frac{\mathbb{P}(b|B)\mathbb{P}(B)}{\mathbb{P}(b|B)\mathbb{P}(B) + \mathbb{P}(b|B^c)\mathbb{P}(B^c)} \\ =& \frac{0.8 \times 0.2}{0.8 \times 0.2 + 0.2 \times 0.8} = 0.500\end{aligned}$$

## Decision Trees — Example

Then work out what each probability should be numerically.



## Decision Trees — Example




Then starting at the RHS calculate expectations and make optimal decisions to determine the solution.

Optimal decision (using EMV rule): "Look at B" (test drilling in B) with expected reward 15.3.
Optimal decision in reduced problem without allowing any test drilling: "Drill B" with expected reward 8.