Institute for Advanced Management Systems Research
Department of Information Technologies
Åbo Akademi University

Decision Making under Uncertainty - Tutorial

Robert Fullér

Directory

October 2, 2010
rfuller@abo.fi

Table of Contents

1. Decision problems

Many decison problems can be represented in the form of a decision table.

		States of nature			
Consequences		θ_1	θ_2	$\dots$	θ_n
	a_1	x_{11}	x_{12}	$\dots$	x_{1n}
	a_2	x_{21}	x_{22}	$\dots$	x_{2n}
Actions	$\dots$			$\dots$	
	$\dots$			$\dots$	
	a_m	x_{m1}	x_{m2}	$\dots$	x_{mn}

The general form of a decision table

The idea underlying this is that the consequence of any action is determined not just by the action itself but also by a number of external factors.

These external factors are both beyond the control of the decision maker and unknown to the decision maker at the time of the decision.

By a state of nature, or simply, state we shall mean a complete description of these extremal factors. Thus, if the decision maker knew the state of nature which would actually hold, he could predict the consequence of any action with certainty.

The state that actually holds will be called the **true state**. We shall assume that only a finite number of mutually exclusive states are possible and label these $\theta_1, \theta_2, \ldots, \theta_n$.

Similarly we assume that only a finite number of actions are available: $a_1, a_2, \ldots, a_m$.

Here x_{ij} stand for complete description of the possible consequences.

If the decision problem involves monetary outcomes, the x_{ij} may be single numbers. Otherwise we shall assume that the decision maker can measure the value of x_{ij} to him through some real-valued function v. By 'measure the value' we mean that

$$v(x_{ij}) > v(x_{kl})$$

if and only if the decision maker would prefer x_{ij} to x_{kl}.

Letting $v_{ij} = v(x_{ij})$, the general form of a decision table becomes

Consequences		States of nature			
		θ_1	θ_2	$\dots$	θ_n
	a_1	v_{11}	v_{12}	$\dots$	v_{1n}
	a_2	v_{21}	v_{22}	$\dots$	v_{2n}
Actions	$\dots$			$\dots$	
	$\dots$			$\dots$	
	a_m	v_{m1}	v_{m2}	$\dots$	v_{mn}

2. Decision under certainty

Here is assumed that the true state is known to the decision maker before he has to make his choice; i.e. he can predict the consequence of his action with certainty.

		State: θ
	a_1	v_1
	a_2	v_2
Actions	$\dots$	
	$\dots$	
	a_m	v_m

Because the v_i increase with the increasing value of the consequences to the decision maker, the optmal choice is, of course, to pick an action with the highest numerical value of v_i.

3. Decisions with risk

Although the decision maker does not know the true state of nature for certain, he can quantify his uncertainty through a possibility distribution:

$$(P(\theta_1), P(\theta_2), \ldots, P(\theta_n)).$$

We shall show later that, if the v_{ij} are measured in a certain way and if the decision maker is prepared to accept a certain definition of rationality, he should choose a_i to maximize the sum

$$\sum_{j=1}^{n} P(\theta_j) v_{ij}.$$

Thus sum is known as the *expected utility* of a_i.

4. Decisions under strict uncertainty

Here the decision maker feels that he can say nothing at all about the true state of nature.

Not only is he unaware of the true state, but he can not quantify his uncertainty in any way.

He is only prepared to say that each θ_j describes a possible state of the world and $\theta_1, \theta_2, \ldots, \theta_n$ is an exhaustive list of the possibilities.

How should a decision maker choose in a situation of strict uncertainty?

4.1. Wald's maximin return

Under the action a_i the worst possible consequence that can occur has a value to the decision maker of

$$s_i = \min_{j=1,\ldots,n} v_{ij}$$

We shall call s_i the security level of a_i, i.e. a_i guarantees the decision maker a return of at least s_i.

Wald (1950) suggested that the decision maker should choose a_k so that it has as large security level as possible. Thus the Wald maximin return criterion is:

$$choose\ a_k\ such\ that\ s_k = \max_{i=1,\ldots,m} s_i = \max_i \min_j v_{ij}.$$

We note that it is a very pessimistic criterion of choice; for its general philosophy is to assume that the worst will happen.

4.2. Hurwicz's optimism-pessimism index

Define the optimism level of a_i to be

$$o_i = \max_{j=1,\ldots,n} v_{ij}.$$

Thus o_i is the value of the best consequence that can result if a_i is taken.

The maximax return criterion is: choose a_k such that

$$o_k = \max_{i=1,\ldots,m} o_i = \max_i \max_j v_{ij}.$$

Hurwicz (1951) argued that a decision maker should rank actions according to the weighted average of the security and optimism levels: $\alpha s_i + (1-\alpha) o_i$, where, $0 \leq \alpha \leq 1$, is the optimism-pessimism index of the decision maker.

Hurwicz recommends the decision rule: choose a_k such that

$$\alpha s_k + (1-\alpha)o_k = \max_i\{\alpha s_i + (1-\alpha)o_i\}.$$

4.3. Savage's minimax regret

Savage defined the regret of a consequence as

$$r_{ij} = \max_{l=1,\ldots,m}\{v_{lj}\} - v_{ij}$$

that is the difference between the value resulting from the best action given that θ_j is the true state of nature and the value resulting from a_i under θ_j.

He argued that the r_{ij} should replace the v_{ij} in the decision table and that in this new regret table the decision maker should choose by following Wald's pessimistic approach, but remembering that regrets are 'losses' not 'gains'.

Each action is assigned the index

$$\rho_i = \max_{j=1,\ldots,n} r_{ij}$$

that is, the worst regret that can result from action a_i, and then an action should be chosen to minimize ρ_i, i.e., choose a_k such that

$$\rho_k = \min_{i=1,\ldots,m} \rho_i = \min_i \max_j r_{ij}.$$

4.4. Laplace's 'principle of insufficient reason'

Laplace (1825) argued that 'knowing nothing at all about the true state of nature' is equivalent to 'all states having equal probability'.

If action a_i is selected and if all states have equal probability, then the decision maker faces an expected value from the uncertain consequences, hence he should seek to maximize the expected value of his choice i.e., choose a_k such that

$$\sum_{j=1}^{n}(1/n)\, v_{kj} = \max_{i=1,\ldots,m} \sum_{j=1}^{n}(1/n)\, v_{ij}.$$

In comparing these four criteria of choice, we emphasize that each may suggest a different choice in the same problem.

	θ_1	θ_2	θ_3	θ_4	s_i	o_i	$\sum_j (1/n) v_{ij}$
a_1	2	2	0	1	0	2	5/4
a_2	1	1	1	1	1	1	1
a_3	0	4	0	0	0	4	1
a_4	1	3	0	0	0	3	1

Milnor's example.

	θ_1	θ_2	θ_3	θ_4	ρ_i
a_1	0	2	1	0	2
a_2	1	3	0	0	3
a_3	2	0	1	1	2
a_4	1	1	1	1	1

Regret table for Milnor's example.

For example,

$$r_{22} = \max\{2, 1, 4, 3\} - 1 = 4 - 1 = 3.$$

Laplace's criterion will pick a_1, Wald's criterion will choose a_2, Hurwicz's criterion assigns the indices $2(1-\alpha)$, 1, $4(1-\alpha)$ and $3(1-\alpha)$ and will pick a_3 for any $\alpha < 3/4$. Savage's minimax regret will pick a_4.

4.5. The St. Petersburg paradox

Daniel Bernoulli spent the years 1725 through 1733 in St. Petersburg, where he analyzed the following coin-toss problem: A fair coin will be repeteadly tossed until it falls 'heads'. If this occurs on the nth toss, you will receive $\$2^n$. How much would you be prepared to pay to play this game once?

Consider the expected payoff. The probability of a fair coin for landing 'heads' on the nth toss is

$$\left(\frac{1}{2}\right)^n.$$

So your expected payoff is

$$\sum_{n=1}^{\infty}\left(\frac{1}{2}\right)^n 2^n = \sum_{n=1}^{\infty} 1 = \infty.$$

Thus, however much you pay to enter the game, you may expect to win more. It is why, you should be willing to risk everything that you posses for the opportunity of playing this game just once. *But no one would consider such a course of action to be rational.* The message of this example is that expected monetary return is not necessarily the appropriate criterion to use in valuing uncertain prospects.

4.6. Blood-bank inventory problem

A hospital has a blood-bank from which it wishes to satisfy its day to day needs for blood transfusion. Two criteria are important for evaluating the effectiveness of a control policy: blood shortage and blood outdating.

Uncertainty is present because the demand for any blood type in any given period is unknown, although past records will enable informed predictions to be made.

4.7. Airport-siting decisions

Given that a decision has been made to build it, where should a government select as a site for a new major international airport?

Typically the choice of site will be limited to a relatively few candidate sites by reasons of geographical requirement, by the location of pre-existing airports and air corridors, and by engineering feasibility.

But, although the choice is very limited, the decision is a very complex and very controversial one. The government must find a suitable compromise between many, often diametrically opposed objectives.

To minimize inconvenience to the traveller, the airport should be sited as near as possible to the capital; yet to maximize safety and minimize inconvenience to the domestic population the airport should be as far as possible from the capital.

4.8. Preferences over time

Numerous decision problems concern projects in which the cost and benefits accrue over a number of years. If costs and benefits are entirely monetary then we talk about *capital budgeting* or *capital investment decisions*. Two kinds of decisions may arise: **accept-or-reject** and **ranking**.

In accept-or-reject decisions each project is considered independently of all other projects. Thus a portfolio of accepted projects is built up from several independent decisions.

In ranking decisions, all the available projects are compared and ranked in order of favourability with the intention of adopting a single project: the most favourable. Several decision rules have been suggested to help decision makers rank projects which involve timestreams of costs and benefits. The simplest rule is to compare projects according to the time that they take to break even, that is according to their payback period.

Year	A	B	C	D	E	F
0	-10	-10	-10	-1	-16	-16
1	5	5	2	0.5	16	3.2
2	5	5	8	0.5	5	19.2
3	0.0	5	5	0.5	0.0	0.0
4	0.0	5	5	0.5	0.0	0.0

Cash-flow from projects.

The **payback method** would consider E the most favourable project and would be unable to distuingish between A,B,C,D and F. This method has a number of faults:

1. No account is taken of the total profit after payback.
2. No account is taken of the size of the investment.
3. No account is taken of the distribution of income and expenditure (compare B and C).
4. The payback period is not necessarily clearly defined if a project involves investments in later years.

The **accounting rate of return** (ARR) may be defined as

$$\frac{\text{The average profit per annum of a project}}{\text{Its capital layout}} \times 100\%$$

That is,

$$
\begin{aligned}
\mathrm{A}RR(A) &= \frac{(5+5-10)/2}{10} \times 100\% = 0\% \\
\mathrm{A}RR(B) &= \frac{(5+5+5+5-10)/4}{10} \times 100\% = 25\% \\
\mathrm{A}RR(C) &= \mathrm{A}RR(D) = 25\% \\
\mathrm{A}RR(E) &= 15.6\% \\
\mathrm{A}RR(F) &= 20\%.
\end{aligned}
$$

So, according to the accounting rate of return, projects B, C and D are jointly best, and the remaining projects are ranked in the order: F,E,A.

We note that

1. No account is taken of the size of the investment.
2. No account is taken of the distribution of income and expenditure (compare B and C).

Neither of the above methods involves discount factors.

Suppose that \$1 now is worth $\$(1 + r)$, $r > 0$, in a year's time.

Thus we can reduce a timestream of cash-flows to the **net present value** (NPV) of a project:

$$\mathrm{N}PV(A) = -10 + \frac{5}{1+r} + \frac{5}{(1+r)^2}$$

where r is generally known as the **discount rate**, or **marginal**

time preference rate, or the **minimum required rate of return**.

We shall now assume that r has been determined for our problem as 10%, i.e. $r = 0.1$. Hence,

$$\begin{aligned}
\text{N}PV(A) &= -1.32 \\
\text{N}PV(B) &= 5.850 \\
\text{N}PV(C) &= 5.601 \\
\text{N}PV(D) &= 0.585 \\
\text{N}PV(E) &= 2.678 \\
\text{N}PV(F) &= 2.777
\end{aligned}$$

So the projects should be ranked as: B, C, F, E, D and A.

The **internal rate of return** (IRR) is defined to be the value of

r such that the NPV of a project is zero. Thus find the IRR of A we need to solve

$$-10 + \frac{5}{1+r} + \frac{5}{(1+r)^2} = 0.$$

Setting $x = 1/(1+r)$ gives

$$-2 + x + x^2 = 0,$$
$$(x+2)(x-1) = 0.$$

Clearly $x > 0$. That is,

$$\frac{1}{1+r} = 1,$$

and , therefore,

$$r = 0.$$

For the other projects we find

$$\text{I}RR(A) = 0$$
$$\text{I}RR(B) = 35$$
$$\text{I}RR(C) = 32$$
$$\text{I}RR(D) = 35$$
$$\text{I}RR(E) = 25$$
$$\text{I}RR(F) = 20$$

So this rule suggests that B and D are the most attractive projects.

5. Preference orders and value functions

We shall write $a \succ b$ to mean that the decision maker strictly prefers object 'a' to object 'b'. We shall write $a \succeq b$ to mean the decision maker weakly prefers object 'a' to object 'b'. We will make four specif ic demands of the consistency that we expect of a rational man's use of $\succeq$.

Axiom 1. *Comparability. $\succeq$ is comparable: viz. $\forall\ a,\ b \in A$, $a \succeq b$ or $b \succeq a$ or both hold.*

Axiom 2. *Transitivity. $\succeq$ is transitive: viz. $\forall\ a,\ b, c \in A$, if $a \succeq b$ and $b \succeq c$, then $a \succeq c$.*

Axiom 3. *Consistency of indifference and weak preference. For any pair of objects $a, b \in A$*

$$a \sim b \iff (a \succeq b \ \textit{and} \ b \succeq a)$$

Axiom 4. *Consistency of strict preference and weak preference. For any pair of objects* $a, b \in A$

$$a \succ b \iff b \not\succeq a.$$

Definition 5.1. *Let* A *be a set of alternatives and* $\succeq$ *the decision maker's preference relation over them. Then*

$$v \colon A \to \mathbb{R},$$

is an ordinal value function representing these preferences if

$$v(a) \geq v(b) \iff a \succeq b.$$

And we say that v *agrees with or represents* $\succeq$ *over* A*.*

Theorem 5.1. *For a finite set of objects*

$$A = \{a_1, a_2, \ldots, a_n\}$$

with a weak order $\succeq$ *obeying Axioms 1-4, we may always construct an agreeing ordinal value function.*

Let $v(a_i)$ = no. of objects $a_j \in A$ such that $a_i \succeq a_j$.

Theorem 5.2. *Let $\succeq$ be a weak order. Then v and w are ordinal value functions both agreeing with $\succeq$ if and only if there exists $\phi \colon \mathbb{R} \to \mathbb{R}$, a strictly increasing function, such that*

$$w(a) = \phi[v(a)],\ \forall a \in A.$$

Proof. Suppose that v is an ordinal value function agreeing with $\succeq$ and that ϕ is a strictly increasing function. Then

$$\begin{aligned} a \succeq b &\iff v(a) \geq v(b) \\ &\iff \phi(v(a)) \geq \phi(v(b)) \iff w(a) \geq w(b). \end{aligned}$$

Hence w is an ordinal value function agreeing with $\succeq$. □

We shall say that ordinal value functions are unique *up to strictly* *increasing transformations*.

Strictly increasing transformations are known as the **admissible transformations** of ordinal value functions.

They are admissible because they do not affect the representation of the underlying preference.

The University Grant Authority (UGA) has responsibility for allocating resources to the state's two universities, which both have the same number of students.

It has already been decided to allocate the resources in relation to the academic quality of student admissions at the two universities. The question that concerns the authority is how to measure this quality. School leavers take a single examination, which classfies their abilities into five grades: A, B, C, D and

E. It is suggested that these grades should betranslated numerically as, 5, 4, 3, 2 and 1.

grade	*Roventie University*	*Altentie University*
A	5	30
B	65	5
C	5	25
D	5	35
E	20	5

Breakdown of admissions.

It is further suggested that resources should be allocated in the ratio of the mean points attained by the student intake at each university.

The mean points attained by the student intakes are, therefore,

Roventie, mean points per student

$$= 0.05 \times 5 + 0.65 \times 4 + 0.05 \times 3 + 0.05 \times 2 + 0.20 \times 1 = 3.3$$

Altentie, mean points per student

$$= 0.30 \times 5 + 0.05 \times 4 + 0.25 \times 3 + 0.35 \times 2 + 0.05 \times 1 = 3.2$$

Since $3.3 > 3.2$, Roventie will receive a higher allocation than Altentie.

However, suppose the UGA had assigned points to grades rather differently. Perhaps they might have used the minimum marks necessary to have attained each grade in the school-leaving examination.

Thus the grades might translate numerically as: 90, 75, 65, 50, and 30.

Under this scheme the mean points attained by the student intakes are:

Roventie, mean points per student

$= 0.05\times90+0.65\times75+0.05\times65+0.05\times50+0.20\times30 = 65.$

Altentie, mean points per student

$= 0.30\times90+0.05\times75+0.25\times65+0.35\times50+0.05\times30 = 66.$

In this case Altentie will receive a higher allocation than Roventie.

The allocation of resources by the UGA would seem to depend upon the rather arbitrary choice of a numerical scale with which to represent the examination performance.

Here we have an example that shows that the ordering of the mean value of v over one group and that over another group is not necessarily preserved by a strictly increasing transformation. In general, let $\{a_i | i = 1, 2, \ldots, n\}$ be a set of n objects. Consider the two sums

$$\sum_{i=1}^{n} \lambda_i v(a_i), \quad \sum_{i=1}^{n} \mu_i v(a_i),$$

for any two sets of coefficients $\{\lambda_i\}$ and$\{\mu_i\}$. Then

$$\sum_{i=1}^{n} \lambda_i v(a_i) \geq \sum_{i=1}^{n} \mu_i v(a_i),$$

does not imply that for any arbitrary strictly increasing ϕ

$$\sum_{i=1}^{n} \lambda_i \phi(v(a_i)) \geq \sum_{i=1}^{n} \mu_i \phi(v(a_i)).$$

6. Lotteries

Suppose that you are offered the choice between the following gambles,

A : $p = 1/2$ to win \$1 and $1 - p = 1/2$ to lose \$0.60

B: $p = 1/2$ to win \$10 and $1 - p = 1/2$ to lose \$5

If you compare these gambles in terms of expected monetary payoff you will choose B: Expected monetary payoff of A

$$1/2 \times 1 + 1/2 \times (-0.6) = 0.20$$

Expected monetary payoff of B

$$1/2 \times 10 + 1/2 \times (-5) = 2.50$$

However a poor man would choose gamble A!

Consider the function

$$u(10) = 1,$$
$$u(1) = 0.2,$$
$$u(-0.6) = -0.1,$$
$$u(-5) = -1.$$

Expectation of u for A is

$$1/2 \times u(1) + 1/2 \times u(-0.6) = 0.2.$$

Expectation of u for B is

$$1/2 \times u(10) + 1/2 \times u(-5) = 0.$$

The expected utility rule can describe poor man's preferences!

A typical lottery may be represented as

$$l = \langle p_1, x_1; p_2, x_2; \ldots; p_r, x_r \rangle$$

where $p_i \geq 0$ is the probablity of winning x_i, $i = 1, 2, \ldots, r$, and

$$\sum_{i=1}^{r} p_i = 1.$$

Of course, it is quite possible that several $p_i = 0$, indicating that certain prizes are not possible in a particular lottery.

This lottery is a simple lottery because the prize is determined. In a compoud lottery some or all the 'prizes' may be entries into further lotteries.

For instance, the compound lottery,

$$\langle q_1, l_1; q_2, l_2; \ldots; q_s, l_s \rangle$$

gives probabilities $q_i \geq 0$ of winning an entry into lottery l_i, $i = 1, 2, \ldots, s$ and

$$q_1 + q_2 + \cdots + q_s = 1.$$

We shall let A be the set of all possible prizes together with a set of simple and finitely compounded lotteries.

We shall make several reasonable assumptions concerning the consistency of the decision maker's preferences if he is to be considered rational.

It can be shown that these assumptions imply the existence of a utility function

$$u\colon X \to \mathbb{R}$$

where $X = \{x_1, x_2, \ldots, x_r\}$ is the set of possible prizes, such that the decision maker holds:

$$x_i \succeq x_j \iff u(x_i) \geq u(x_j)$$

for any pair $x_i, x_j \in X$ and

$$\langle p_1, x_1; p_2, x_2; \ldots; p_r, x_r\rangle \succeq \langle p'_1, x_1; p'_2, x_2; \ldots; p'_r, x_r\rangle$$
$$\iff \sum_{i=1}^{r} p_i u(x_i) \geq \sum_{i=1}^{r} p'_i u(x_i).$$

for any pair of simple lotteries in A. The first condition shows that u is an ordinal value function on the set of prizes X.

The second condition shows that u possesses the expected utility property on the set of simple lotteries; i.e. that it is appropriate to choose betweeen simple lotteries according to the expected utility rule. We shall assume the following

Axiom 1. *Weak ordering. The decision maker's preferences over A form a weak order. Without loss of generality, we shall label the prizes such that $x_1 \succeq x_2 \succeq \cdots \succeq x_r$.*

Axiom 2. *Non-triviality. To avoid triviality we shall assume that he strictly prefers x_1 to x_r. That is, $x_1 \succ x_r$.*

Axiom 3. *Reduction of compound lotteries. Consider a compound lottery $l = \langle q_1, l_1; q_2, l_2; \ldots; q_s, l_s\rangle$ which gives as prizes entries into further simple lotteries $l_1, l_2, \ldots, l_s$, where*

$$l_j = \langle p_{j1}, x_1; p_{j2}, x_2; \ldots; p_{jr}, x_r\rangle$$

for $j = 1, 2, \ldots, s$.

Let l' be the simple lottery $\langle p_1, x_1; p_2, x_2; \ldots; p_r, x_r \rangle$ where

$$p_i = q_1 p_{1i} + q_2 p_{2i} + \cdots + q_s p_{si}$$

for $i = 1, 2, \ldots, r$. Then the decision maker must hold $l \sim l'$.

Axiom 4. *Substitutability. Let $b, c \in A$ such that the decision maker holds $b \sim c$. Let $l \in A$ be any lottery, simple or compounded, such that*

$$l = \langle \ldots; q, b; \ldots \rangle$$

i.e. there is a probability q that b is a direct outcome of l.

Let l' be constructed from l by substituting c for b and leaving all other outcomes and all probabilities unchanged, viz.

$$l' = \langle \ldots; q, c; \ldots \rangle$$

Then the decision maker holds $l \sim l'$.

Axiom 5. *The reference experiment*

$$x_1 p x_r \in A,\ \forall p,\ 0 \le p \le 1.$$

where

$$x_1 p x_r = \langle p, x_1; 0, x_2; \ldots; (1 - p_r), x_r \rangle$$

Axiom 6. *Monotonicity.* $x_1 p x_r \succeq x_1 p' x_r \iff p \ge p'$. *In other words, the more chance he has of winning* x_1, *the more the decision maker prefers the reference lottery.*

Axiom 7. *Continuity:* $\forall x_i \in X \quad \exists u_i$, $0 \le u_i \le 1$, *such that*

$$x_i \sim x_1 u_i x_r.$$

We have chosen to use u_i *rather than* p_i *to denote the probability in the reference lottery which gives indifference with* x_i, *because the utility function whose existence we shall shortly show, is such that* $u(x_i) = u_i$.

Theorem 6.1. *If the decision maker's preferences over A obey Axioms 1-7, there exists a utility function u on X which represents $\succeq$ in the sense*

$$x_i \succeq x_j \iff u(x_i) \geq u(x_j)$$

for any pair x_i, $x_j \in X$ and

$$\langle p_1, x_1; p_2, x_2; \ldots; p_r, x_r \rangle \succeq \langle p'_1, x_1; p'_2, x_2; \ldots; p'_r, x_r \rangle$$
$$\iff \sum_{i=1}^{r} p_i u(x_i) \geq \sum_{i=1}^{r} p'_i u(x_i).$$

for any pair of simple lotteries in A.

Example 6.1. *Consider the following lotteries*

	θ_1	θ_2	θ_3
l_1	*6*	*14*	*8*
l_2	*10*	*8*	*10*

$$P(\theta_1) = 1/4 \quad P(\theta_2) = 1/2 \quad P(\theta_3) = 1/4$$

Suppose that we have the following indifference relations:

$$10 \sim 14(0.8)6, \quad 8 \sim 14(0.6)6,$$

Which lottery is better?

Solution 1. *Without loss of generality we can assume that*

$$u(14) = 1, \quad u(6) = 0.$$

From the utility indifference relationship we find

$$u(10) = 0.8 \times u(14) + 0.2 \times u(6) = 0.8,$$
$$u(8) = 0.6 \times u(14) + 0.4 \times u(6) = 0.6,$$

Applying the expected utility rule we get

$$E(u|l_1) = 1/4 \times u(6) + 1/2 \times u(14) + 1/4 \times u(8) = 0.65$$

and

$$E(u|l_2) = 1/4 \times u(10) + 1/2 \times u(8) + 1/4 \times u(10) = 0.7$$

That is,

$$E(u|l_2) > E(u|l_1)$$

and, therefore, $l_2 \succ l_1$.

Let u be a utility function over a finite X and let $w = \alpha u + \beta$ with $\alpha > 0$. Then w is also an ordinal value function over X, moreover,

$$\sum_{i=1}^{r} p(x_i) w(x_i) \geq \sum_{i=1}^{r} p'(x_i) w(x_i) \iff$$

$$\alpha \left(\sum_{i=1}^{r} p(x_i) u(x_i) \right) + \beta \geq \alpha \left(\sum_{i=1}^{r} p'(x_i) u(x_i) \right) + \beta.$$

Since $\sum_{i=1}^{r} p(x_i) = \sum_{i=1}^{r} p'(x_i) = 1$, we get

$$\sum_{i=1}^{r} p(x_i) w(x_i) \geq \sum_{i=1}^{r} p'(x_i) w(x_i)$$

$$\iff$$

$$\sum_{i=1}^{r} p(x_i) u(x_i) \geq \sum_{i=1}^{r} p'(x_i) u(x_i).$$

Theorem 6.2. *If u is a utility function on X, then*

$$w = \alpha u + \beta$$

($\alpha > 0$) is also a utility function representing the same preferences. Conversely, if u and w are two utility functions on X representing the same preferences, then there exists $\alpha > 0$ and β such that

$$w = \alpha u + \beta.$$

The utility function is unique up to positive affine transformation.

7. Risk attitudes

Associated with any lottery are two expectations: its expected monetary value

$$E(x|p) = \begin{cases} \displaystyle\sum_{i=1}^{r} p(x_i)x_i & \text{if } X \text{ is finite} \\ \displaystyle\int_X p(x)x dx & \text{otherwise} \end{cases}$$

and its expected utility

$$E(u|p) = \begin{cases} \displaystyle\sum_{i=1}^{r} p(x_i)u(x_i) & \text{if } X \text{ is finite} \\ \displaystyle\int_X p(x)u(x)dx & \text{otherwise} \end{cases}$$

The expected monetary value is simply the average payoff in monetary terms that results from the lottery.

Related to the expected utility of a lottery is its certainty equivalent, x_c, which is the monetary value that the decision maker places on the lottery.

If he were offered the choice, the decision maker would be indifferent between accepting the monetary sum x_c for certain and accepting the lottery.

Thus

$$u(x_c) = E(u|p)$$

or, equivalently,

$$x_c = u^{-1}(E(u|p)).$$

The risk premium of a lottery is

$$\pi = E(x|p) - x_c.$$

Consider any lottery with only to possible prizes. Then

$$E(x|p) = px_1 + (1-p)x_2,$$

$$E(u|p) = pu(x_1) + (1-p)u(x_2),$$

where p is the probability of winning x_1.

The risk premium is

$$\pi = px_1 + (1-p)x_2 - u^{-1}(pu(x_1) + (1-p)u(x_2)),$$

For the decision maker to be risk averse the risk premium must be non-negative for all x_1, x_2 and for all $0 \leq p \leq 1$. Equivalently,

$$px_1 + (1-p)x_2 \geq u^{-1}(pu(x_1) + (1-p)u(x_2)),$$

for all $0 \leq p \leq 1$.

On noting that u is strictly increasing

$$u(px_1 + (1-p)x_2) \geq pu(x_1) + (1-p)u(x_2),\ 0 \leq p \leq 1.$$

Thus, if an individual is risk averse, his utility function must be concave. For general lotteries the Jensen's inequality is needed to prove the results:

$$u(x_c) = E(u|p) \leq u(E(x|p)),$$

by Jensen's inequality.

Hence, u being strictly increasing,

$$x_c \leq E(x|p) \Rightarrow \pi = E(x|p) - x_c \geq 0.$$

A concave utility function implies that any lottery has a non-negative risk premium.

Example. Consider a lottery

$$\langle 1/3, \$100; 2/3, -\$25 \rangle$$

Calculate the decision maker's risk premium for this lottery if his utility function is: (i) $u(x) = \ln(x + 200)$; (ii) $\exp(1 + x/200)$.

The expected monetary payoff of the lottery

$$E(x|p) = \$1/3 \times 100 + 2/3 \times (-25) = \$16.67.$$

If $u(x) = \ln(x + 200)$, the certainty equivalent is given by

$$u(x_c) = 1/3 \times u(100) + 2/3 \times u(-25)$$

that is,

$$\ln(x_c + 200) = 1/3 \times \ln(300) + 2/3 \times \ln(175)$$

from which we find $x_c = \$9.44$.

So, the risk premium is

$$\pi = E(x|p) - x_c = \$7.23.$$

If $u(x) = \exp(1 + x/200)$, the certainty equivalent is given by

$$\exp(x_c + 200) = 1/3 \times \exp(300) + 2/3 \times \exp(175)$$

that is, $x_c = \$25.84$. So the risk premium is

$$\pi = E(x|p) - x_c = -\$9.17.$$

Noting that $\ln$ is concave and $\exp$ is convex, we see that these results are comfortingly in accord with the hteory: the risk premiums are, respectively, positive and negative.

Pratt (1964) suggested measuring the risk averseness encoded in a utility function by

$$r(x) = -\frac{u''(x)}{u'(x)} = \frac{d^2u}{dx^2}\bigg/\frac{du}{dx}.$$

8. Fuzzy sets

Fuzzy sets were introduced by Zadeh in 1965 to represent/manipulate data and information possessing nonstatistical uncertainties.

It was specifically designed to mathematically represent uncertainty and vagueness and to provide formalized tools for dealing with the imprecision intrinsic to many problems.

Fuzzy logic provides an inference morphology that enables approximate human reasoning capabilities to be applied to knowledge-based systems.

The theory of fuzzy logic provides a mathematical strength to capture the uncertainties associated with human cognitive processes, such as thinking and reasoning.

Some of the essential characteristics of fuzzy logic relate to the following (Zadeh, 1992):

- In fuzzy logic, exact reasoning is viewed as a limiting case of approximate reasoning.
- In fuzzy logic, everything is a matter of degree.
- In fuzzy logic, knowledge is interpreted a collection of elastic or, equivalently, fuzzy constraint on a collection of variables.
- Inference is viewed as a process of propagation of elastic constraints.
- Any logical system can be fuzzified.

There are two main characteristics of fuzzy systems that give them better performance for specific applications.

- Fuzzy systems are suitable for uncertain or approximate reasoning, especially for the system with a mathematical model that is difficult to derive.
- Fuzzy logic allows decision making with estimated values under incomplete or uncertain information.

In classical set theory, a subset A of a set X can be defined by its characteristic function χ_A

$$\chi_A \colon X \to \{0, 1\}.$$

This mapping may be represented as a set of ordered pairs, with exactly one ordered pair present for each element of X. The first element of the ordered pair is an element of the set X, and the second element is an element of the set $\{0, 1\}$.

The value zero is used to represent non-membership, and the value one is used to represent membership. The truth or falsity of the statement

$$"x \text{ is in } A"$$

is determined by the ordered pair $(x, \chi_A(x))$. The statement is true if the second element of the ordered pair is 1, and the statement is false if it is 0.

Similarly, a *fuzzy subset* A of a set X can be defined as a set of ordered pairs, each with the first element from X, and the second element from the interval $[0, 1]$, with exactly one ordered pair present for each element of X.

Zero is used to represent complete non-membership, one is used to represent complete membership, and values in between are used to represent intermediate degrees of membership.

The set X is referred to as the universe of discourse for the fuzzy subset A. Frequently, the mapping μ_A is described as a function, the membership function of A. The degree to which the statement

$$\text{"}x \text{ is } A\text{"}$$

is true is determined by finding the ordered pair

$$(x, \mu_A(x)).$$

The degree of truth of the statement is the second element of the ordered pair. It should be noted that the terms *membership function* and *fuzzy subset* get used interchangeably.

Definition 8.1. *(Zadeh, 1965) Let X be a nonempty set. A fuzzy set A in X is characterized by its membership function*

$$\mu_A \colon X \to [0, 1]$$

and $\mu_A(x)$ is interpreted as the degree of membership of element x in fuzzy set A for each $x \in X$.

Frequently we will write simply $A(x)$ instead of $\mu_A(x)$.

Let A be a fuzzy subset of X; the *support* of A, denoted $\mathrm{supp}(A)$, is the crisp subset of X whose elements all have nonzero membership grades in A.

8.1. Fuzzy numbers

A fuzzy set A of the real line $\mathbb{R}$ is defined by its membership function (denoted also by A)

$$A\colon \mathbb{R} \to [0, 1].$$

If $x \in \mathbb{R}$ then $A(x)$ is interpreted as the degree of membership of x in A.

A fuzzy set in $\mathbb{R}$ is called normal if there exists an $x \in \mathbb{R}$ such that $A(x) = 1$. A fuzzy set in $\mathbb{R}$ is said to be convex if A is unimodal (as a function). A fuzzy number A is a fuzzy set of the real line with a normal, (fuzzy) convex and continuous membership function of bounded support.

Definition 8.2. *A fuzzy set A is called triangular fuzzy number with peak (or center) a, left width $\alpha > 0$ and right width $\beta > 0$ if its membership function has the following form*

$$A(t) = \begin{cases} 1 - \dfrac{a-t}{\alpha} & \textit{if } a - \alpha \le t \le a \\[2ex] 1 - \dfrac{t-a}{\beta} & \textit{if } a \le t \le a + \beta \\[2ex] 0 & \textit{otherwise} \end{cases}$$

and we use the notation $A = (a, \alpha, \beta)$.

The support of A is $(a - \alpha, b + \beta)$. A triangular fuzzy number with center a may be seen as a fuzzy quantity

"x is close to a"or"x is approximately equal to a".

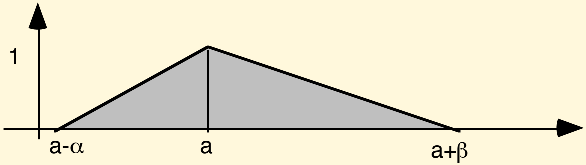

Figure 1: A triangular fuzzy number.

Definition 8.3. *A fuzzy set of the real line given by the membership function*

$$A(t) = \begin{cases} 1 - \dfrac{|a-t|}{\alpha} & \textit{if } |a-t| \leq \alpha, \\ 0 & \textit{otherwise}, \end{cases} \quad (1)$$

Where $\alpha > 0$ will be called a symmetrical triangular fuzzy number with center $a \in \mathbb{R}$ and width 2α and we shall refer to *it by the pair* (a, α).

Let $A = (a, \alpha)$ and $B = (b, \beta)$ be two fuzzy numbers of the form (1), $\lambda \in \mathbb{R}$. Then we have

$$A + B = (a + b, \alpha + \beta), \lambda A = (\lambda a, |\lambda|\alpha). \quad (2)$$

Which can be interpreted as

"x is approximately equal to a"
+
"y is approximately equal to b"
=
"$x + y$ is approximately equal to $a + b$"

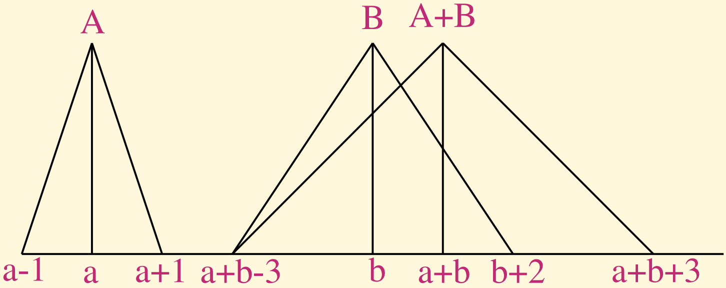

Figure 2: $A = (a, 1), B = (b, 2), A + B = (a + b, 3)$

Definition 8.4. *A fuzzy set A is called trapezoidal fuzzy number with tolerance interval $[a, b]$, left width α and right width β if* *its membership function has the following form*

$$A(t) = \begin{cases} 1 - \dfrac{a-t}{\alpha} & \textit{if } a - \alpha \le t \le a \\ 1 & \textit{if } a \le t \le b \\ 1 - \dfrac{t-b}{\beta} & \textit{if } a \le t \le b + \beta \\ 0 & \textit{otherwise} \end{cases}$$

and we use the notation

$$A = (a, b, \alpha, \beta). \tag{3}$$

The support of A is $(a - \alpha, b + \beta)$. A trapezoidal fuzzy number may be seen as a fuzzy quantity

"x is approximately in the interval $[a, b]$".

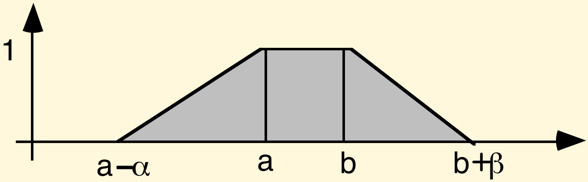

Figure 3: Trapezoidal fuzzy number.

Let A and B are fuzzy subsets of a classical set X. A and B are said to be equal, denoted $A = B$, if $A \subset B$ and $B \subset A$. We note that $A = B$ if and only if $A(x) = B(x)$ for $x \in X$.

We extend the classical set theoretic operations from ordinary set theory to fuzzy sets. We note that all those operations which are extensions of crisp concepts reduce to their usual meaning when the fuzzy subsets have membership degrees that are drawn from $\{0, 1\}$.

For this reason, when extending operations to fuzzy sets we use the same symbol as in set theory. Let A and B are fuzzy subsets of a nonempty (crisp) set X.

Definition 8.5. *The intersection of A and B is defined as*

$$(A \cap B)(t) = \min\{A(t), B(t)\} = A(t) \wedge B(t),$$

for all $t \in X$.

Definition 8.6. *(union) The union of A and B is defined as*

$$(A \cup B)(t) = \max\{A(t), B(t)\} = A(t) \vee B(t),$$

for all $t \in X$.

8.2. Bellman and Zadeh's principle to fuzzy decision making

A classical MADM problem can be expressed in a matrix format. The decision matrix is an $m \times n$ matrix whose element x_{ij} indicates the performance rating of the i-th alternative, a_i, with respect to the j-th attribute, X_j:

$$\begin{matrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{matrix} \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \vdots & \vdots & & \vdots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{pmatrix}$$

The classical maximin method is defined as:

$$\text{choose } a_k \text{ such that } s_k = \max_{i=1,\dots,m} s_i = \max_i \min_j x_{ij}.$$

i.e., s_i the security level of a_i, i.e. a_i guarantees the decision maker a return of at least s_i

Example 8.1.

	$Mathematics$	$English$	$History$
$John$	5	8	10
$Mary$	6	6	7
$Kate$	10	7	5

Mary is selected if we use the maximin method, because her minimal performance is the maximal.

The overall performance of an alternative is determined by its weakest performance: *A chain is only as strong as its weakest link.*

The classical maximax return criterion is:

$$\text{choose } a_k \text{ such that } o_k = \max_{i=1,\dots,m} o_i = \max_i \max_j x_{ij}.$$

is the best performance of a_i.

Example 8.2.

	Math	*English*	*History*
John	5	8	10
Mary	6	6	7
Kate	10	7	5

John or Kate are selected if we use the maximax method, because their maximal performances provide the global maximum.

In fuzzy case the values of the decision matrix are given as degrees of "how an alternative satisfies a certain attribute". For each attribute X_j we are given a fuzzy set μ_j measuring the degrees of satisfaction to the j-th attribute.

$$\mu_{Math}(t) = \begin{cases} 1 - (8-t)/8 & \text{if } t \leq 8 \\ 1 & \text{otherwise} \end{cases}$$

$$\mu_{English}(t) = \begin{cases} 1 - (7-t)/7 & \text{if } t \leq 7 \\ 1 & \text{otherwise} \end{cases}$$

$$\mu_{History}(t) = \begin{cases} 1 - (6-t)/6 & \text{if } t \leq 6 \\ 1 & \text{otherwise} \end{cases}$$

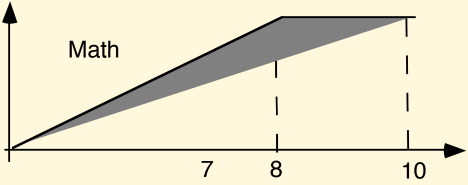

Figure 4: Membership function for attribute 'Math'.

	Math	*English*	*History*
John	5/8	1	1
Mary	6/8	6/7	1
Kate	1	1	5/6

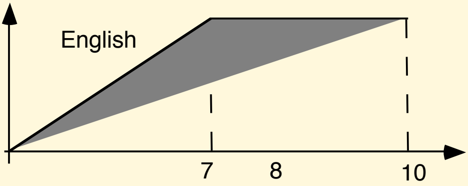

Figure 5: Membership function for attribute 'English'.

Bellman and Zadeh's principle to fuzzy decision making chooses the "best compromise" alternative using the maximin method:

$$score(\text{J}ohn) = \min\{5/8, 1, 1\} = 5/8$$
$$score(\text{M}ary) = \min\{6/8, 6/7, 1\} = 6/8$$
$$score(\text{K}ate) = \min\{1, 1, 5/6\} = 5/6$$

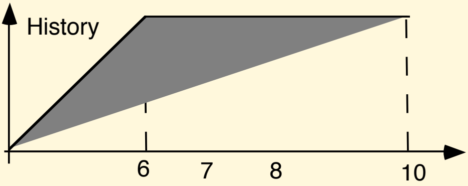

Figure 6: Membership function for attribute 'History'.

Because,

$$score(\mathrm{K}ate) = \max\{5/8, 6/8, 5/6\} = 5/6.$$

Therefore, Kate is choosen as the best student.

Let x be an object such that for any criterion C_j, $C_j(x) \in [0, 1]$ indicates the degree to which the j-th criterion is satisfied by x.

If we want to find out the degree to which x satisfies

'all the criteria'

denoting this by $D(x)$, we get following

$$D(x) = \min\{C_1(x), \ldots, C_n(x)\}.$$

In this case we are essentially requiring x to satisfy

$$C_1 \text{ and } C_2 \text{ and } \ldots \text{ and } C_n.$$

The best alternative, denoted by x^*, is determined from the relationship

$$D(x^*) = \max_{x \in X} D(x) = \max_{x \in X} \min\{C_1(x), \ldots, C_n(x)\}$$

This method is called *Bellman and Zadeh's principle to fuzzy decision making*.

Example 8.3. *The first attribute called "x should be close to 3" is represented by the fuzzy set C_1 defined by*

$$C_1(x) = \begin{cases} 1 - |3-x|/2 & \textit{if } |3-x| \le 2 \\ 0 & \textit{otherwise} \end{cases}$$

The second attribute called "x should be close to 5" is represented by the fuzzy set C_2 defined by

$$C_2(x) = \begin{cases} 1 - |5-x|/2 & \textit{if } |5-x| \le 2 \\ 0 & \textit{otherwise} \end{cases}$$

The fuzzy decision D is defined by the intersection of C_1 and C_2 and the membership function of D is

$$D(x) = \begin{cases} 1/2(1 - |4-x|) & \textit{if } |4-x| \le 1 \\ 0 & \textit{otherwise} \end{cases}$$

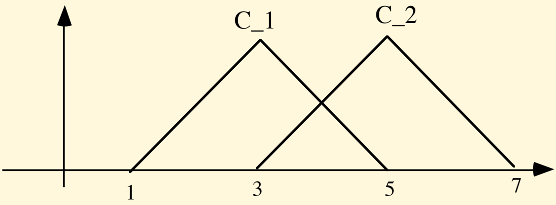

Figure 7: Two fuzzy attributes.

The optimal solution is

$$x^* = 4$$

because

$$D(4) = 1/2 = \max_x D(x).$$