Decision Methods and Models

Master's Degree in Computer Science

Roberto Cordone

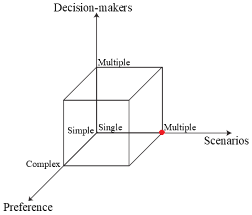
DI - Università degli Studi di Milano

Schedule:	Thursday 16.30 - 18.30 in Aula Magna (CS department)
	Friday 12.30 - 14.30 in classroom 301
Office hours:	on appointment
E-mail:	roberto.cordone@unimi.it
Web page:	https://homes.di.unimi.it/cordone/courses/2024-mmd/2024-mmd.html
Ariel site:	https://myariel.unimi.it/course/view.php?id=4467

Decision-making under risk

We assume

- a preference relation Π that is a weak order
  with a known consistent value function $u(f)$ (replaced by a cost f)
- a uncertain environment: $|\Omega| > 1$ with probabilistic information
- a single decision-maker: $|D| = 1 \Rightarrow \Pi_d$ reduces to Π



- if Ω is discrete, a probability function:

$$\pi : \Omega \to [0,1] \text{ such that } \sum_{\omega \in \Omega} \pi(\omega) = 1$$

- if Ω is continuous, a probability density function:

$$\pi : \Omega \to \mathbb{R}^+ \text{ such that } \int_{\omega \in \Omega} \pi(\omega) = 1$$

We will consider only finite spaces (therefore, discrete)

Approaches to the modelling of risk

Exactly as for the impact function $f(x, \omega)$

- the values of $\pi(\omega)$ derive from descriptive models
- we will assume that they are given
- they might be more or less reliable and this should be taken into account (for example, with sensitivity studies)

Given a probabilistic model (sample set Ω and probability function $\pi(\omega)$)

- the impact $f(x, \omega)$ of any fixed solution x is a random variable

As for the case of ignorance, we will aim to define a choice criterium, removing the dependence of the impact from the scenario:

$$f(x, \omega) \rightarrow \phi_{\Omega,\pi}(x)$$

Several approaches exist, but we will focus on

1. the expected value criterium, that is the first historical proposal with a number of important limitations
2. the stochastic utility theory, that is a successful approach, though not a perfect model of human attitudes with respect to risk

The definition of probability: classical

Contrary to the impact, probability is a debated concept with definitions

- obtained from different sources
- with different meanings
- possibly contradictory

Let us make a quick survey

Classical probability: scenarios are composed by elementary cases

The probability of a scenario is the number of its elementary cases divided by the overall number of elementary cases

$$\pi(\omega) = \frac{n(\omega)}{n(\Omega)}$$

Problems

- what is an elementary case?
- why are the elementary cases equally likely?

It looks like a circular definition

The classical definition works well for games (everything is under control)

The definitions of probability: frequentist

Frequentist probability: the probability of a scenario is the limit of its frequency in a number of observations increasing to infinity

$$\pi(\omega) = \lim_{n \to +\infty} \frac{n(\omega)}{n}$$

similar to the classical definition, but

- it does not require to identify elementary cases
- it allows nonuniform probabilities
- it requires a historical set of observations (Big Data)

The definition of probability: frequentist

Problems of the frequentist definition

- the empirical information must be of good quality
- the quantity of observations must be very large

We are approximating a limit with a finite number

- the future behaviour must be similar to the past

We are using past data to take decisions for the future

In financial applications these requirements are rarely met

Example: what is the probability of default when investing in...

- ...Argentinian treasury bonds? (8 defaults in about 200 years)
- ...German treasury bonds? (3 – 4 defaults in about 150 years)
- ...Italian treasury bonds? (no default in about 160 years)

Markets (correctly) do not use frequentist probabilities

The definition of probability: subjective

Subjective probability: the probability of a scenario is the price considered fair to pay in order to gain

- 1 if the scenario occurs
- 0 if the scenario does not occur

$$\pi(\bar{\omega}) = u(f(\omega)) \text{ where } f(\omega) = \begin{cases} 1 & \text{if } \omega = \bar{\omega} \\ 0 & \text{if } \omega \in \Omega \setminus \{\bar{\omega}\} \end{cases}$$

It is the result of a gamble, based on economic reasoning

- it does not require repeated experiments in identical conditions
- it can exploit any kind of information

Problem

- it depends on the personal subjective opinion of the modeller

The personal attitudes are no longer limited to the preference

The definition of probability: axiomatic

Axiomatic probability: the probability function is any function that respects the basic axioms of probability theory (Kolmogorov)

- being restricted in $[0, 1]$ for all $\omega \in \Omega$
- summing to 1 over all $\omega \in \Omega$
- being additive over sets of disjoint scenarios

Probability theory

- only guarantees that the theorems work
- does not provide numerical values to the probabilities

In practice, one adopts frequentist or subjective values (or a mix) on a case-by-case basis and gets prepared to handle estimation errors

The expected value criterium

One of the founders of probability theory, Blaise Pascal, proposed the expected value criterium, that sums up the impacts of a solution over all scenarios with the convex combination of impacts with probabilities

$$\phi_{EV}(x) = E[f(x,\omega)] = \sum_{\omega \in \Omega} \pi(\omega) f(x,\omega)$$

In the continuous case, this becomes $\int_{\omega \in \Omega} \pi(\omega) f(x,\omega) \, d\omega$

Given the evaluation matrix $U = \{f(x,\omega)\}$, this choice criterium can also be expressed as

$$\phi_{EV}(x) = U \, \pi$$

It is a "position measure" of the distribution of impact values:

- it lies in the range of these values
- it tends to lie in the "middle" of this range
- it is closer to the values with larger probability

Example

$f(x,\omega)$	$\omega^{(1)}$	$\omega^{(2)}$	$\omega^{(3)}$	$\omega^{(4)}$
$x^{(1)}$	1	3	4	6
$x^{(2)}$	2	2	2	4
$x^{(3)}$	3	2	1	9
$x^{(4)}$	6	6	1	3

$\pi(\omega)$	0.20	0.25	0.50	0.05

The expected value criterium values are

	$x^{(1)}$	$x^{(2)}$	$x^{(3)}$	$x^{(4)}$
$\phi(x)$	3.25	2.10	2.05	3.35

and imply the following ranking

$$x^{(3)} \prec x^{(2)} \prec x^{(1)} \prec x^{(4)}$$

Probability space

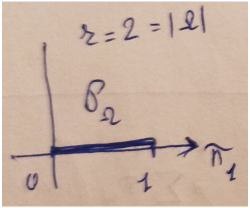
What if the probabilities are wrong?

A sensitivity analysis can tell if the ranking is stable or not

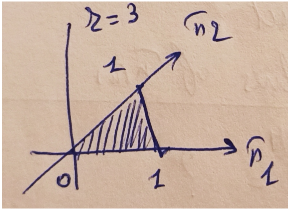
Probability space if the set of all possible probability vectors

$$\mathcal{P}(\Omega) = \left\{ \pi_\omega \in [0,1]^{|\Omega|} : \sum_{\omega \in \Omega} \pi_\omega = 1 \right\}$$

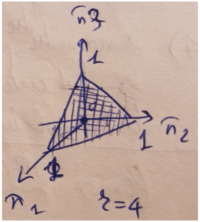
It is defined by $r = |\Omega|$ parameters summing to 1 ($r - 1$ are independent)



$r = 2 \Rightarrow 1D$



$r = 3 \Rightarrow 2D$



$r = 4 \Rightarrow 3D$

Every probability vector $\pi \in \mathcal{P}(\Omega)$ corresponds to optimal solution x^*_π

Probability support and sensitivity analysis

Probabilistic support of a solution $x \in X$ is the set of probability vectors π for which x is the best according to the expected value criterium

$$\mathrm{Supp}\,(x) = \left\{ \pi \in \mathcal{P}(\Omega) : x \in \arg\min_{x' \in X} \phi_{EV}\,(x') \right\}$$

A complete sensitivity analysis partitions $\mathcal{P}(\Omega)$ into the various supports

Of course, it becomes quite complex for $r > 2$

We consider a simplified analysis

- one scenario $\bar{\omega}$ has a very uncertain probability
- the other scenarios have an overall very uncertain probability, but reliable ratios to each other

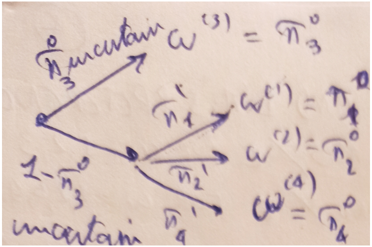
For example, consider

- scenario set $\Omega = \{$ "cloudy", "rainy", "sunny", "snowy" $\}$
- assume an uncertain value for P["sunny"]
- consequently have an uncertain value for P["bad weather"], where "bad weather" $= \{$ "cloudy", "rainy", "snowy" $\}$
- assume precise conditional probabilities for P["cloudy" | "bad weather"], P["rainy" | "bad weather"] and P["snowy" | "bad weather"]

Probability support and sensitivity analysis

Assume a nominal probability value $\pi^\circ(\omega)$ for each scenario $\omega \in \Omega$ and study the sensitivity to the variations of $\pi\left(\omega^{(3)}\right) = \alpha$

$f(x,\omega)$	$\omega^{(1)}$	$\omega^{(2)}$	$\omega^{(3)}$	$\omega^{(4)}$
$x^{(1)}$	1	3	4	6
$x^{(2)}$	2	2	2	4
$x^{(3)}$	3	2	1	9
$x^{(4)}$	6	6	1	3
$\pi^\circ(\omega)$	0.20	0.25	0.50	0.05



For the sake of simplicity, let us denote as

- π_k° the nominal probability $\pi^\circ\left(\omega^{(k)}\right)$
- π_k the unknown real probability $\pi\left(\omega^{(k)}\right)$
- π_k' the conditional probability $\pi\left(\omega^{(k)} \mid \left(\Omega \setminus \{\omega^{(3)}\}\right)\right)$ (for $k \neq 3$)

Example: the algebraic way

Setting $\pi_3 = \alpha$, we obtain that

$$\begin{cases} \pi_k^\circ = (1 - \pi_3^\circ)\pi_k' & \text{for } k \neq 3 \\ \pi_k = (1 - \alpha)\pi_k' & \text{for } k \neq 3 \end{cases} \Rightarrow \pi_k = \pi_k^\circ \frac{1 - \alpha}{1 - \pi_3^\circ} \text{ for } k \neq 3$$

All probabilities are linear functions in α

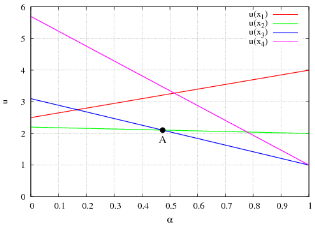
Consequently, also $\phi_{EV}(x)$ is linear in α

In the present case, $1 - \pi_3^\circ = (1 - 0.5)$ from which:

	$\omega^{(1)}$	$\omega^{(2)}$	$\omega^{(3)}$	$\omega^{(4)}$
$\phi_{EV}(x)$	$(2.5 + 1.5\alpha)$	$(2.2 - 0.2\alpha)$	$(3.1 - 2.1\alpha)$	$(5.7 - 4.7\alpha)$

Example: the algebraic way

	$\omega^{(1)}$	$\omega^{(2)}$	$\omega^{(3)}$	$\omega^{(4)}$
$\phi_{EV}(x)$	$(2.5 + 1.5\alpha)$	$(2.2 - 0.2\alpha)$	$(3.1 - 2.1\alpha)$	$(5.7 - 4.7\alpha)$



Two solutions have nonempty support, and the threshold is obtained by intersecting their impact values

$$2.2 - 0.2\alpha = 3.1 - 2.1\alpha \Rightarrow 1.9\alpha = 0.9 \Rightarrow \alpha = 9/19$$

so that $\text{Supp}\left(x^{(2)}\right) = [0, 9/19]$ and $\text{Supp}\left(x^{(3)}\right) = [9/19, 1]$

Example: the geometric way

Since all probabilities are linear in α, the profiles of the impacts for all alternatives can be drawn from two given values:

- $\alpha = 1$ trivially implies $\phi_{EV}(x) = f(x, \omega^{(3)}) = [\ 4 \quad 2 \quad 1 \quad 1\ ]$
- $\alpha = 0$ makes $\omega^{(3)}$ impossible, so that the other three scenarios divide the original probabilities by $(1 - \pi_3^\circ)$

$$[\ 0.20 \quad 0.25 \quad 0.50 \quad 0.05] \rightarrow [\ 0.40 \quad 0.50 \quad - \quad 0.10\ ]$$

$f(x,\omega)$	$\omega^{(1)}$	$\omega^{(2)}$	$\omega^{(3)}$	$\omega^{(4)}$
$x^{(1)}$	1	3	4	6
$x^{(2)}$	2	2	2	4
$x^{(3)}$	3	2	1	9
$x^{(4)}$	6	6	1	3

$\pi^\circ(\omega)$	0.40	0.50	-	0.10

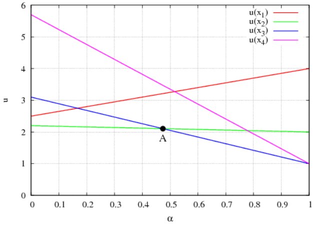
which implies

	$x^{(1)}$	$x^{(2)}$	$x^{(3)}$	$x^{(4)}$
$\phi_{EV}(x)$	2.50	2.20	3.10	5.70

Example: the geometric way

Combining the two values for each alternatives yields the same profiles

$\phi_{EV}(x)$	$x^{(1)}$	$x^{(2)}$	$x^{(3)}$	$x^{(4)}$
$\alpha = 0$	2.50	2.20	3.10	5.70
$\alpha = 1$	4	2	1	1



Formal defects of the expected value criterium

The expected value criterium, however, has strong defects

1. the actual preferences are inconsistent with the expected values
2. extreme values of probability and impact have paradoxical effects

Concerning the inconsistencies, consider the following four alternatives

1. throw a die and gain 100 Euros for all outcomes
2. throw a die and gain 200 Euros for $4, 5, 6$
3. throw a die and gain 600 Euros for 6
4. throw a die and gain 200 Euros for $2, 3, 4, 5, 6$ and -400 Euros for 1

All four alternatives have equal expected value: $\phi_{EV}(x) = 100, \forall x \in X$

Most decision-maker, however, would consider them as very different

Simply prescribing that people should consider them as equivalent
and labelling them as irrational and inconsistent if they do not,
seems very questionable

Formal defects of the expected value criterium

Pascal's wager is a famous thought experiment: it has the form of a conversation with a gambler to convince him to repent and convert

- humans have two alternative choices: believe and not believe
- there are two scenarios: God exists or does not exist

Someone considers it a proof of God's existence (Pascal is very nuanced)

The four configurations have the following impacts

$f(x, \omega)$	$\exists$ God	$\nexists$ God
Believe	A	b
Disbelieve	c	d

where $d > b$ (earthly enjoyment), but A is a huge prize (eternal life)

The probability of the two scenarios are α and $1 - \alpha$

Repeated experiments? Subjective probabilities?

Formal defects of the expected value criterium

If we apply the expected value criterium, we obtain

- $\phi_{EV}(\text{Believe}) = \alpha A + (1-\alpha) b$
- $\phi_{EV}(\text{Disbelieve}) = \alpha c + (1-\alpha) d$

Consequently, believing is the right choice as long as

$$\alpha > \frac{1}{1 + \frac{(A-c)}{(d-b)}}$$

For a sufficiently large reward A, this is always true, because

$$\lim_{A \to +\infty} \frac{1}{1 + \frac{(A-c)}{(d-b)}} = 0$$

However, the reasoning is not fully convincing: in particular, combining small probabilities with large impacts looks problematic

Saint Petersburg's paradox

Daniel Bernouilli provided a more down-to-earth example of this defect

The situation concerns once again a gamble:

- the gambler pays P to take part to the game
- we flip a coin until a tail is obtained
- the gambler wins a sum depending on the number of flips before the end of the game

The model of the game is the following:

- alternatives: do not play or play
- scenarios: the coin is flipped k times obtaining heads before the first tail ($\Omega = \mathbb{N}$)
- probability function: $\pi_k = \dfrac{1}{2^{k+1}}$
- impact function:
    - if we do not play, $f(P, k) = 0$
    - if we do play, $f(P, k) = 2^k - P$

What is the largest P one should be willing to pay to play the game?

Saint Petersburg's paradox

Let us apply the expected value criterium

- if we do not play, $\phi_{EV}(P) = 0$
- if we do play, $\phi_{EV}(P) = \sum_{k \in \mathbb{N}} \frac{2^k - P}{2^{k+1}} = \sum_{k \in \mathbb{N}} \frac{1}{2} - \sum_{k \in \mathbb{N}} \frac{P}{2^{k+1}} = +\infty$

Playing pays out an infinite gain: any cost P is justified

However, in real life different people would pay different amounts and nobody would pay a large sum

Why doesn't the criterium represent real choices?

Combining small probabilities with large impacts looks problematic (*even without the complicating theological aspects of Pascal's wager*)

A possible way out

Bernouilli suggested not to multiply probabilities and impacts:

- for finite creatures the utility of a gain is not proportional to it
- the marginal increase in utility is decreasing

For example, consider the case of a gourmand

- a 1 000 salary allows to dine out 1 evening a week (say)
- a 2 000 salary might mean dining out 2 evenings a week
- a 4 000 salary probably does not mean dining out 4 evenings a week, but maybe 2 evenings in a more expensive place
- a 8 000 salary certainly does not mean dining out 8 evenings a week, and probably also not 4 evenings in a more expensive place
- ...

As the gain increases, it becomes progressively harder to spend it (depending on one's personality)

A possible way out

Bernouilli suggested that utility increases logaritmically with the gain

$$u = \log(f)$$

This solves the original Saint Petersburg's paradox, but not a variant with gains f increasing more than exponentially at each coin flip

The paradox disappears only when the utility function is upper bounded

The stochastic utility theory will provide a sound solution

- introducing axioms to describe desirable properties
- a constructive theorem to build exactly one function to satisfy them