GSOE9210 Engineering Decisions

Victor Jauregui

vicj@cse.unsw.edu.au
www.cse.unsw.edu.au/~gs9210

Bayes decisions

1. Decisions under risk
2. *Bayes* decisions

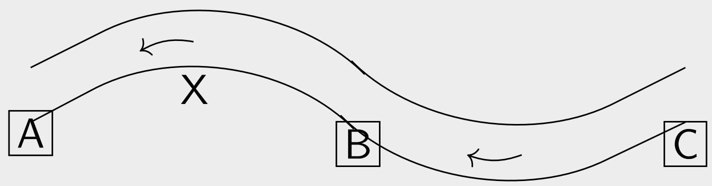
Outline

Decision problem classes

Decision problems can be classified based on an agent's epistemic state:

- Decisions under *certainty*: the agent knows the actual state
- Decisions under *uncertainty*:
  - Decisions under *ignorance* (full uncertainty): the agent believes multiple states/outcomes are possible; likelihoods unknown
  - Decisions under *risk*: the agent believes multiple states/outcomes are possible; likelihood information available

River example



Example (River logistics)

Alice's warehouse is located at X on a river that flows down-stream from C to A. She delivers goods to a client at C via motor boats. On some days a (free) goods ferry travels up the river, stopping at A then B and C, but not at X.

The fuel required (litres) to reach C from each starting point:

	A	X	B	C
To C from:	4	3	2	0

Alice wants to minimise fuel consumption (in litres).

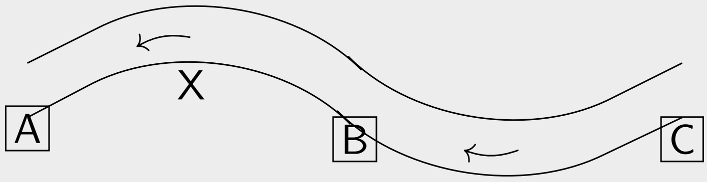
Decisions under incomplete information: risk

Example (Ferry likelihood)

In the river logistics problem, suppose Alice has received an order requiring a package to be delivered every day for the next eight days. Her records show that out of the last 100 days, the ferry was operating on 75.

- Additional information (Alice's records) can be used to estimate likelihood of ferry being operational on any given day
- *Maximum likelihood estimation* assumption: ferry operates in three out of every four days
- Is the *Maximin* strategy (B) still the most rational choice?

River example



	f	$\overline{f}$
A	4	0
B	3	1
C	1	1

Alice considers three possible ways to get to C (from starting point X):

A : via A, by floating down the river

B : via B, by travelling up-stream to B

C : by travelling all the way to C

Outcomes are measured in *litres left* in a four-litre tank.

Exercise

Let $w : \Omega \to \mathbb{R}$ denote fuel consumption in litres. What transformation $f : \mathbb{R} \to \mathbb{R}$ is responsible for the values $v : \Omega \to \mathbb{R}$ in the decision table?

Single decision; multiple trials

- Fuel savings for delivering one package per day over eight days when ferry operates on six of those days:

	f	$\overline{f}$	$\sum$	Avg	min
A	24	0	24	3	0
B	18	2	20	2.5	2

where, *e.g.*, $24 = 6 \times 4 + 2 \times 0$, $20 = 6 \times 3 + 2 \times 1$, *etc.*

- Can we assume the ferry will operate in six of the eight days?
- *Maximin* chooses based on least favourable state $(\overline{f})$
- Given information about likelihood of f, is *Maximin* suitable?

Single decision; multiple trials

Alternatively:

- In how many of the next eight days will ferry operate: Six? Five? Eight? None?
- Assume long sequence of days ... or *maximum likelihood estimate* (six out of eight)
- Proportion of days in which ferry operational: $p = \frac{6}{8} = \frac{3}{4}$

	$\frac{3}{4}$	$\frac{1}{4}$		
	f	$\overline{f}$	E	min
A	4	0	3	0
B	3	1	2.5	1

- Is p *probability* that ferry will operate on any given day?

Frequency interpretation of probability

Outcomes:

$$\underbrace{t, h, h, t, t, h, t}_{n}, h, t, h, \ldots$$

Definition (Frequency interpretation of probability)

The *probability* of an event, E, in an experiment of chance, is the limit of the average occurrences of E over any sequence of indefinitely many trials; *i.e.*,

$$P(E) = \lim_{n \to \infty} \frac{n_E}{n}$$

where n_E is the number of occurrences of event E in the first n trials.

- *e.g.*, For event H: $\frac{0}{1}, \frac{1}{2}, \frac{2}{3}, \frac{2}{4}, \frac{2}{5}, \frac{3}{6}, \frac{3}{7}, \ldots, \frac{n_H}{n}, \ldots$
- What is $P(H)$ for this experiment?

Expected values

Definition (Expected value)

The *expected value* of a random variable $X : \Omega \to \mathbb{R}$ with probability distribution $P : \Omega \to \mathbb{R}$ is given by:

$$E(X) = \sum_{\omega \in \Omega} P(\omega)X(\omega)$$

Definition

The event corresponding to value $x \in \mathbb{R}$, denoted X_x, is defined as:

$$X_x = X^{-1}[x] = \{\omega \in \Omega \mid X(\omega) = x\}$$

More generally, for $A \subseteq \mathbb{R}$:

$$X_A = X^{-1}[A] = \{\omega \in \Omega \mid X(\omega) \in A\}$$

Expected values

For a random variable (real-valued function from Ω to $\mathbb{R}$) X:

- $E(X)$ is also called the limiting (or long run) *average* of X
- $E(X)$ may not be any actual value in $\operatorname{ran} X$
- $E(X)$ is a measure of the 'centre', or *centroid*, of the values of the outcomes
- Natural correspondence with the 'centre of gravity/mass' of a distribution of point masses on a line, where $P(X = x_i)$ corresponds to the proportion of the total mass positioned at x_i

Multiple random trials

- In this situation there are multiple trials (days) of some random process: 100 days
- In each trial (day) different states may occur: ferry (f) or no ferry ($\overline{f}$)
- Information exists about the 'likelihood' of occurrence of states: 75% ferry to 25% no ferry
- *Maximin* assumes worst case for each action even when the worst case (no ferry) is unlikely; *i.e.*, it ignores likelihood information
- Over 100 working days, Alice's total value is greater via A than B
- A decision rule which takes likelihood information into account would be preferable

Probabilistic lotteries

Definition (Probabilistic lottery)

A *probabilistic lottery* over a finite set of outcomes, or *prizes*, Ω, is a pair $\ell = (\Omega, P)$, where $P : \Omega \to \mathbb{R}$ is a probability function. The lottery ℓ is written:

$$\ell = [p_1 : c_1 | p_2 : c_2 | \ldots | p_n : c_n]$$

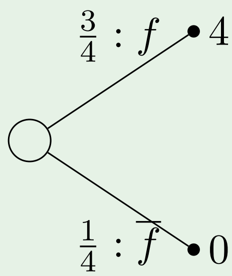
where for each $s_i \in \mathcal{S} \subseteq \mathbb{P}(\Omega)$, $p_i = P(s_i) = P(c_i)$.

Example (To C via A)

Alice's decision to travel via A corresponds to:

$$\ell_{\mathsf{A}} = [\tfrac{3}{4} : 4 | \tfrac{1}{4} : 0]$$

where outcomes have been replaced by their values.



Value of a lottery

Definition (Value of a lottery)

The value of a probabilistic lottery (Ω, P, v) is the expected value over its outcomes:

$$V_v(\ell) = E(v) = \sum_{\omega \in \Omega} P(\omega) v(\omega)$$

- For strategy A:

$$V(\ell_\mathsf{A}) = \tfrac{3}{4}(4) + \tfrac{1}{4}(0) = 3 + 0 = 3$$

- Note: not value of any outcome of strategy A: $4, 0$
- Frequency interpretation: $V(\ell_\mathsf{A})$ is the average value of A over many days

Outline

Bayes decisions

Under risk, each strategy in a decision problem corresponds to a probabilistic lottery.

Definition (Bayes value)

Given a probability distribution over states, the *Bayes value*, V_B, of a strategy is the expected value of its outcomes.

Definition (Bayes strategy)

A *Bayes strategy* is a strategy with maximal *Bayes* value.

Definition (Bayes decision rule)

The *Bayes decision rule* is the rule which selects all the *Bayes* strategies.

Bayes strategies

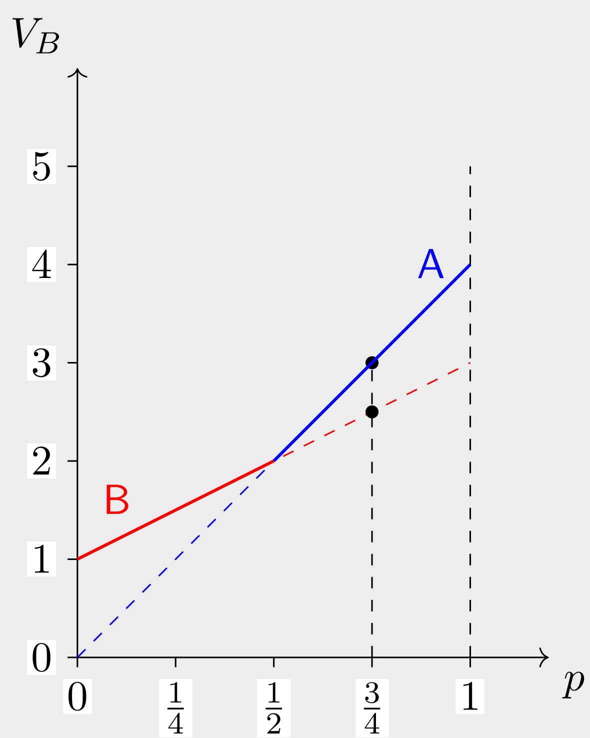
For the river problem, assume the proportion of days in which the ferry operates is $p_f = p$:

	p	$1-p$	
	f	$\overline{f}$	V_B
A	4	0	$4p$
B	3	1	$2p+1$

Bayes values for each strategy plotted for all values of $p \in [0, 1]$.

Exercise

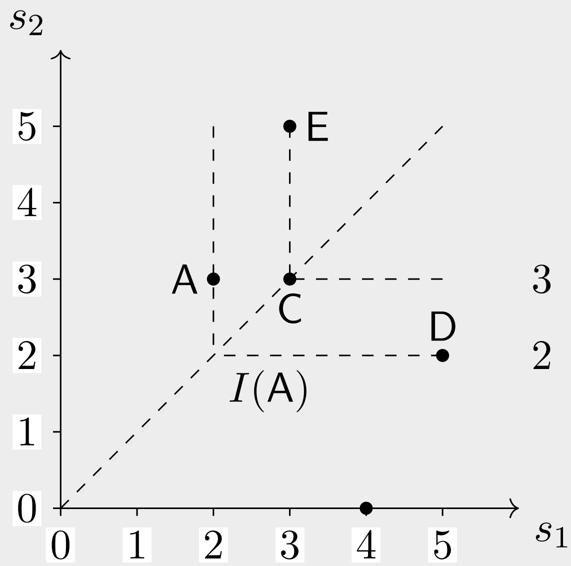
For what values of p will the *Bayes* decision rule prefer A to B?



Indifference curves: *Maximin*

For the pure actions below:

	s_1	s_2
A	2	3
B	4	0
C	3	3
D	5	2
E	3	5

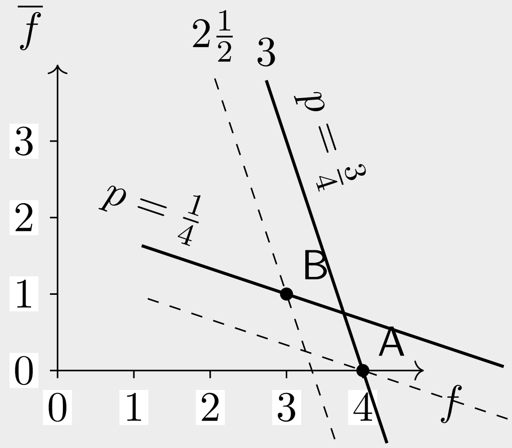


Consider curves of all points representing strategies with same *Maximin* value; *i.e.*, *Maximin indifference curves*.

Indifference curves: *Bayes*

Bayes decision rule indifference curves are linear:

	p	$1-p$	
	f	$\overline{f}$	V_B
A	4	0	$4p$
B	3	1	$2p+1$
a	v_1	v_2	$pv_1 + (1-p)v_2$



Indifference curves:

$$V_B(a) = pv_1 + (1-p)v_2 = u$$

- In gradient-intercept form, $v_2 = \frac{u}{1-p} - \frac{p}{1-p}v_1$, where $m = -\frac{p}{1-p}$; *e.g.*, for $p = \frac{3}{4}$, $m = -\frac{3}{4}/\frac{1}{4} = -\frac{3}{1}$
- Because $v_2 \propto u$; *i.e.*, 'higher' lines receive greater *Bayes* values

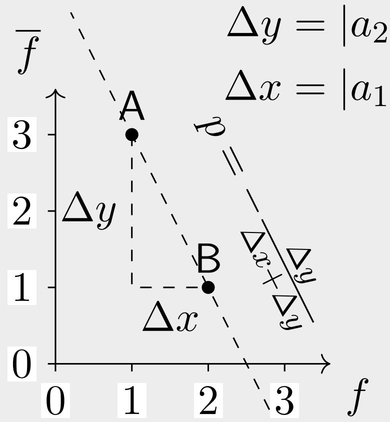
Indifference curves: *Bayes*

In general, for two actions:

	p	$1-p$
	s_1	s_2
A	a_1	a_2
B	b_1	b_2

$$p = \frac{\Delta y}{\Delta x + \Delta y} = \frac{m}{m-1}$$

where m is the gradient of line AB.



$$\Delta y = |a_2 - b_2|$$
$$\Delta x = |a_1 - b_1|$$

For example: if A is $(1,3)$ and B is $(2,1)$ then:

$$p = \frac{3-1}{(2-1)+(3-1)} = \frac{2}{1+2} = \frac{2}{3}$$

Indifference classes and *Bayes* decisions

Exercises

- Prove the expression for p
- For the river problem, what is the slope of the line joining the two actions?
- For what probability are the two actions of equal *Bayes* value?
- What is the *Bayes* value associated with this line?
- Repeat the above exercises for regret

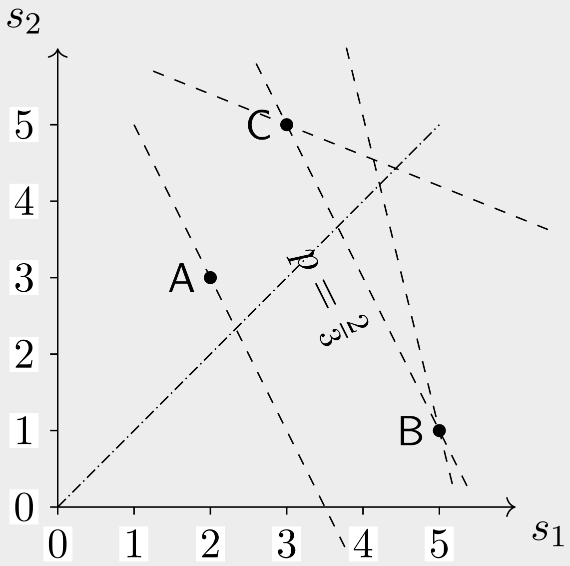
Bayes strategies

For the pure actions below with $P(s_1) = p$:

	s_1	s_2	V_B
A	2	3	$3 - p$
B	5	1	$1 + 4p$
C	3	5	$5 - 2p$

Slope of BC: $m = \frac{5-1}{3-5} = -2$.
$\therefore\ p = \frac{2}{2+1} = \frac{2}{3}$.

Note: $p \propto -m$.

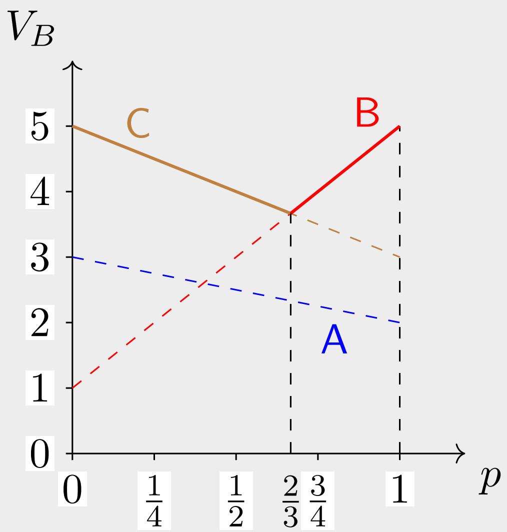


Bayes strategies

For the pure actions below with $P(s_1) = p$:

	s_1	s_2	V_B
A	2	3	$3 - p$
B	5	1	$1 + 4p$
C	3	5	$5 - 2p$

For $p = \frac{2}{3}$, the value of the *Bayes* action(s) is least.



Definition

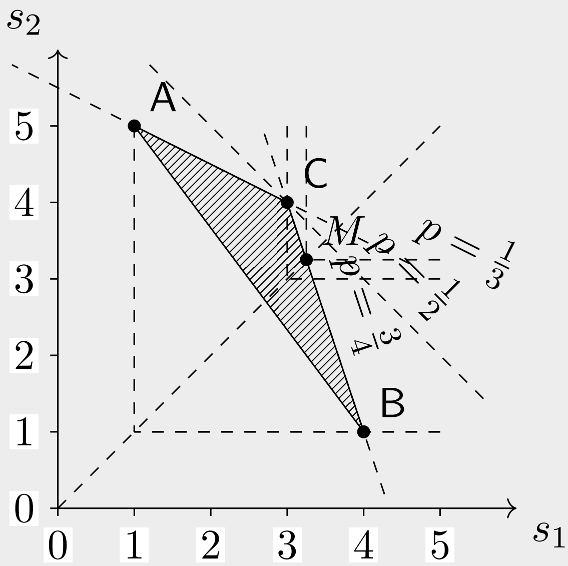
The *least favourable probability distribution* on the states/outcomes is the probability distribution for which *Bayes* strategies have minimal values.

Bayes solutions

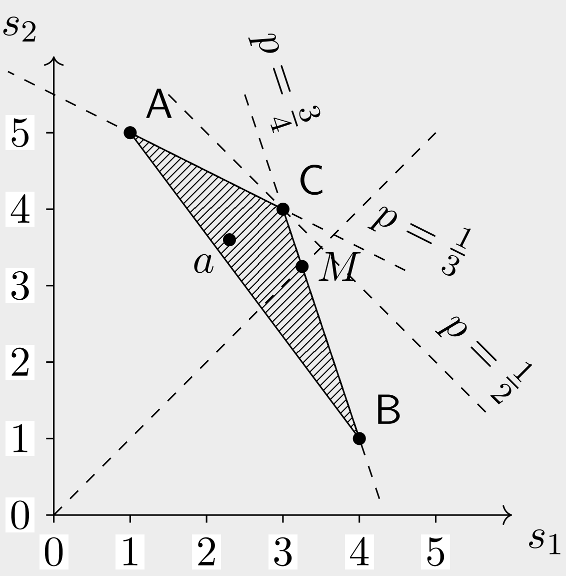
For the pure actions below with $P(s_1) = p$:

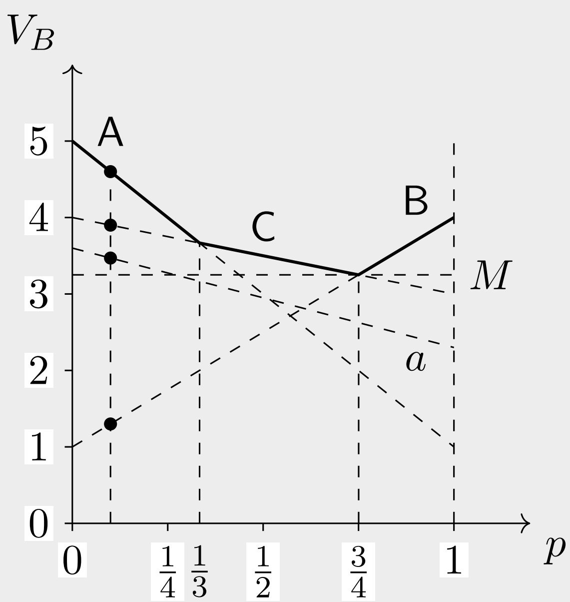
	s_1	s_2	V_B
A	1	5	$5-4p$
B	4	1	$1+3p$
C	3	4	$4-p$

Slope of BC: $m = \frac{4-1}{3-4} = -3$.
$\therefore\ p = \frac{3}{4}$.
Slope of AC: $m = \frac{-1}{2}$.
$\therefore\ p = \frac{1}{3}$.



Bayes strategies





- Note that the *Maximin* action is a *Bayes* action for $p = \frac{3}{4}$
- Note that the internal mixed strategy $a \sim 0.5$A0.3B0.2C is not *Bayes*

Bayes summary

Theorem

Results about Bayes decision rule:

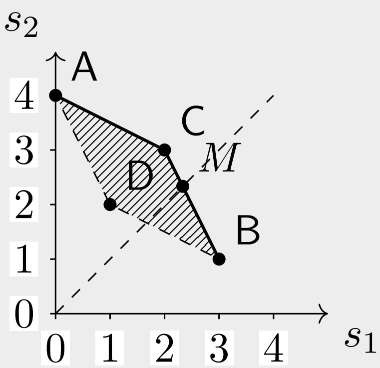
- *Mixing can improve upon the Maximin value of pure strategies, but it does not improve upon the Bayes value of pure strategies*
- *Bayes strategies are invariant/preserved under regret; i.e., the same strategy is chosen under regret as otherwise*

Exercise

Prove the theorems above.

Admissible mixed strategies

	s_1	s_2
A	0	4
B	3	1
C	2	3
D	1	2



Exercises

- Which mixed strategies above are admissible?
- Are *Maximin* mixed strategies always admissible?
- Are *Bayes* mixed strategies always admissible?
- Are *Maximin* mixed strategies always *Bayes* for some p?
- Are admissible mixed strategies *Bayes* for some p?

Bayes summary

- Partial information situations (*risk*)
- Information can affect degree of likelihood/belief (Bayesian probability)
- *Bayes* rule more appropriate when partial information present
- *Bayes* values, *Bayes* decision rule, *Bayes* strategies
- Graphical representation of *Bayes* values
- *Bayes* indifference curves