Wednesday 9:00-12:30

Part 6
Decision making under severe uncertainty & applications in classification and risk analysis

by Matthias C. M. Troffaes

Decision making under severe uncertainty & applications in classification and risk analysis

Outline

Decision making under severe uncertainty & applications in classification and risk analysis

Outline

What is Decision Making: Offshore Wind Example

ECOLOGIST
SETTING THE ENVIRONMENTAL AGENDA SINCE 1970

CLIMATE CHANGE | FOOD + FARMING | ENERGY | HEALTH | POLITICS + ECONOMICS | SCIENCE + TECHNOLOGY | NATURAL WORLD | SOCIET

NEWS | GREEN LIVING | CAMPAIGNING | BLOGS & COMMENTS | INTERVIEWS | MAGAZINE | VIDEO | ARCHIVE | EXTRAS | DIRECTOR

Home | Blogs and Comments | Comment

SEARCH Go

Offshore wind is too expensive!

Jim Platts

7th December 2013

Tweet 28 | +1 0 | Like 35

Prices paid in the UK to solar and wind generators will change to favour offshore wind at the expense of the others. Jim Platt warns that the policy is doomed to failure - offshore wind is just too expensive, and likely to remain so.

> Building things in the middle of the sea is very, very expensive, and that isn't going to change.

In 2012 the global wind industry manufactured and installed more than 20,000 turbines generating 45GW of energy. The leading firm alone, Danish company Vestas, installed more than 6GW of this via 2,500 turbines.

Offshore wind on the rocks? Photo: Chris Lishman / Shutterstock.com.

More articles about

wind | renewables | energy | oceans |

Related Articles

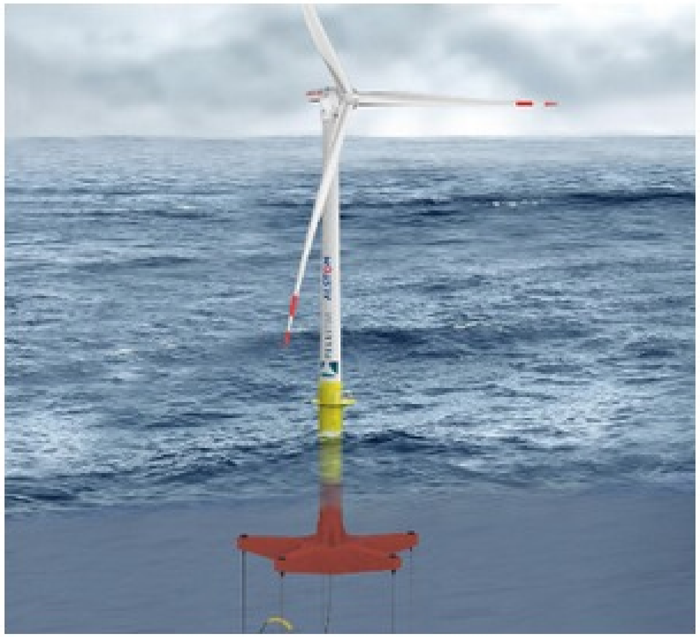
- **Brazil and FIFA have failed to protect their World Cup mascot**

What is Decision Making: Offshore Wind Example

Floating turbines could cut offshore wind energy costs: study

11 July 2014

Print Email



Floating turbines could cut the cost of offshore wind power to below £85/MWh by the mid-2020s, according to an engineering design study by The Glosten Associates for the Energy Technologies Institute (ETI).

The new study revealed that the company's PelaStar tension leg floating platform (TLP) could deliver further reductions as the technology matures and is designed to provide high capacity factors in wind speeds exceeding 10m per second in water between 60m and 1,200m deep.

The UK is said to have over a third of Europe's potential offshore wind resource, which is enough to meet the power demand of the country nearly three times over.

The FEED study has shown that Glosten's PelaStar TLP design could play a major role in reducing UK offshore wind energy costs.

The company said that the TLP technology is suitable for water depths from as low as 55m up to several hundred metres.

What is Decision Making: Offshore Wind Example

Operations & Maintenance of Offshore Wind

30% of cost of offshore wind is operations & maintenance
= **huge** chunk of money

Types of Maintenance

- preventive (prevent future failures)
- corrective (fix after failure)

What is Decision Making: Offshore Wind Example

Decisions

criterion: minimize cost

- when to perform maintenance?
- what is a good preventive/corrective balance?

limiting factor = wind speed & wave height for boarding

Uncertainties

Enormous potential for saving costs by making accurate predictions of:

- wind & waves at different time scales
  avoid missing maintenance opportunities
  avoid costly transport when turbine cannot be boarded
- forecast failures before they happen
  cost of preventing $\lll$ cost of fixing

What is Decision Making: Offshore Wind Example

drastically different issues at different time scales:

Short Term: Optimize Actual Operations

what data on the wind farm should we collect
how to use it?

Medium Term: Business Case

how to convince investors to invest in offshore wind
may not have very much data to go from!

Long Term: Policy & Politics

should we encourage offshore, or look at other technologies?
very little data to go by, enormous uncertainty concerning future climate change, attitude of electorate, etc.
not just about money

What is Decision Making: Offshore Wind Example

Why Use Imprecise Probability for Decision Making?

- increases confidence in analysis based on sparse data
  may help at all levels/time horizons
- risk-averse industries: rare events with large impact

Why **NOT** Use Imprecise Probability for Decision Making?

- computational expense
- abundant data, non-critical decisions
  standard statistical treatment works as well

Communication!

how to communicate uncertainty?
uncertainty analysis only useful if results can be communicated

Decision making under severe uncertainty & applications in classification and risk analysis

Outline

Review of Classical Decision Theory: Example

Example: Visit Offshore Turbine by Boat in the Next Hour?

- parameter: average wave height X for next hour: unknown!
  assume only possible values are $x = 0.5$ and $x = 2$
- data: observation Y, say average wave height in last hour
  assume only possible values are $y = 0.5$ and $y = 2$
- decision: $d =$ take boat, or $d =$ do not take boat
- decision strategy δ:
  which decision to make based on data y?

Review of Classical Decision Theory: Example

Example: Visit Offshore Turbine by Boat in the Next Hour?

- utility function $U(d, x)$: each combination of decision & parameter leads to a different final reward value
  - can only board offshore turbine for maintenance if $X < 1$
  - taking boat costs €1000
  - doing maintenance saves €4000

  for example, expressed in units of €1000:

$U(d, x)$	$x = 0.5$	$x = 2$
$d =$ boat	3	-1
$d =$ no boat	0	0

- likelihood: probability of data given parameter $p(y \mid x)$

$p(y \mid x)$	$y = 0.5$	$y = 2$
$x = 0.5$	0.9	0.1
$x = 2$	0.3	0.7

- prior: probability of parameter $p(x)$ before you see the data

$p(x)$	$x = 0.5$	$x = 2$
	0.4	0.6

Review of Classical Decision Theory: Example

Frequentist Solution: Wald's Expected Utility, Admissibility

frequentist = only use likelihood

1. for every possible strategy δ
   and for every possible value x of X
   calculate Wald's expected utility   expected utility = −risk

$$U(\delta|x) := E(U(\delta(Y), x)|x) = \sum_y U(\delta(y), x)p(y|x) \qquad (6)$$

2. a strategy δ is inadmissible if there is a strategy δ' such that
   $U(\delta'|x) \geq U(\delta|x)$ for all x, and
   $U(\delta'|x) > U(\delta|x)$ for at least one x partial ordering of strategies
3. **optimal Wald strategy**
   all admissible strategies   maximal elements w.r.t. partial ordering

Review of Classical Decision Theory: Example

Bayesian Solution: Maximize Posterior Expected Utility

Bayesian = only use posterior ($\propto$ likelihood $\times$ prior)

1. calculate the posterior

$$p(x|y) = \frac{p(y|x)p(x)}{\sum_{x'} p(y|x')p(x')} \tag{7}$$

2. for every possible observation y
   and every possible decision d
   calculate the posterior expected utility:

$$U(d|y) = E(U(d, X)|y) = \sum_x U(d, x)p(x|y) \tag{8}$$

3. **optimal Bayes strategy** δ^*: max posterior expected utility

$$\delta^*(y) = \arg\max_d U(d|y) \tag{9}$$

much easier to calculate than Wald's admissible strategies! (why?)

Review of Classical Decision Theory: Wald's Theorem

Wald's Theorem (1939 [25])

The set of Wald admissible strategies can always be recovered from a Bayesian analysis, simply by varying the prior over all possible distributions.

[Technical details omitted!]

'equivalence' of $\underbrace{\text{robust Bayesian statistics}}_{\text{sets of priors}}$ and frequentist statistics

Plan

- develop decision making directly from sets of distributions
- look at some practical examples

Decision making under severe uncertainty & applications in classification and risk analysis

Outline

Decision making under severe uncertainty & applications in classification and risk analysis

Outline

Robust Decision Making: R Code Preparation

1. start R
2. visit with browser: `https://raw.githubusercontent.com/mcmtroffaes/improb-redux/master/improb-redux.r`
3. select and copy all R code from browser: CTRL-A, CTRL-C
4. go to R console
5. paste code into R console: CTRL-V, ENTER
6. keep browser window open, so you can rinse & repeat steps 3–5 every time you start a new R session

Robust Decision Making: Aim & Assumptions

Can we develop a decision theory based on just the desirability axioms, without the 'fair price' assumption, discussed on Monday?

Simple setting:

- Set $\mathcal{M}$ of probability mass functions on Ω.
- Consider gambles as functions on Ω (random reward expressed in a utility scale).
- How should we choose among gambles?

Notation for any gamble $g \colon \Omega \to \mathbb{R}$

$$E_p(g) := \sum_{\omega \in \Omega} p(\omega) g(\omega) \qquad \text{for any } p \in \mathcal{M} \tag{10}$$

$$\underline{P}(g) := \min_{p \in \mathcal{M}} E_p(g) \qquad \text{lower prevision} \tag{11}$$

$$\overline{P}(g) := \max_{p \in \mathcal{M}} E_p(g) \qquad \text{upper prevision} \tag{12}$$

Decision making under severe uncertainty & applications in classification and risk analysis

Outline

A Very Simple Example

Example (Machinery, Overtime, or Nothing?)

A company makes a product, and believes in increasing future demand. The manager asks you, the decision expert, whether he should buy new machinery, use overtime, or do nothing. The upcoming year, demand can either increase or remain the same.

If we buy new machinery, then the profit at the end of the year will be 440 (in thousands of pounds) if demand increases, and 260 otherwise. On the other hand, if we use overtime, then the profit will be 420 if demand increases, and 300 otherwise. If we do nothing, profit will be 370.

According to our best current judgement, demand will increase with probability at least 0.5, and at most 0.8:

$\mathcal{M} =$

	p_1	p_2
increase	0.5	0.8
stay	0.5	0.2

(each column is a probability mass function)

What advice can we give the manager?

A Very Simple Example: Choice

each row is a gamble

set of decisions	set of gambles: increase	stay
machinery	440	260
overtime	420	300
nothing	370	370

choice function

what is a good choice function, under severe uncertainty?

set of optimal decisions	set of optimal gambles: increase	stay
overtime	420	300
nothing	370	370

Decision making under severe uncertainty & applications in classification and risk analysis

Outline

Γ-Maximin

(Wald 1945 [26], Gilboa & Schmeidler 1989 [11])

Definition (Γ-Maximin Optimality Criterion)

Choose any gamble whose lower prevision is maximal.

Recipe (Γ-Maximin Optimality Criterion)

1. set up the table with gambles and probabilities
2. calculate the expectation of each gamble with respect to each probability mass function — matrix multiplication
3. calculate minimum expectation of each gamble — minimum of each row
4. choose decision with highest minimum expectation — maximize

$$\arg\max_{d \in D} \underline{P}(g_d) \tag{13}$$

Γ-Maximin: Example

Example (Machinery, Overtime, or Nothing)

	increase	stay	p_1	p_2	$\underline{P}$
increase			0.5	0.8	
stay			0.5	0.2	
machinery	440	260			
overtime	420	300			
nothing	370	370			
	(1)		(2)		(3) & (4)

```
pmfs = c(
  0.5, 0.5,
  0.8, 0.2)
rvars = c(
  440, 260,
  420, 300,
  370, 370)
getexpectations = getexpectationsfunc(2, pmfs)
getlowerprevisions = getlowerprevisionsfunc(getexpectations)
isgammamaximin = isgammamaxisomethingfunc(getlowerprevisions)
isgammamaximin(rvars)
```

Γ-Maximax

(Satia and Lave 1973 [21], probably others as well)

- Γ-maximin seems overly pessimistic; something more optimistic?

Definition (Γ-Maximax Optimality Criterion)

Choose any gamble whose *upper* prevision is maximal.

Recipe (Γ-Maximax Optimality Criterion)

1. set up the table with gambles and probabilities
2. calculate the expectation of each gamble with respect to each probability mass function — matrix multiplication
3. calculate *maximum* expectation of each gamble — maximum of each row
4. choose decision with highest maximum expectation — maximize

$$\arg\max_{d \in D} \overline{P}(g_d) \tag{14}$$

Γ-Maximax: Example

Example (Machinery, Overtime, or Nothing)

	increase	stay	p_1	p_2	$\overline{P}$
increase			0.5	0.8	
stay			0.5	0.2	
machinery	440	260			
overtime	420	300			
nothing	370	370			
	(1)		(2)		(3) & (4)

```
pmfs = c(
  0.5, 0.5,
  0.8, 0.2)
rvars = c(
  440, 260,
  420, 300,
  370, 370)
getexpectations = getexpectationsfunc(2, pmfs)
getupperprevisions = getupperprevisionsfunc(getexpectations)
isgammamaximax = isgammamaxisomethingfunc(getupperprevisions)
isgammamaximax(rvars)
```

Interval Maximality

liteture: 'interval dominance'

(Condorcet 1785 [9], Sen 1977 [22], Satia and Lave 1973 [21], Kyburg 1983 [12], *many* others)

- get all reasonable options (from pessimistic to optimistic) at once?

Definition (Partial Ordering by Interval Comparison)

We say that a gamble f interval dominates g, and write

$$f \sqsupset g \tag{15}$$

whenever

$$\underline{P}(f) > \overline{P}(g) \tag{16}$$

$[\overline{P}(f), \underline{P}(f)]$ dominates $[\overline{P}(g), \underline{P}(g)]$

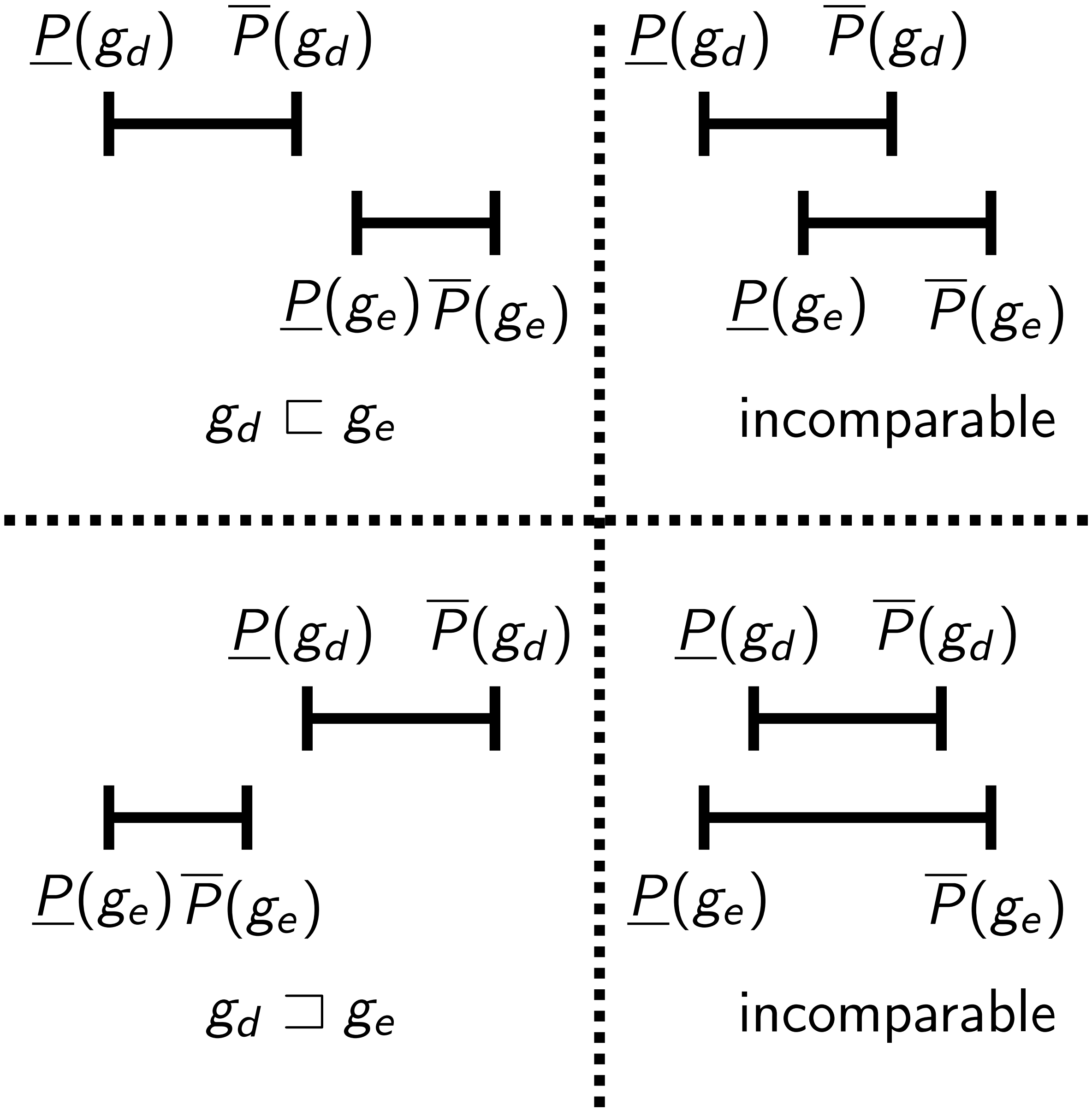
Definition (Interval Maximality Optimality Criterion)

Choose any gamble which is undominated with respect to $\sqsupset$.

$$\{d\colon (\forall e \in D)(g_d \not\sqsubset g_e)\} \tag{17}$$

Interval Maximality: Partial Ordering

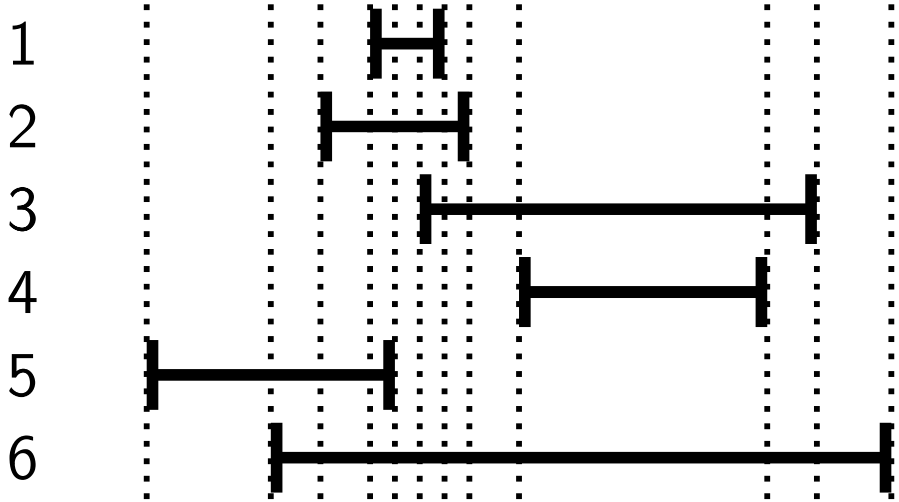
- $\sqsubset$ determines a partial ordering between gambles



Interval Maximality: Hasse Diagram & Algorithm

maximal elements with partial ordering = undominated elements

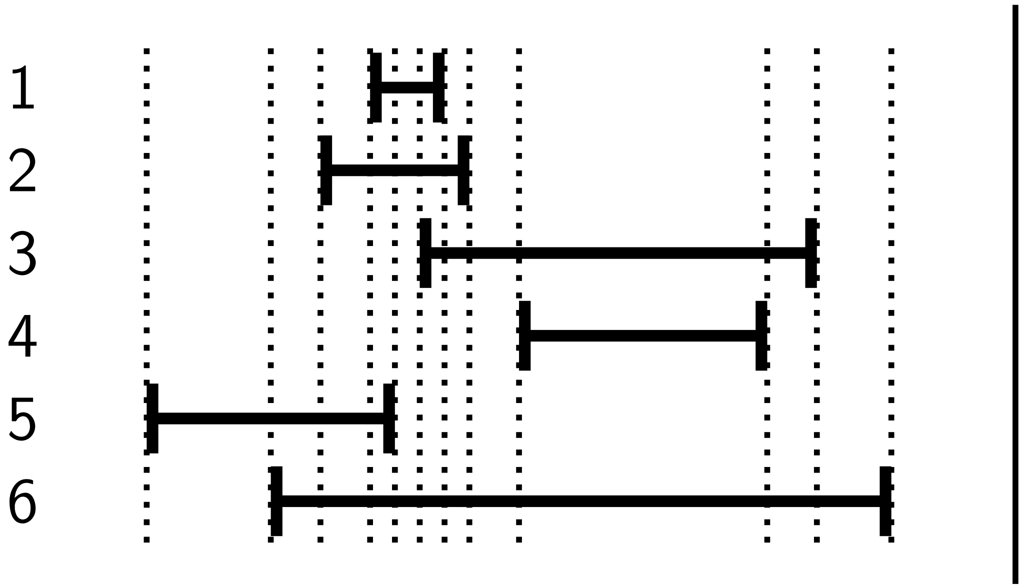
example:



Interval Maximality: Hasse Diagram & Algorithm

maximal elements with partial ordering = undominated elements

example:

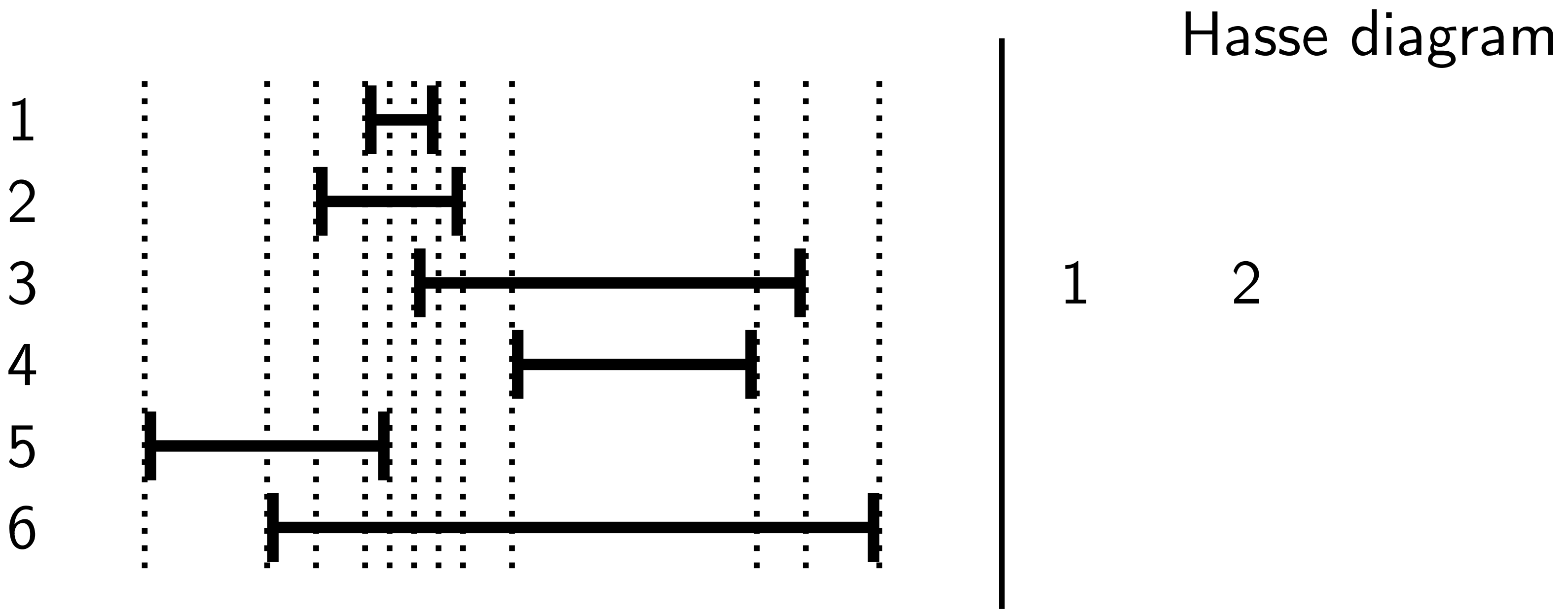


Hasse diagram

Interval Maximality: Hasse Diagram & Algorithm

maximal elements with partial ordering = undominated elements

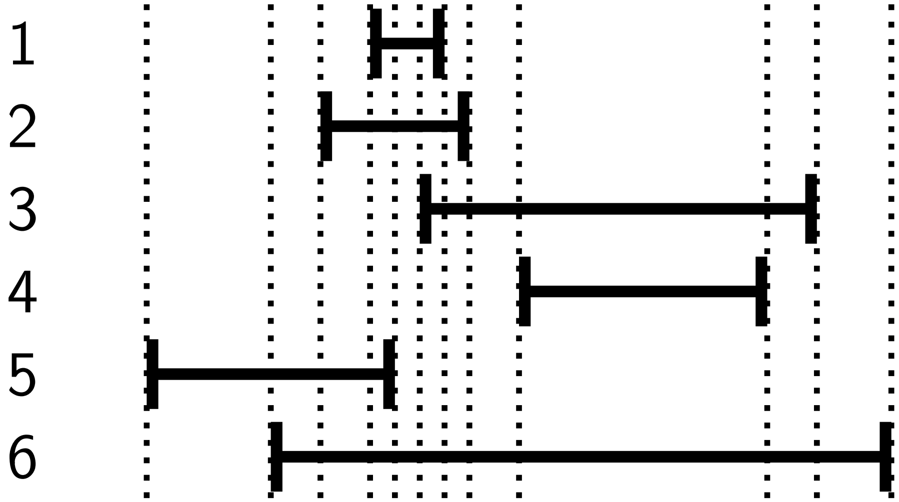
example:



Interval Maximality: Hasse Diagram & Algorithm

maximal elements with partial ordering = undominated elements

example:

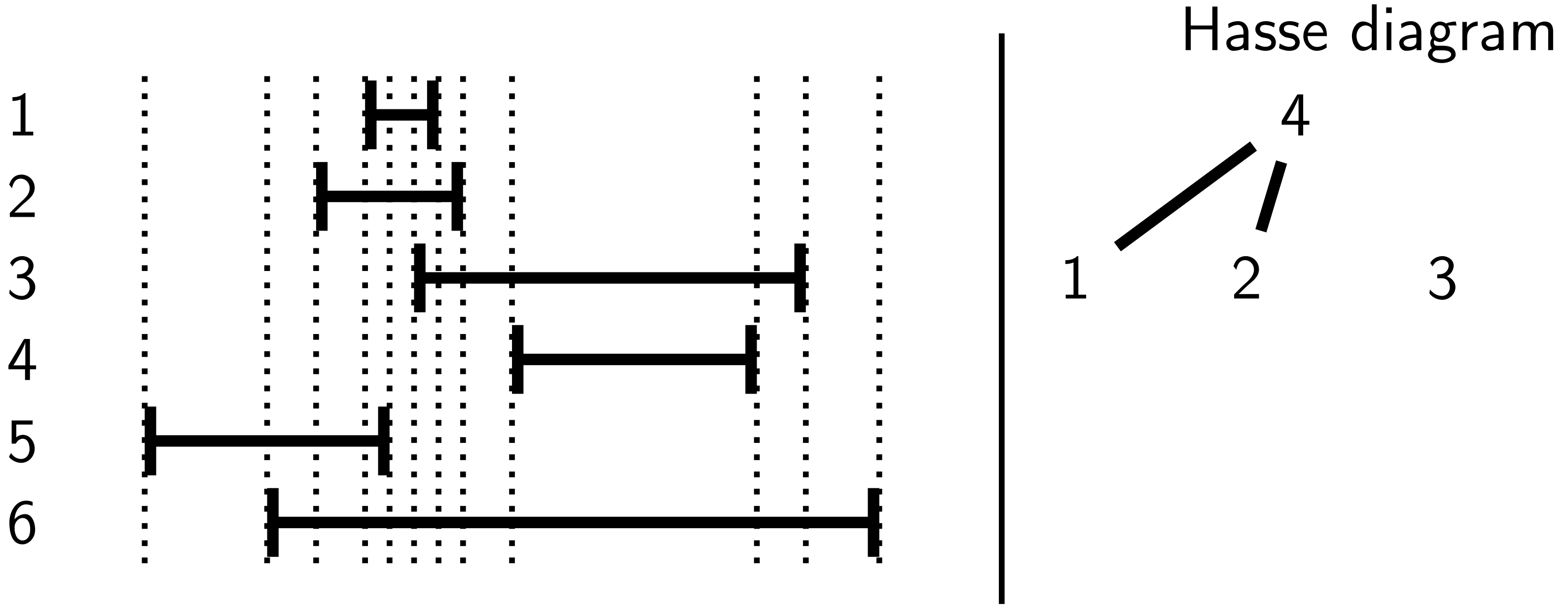


Hasse diagram

Interval Maximality: Hasse Diagram & Algorithm

maximal elements with partial ordering = undominated elements

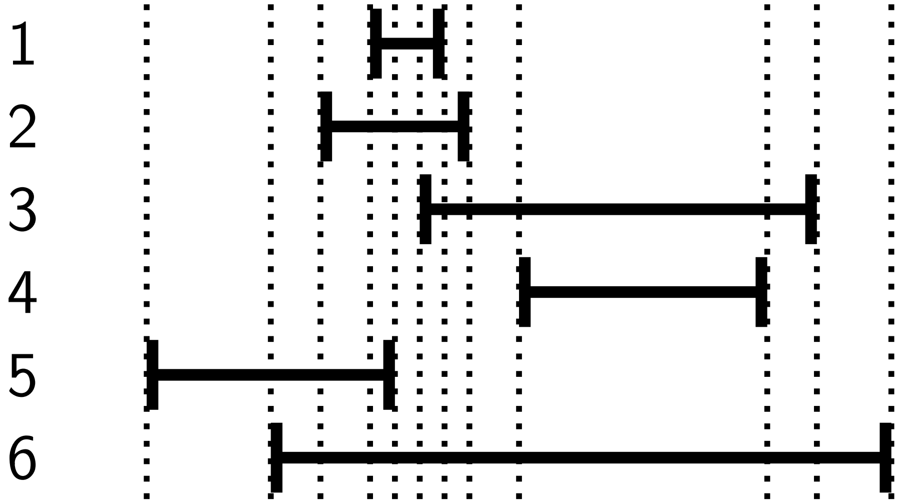
example:



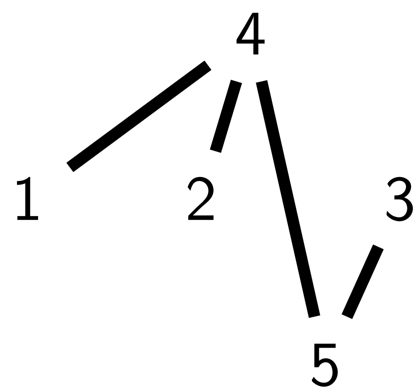
Interval Maximality: Hasse Diagram & Algorithm

maximal elements with partial ordering = undominated elements

example:



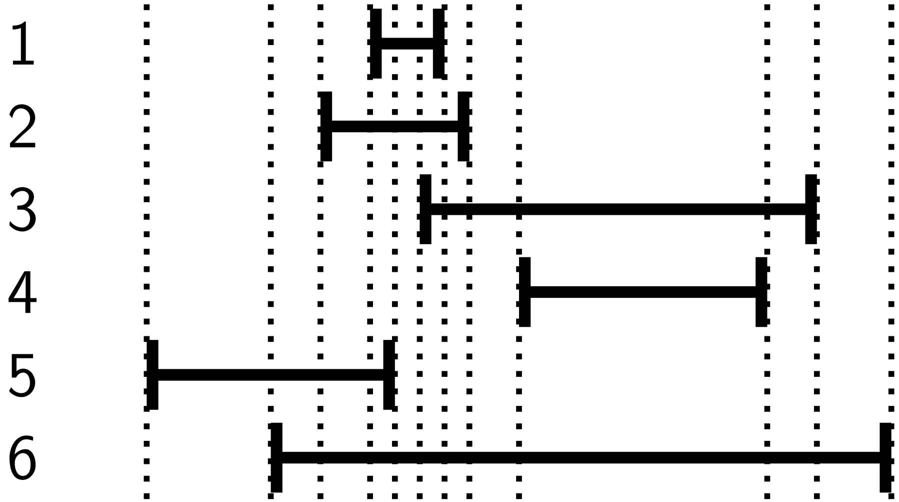
Hasse diagram



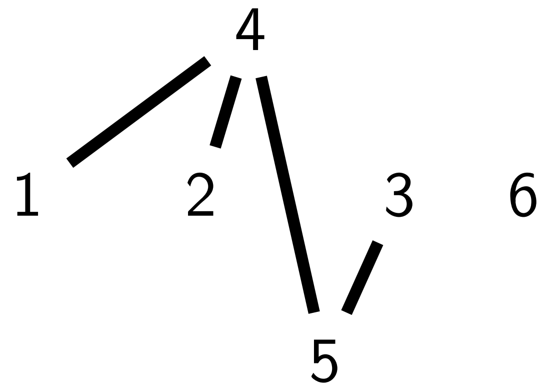
Interval Maximality: Hasse Diagram & Algorithm

maximal elements with partial ordering = undominated elements

example:



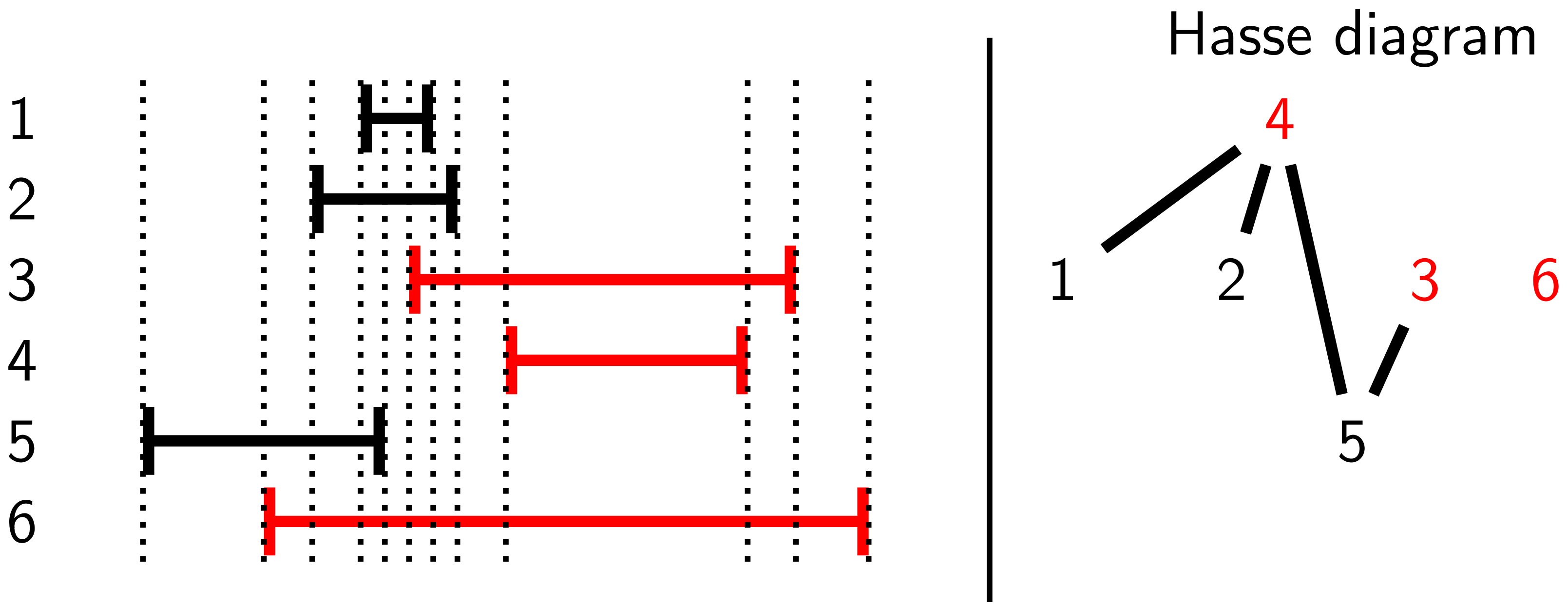
Hasse diagram



Interval Maximality: Hasse Diagram & Algorithm

maximal elements with partial ordering = undominated elements

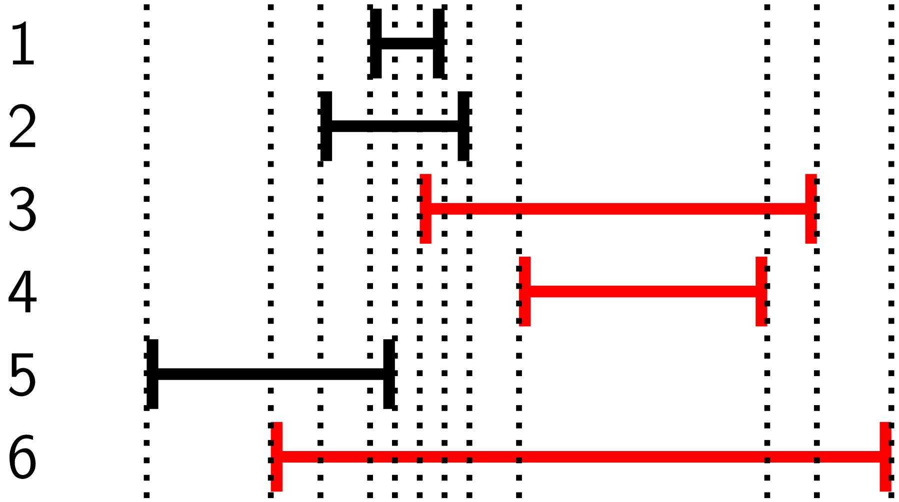
example:

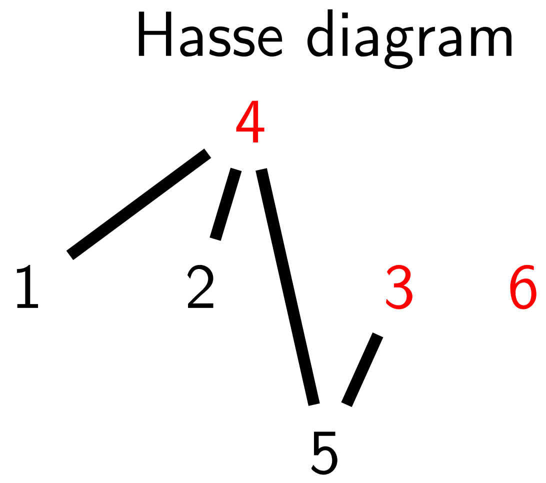


Interval Maximality: Hasse Diagram & Algorithm

maximal elements with partial ordering = undominated elements

example:





Theorem

All non-interval-maximal elements are dominated by the interval that has the highest lower bound.

$\Longrightarrow$ no need for Hasse diagram to find interval maximal elements

Interval Maximality: Practical Implementation

Recipe (Interval Maximality Optimality Criterion)

1. set up the table with gambles and probabilities
2. calculate the expectation of each gamble with respect to each probability mass function — matrix multiplication
3. calculate *minimum and maximum* expectation of each gamble = interval expectation — minimum & maximum of each row
4. choose the decisions whose maximum expectation exceeds the overall largest minimum expectation — undominated intervals

$$\left\{ d \colon \overline{P}(g_d) \geq \max_{e \in D} \underline{P}(g_e) \right\} \tag{18}$$

Interval Maximality: Example

Example (Machinery, Overtime, or Nothing)

	increase	stay	p_1	p_2	$\underline{P}$	$\overline{P}$
increase			0.5	0.8		
stay			0.5	0.2		
machinery	440	260				
overtime	420	300				
nothing	370	370				
	(1)		(2)		(3) & (4)	

```
pmfs = c(
  0.5, 0.5,
  0.8, 0.2)
rvars = c(
  440, 260,
  420, 300,
  370, 370)
getexpectations = getexpectationsfunc(2, pmfs)
isintervalmaximal = ismaximalfunc(getexpectations, intervalcompare
isintervalmaximal(rvars)
```

Robust Bayes Maximality

literature: 'maximality'

(Condorcet 1785 [9], Sen 1977 [22], Walley 1991 [27])

- exploits the behavioural interpretation of lower previsions
- refines interval maximality (see Exercise 3 later!)

Definition (Partial Ordering by Robust Bayesian Comparison)

We say that f robust Bayes dominates g, and write

$$f \succ g \tag{19}$$

whenever any of the following equivalent conditions hold:

$$(\forall p \in \mathcal{M})\,(E_p(f) > E_p(g)) \tag{20}$$

$$\underline{P}(f - g) > 0 \tag{21}$$

(willing to pay a small amount in order to trade g for f)
($f - g + \epsilon$ is desirable for some $\epsilon > 0$)

Remember, for any probability mass function p and any gamble f:

$$E_p(f) := \sum_{\omega\in\Omega} p(\omega)f(\omega) \tag{22}$$

Robust Bayes Maximality: Hasse Diagram & Algorithm

Definition (Robust Bayes Maximality Optimality Criterion)

Choose any gamble which is undominated with respect to $\succ$.

example:

	E_{p_1}	E_{p_3}	E_{p_3}
g_1	1	0	-1
g_2	0	0	0
g_3	0.5	-1	-2
g_4	0.2	-2	-3
g_5	2	1	-0.5

for browny points: interval maximal gambles?

Robust Bayes Maximality: Hasse Diagram & Algorithm

Definition (Robust Bayes Maximality Optimality Criterion)

Choose any gamble which is undominated with respect to $\succ$.

example:

Hasse diagram

	E_{p_1}	E_{p_3}	E_{p_3}
g_1	1	0	-1
g_2	0	0	0
g_3	0.5	-1	-2
g_4	0.2	-2	-3
g_5	2	1	-0.5

1

for browny points: interval maximal gambles?

Robust Bayes Maximality: Hasse Diagram & Algorithm

Definition (Robust Bayes Maximality Optimality Criterion)

Choose any gamble which is undominated with respect to $\succ$.

example:

Hasse diagram

	E_{p_1}	E_{p_3}	E_{p_3}
g_1	1	0	-1
g_2	0	0	0
g_3	0.5	-1	-2
g_4	0.2	-2	-3
g_5	2	1	-0.5

1 2

for browny points: interval maximal gambles?

Robust Bayes Maximality: Hasse Diagram & Algorithm

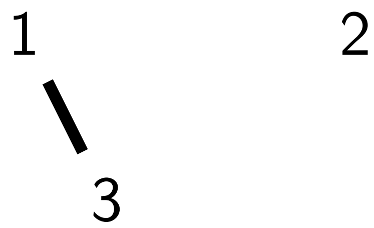
Definition (Robust Bayes Maximality Optimality Criterion)

Choose any gamble which is undominated with respect to $\succ$.

example:

Hasse diagram

	E_{p_1}	E_{p_3}	E_{p_3}
g_1	1	0	-1
g_2	0	0	0
g_3	0.5	-1	-2
g_4	0.2	-2	-3
g_5	2	1	-0.5



for browny points: interval maximal gambles?

Robust Bayes Maximality: Hasse Diagram & Algorithm

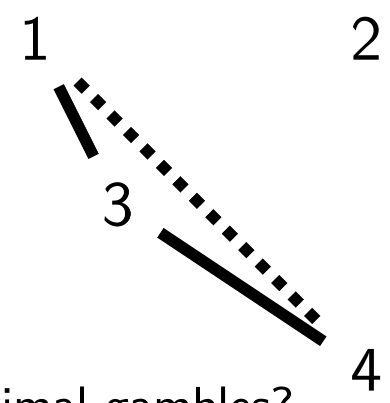
Definition (Robust Bayes Maximality Optimality Criterion)

Choose any gamble which is undominated with respect to $\succ$.

example:

	E_{p_1}	E_{p_3}	E_{p_3}
g_1	1	0	-1
g_2	0	0	0
g_3	0.5	-1	-2
g_4	0.2	-2	-3
g_5	2	1	-0.5

Hasse diagram



for browny points: interval maximal gambles?

Robust Bayes Maximality: Hasse Diagram & Algorithm

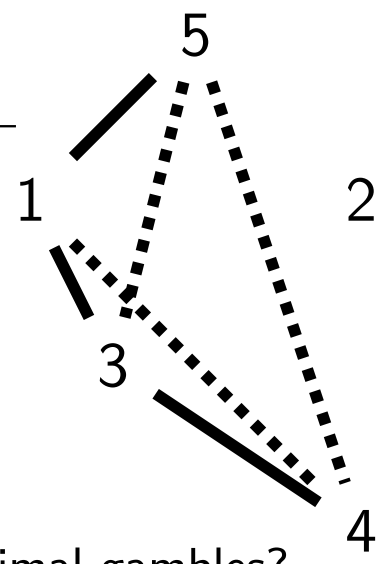
Definition (Robust Bayes Maximality Optimality Criterion)

Choose any gamble which is undominated with respect to $\succ$.

example:

	E_{p_1}	E_{p_3}	E_{p_3}
g_1	1	0	-1
g_2	0	0	0
g_3	0.5	-1	-2
g_4	0.2	-2	-3
g_5	2	1	-0.5

Hasse diagram



for browny points: interval maximal gambles?

Robust Bayes Maximality: Hasse Diagram & Algorithm

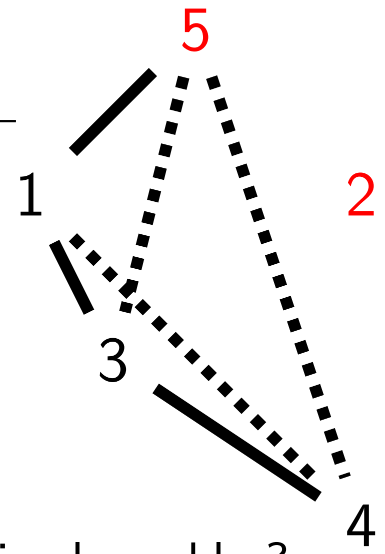
Definition (Robust Bayes Maximality Optimality Criterion)

Choose any gamble which is undominated with respect to $\succ$.

example:

	E_{p_1}	E_{p_3}	E_{p_3}
g_1	1	0	-1
g_2	0	0	0
g_3	0.5	-1	-2
g_4	0.2	-2	-3
g_5	2	1	-0.5

Hasse diagram



for browny points: interval maximal gambles?

Robust Bayes Maximality: Hasse Diagram & Algorithm

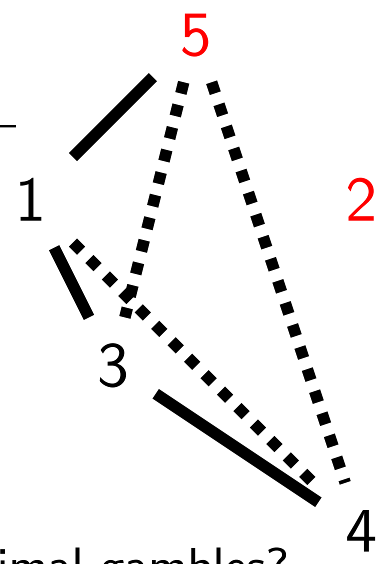
Definition (Robust Bayes Maximality Optimality Criterion)

Choose any gamble which is undominated with respect to $\succ$.

example:

	E_{p_1}	E_{p_3}	E_{p_3}
g_1	1	0	-1
g_2	0	0	0
g_3	0.5	-1	-2
g_4	0.2	-2	-3
g_5	2	1	-0.5

Hasse diagram



for browny points: interval maximal gambles?

Theorem

Every non-maximal element is dominated by a maximal element.

holds for arbitrary partial orderings!

$\Longrightarrow$ no need for Hasse diagram to find maximal elements:
once non-maximal element removed, no need to consider further!

Robust Bayes Maximality: Hasse Diagram & Algorithm

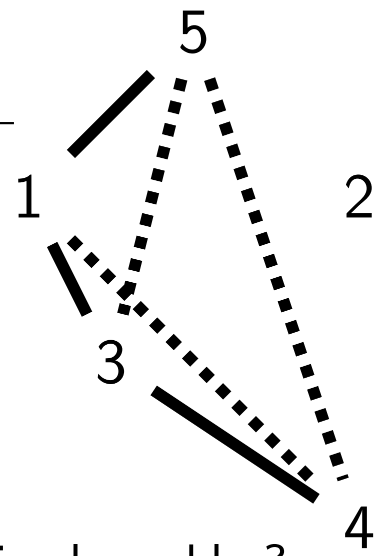
Definition (Robust Bayes Maximality Optimality Criterion)

Choose any gamble which is undominated with respect to $\succ$.

example:

	E_{p_1}	E_{p_3}	E_{p_3}
g_1	1	0	-1
g_2	0	0	0
g_3	0.5	-1	-2
g_4	0.2	-2	-3
g_5	2	1	-0.5

Hasse diagram



Maximality

1

for browny points: interval maximal gambles?

Theorem

Every non-maximal element is dominated by a maximal element.

holds for arbitrary partial orderings!

$\Longrightarrow$ no need for Hasse diagram to find maximal elements: once non-maximal element removed, no need to consider further!

Robust Bayes Maximality: Hasse Diagram & Algorithm

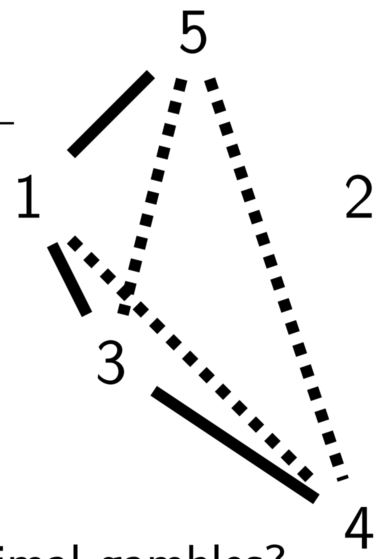
Definition (Robust Bayes Maximality Optimality Criterion)

Choose any gamble which is undominated with respect to $\succ$.

example:

	E_{p_1}	E_{p_3}	E_{p_3}
g_1	1	0	-1
g_2	0	0	0
g_3	0.5	-1	-2
g_4	0.2	-2	-3
g_5	2	1	-0.5

Hasse diagram



Maximality

for browny points: interval maximal gambles?

Theorem

Every non-maximal element is dominated by a maximal element.

holds for arbitrary partial orderings!

$\Longrightarrow$ no need for Hasse diagram to find maximal elements:
once non-maximal element removed, no need to consider further!

Robust Bayes Maximality: Hasse Diagram & Algorithm

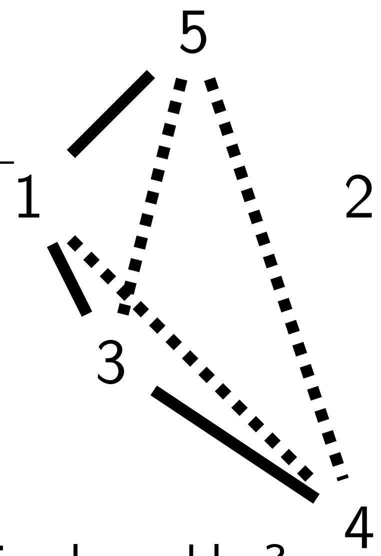
Definition (Robust Bayes Maximality Optimality Criterion)

Choose any gamble which is undominated with respect to $\succ$.

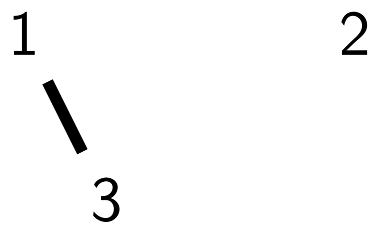
example:

	E_{p_1}	E_{p_3}	E_{p_3}
g_1	1	0	-1
g_2	0	0	0
g_4	0.2	-2	-3
g_5	2	1	-0.5

Hasse diagram



Maximality



for browny points: interval maximal gambles?

Theorem

Every non-maximal element is dominated by a maximal element.

holds for arbitrary partial orderings!

$\Longrightarrow$ no need for Hasse diagram to find maximal elements:
once non-maximal element removed, no need to consider further!

Robust Bayes Maximality: Hasse Diagram & Algorithm

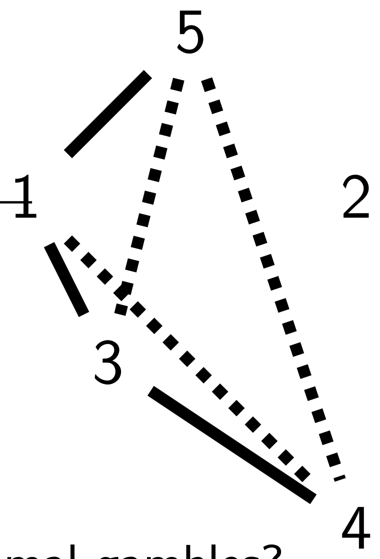
Definition (Robust Bayes Maximality Optimality Criterion)

Choose any gamble which is undominated with respect to $\succ$.

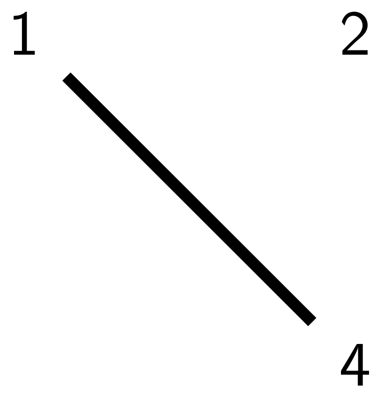
example:

	E_{p_1}	E_{p_3}	E_{p_3}
g_1	1	0	-1
g_2	0	0	0
g_5	2	1	-0.5

Hasse diagram



Maximality



for browny points: interval maximal gambles?

Theorem

Every non-maximal element is dominated by a maximal element.

holds for arbitrary partial orderings!

$\Longrightarrow$ no need for Hasse diagram to find maximal elements:
once non-maximal element removed, no need to consider further!

Robust Bayes Maximality: Hasse Diagram & Algorithm

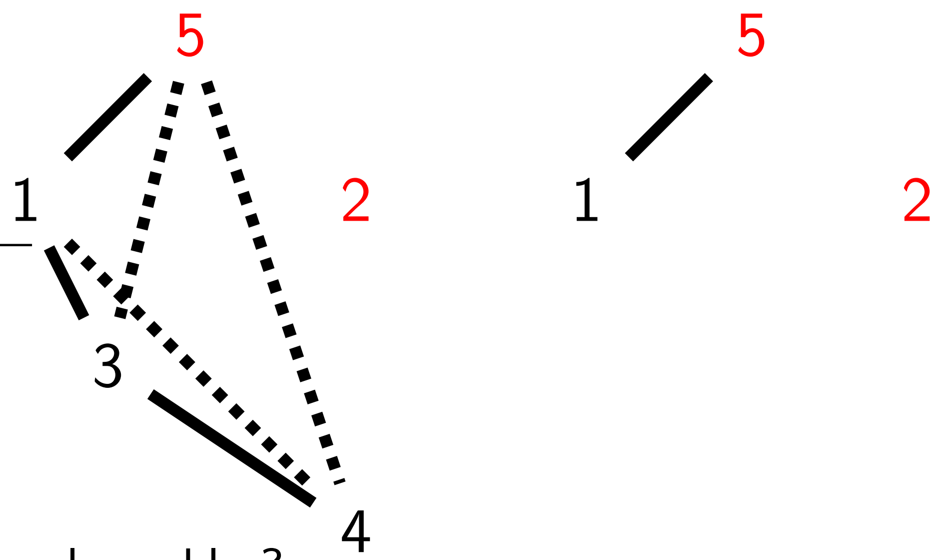
Definition (Robust Bayes Maximality Optimality Criterion)

Choose any gamble which is undominated with respect to $\succ$.

example:

	E_{p_1}	E_{p_3}	E_{p_3}
g_2	0	0	0
g_5	2	1	-0.5

Hasse diagram — Maximality



for browny points: interval maximal gambles?

Theorem

Every non-maximal element is dominated by a maximal element.

holds for arbitrary partial orderings!

$\Longrightarrow$ no need for Hasse diagram to find maximal elements:
once non-maximal element removed, no need to consider further!

Robust Bayes Maximality: Practical Implementation

Recipe (Robust Bayes Maximality Optimality Criterion)

1. set up the table with gambles and probabilities
2. calculate the expectation of each gamble
   with respect to each probability mass function   matrix multiplication
3. sequentially remove all decisions
   whose expectation rows are point-wise dominated

Robust Bayes Maximality: Example

Example (Machinery, Overtime, or Nothing)

	increase	stay	p_1	p_2
increase			0.5	0.8
stay			0.5	0.2
machinery	440	260		
overtime	420	300		
nothing	370	370		
	(1)		(2)	

```
pmfs = c(
  0.5, 0.5,
  0.8, 0.2)
rvars = c(
  440, 260,
  420, 300,
  370, 370)
getexpectations = getexpectationsfunc(2, pmfs)
isrbayesmaximal = ismaximalfunc(getexpectations, rbayescompare)
isrbayesmaximal(rvars)
```

Robust Bayes Admissibility

literature: 'E-admissibility'

(Pascal 1662 [16], Levi 1980 [13], Berger 1984 [6], Walley 1991 [27])

- refines robust Bayes maximality

Definition (Robust Bayes Admissibility Optimality Criterion)

Choose any gamble which maximizes expectation with respect to some $p \in \mathcal{M}$.

example:

	E_{p_1}	E_{p_3}	E_{p_3}
g_1	1	0	-1
g_2	0	0	0
g_3	0.5	-1	-2
g_4	0.2	-2	-3
g_5	2	1	-0.5

notes:

- computational challenge if $\mathcal{M}$ is large
- not invariant under convex hull operation: not enough just to look at extreme points

Robust Bayes Admissibility

literature: 'E-admissibility'

(Pascal 1662 [16], Levi 1980 [13], Berger 1984 [6], Walley 1991 [27])

- refines robust Bayes maximality

Definition (Robust Bayes Admissibility Optimality Criterion)

Choose any gamble which maximizes expectation with respect to some $p \in \mathcal{M}$.

example:

	E_{p_1}	E_{p_3}	E_{p_3}
g_1	1	0	-1
g_2	0	0	0
g_3	0.5	-1	-2
g_4	0.2	-2	-3
g_5	2	1	-0.5

notes:

- computational challenge if $\mathcal{M}$ is large
- not invariant under convex hull operation: not enough just to look at extreme points

Robust Bayes Admissibility: Practical Implementation

Recipe (Robust Bayes Admissibility Optimality Criterion)

1. set up the table with gambles and probabilities
2. calculate the expectation of each gamble
   with respect to each probability mass function   matrix multiplication
3. take all decisions that achieve a maximum
   in some expectation column

Robust Bayes Admissibility: Example

Example (Machinery, Overtime, or Nothing)

	increase	stay	p_1	p_2
increase			0.5	0.8
stay			0.5	0.2
machinery	440	260		
overtime	420	300		
nothing	370	370		
	(1)		(2) & (3)	

```
pmfs = c(
  0.5, 0.5,
  0.8, 0.2)
rvars = c(
  440, 260,
  420, 300,
  370, 370)
getexpectations = getexpectationsfunc(2, pmfs)
isrbayesadmissible = isrbayesadmissiblefunc(getexpectations)
isrbayesadmissible(rvars)
```

Robust Bayes Admissibility: Extreme Points Issue

Example (Machinery, Overtime, or Nothing)

	increase	stay	p_1	p_2	p_3
increase			0.5	0.8	0.65
stay			0.5	0.2	0.35
machinery	440	260			
overtime	420	300			
nothing	370	370			

```
pmfs = c(
  0.5, 0.5,
  0.8, 0.2,
  0.65, 0.35)
rvars = c(
  440, 260,
  420, 300,
  370, 370)
getexpectations = getexpectationsfunc(2, pmfs)
isrbayesadmissible = isrbayesadmissiblefunc(getexpectations)
isrbayesadmissible(rvars)
```

Decision making under severe uncertainty & applications in classification and risk analysis

Outline

Exercises

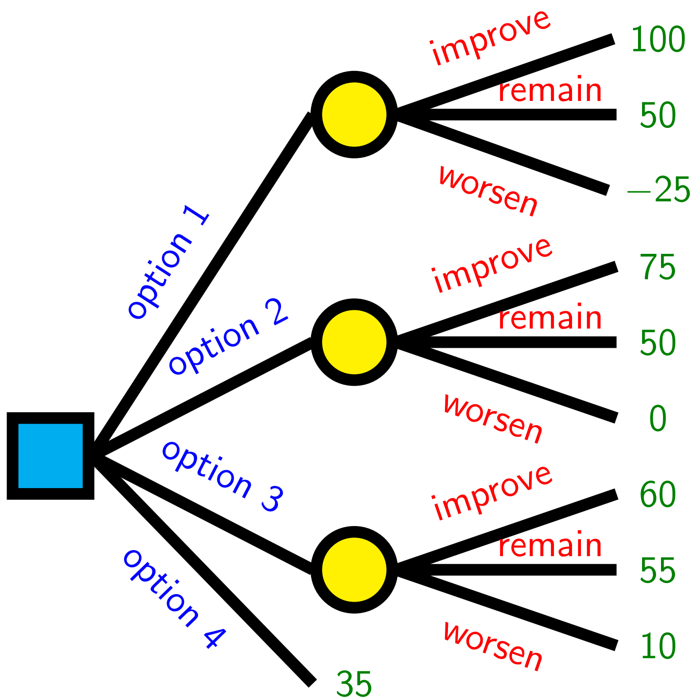
- Consider again the same very simple example. We have done additional market research, and we now know that demand will increase with probability at least 0.6, and at most 0.65. What advice can we give the manager now? Investigate with each optimality criterion.

Hint: $\mathcal{M} =$

	p_1	p_2
increase	0.6	0.65
stay	0.4	0.35

Exercises

- You have the option to invest some money. The market can either improve, remain, or worsen. Your set of probabilities are tabulated below. You have the choice between 4 options, summarized in the decision tree below.



$$\mathcal{M} = \begin{array}{c|cc} & p_1 & p_2 \\ \hline \text{improve} & 0.0 & 0.3 \\ \text{remain} & 0.6 & 0.3 \\ \text{worsen} & 0.4 & 0.4 \end{array}$$

Which options should you definitely not consider? First consider interval maximality, then consider robust Bayes maximality. Which of these two criteria gives the better answer?

A ranking problem

In an environmental problem, three possible decisions can be made. The table below lists the suitability of each of the options, as a 'best estimate', and also giving lower and upper bounds:

option	best estimate	lower bound	upper bound
1	6	5	12
2	10	3	11
3	8	4	10

You may assume that there is a possibility space Ω for this problem, and that each option i induces some gamble g_i on Ω.

A ranking problem (continued)

- Assume the best estimate corresponds to the (precise) expectation of g_i, so for example, $\underline{P}(g_1) = \overline{P}(g_1) = 6$. What is the optimal decision according to each of the decision criteria?
- We are not really sure whether we can interpret the best estimates as precise expectations, so we propose the following lower prevision, where $E(g_i)$ is the best estimate of g_i, and β is a parameter between 0 and 1:

$$\underline{P}\left(\alpha_0 + \sum_{i=1}^{3} \alpha_i g_i\right) := \alpha_0 + \beta \sum_{i=1}^{3} \alpha_i E(g_i) \tag{23}$$

$$+ (1-\beta) \sum_{i=1}^{3} \min\{\alpha_i \min(g_i), \alpha_i \max(g_i)\} \tag{24}$$

for any values of $\alpha_0, \ldots, \alpha_3 \in \mathbb{R}$. Try to interpret the above formula as well as the β parameter. Identify the optimal decisions for $\beta = 0$, $\beta = 1$, and $\beta = 0.5$ according to Γ-maximin, interval dominance, and maximality.

Decision making under severe uncertainty & applications in classification and risk analysis

Outline

Decision making under severe uncertainty & applications in classification and risk analysis

Outline

Credal Classification: What is Classification?

- actual class c (unknown), attributes $a_1, \ldots, a_k$
- decided class d
- $U(d, c)$ utility for deciding class is d if real class is c
  typical choice: $U(d, c) = 1$ if $d = c$ and $U(d, c) = 0$ if $d \neq c$
- aim: choose the best class given attributes

$$
\begin{aligned}
d^* &= \arg\max_d \sum_c U(d, c) p(c|a) && (25) \\
&= \arg\max_c p(c|a) = \arg\max_c p(c, a)/p(a) && (26) \\
&= \arg\max_c p(c, a) && (27)
\end{aligned}
$$

Open issues:

- How do we estimate the probabilities?
- Dealing with scarce data?
- Dealing with missing data?

Credal Classification: The Naive Bayes Classifier

Naive Bayes Classifier

Assume attributes are independent conditional on class:

$$p(c, a) = p(c)p(a|c) = p(c)\prod_{i=1}^{k} p(a_i|c) \tag{28}$$

Estimation of $p(c)$ and $p(a|c)$?

- maximum likelihood:

$$p(c) = \frac{n(c)}{N} \qquad p(a_i|c) = \frac{n(a_i, c)}{n(c)} \tag{29}$$

- Bayesian estimate with Dirichlet prior:

$$p(c) = \frac{n(c) + st(c)}{N + s} \qquad p(a_i|c) = \frac{n(a_i, c) + st(a_i, c)}{n(c) + st(c)} \tag{30}$$

(where $\sum_{a_i} t(a_i, c) = t(c)$)

Credal Classification: The Naive Credal Classifier

Estimation of $p(c)$ and $p(a|c)$?

- robust Bayesian estimate with imprecise Dirichlet model: as with Bayesian estimate but with sensitivity analysis over all possible $t(c)$ and $t(a_i, c)$
- Bounds for probabilities/expected utilities via optimisation.
- Use any of the decision criteria we discussed (interval dominance, robust Bayes maximality, robust Bayes admissibility, ...)

Case that we will study here:

- Simple approximate probability intervals.
- Interval dominance criterion.

Credal Classification: The Naive Credal Classifier

Bounds

$$\underline{p}(c, a) = \inf_t \frac{n(c) + st(c)}{N + s} \prod_{i=1}^{k} \frac{n(a_i, c) + st(a_i, c)}{n(c) + st(c)} \tag{31}$$

$$\geq \frac{n(c)}{N + s} \prod_{i=1}^{k} \frac{n(a_i, c)}{n(c) + s} = \underline{p}(c) \prod_{i=1}^{k} \underline{p}(a_i | c) \tag{32}$$

$$\overline{p}(c, a) = \sup_t \frac{n(c) + st(c)}{N + s} \prod_{i=1}^{k} \frac{n(a_i, c) + st(a_i, c)}{n(c) + st(c)} \tag{33}$$

$$\leq \frac{n(c) + s}{N + s} \prod_{i=1}^{k} \frac{n(a_i, c) + s}{n(c) + s} = \overline{p}(c) \prod_{i=1}^{k} \overline{p}(a_i | c) \tag{34}$$

Credal Classification: The Naive Credal Classifier

Interval Dominance

Consider the set of all classes c for which

$$\overline{p}(c) \prod_{i=1}^{k} \overline{p}(a_i|c) \geq \max_{c'} \underline{p}(c') \prod_{i=1}^{k} \underline{p}(a_i|c') \qquad (35)$$

classifier can return multiple classes if it is unsure about the probabilities!

Credal Classification

Decision making under severe uncertainty & applications in classification and risk analysis

Outline

Breast Cancer Example: R Code Preparation

1. start R
2. visit course webpage with browser
3. download the data.txt and code.r files
4. select and copy all R code from your favourite editor: CTRL-A, CTRL-C
5. go to R console
6. paste code into R console: CTRL-V, ENTER
7. keep browser window open, so you can rinse & repeat steps 3–5 every time you start a new R session

Breast Cancer Example

```
mammo = getdata()
myclassifier = classifier.naive2(0)
model = myclassifier$trainer(mammo, 1:5, 6)
testrow = mammo[6,]
print(testrow)
print(myclassifier$tester(model, testrow))
testrow = mammo[5,]
print(testrow)
print(myclassifier$tester(model, testrow))
```

```
myclassifier = classifier.composed(
  list(classifier.naive2(0),
       classifier.naive2(1),
       classifier.credal(2)))
mammo = getdata()[1:30,]
print(kfcv.classifier(mammo, 1:5, 6, myclassifier))
print(kfcv.classifier(mammo, 2:5, 6, myclassifier))
print(kfcv.classifier(mammo, 1, 6, myclassifier))
mammo = getdata()
print(kfcv.classifier(mammo, 2:5, 1, myclassifier))
```

Exercises

- Try to run the code for the credal classifier.
- What do you observe if you increase the amount of data that is used to train the classifier? Compare the results you get from the traditional classifier with the results you get from the credal classifier.
- What happens if you increase the s value of the credal classifier? Confirm your intuition by running the code.
- How would the formulas for the credal classifier (based on interval dominance) change if the utilities were not 0–1 valued?
- Zaffalon's 2001 paper discusses how the problem can be solved using robust Bayes maximality. Try to implement his algorithm in R by modifying the existing code for interval dominance.

Decision making under severe uncertainty & applications in classification and risk analysis

Outline