Meeting 6:
Decision-theoretic components…

Inferential approaches within statistics

- Descriptive/Explorative
  - Tables, Diagrams, Sample statistics
  - No quantification of the uncertainty
- Explanatory
  - Application of statistical models to data
  - Estimation and interpretation of parameters
- Predictive
  - Application of statistical models to data
  - Estimation and tuning (learning) of parameters for the prediction of new cases
- Decisive
  - Main objective is to make decisions under uncertainty
  - Estimation (and prediction) with statistical models is made by probabilistic (or non-probabilistic) criteria

Frequentist or Bayesian?

- Statistical decision theory is not per definition limited to a specific paradigm (frequentist vs. Bayesian)

- However, evolving from the explanatory or predictive approach to a decisive approach is very often coupled with the application of Bayesian statistical thinking

- Two main types of criteria:

    - Probabilistic criteria: Maximise the expected *utility* (or maximise expected *payoff* $\Leftrightarrow$ minimise expected *loss*) – expected with respect to the probability distribution of the state of nature (*Bayesian*)
    - Non-probabilistic criteria: Maximin principle, Maximax principle, Minimax principle (*non-Bayesian*)

Decision-theoretic elements

1. One of a number of decisions (or actions) should be chosen
2. State of world/nature: A number of states possible – can be an infinite number. Usually represented by a parameter $\boldsymbol{\theta}$
3. The consequence of taking a particular action given a certain state of world/ nature is known (for all combinations of states and actions)
4. For each state of nature the relative desirability of each of the different actions possible can be quantified
5. Prior information for the different states of nature *may* be available: Prior distribution of $\boldsymbol{\theta}$
6. Data *may* be available. Usually represented by $\boldsymbol{x}$. Can be used to update the knowledge about the relative desirability of (each of) the different actions

The classical approach

True state of world: /nature $\boldsymbol{\theta}$ — Unknown. The Bayesian description of this uncertainty is in terms of a random variable $\tilde{\boldsymbol{\theta}}$ with prior density $p(\boldsymbol{\theta})$

Data: $\boldsymbol{x}$ — Observation of $\tilde{\boldsymbol{x}}$, whose pdf (or pmf) depends on $\boldsymbol{\theta}$ (data is thus assumed to be available)

Decision rule: δ

Action: $\delta(\boldsymbol{x})$ — The decision rule becomes an action when applied to given data $\boldsymbol{x}$

Loss function: $L(\delta(\boldsymbol{x}), \boldsymbol{\theta})$ — measures the *loss* from taking action $\delta(\boldsymbol{x})$ when $\boldsymbol{\theta}$ holds

Payoff function $R(\delta(\boldsymbol{x}), \boldsymbol{\theta})$ — measures the *payoff* from taking action $\delta(\boldsymbol{x})$ when $\boldsymbol{\theta}$ holds

Risk function:

$$D(\delta, \boldsymbol{\theta}) = \int_{\boldsymbol{x}} L(\delta(\boldsymbol{x}), \boldsymbol{\theta}) \underbrace{f(\boldsymbol{x}|\boldsymbol{\theta})}_{\text{Likelihood}} d\boldsymbol{x} = E_{\tilde{\boldsymbol{x}}}\left(L(\delta(\tilde{\boldsymbol{x}}), \boldsymbol{\theta})\right)$$

Chance function

$$C(\delta, \boldsymbol{\theta}) = \int_{\boldsymbol{x}} R(\delta(\boldsymbol{x}), \boldsymbol{\theta}) f(\boldsymbol{x}|\boldsymbol{\theta}) d\boldsymbol{x} = E_{\tilde{\boldsymbol{x}}}\left(R(\delta(\tilde{\boldsymbol{x}}), \boldsymbol{\theta})\right)$$

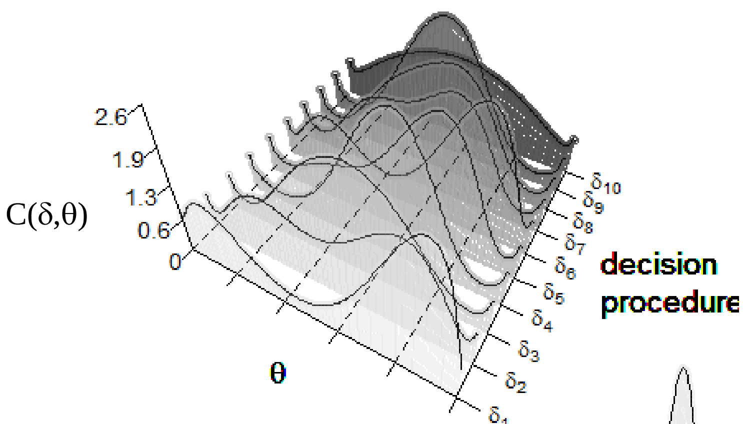
<u>Expected</u> loss/payoff <u>with respect to</u> variation in $\boldsymbol{x}$

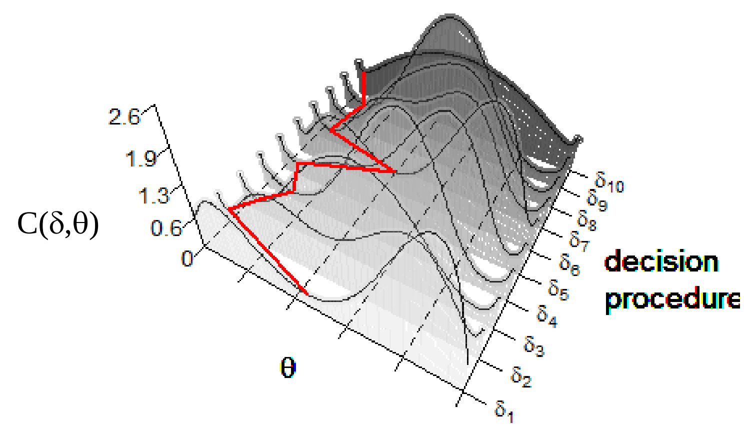
Functions of the state of world/nature and the *decision rule* (and not the action)

Maximin, Maximax and Minimax decision rules [non-probabilistic criteria]

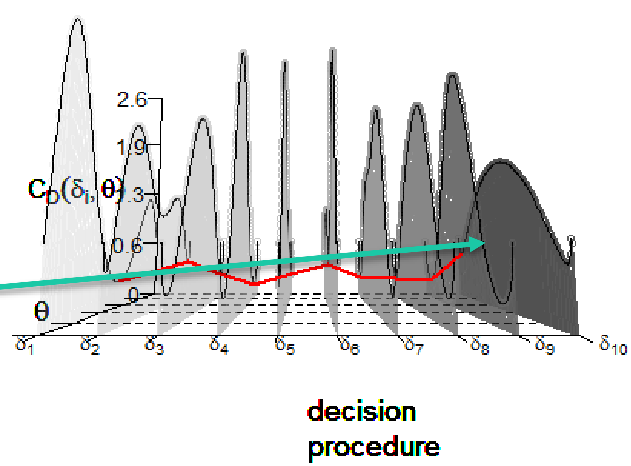
A procedure δ^* is a *maximin* decision rule if $C\left(\delta^*,(\boldsymbol{\theta})\right) = \max_{\delta}\left\{\min_{\boldsymbol{\theta}} C\left(\delta,\boldsymbol{\theta}\right)\right\}$

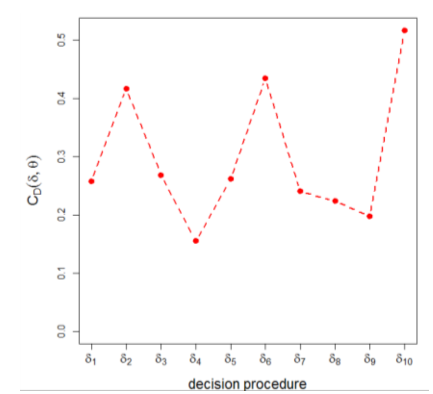
i.e. $\boldsymbol{\theta}$ is for each decision rule is chosen to be the "worst" possible value, and under that value the decision rule that gives the highest possible chance is chosen.





$\delta^* = \delta_{10}$

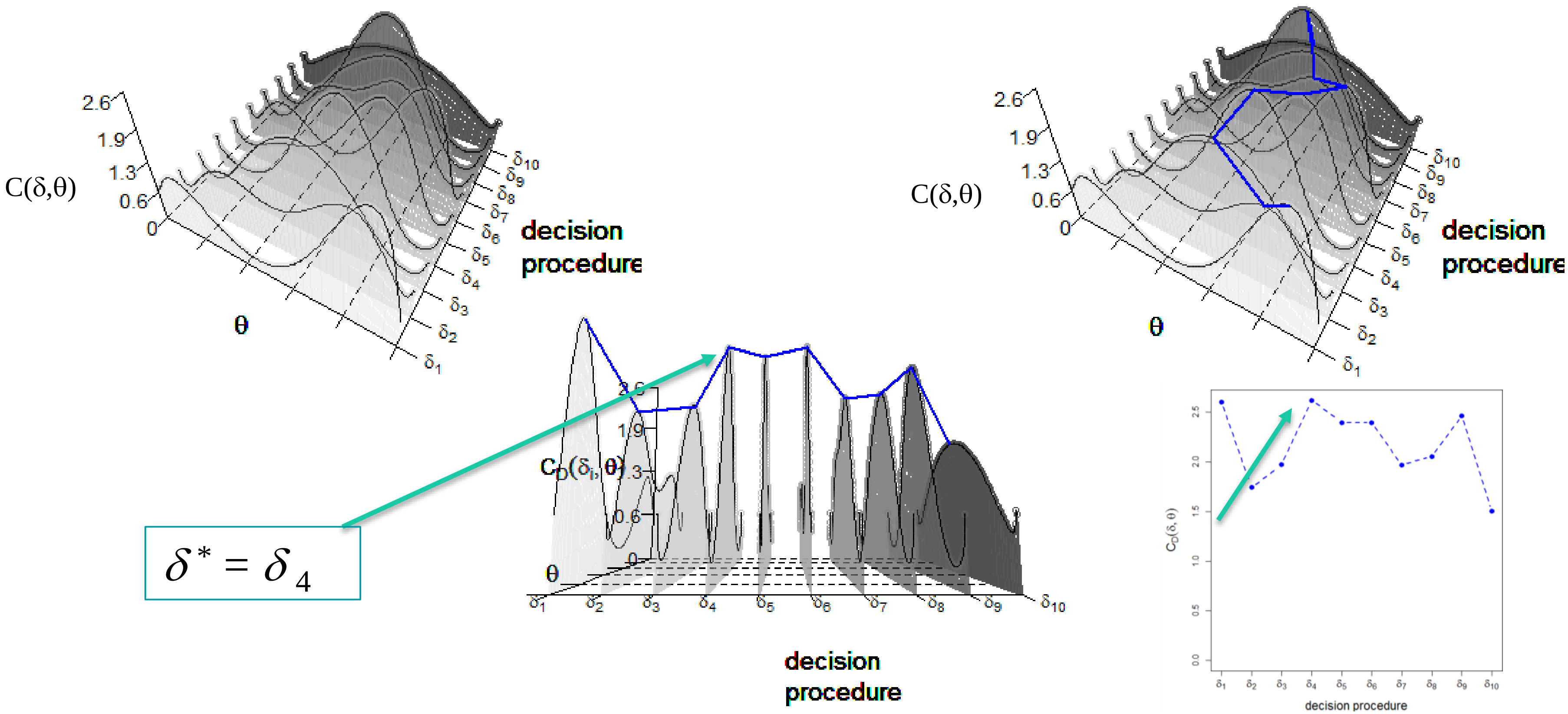




A procedure δ^* is a *maximax* decision rule if $C\left(\delta^*,(\boldsymbol{\theta})\right) = \max\limits_{\delta}\left\{\max\limits_{\boldsymbol{\theta}} C(\delta,\boldsymbol{\theta})\right\}$

i.e. $\boldsymbol{\theta}$ is for each decision rule chosen to be the “best” possible value, and under that value the decision rule that gives the highest possible chance is chosen.

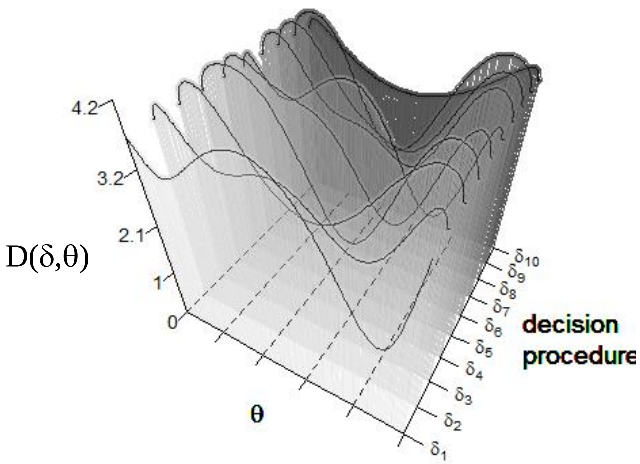
The maximax rule is a typical *optimistic* rule

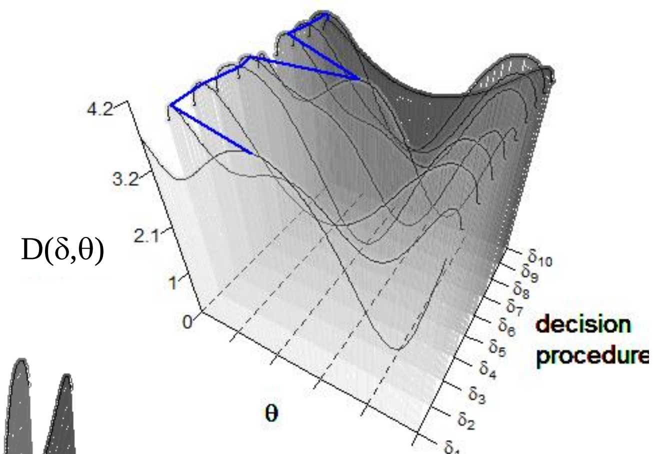


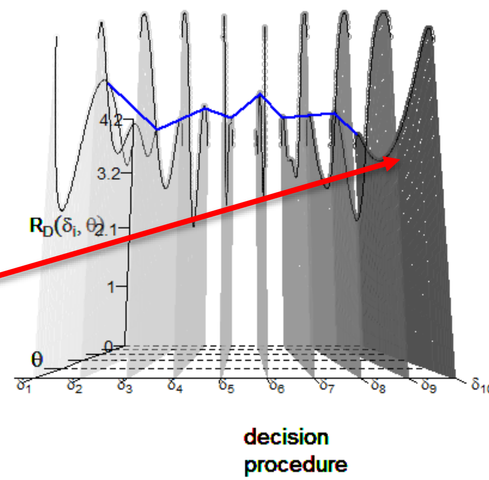
A procedure δ^* is a *minimax* decision rule if $D\left(\delta^*,(\boldsymbol{\theta})\right)=\min_{\delta}\left\{\max_{\boldsymbol{\theta}} D(\delta,\boldsymbol{\theta})\right\}$

i.e. $\boldsymbol{\theta}$ is chosen to be the "worst" possible value, and under that value the decision rule that gives the lowest possible risk is chosen.

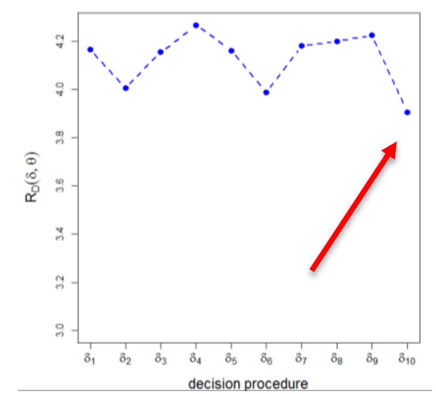
The minimax rule is a typical *pessimistic* rule







$\delta^* = \delta_{10}$



Maximin $$C\left(\delta^{*},(\boldsymbol{\theta})\right)=\max_{\delta}\left\{\min_{\boldsymbol{\theta}} C\left(\delta,\boldsymbol{\theta}\right)\right\}$$

Maximax $$C\left(\delta^{*},(\boldsymbol{\theta})\right)=\max_{\delta}\left\{\max_{\boldsymbol{\theta}} C\left(\delta,\boldsymbol{\theta}\right)\right\}$$

Minimax $$R\left(\delta^{*},(\boldsymbol{\theta})\right)=\min_{\delta}\left\{\max_{\boldsymbol{\theta}} R\left(\delta,\boldsymbol{\theta}\right)\right\}$$

The maximin, maximax and minimax rules use no prior information about $\boldsymbol{\theta}$, thus are not Bayesian rules.

Note! In this description the rules work with expected payoffs and losses with respect to data (Chance and Risk functions). All three rules may however be applied directly to payoff or loss tables for different states of the world/nature.

Example

Suppose you are about to make a decision on whether you should buy or rent a new TV to have for two years = 24 months.

➔ δ_1 = "Buy the TV" δ_2 = "Rent the TV"

Now, assume θ is the mean time until the TV breaks down for the first time.

Let θ assume three possible values: 6, 12 and 24 months.

The cost of the TV is $500 if you buy it and $30 per month if you rent it.

If the TV breaks down after 12 months (length of warranty) you'll have to replace it for the same cost as you bought it if you bought it. If you rented it you will get a new TV for no cost provided you proceed with your contract.

Let X be the time in months until the TV breaks down and assume this variable is exponentially distributed with mean θ.

➔ A loss function for an ownership of maximum 24 months may be defined as

$L(\delta_1(X), \theta) = 500 + 500 \cdot \mathbf{1}_{\{X-12\}}$ and

$L(\delta_2(X), \theta) = 30 \cdot 24 = 720$

where $\mathbf{1}_{\{y\}} = \begin{cases} 0 & y < 0 \\ 1 & y \geq 0 \end{cases}$

Then $$D(\delta_1,\theta)=E\left(500+500\cdot\mathbf{1}_{\{X-12\}}\right)=500+500\int_{12}^{\infty}\theta^{-1}e^{-\theta^{-1}x}=$$

$$=500\cdot\left(1+e^{-12/\theta}\right)$$

$$D(\delta_2,\theta)=720$$

Now compare the risks for the three possible values of θ.

θ	$D(\delta_1,\theta)$	$D(\delta_2,\theta)$
6	$500\cdot(1+e^{-12/6}) = 568$	720
12	$500\cdot(1+e^{-12/12}) = 684$	720
24	$500\cdot(1+e^{-12/24}) = 803$	720

Clearly the risk for the first rule increases with θ while the risk for the second is constant. In searching for the minimax rule we therefore focus on the largest possible value of θ and there δ_2 has the smallest risk.

Minimax decision rule = δ_2

For the maximin and maximax rules, let $C(\delta_i,\theta) = K - D(\delta_i,\theta)$

Details about the relation will come

$\Rightarrow$ Maximin decision rule = δ_2 and Maximax decision rule = δ_1

Bayes decision rule(s) [probabilistic criteria]

To make formulas a bit simpler we use the notation

$p(\boldsymbol{\theta}) = f'(\boldsymbol{\theta})$ prior probability density (or mass) function for $\tilde{\boldsymbol{\theta}}$

$f(\boldsymbol{x}|\boldsymbol{\theta})$ likelihood

$q(\boldsymbol{\theta}|\boldsymbol{x}) = f''(\boldsymbol{\theta}|\boldsymbol{x})$ posterior probability density (or mass) function for $\tilde{\boldsymbol{\theta}}$

Bayes risk of procedure δ: $$B(\delta) = \int_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} D(\delta,\boldsymbol{\theta})p(\boldsymbol{\theta})d\boldsymbol{\theta}$$

i.e. the risk function averaged over the prior distribution.

Note! The integral is a sum when $p(\boldsymbol{\theta})$ is a pmf.

A *Bayes rule* is a procedure that minimizes the Bayes risk

$$\delta_B = \arg\min_{\delta}\left\{\int_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} D(\delta,\boldsymbol{\theta})p(\boldsymbol{\theta})d\boldsymbol{\theta}\right\}$$

Note! This is about <u>the decision rule</u>, not a specific action

However,

$$\int_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} D(\delta,\boldsymbol{\theta})p(\boldsymbol{\theta})d\boldsymbol{\theta} = \int_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \left(\int_{\boldsymbol{x}} L(\delta(\boldsymbol{x}),\boldsymbol{\theta}) f(\boldsymbol{x}|\boldsymbol{\theta})d\boldsymbol{x} \right) p(\boldsymbol{\theta})d\boldsymbol{\theta}$$

$$= \int_{\boldsymbol{x}} \int_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} L(\delta(\boldsymbol{x}),\boldsymbol{\theta}) f(\boldsymbol{x}|\boldsymbol{\theta})p(\boldsymbol{\theta})d\boldsymbol{\theta}d\boldsymbol{x} = \int_{\boldsymbol{x}} \int_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} L(\delta(\boldsymbol{x}),\boldsymbol{\theta}) q(\boldsymbol{\theta}|\boldsymbol{x})h(\boldsymbol{x})d\boldsymbol{\theta}$$

where $h(\boldsymbol{x})$ is the marginal density of $\boldsymbol{X}$, i.e

$$h(\boldsymbol{x}) = \int_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} f(\boldsymbol{x}|\boldsymbol{\theta})p(\boldsymbol{\theta})d\boldsymbol{\theta}$$ (prior predictive distribution)

$$\boxed{q(\boldsymbol{\theta}|\boldsymbol{x}) = \frac{f(\boldsymbol{x}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{h(\boldsymbol{x})}}$$

$$\Rightarrow \int_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} D(\delta,\boldsymbol{\theta})p(\boldsymbol{\theta})d\boldsymbol{\theta} = \int_{\boldsymbol{x}} h(\boldsymbol{x}) \int_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} L(\delta(\boldsymbol{x}),\boldsymbol{\theta}) q(\boldsymbol{\theta}|\boldsymbol{x})d\boldsymbol{\theta}d\boldsymbol{x}$$

$$= \int_{\boldsymbol{x}} h(\boldsymbol{x}) E_{\tilde{\boldsymbol{\theta}}|\boldsymbol{x}}\left(L\left(\delta(\boldsymbol{x}),\tilde{\boldsymbol{\theta}}\right)\right)d\boldsymbol{x}$$

$$\Rightarrow \arg\min_{\delta}\left\{ \int_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} D(\delta,\boldsymbol{\theta})p(\boldsymbol{\theta})d\boldsymbol{\theta} \right\} = \arg\min_{\delta}\left\{ E_{\tilde{\boldsymbol{\theta}}|\boldsymbol{x}}\left(L\left(\delta(\boldsymbol{x}),\tilde{\boldsymbol{\theta}}\right)\right)\right\}$$

for any given value of $\boldsymbol{x}$

$$\delta_B = \arg\min_{\delta}\left\{E_{\tilde{\theta}|\boldsymbol{x}}\left(L\left(\delta(\boldsymbol{x}),\tilde{\theta}\right)\right)\right\}$$

Hence, a procedure that minimises the posterior expected loss is a Bayes decision rule (a probabilistic criterion).

Example: Rent or buy TV cont.

$L(\delta_1(X),\theta) = 500 + 500 \cdot \mathbf{1}_{\{X-12\}}$

$L(\delta_2(X),\theta) = 30 \cdot 24 = 720$

$f(\mathrm{x}|\theta) = \theta^{-1}\mathrm{e}^{-x/\theta}$

Assume the three possible values of θ (6, 12 and 24) have the prior probabilities 0.2, 0.3 and 0.5 respectively.

$$p(\theta) = \begin{cases} 0.2 & \theta = 6 \\ 0.3 & \theta = 12 \\ 0.5 & \theta = 24 \\ 0 & \text{otherwise} \end{cases} \qquad (\text{pmf})$$

Using the Bayes risks directly

$$B(\delta_1) = \left\langle \begin{matrix} \theta \text{ is discrete -} \\ \text{valued} \end{matrix} \right\rangle = \sum_{\theta\in\boldsymbol{\Theta}} D(\delta_1,\theta)p(\theta) = \left\langle \begin{matrix} \text{from previous} \\ \text{example} \end{matrix} \right\rangle = \sum_{\theta\in\boldsymbol{\Theta}} 500\left(1-e^{-12/\theta}\right)p(\theta)$$

$$= 500\left(1-e^{-12/6}\right)\cdot 0.2 + 500\left(1-e^{-12/12}\right)\cdot 0.3 + 500\left(1-e^{-12/24}\right)\cdot 0.5 = 280$$

$$B(\delta_2) = \sum_{\theta\in\boldsymbol{\Theta}} D(\delta_2,\theta)p(\theta) = \left\langle R_D(\delta_2,\theta) \equiv 720 \right\rangle = 720\cdot 0.2 + 720\cdot 0.3 + 720\cdot 0.5 = 720$$

Thus, the minimal Bayes risk is with procedure δ_1 and therefore δ_1 is the Bayes decision rule (among δ_1 and δ_2).

Working with payoffs instead of losses

How do we define payoff and how do we define loss?

The **payoff** represents the *net change* in your total “wealth” as a function of your action and the actual state of the world/nature.

The payoff can then be seen as the *consequence* of your decision with the actual state of the world/nature.

The term “wealth” is not necessarily to be interpreted in monetary units, but very often it should be possible – but perhaps difficult – to translate non-monetary consequences to cash equivalents

Defining payoff as the net change means that all costs involved are taken into consideration. Hence the payoff may be negative.

Example

Suppose you are about to sell apples on an open market a Saturday in October. You are choosing between selling one of two kinds of apples. For the first kind – a lower quality apple - you can buy 100 kg apples for SEK 1000, and you deem a reasonable highest selling price to be SEK 18 per kg. For the second kind – a higher quality apple – you can buy 80 kg for SEK 1200, and you deem a reasonable highest selling price to be 25 per kg.

Assume the total demand for your apples that day is 50 kg (no matter what kind of apple).

Your payoff with the decision to sell apples of the lower quality will be SEK 50×18 – 1000 = –100.

Your payoff with the decision to sell apples of the higher quality will be SEK 50×25 – 1200 = 50.

If the total demand is 100 kg? If your decision is to sell apples of the lower quality, your payoff will be SEK 100×18 – 1000 = 800, and if your decision is to sell apples of the higher quality, your payoff will be SEK 80×25 – 1200 = 800.

The **loss** of an action is always to be interpreted in terms of *opportunity* loss, i.e. it is the difference between the maximal payoff that can be obtained with a certain state of the world and the payoff of the particular action.

Hence, the relation between loss (L) and payoff (R) for an action $\delta^*(\boldsymbol{x})$ is

$$L\left(\delta^*(\boldsymbol{x}),\boldsymbol{\theta}\right)=\max_{\delta}\left\{R\left(\delta(\boldsymbol{x}),\boldsymbol{\theta}\right)\right\}-R\left(\delta^*(\boldsymbol{x}),\boldsymbol{\theta}\right)$$

This implies that the loss can never be negative , while the payoff can be positive, negative or zero.

Finite-action problems

Applied to a payoff table with m rows and n column, where each row represents an action and each column a state of the world:

	θ_1	θ_2	$\ldots$
δ_1	L_{11} , R_{11}	L_{12} , R_{12}	
δ_2	L_{21} , R_{21}	L_{22} , R_{22}	
$\ldots$			

$$L_{ij}=\max_{k}\left\{R_{kj}\right\}-R_{ij}\quad i=1,\ldots,m\,,\ j=1,\ldots,n$$

where L_{ij} and R_{ij} stand for the loss and the payoff respectively when action i is chosen with state j of the world.

Now,
$$\delta_B = \arg\min_{\delta}\left\{E_{\tilde{\boldsymbol{\theta}}|\boldsymbol{x}}\left(L\left(\delta(\boldsymbol{x}),\tilde{\boldsymbol{\theta}}\right)\right)\right\}$$
$$= \arg\min_{\delta}\left\{E_{\tilde{\boldsymbol{\theta}}|\boldsymbol{x}}\left(\underbrace{\max_{d}\left\{R\left(d(\boldsymbol{x}),\tilde{\boldsymbol{\theta}}\right)\right\}}_{\text{not depending on } \delta} - R\left(\delta(\boldsymbol{x}),\tilde{\boldsymbol{\theta}}\right)\right)\right\}$$
$$= \arg\min_{\delta}\left\{-E_{\tilde{\boldsymbol{\theta}}|\boldsymbol{x}}\left(R\left(\delta(\boldsymbol{x}),\tilde{\boldsymbol{\theta}}\right)\right)\right\} = \arg\max_{\delta}\left\{E_{\tilde{\boldsymbol{\theta}}|\boldsymbol{x}}\left(R\left(\delta(\boldsymbol{x}),\tilde{\boldsymbol{\theta}}\right)\right)\right\}$$

Hence, a procedure that maximises the posterior expected payoff is (also) a Bayes procedure.

Important, though: Payoff is not synonymous with *utility*

Note that implicit in the derivation above is:

$$E_{\tilde{\boldsymbol{\theta}}|\boldsymbol{x}}\left(L\left(\delta(\boldsymbol{x}),\tilde{\boldsymbol{\theta}}\right)\right) = E_{\tilde{\boldsymbol{\theta}}|\boldsymbol{x}}\left(\max_{d}\left\{R\left(d(\boldsymbol{x}),\tilde{\boldsymbol{\theta}}\right)\right\}\right) - E_{\tilde{\boldsymbol{\theta}}|\boldsymbol{x}}\left(R\left(\delta(\boldsymbol{x}),\tilde{\boldsymbol{\theta}}\right)\right) = T - E_{\tilde{\boldsymbol{\theta}}|\boldsymbol{x}}\left(R\left(\delta(\boldsymbol{x}),\tilde{\boldsymbol{\theta}}\right)\right)$$

where T is a quantity that does not depend on the action $\delta(\boldsymbol{x})$

(In)Admissibility

Decision procedures

A decision procedure δ_1 is ***inadmissible*** if there exists another decision procedure δ_2 such that

$$D(\delta_1, \boldsymbol{\theta}) \geq D(\delta_2, \boldsymbol{\theta})$$ for all states of the world/nature θ <u>and</u>

$$D(\delta_1, \boldsymbol{\theta}) > D(\delta_2, \boldsymbol{\theta})$$ (strict inequality) for at least one state θ

i.e. the risk function of procedure δ_2 is never larger than the risk function of procedure δ_1, and strictly less for at least one state of the world/nature.

The procedure δ_2 is then said to *dominate* procedure δ_1.

If a procedure is inadmissible, it should never be considered in a decision problem

If a procedure is not dominated by any other procedure, it is ***admissible*** .

Actions

An action $\delta_1(\boldsymbol{x})$ is ***inadmissible*** if there exists another action $\delta_2(\boldsymbol{x})$ such that

$$R(\delta_2(\boldsymbol{x}),\boldsymbol{\theta}) \geq R(\delta_1(\boldsymbol{x}),\boldsymbol{\theta}) \text{ or } L(\delta_2(\boldsymbol{x}),\boldsymbol{\theta}) \leq L(\delta_1(\boldsymbol{x}),\boldsymbol{\theta})$$

for all states of the world/nature θ and

$$R(\delta_2(\boldsymbol{x}),\boldsymbol{\theta}) > R(\delta_1(\boldsymbol{x}),\boldsymbol{\theta}) \text{ or } L(\delta_2(\boldsymbol{x}),\boldsymbol{\theta}) < L(\delta_1(\boldsymbol{x}),\boldsymbol{\theta})$$

(strict inequality) for at least one state θ

The action $\delta_2(\boldsymbol{x})$ is then said to *dominate* action $\delta_1(\boldsymbol{x})$.

Note that for actions we compare the actual payoffs (or the losses) [and not their expectations over all possible data]

If an action is inadmissible, it should not be considered.

If an action is not dominated by any other action, it is ***admissible*** .

Note! An action $\delta(\boldsymbol{x})$ can be inadmissible even if the procedure δ is admissible.

Example, apples cont.

We can form a payoff table as

	Demand is 50 kg	Demand is 100 kg
δ_1 = Sell lower quality apples	$R_{11} = -100$	$R_{12} = 800$
δ_2 = Sell higher quality apples	$R_{21} = 50$	$R_{22} = 800$

Since the decision to sell higher quality apples (δ_2) would give payoffs that do not fall short of the payoffs given by selling lower quality apples (δ_1), <u>and</u> the payoff with decision δ_2 when the demand is 50 kg is higher than the payoff with decision $\delta_{1,}$ the decision to sell lower quality apples is inadmissible.

provided the only states of the world (possible demands) considered are 50 kg and 100 kg

The presentation in "Wrinkler: An introduction to Bayesian Inference and Decision" is slightly different from what has been presented here.

We have seen that one way of obtaining a (*n.b "a" not "the"*) Bayes rule for decision making (a probabilistic rule) is either to…

…minimise the posterior expected loss, i.e. $E_{\tilde{\boldsymbol{\theta}}|\boldsymbol{x}}\left(L\left(\delta(\boldsymbol{x}),\tilde{\boldsymbol{\theta}}\right)\right)$, or to…

…maximise the posterior expected payoff, i.e. $E_{\tilde{\boldsymbol{\theta}}|\boldsymbol{x}}\left(R\left(\delta(\boldsymbol{x}),\tilde{\boldsymbol{\theta}}\right)\right)$

If no data is available we would still be able to use the quantified uncertainty about the state of the world ($\boldsymbol{\theta}$), but now in terms of a *prior distribution*. Hence a Bayes rule can be obtained by

…minimising the prior expected loss, i.e. $E_{\tilde{\boldsymbol{\theta}}}\left(L\left(\delta,\tilde{\boldsymbol{\theta}}\right)\right)$, or…

…maximising the prior expected payoff, i.e. $E_{\tilde{\boldsymbol{\theta}}}\left(R\left(\delta,\tilde{\boldsymbol{\theta}}\right)\right)$

Now, Wrinkler writes…

- $ER(\delta)$ for the expected payoff from making decision δ
- $EL(\delta)$ for the expected loss from making decision δ

…with both expectations under an anticipated probability distribution for the state of the world.

However, this distribution can be either a prior distribution or a posterior distribution upon obtaining some data.

ER criterion for finding a probabilistic rule:
Maximise the expected payoff with respect to a probability distribution for the state of the world.

EL criterion for finding a probabilistic rule:
Minimise the expected loss with respect to a probability distribution for the state of the world.

The *ER* criterion always leads to the same decision as the *EL* criterion and vice versa since (from previous derivation) $EL = T - ER$.

Exercise 5.3

Loss table:

	I	*II*	*III*
1	0	3	6
2	1	1	0
3	4	0	1

The Payoff table $\downarrow$ should be completed

	I	*II*	*III*
1	12	7	9
2			
3			

The relationship between Loss and Payoff for action i, state j, here L_{ij} and R_{ij} is

$$L_{ij} = \max_k R_{kj} - R_{ij} \Leftrightarrow R_{ij} = \max_k R_{kj} - L_{ij}$$

$$\Rightarrow \max_k R_{k1} - R_{11} = L_{11} = 0 \Rightarrow \max_k R_{k1} = R_{11} = 12\,;\ \max_k R_{k2} - R_{12} = L_{12} = 3 \Rightarrow \max_k R_{k2} = 3 + R_{12} = 10$$

$$\max_k R_{k3} - R_{13} = L_{13} = 6 \Rightarrow \max_k R_{k3} = 6 + R_{13} = 15$$

$\Rightarrow$The payoff table is

	I	*II*	*III*
1	12	7	9
2	12 – 1 = 11	10 – 1 = 9	15 – 0 = 15
3	12 – 4 = 8	10 – 0 = 10	15 – 1 = 14

$\Leftrightarrow$

	I	*II*	*III*
1	12	7	9
2	11	9	15
3	8	10	14

Using the loss table

	I	*II*	*III*
1	0	3	6
2	1	1	0
3	4	0	1

we see that no action gives a higher or equal loss compared to any of the other actions for all three states of the world. Hence, there is no inadmissible action.

Using the payoff table

	I	*II*	*III*
1	12	7	9
2	11	9	15
3	8	10	14

we see that no action gives a lower or equal payoff compared to any of the other actions for all three states of the world .