Decision Theory

Mário S. Alvim
(msalvim@dcc.ufmg.br)

Information Theory

DCC-UFMG
(2017/02)

Decision Theory

- Decision theory is a very simple concept:

  1. You have a choice of various **actions** $a \in \mathcal{A}$.
  2. The **world** may be in one of many **states** $x \in \mathcal{X}$; which one occurs may be in influenced by your action.

     The world's state has a **probability distribution** $p(x \mid a)$.
  3. There is a **utility function** $U(x, a)$ which specifies the payoff you receive when the world is in state $x \in \mathcal{X}$ and you choose action $a \in \mathcal{A}$.
  4. The task of decision theory is to select the action a that <u>maximizes the expected utility</u>

$$E[U \mid a] = \sum_{x} p(x \mid a) U(x, a).$$

- The computational problem is to maximize $E[U \mid a]$ over $a \in \mathcal{A}$.

- $\boxed{\text{Example 1}}$ **(Should you bet on the Lottery?)** Consider a lottery that randomly draws 6 numbers from a set of 60.

  A bet in this lottery consists in picking a subset of the set of 60 numbers, and there are two types of bets allowed:

  - picking a set of 6 numbers at a cost of \$3.50, or
  - picking a set of 7 numbers at a cost of \$24.50.

  The lottery distributes three types of mutually exclusive prizes:

  - \$23 221 167.26 for each bet that got all 6 numbers right ("Mega-Sena"),
  - \$51 651.09 for each bet that got exactly 5 numbers right ("Quina"), and
  - \$788.76 for each bet that got exactly 4 number right ("Quadra").

  (Coincidentally, the values in this lottery are the same as the ones from the "Mega Sena" lottery of October 03rd, 2015.)

Decision Theory - Should you bet on the Lottery?

- $\boxed{\text{Example 1}}$ (Continued)

The probabilities of bets getting prizes in this lottery are the following.

Bet \ Prize	Mega-Sena	Quina	Quadra
6 numbers	$\frac{1}{50\,063\,860}$	$\frac{1}{154\,518}$	$\frac{1}{2\,332}$
7 numbers	$\frac{1}{7\,151\,980}$	$\frac{1}{44\,981}$	$\frac{1}{1\,038}$

Assume that prizes are non-cumulative, so if no bet got 6, 5, or 4 numbers right, the money of the prizes stays with the lottery manager.

You have the option of picking a 6-number bet, picking a 7-number bet, or not betting at all on this Lottery.

If you want to maximize your expected gain, what is the best decision you can make?

Decision Theory - Should you bet on the Lottery?

- $\boxed{\text{Example 1}}$ (Continued)

Solution. Let's model this problem as a decision-theory problem.

The set of actions available to you is $\mathcal{A} = \{b_\emptyset, b_6, b_7\}$, where

- $b_\emptyset$ represents not betting at all,
- b_6 represents betting on 6 numbers, and
- b_7 represents betting on 7 numbers.

The set of states of the world is $\mathcal{X} = \{w_6, w_5, w_4, \ell\}$, where

- w_6 represents that your bet got all 6 numbers right,
- w_5 represents that your bet got exactly 5 numbers right,
- w_4 represents that your bet got exactly 4 numbers right, and
- ℓ represents that your bet got fewer than 4 numbers right (a loser's choice).

Decision Theory - Should you bet on the Lottery?

- $\boxed{\text{Example 1}}$ (Continued)

  Your job is to pick an action $a \in \mathcal{A}$ that maximizes the expectation

  $$E[U \mid a] = \sum_x p(x \mid a) U(a, x).$$

  Before we can compute that, we need to determine the probability $p(x \mid a)$ of the world being in state x if you take action a.

  These probabilities were given, and we here we convert them to decimals.

States of the World

	$p(x \mid a)$	w_6	w_5	w_4	ℓ
Actions	$b_\emptyset$	0	0	0	1
	b_6	0.00000002	0.00000647	0.00042882	0.99956469
	b_7	0.00000014	0.00002223	0.00096339	0.99901424

- Example 1 (Continued)

  We also need to determine the value $U(a, x)$ of the utility function describing your payoff when you take action a and the world turns up to be in state x.

  These utilities are calculated as the difference between the prize you get in each state minus the cost you incur for each action (every bet has a price).

  For instance, if you pick a 6-number bet you pay \$3.50 and if you get 4 numbers right you win \$788.76, so $U(b_6, w_4) = \$788.76 - \$3.50 = \$785.26$.

  The values of utility are shown in the next table.

States of the World

	$u(a, x)$	w_6	w_5	w_4	ℓ
Actions	$b_\emptyset$	\$23 221 167.26	\$51 651.09	\$788.76	\$0
	b_6	\$23 221 163.76	\$51 647.59	\$785.26	−\$3.50
	b_7	\$23 221 142.76	\$51 626.59	\$764.26	−\$24.50

- $\boxed{\text{Example 1}}$ (Continued)

Now we are ready to compute the expected utility

$$E[U \mid a] = \sum_{x} p(x \mid a) U(a, x)$$

of each possible action a.

$$\begin{aligned} E[U \mid b_\emptyset] &= p(w_6 \mid b_\emptyset) U(b_\emptyset, w_6) + p(w_5 \mid b_\emptyset) U(b_\emptyset, w_5) + \\ &\quad p(w_4 \mid b_\emptyset) U(b_\emptyset, w_4) + p(\ell \mid b_\emptyset) U(b_\emptyset, \ell) \\ &= 0 \cdot \$23\,221\,167.26 + 0 \cdot \$51\,651.09 + 0 \cdot \$788.76 + 1 \cdot \$0 \\ &= \$0. \end{aligned}$$

Decision Theory - Should you bet on the Lottery?

- $\boxed{\text{Example 1}}$ (Continued)

$$
\begin{aligned}
E[U \mid b_6] &= p(w_6 \mid b_6)U(b_6, w_6) + p(w_5 \mid b_6)U(b_6, w_5) + \\
&\quad p(w_4 \mid b_6)U(b_6, w_4) + p(\ell \mid b_6)U(b_6, \ell) \\
&= 0.00000002 \cdot \$23\,221\,163.76 + 0.00000647 \cdot \$51\,647.59 + \\
&\quad 0.00042882 \cdot \$785.26 + 0.99956469 \cdot (-\$3.50) \\
&= -\$2.36.
\end{aligned}
$$

$$
\begin{aligned}
E[U \mid b_7] &= p(w_6 \mid b_7)U(b_7, w_6) + p(w_5 \mid b_7)U(b_7, w_5) + \\
&\quad p(w_4 \mid b_7)U(b_7, w_4) + p(\ell \mid b_7)U(b_7, \ell) \\
&= 0.00000014 \cdot \$23\,221\,142.76 + 0.00002223 \cdot \$51\,626.59 + \\
&\quad 0.00096339 \cdot \$764.26 + 0.99901424 \cdot (-\$24.50) \\
&= -\$19.34.
\end{aligned}
$$

Decision Theory - Should you bet on the Lottery?

- $\boxed{\text{Example 1}}$ (Continued)

  As we can see:

  - if you pick a 6-number bet you are expected to lose \$2.36,
  - if you pick a 7-number bet you are expected to lose \$19.34,
  - if you don't bet at all you gain nothing and lose nothing.

  Therefore, the best decision is not to bet on the lottery at all!

  In fact, if you are worried about expected utility, not betting is always the best decision in real world lotteries (in which the manager is guaranteed not to lose money).

  However, many people would still be optimistic, and make decisions based on the best case utility, and not on the expected case of utility. We discuss such decision makers in the Appendix at the end of this lecture.

  ◁

Decision Theory - "To cheat or not to cheat?"

- $\boxed{\text{Example 2}}$ **(Good students don't cheat.)** Prof. Mário Alvim is applying a very reasonable Information Theory final exam today.

  A well-prepared student who has read the text-book, studied hard, and carefully done every homework assignment has a probability of 0.9 of answering any given question correctly.

  A badly-prepared student who has not worked hard enough[1] has a probability of 0.20 of answering any particular question correctly.

  The grading for this exam is binary:

  i) either a question is considered correct, granting the student 1 full mark, or

  ii) the question is considered incorrect, granting the student 0 marks.

  Let us reason about some aspects of this problem.

[1] The existence of such a student is only a hypothesis: every student works hard in this class!

- $\boxed{\text{Example 2}}$ (Continued)

Question A: What is the expected grade per question for a well-prepared student? And for a badly-prepared student?

Solution.

A well-prepared student has a probability of 0.9 of answering a question correctly, in which case they earn a grade of 1 mark, and a 0.1 probability of answering it wrongly, in which case they earn a grade of 0 marks.

Hence, the expected grade per question for the well-prepared student is $0.9 \cdot 1 + 0.1 \cdot 0 = 0.9$ marks.

As for the badly-prepared student, they have a probability of 0.2 of answering a question correctly, in which case they earn a grade of 1 mark, and a 0.8 probability of answering wrongly, in which case they earn a grade of 0 marks.

Hence, the expected grade per question for the badly-prepared student is $0.2 \cdot 1 + 0.8 \cdot 0 = 0.2$ marks.

Decision Theory - "To cheat or not to cheat?"

- $\boxed{\text{Example 2}}$ (Continued)

  Question B: Assume that a badly-prepared student comes up with a very "unorthodox" idea: to cheat on the exam!

  For each question, this badly-prepared student has two options:

  i) either answer the question honestly, with a probability of 0.2 of getting the answer correctly; or

  ii) to cheat, with a probability of 1 of getting the answer correctly (their method of cheating is very accurate).

  Assuming that cheating is safe (i.e., Mário is not paying attention to who is cheating and who is not), if the badly-prepared student cheats on a question, what's their expected grade per question?

  In this case, what is the badly-prepared student's best strategy (to cheat or not to cheat)?

- $\boxed{\text{Example 2}}$ (Continued)

Solution.

If the student doesn't cheat, their grade per question will be the same as we calculated in Question (A), that is, 0.20 marks.

If the student cheats, there is a probability 1 of answering the question correctly, thus earning 1 mark, and a probability 0 of answering the question wrongly, thus earning 0 marks. Hence, in this case their expected grade per question will be $1 \cdot 1 + 0 \cdot 0 = 1$ mark.

Therefore, if there is no chance of a cheating student being caught, the best strategy for the badly-prepared student is to cheat on every question.

- Example 2 (Continued)

Question C: Fortunately, Mário knows a thing or two about decision theory, and is ready to use it to discourage bad behavior on his students.

Mário introduces a system of punishments for students who are caught cheating as follows.

1. If a student is caught cheating on a question, the student doesn't get any marks for this particular question and, instead, gets a punishment of k marks subtracted from their grade.
2. If a student is not caught cheating on a question, the usual case applies, and they will be granted 1 mark for a correct answer, and 0 marks for a wrong answer.

- Example 2 (Continued)

Question C: (Continuation)

Assume that if a student is cheating on a question, there is a probability p that they are caught cheating on that particular question.

Assume also that if a student is not cheating, they won’t be wrongly accused of cheating.

Find a formula for the value for the minimum punishment k necessary so students will be completely discourage from cheating.

Write this formula as a function of the probability p of the student being caught cheating.

Decision Theory - "To cheat or not to cheat?"

- $\boxed{\text{Example 2}}$ (Continued)

  Solution.

  Let us model the problem as a decision theory problem.

  We will first identify the actions available to the student, the states of the world, and the probabilities and utilities of each pair of state/action.

  Then we will find a utility function in terms of k that, for a badly-prepared student, makes the expected utility of being honest surpass the expected utility of cheating.

- $\boxed{\text{Example 2}}$ (Continued)

  For each question, the set of actions available for the badly-prepared student is $\mathcal{A} = \{h, \overline{h}\}$, where

  - h stands for answering the question honestly, and
  - $\overline{h}$ stands for being dishonest and cheating on that particular question

  For each question, the state of the world is given by a pair of values:

  - the first element in the pair indicates whether the student was caught cheating (c) or not caught ($\overline{c}$), and
  - the second element in the pair indicates whether the student got the answer right (r) or not ($\overline{r}$).

  Hence, each state of the world is a tuple, and the set of all possible states is $\mathcal{X} = \{(r, c), (r, \overline{c}), (\overline{r}, c), (\overline{r}, \overline{c})\}$.

- $\boxed{\text{Example 2}}$ (Continued)

  We need to calculate the conditional probability $p(x \mid a)$ of each state x given an action a.

  Let us assume that for any student, the probability of cheating and the probability of getting the answer correctly are independent, given that they are honest or dishonest.

  Then, if the student is honest, we have

  - $p((r,c) \mid h) = p(r \mid h) \cdot p(c \mid h) = 0.2 \cdot 0 = 0,$
  - $p((r,\overline{c}) \mid h) = p(r \mid h) \cdot p(\overline{c} \mid h) = 0.2 \cdot 1 = 0.2,$
  - $p((\overline{r},c) \mid h) = p(\overline{r} \mid h) \cdot p(c \mid h) = 0.8 \cdot 0 = 0,$ and
  - $p((\overline{r},\overline{c}) \mid h) = p(\overline{r} \mid h) \cdot p(\overline{c} \mid h) = 0.8 \cdot 1 = 0.8.$

Decision Theory - "To cheat or not to cheat?"

- $\boxed{\text{Example 2}}$ (Continued)

  On the other hand, if the student is dishonest we have

  - $p((r, c) \mid \overline{h}) = p(r \mid \overline{h}) \cdot p(c \mid \overline{h}) = 1 \cdot p = p,$
  - $p((r, \overline{c}) \mid \overline{h}) = p(r \mid \overline{h}) \cdot p(\overline{c} \mid \overline{h}) = 1 \cdot (1 - p) = 1 - p,$
  - $p((\overline{r}, c) \mid \overline{h}) = p(\overline{r} \mid \overline{h}) \cdot p(c \mid \overline{h}) = 0 \cdot p = 0$, and
  - $p((\overline{r}, \overline{c}) \mid \overline{h}) = p(\overline{r} \mid \overline{h}) \cdot p(\overline{c} \mid \overline{h}) = 0 \cdot (1 - p) = 0.$

  These probabilities are summarized in the table below.

States of the World

	$p(x \mid a)$	(r, c)	$(r, \overline{c})$	$(\overline{r}, c)$	$(\overline{r}, \overline{c})$
Actions	h	0	0.2	0	0.8
	$\overline{h}$	p	$1 - p$	0	0

- $\boxed{\text{Example 2}}$ (Continued)

  As for the utility, note that it is:

  - $-k$ whenever the student is caught cheating,
  - 1 whenever the student is not caught cheating and the answer is correct, and
  - 0 whenever the student is not caught cheating and the answer is wrong.

  The following table presents the values $U(a, x)$ for the utility function.

		States of the World			
	$U(a, x)$	(r, c)	$(r, \overline{c})$	$(\overline{r}, c)$	$(\overline{r}, \overline{c})$
Actions	h	$-k$	1	$-k$	0
	$\overline{h}$	$-k$	1	$-k$	0

- $\boxed{\text{Example 2}}$ (Continued)

  Now we can calculate the expected utility of each action the student can take.

  If the student decides to be honest, the expected utility is

$$
\begin{aligned}
E[U \mid h] &= \sum_{x \in \mathcal{X}} p(x \mid h) \cdot u(h, x) \\
&= p((r, c) \mid h) \cdot u(h, (r, c)) + p((r, \overline{c}) \mid h) \cdot u(h, (r, \overline{c})) + \\
&\quad p((\overline{r}, c) \mid h) \cdot u(h, (\overline{r}, c)) + p((\overline{r}, \overline{c}) \mid h) \cdot u(h, (\overline{r}, \overline{c})) \\
&= 0 \cdot (-k) + 0.2 \cdot 1 + 0 \cdot (-k) + 0.8 \cdot 0 \\
&= 0.2.
\end{aligned}
$$

- $\boxed{\text{Example 2}}$ (Continued)

If the student decides to be dishonest, the expected utility is

$$
\begin{aligned}
E[U \mid \overline{h}] &= \sum_{x \in \mathcal{X}} p(x \mid \overline{h}) \cdot u(\overline{h}, x) \\
&= p((r,c) \mid \overline{h}) \cdot u(\overline{h}, (r,c)) + p((r,\overline{c}) \mid \overline{h}) \cdot u(\overline{h}, (r,\overline{c})) + \\
&\quad\; p((\overline{r},c) \mid \overline{h}) \cdot u(\overline{h}, (\overline{r},c)) + p((\overline{r},\overline{c}) \mid \overline{h}) \cdot u(\overline{h}, (\overline{r},\overline{c})) \\
&= p \cdot (-k) + (1-p) \cdot 1 + 0 \cdot (-k) + 0 \cdot 0 \\
&= 1 - kp - p.
\end{aligned}
$$

Decision Theory - "To cheat or not to cheat?"

- $\boxed{\text{Example 2}}$ (Continued)

  Since want to discourage cheating, we must enforce

  $$E[U \mid h] > E[U \mid \overline{h}],$$

  which means that we must have

  $$0.2 > 1 - kp - p,$$

  which implies that

  $$k > \frac{0.8 - p}{p}.$$

  Hence, to discourage cheating, the punishment k should be, at least, $(0.8 - p)/p$ marks per question to any student caught cheating.

- $\boxed{\text{Example 2}}$ (Continued)

  Question D: Using the result from the previous item, calculate the minimum value of the punishment k when $p = 0$, $p = 0.2$, $p = 0.5$, $p = 0.8$, and $p = 1$.

  Analyze whether the values found make sense to you, explaining each of them.

Decision Theory - "To cheat or not to cheat?"

- $\boxed{\text{Example 2}}$ (Continued)

Solution.

The minimum value of the punishment k for each probability p of being caught cheating are shown in the following table.

p	0	0.2	0.5	0.8	1
k	∞	3	0.6	0	-0.2

These minimum values of the punishment k for each probability p of being caught can be interpreted as follows.

- If the student is never caught cheating ($p = 0$), it would be necessary an infinite punishment to discourage them from cheating.

  (That's a decision-theoretic explanation of why hell is usually portrayed as infinite punishment: people in general tend to find the concept of hell quite implausible.)

Decision Theory - "To cheat or not to cheat?"

- $\boxed{\text{Example 2}}$ (Continued)

  - For the values $p = 0.2$ and $p = 0.5$ the minimum k is not surprising: the more likely it is that someone will be caught, the smaller the punishment necessary (respectively $k = 3$ and $k = 0.6$ marks).

  - For $p = 0.8$, it is enough to punish simply by not adding or subtracting any marks from the student $k = 0$.

    This is because this is a point of equilibrium: the advantage of cheating is completely canceled by the risk of being caught, hence no negative punishment is necessary.

Decision Theory - "To cheat or not to cheat?"

- $\boxed{\text{Example 2}}$ (Continued)

  - When $p = 1$, the student will be certainly caught if cheating, and any punishment $k > -0.2$ is enough to discourage cheating.

    To see why, notice that if the student is caught cheating and we subtract a punishment $k > -0.2$, we are indeed adding to his grade less than 0.2 marks per question.

    However, if the student didn't cheat, they'd get an expected addition of 0.2 marks per question, so cheating does not pay off!

    Hence, if it is certain that a student will be caught cheating, it is enough to punish them by simply giving them a smaller grade than they'd be expected to get when being honest.

$\triangleleft$

Appendix - Alternative types of decision makers

Alternative types of decision makers

- We formalized the decision problem in the case the decision maker is worried about to maximizing the expected value of utility over all possible states of the world.

- But in many cases people don't make decisions based on the expected value of utility, but rather on the the best case (e.g., the lottery), or worst-case utility, for instance.

  There are various formulations of the decision-making problem depending on the type of decision maker.

Alternative types of decision makers

- Alternative types of decision makers:

  1. **Expected Value (Realist)**

     This decision maker computes the expected value under each action and then picks the action with the largest expected value.

     This is the only decision maker of the four that incorporates the probabilities of all possible states of nature.

     The expected value criterion is also called the **Bayesian principle**.

Alternative types of decision makers

- Alternative types of decision makers:

  2. **Maximax (Optimist)**

     The maximax decision maker looks at the best that could happen under each action and then chooses the action with the largest value.

     They assume that they will get the most possible and then they take the action with the best case scenario. They are guided by the **maxi**mum of the **max**imums, or the "best of the best".

     These are the "lotto players"; they seek large payoffs and ignore probabilities (as long as they are at least non-zero).

Alternative types of decision makers

- Alternative types of decision makers:

3. **Maximin (Pessimist)**

   The maximin decision maker looks at the worst that could happen under each action and then choose the action with the largest payoff.

   They assume that the worst that can happen will, and then they take the action with the best worst case scenario. They are guided by the **maxi**mum of the **min**imums, or the "best of the worst".

   These are the people who put their money into a savings account because they could lose money at the stock market.

- Alternative types of decision makers:

4. **Minimax (Opportunist)**

   Minimax decision making is based on opportunistic loss.

   These decision makers are the kind that look back after the state of nature has occurred and say *"Now that I know what happened, if I had only picked this other action instead of the one I actually did, I could have done better"*.

   So, to make their decision (before the event occurs), they create an opportunistic loss (or regret) table. Then they take the **mini**mum of the **max**imum.

   That sounds backwards, but remember, this is a loss table. This similar to the maximin principle in theory; they want the best of the worst losses.