

Decision Theory

I. Some Basics

The notion of expected value: decision theory helps you make decisions by telling you to choose the act with the greatest expected value. What is the expected value of some act A? Consider the different outcomes that could come about if you do A: $O_1 \dots O_n$. Now consider how good or bad these outcomes are: $V(O_1), V(O_2) \dots V(O_n)$. (Note that to do this we need to assign numbers to the values different outcomes – the values of outcomes are measured in utiles. The more utiles something has, the better it is.) Finally, consider how likely these outcomes are given that you do A. $Pr(O_1), Pr(O_2) \dots Pr(O_n)$.

$$EV(A) = Pr(O_1)V(O_1) + Pr(O_2)V(O_2) \dots + Pr(O_n)V(O_n)$$

Example:

I'm trying to decide whether to go to order dessert at a fancy restaurant. It's a surprise dessert and I only like chocolate desserts. I think there's a 70% chance that it will be a chocolate dessert, and ordering the surprise dessert costs 10 dollars. What should I do?

(1) There are four outcomes:

- O1- I order and it's chocolate (Yay – yummy dessert, though I'm down ten dollars).
- O2 – I order and it's not chocolate (I'm a little bummed – because I'm down ten dollars)
- O3 – I don't order dessert and the dessert is a chocolate dessert (I'm really bummed that I didn't order it!)
- O4 – I don't order dessert and the dessert is not a chocolate dessert (I'm glad I didn't order it!)

(2) Make a decision table: (the numbers in the chart are made up, you could change them around)

	Dessert is Chocolate (.7)	Dessert is Not Chocolate (.3)
Order Dessert	100-10 = 90 outcome 1	-10 outcome 2
Don't Order Dessert	-15 outcome 3	5 outcome 4

To calculate the expected value of ordering dessert, consider the two outcomes that could come about if I order dessert: O1 and O2.

$$EV(\text{order dessert}) = Pr(O_1)V(O_1) + Pr(O_2)V(O_2) = (0.7)(90) + (-10)(0.3) = 60$$

To calculate the expected value of not ordering dessert, consider the two outcomes that could about if I don't order dessert O3 and O4.

$$EV(\text{don't order dessert}) = Pr(O_3)V(O_3) + Pr(O_4)V(O_4) = (0.7)(-15) + (0.3)(5) = -9$$

Since $60 > -9$, I should order dessert.

II. Pascal's Wager

Suppose you have the following decision table.

	God exists (0.000000...1)	God doesn't exist (0.99999999..)
Believe	∞	-100
Don't Believe	-1,000,000	100

If we use expected value theory we get the result that the practically rational thing to do is believe in God. Of course, it's not that simple because we don't have voluntary control over what we believe. However, we can argue that we should do things that will increase our chances of believing (like going to church, behaving like a believer, etc).

III. The Argument for Behaving like a Believer

1. The practically rational thing to do is the act with the highest expected value. (Assp)
2. Behaving like a believer will increase your chances of believing. (Assp)
3. The expected value of believing is greater than the expected value of not believing.
4. The expected value of behaving like a believer is higher than that of not. (2,3)
5. Behaving like a believer and not behaving like a believer are the only options. (Assp)
6. The practically rational thing to do is behave like a believer .(1,4,5)

IV. Objections

Denying (1): We can deny that it's practically rational to do the thing with the highest expected value.

Denying (3): Most ways to object to the argument involve denying the second premise by objecting to the decision matrix.

- (a) Probability of God existing is 0. Problem – this is a very strong claim. Though if you think the notion of God is incoherent (for example, because of the problem of evil) – this might be a possibility.
- (b) There can be no infinite utilities. Whether this is true is up for debate.
- (c) There are not enough outcomes – maybe the Christian God is shy rewards infinite utility to those who believe he doesn't exist, or a purple loving god who rewards infinite utility to those who spend their entire life painting the world purple, or Greek gods, the Islamic god, Hindu gods, Jewish God etc.

	Non-shy God Exists (0.09)	Shy God Exists (0.01)	No God Exists (0.9)
Believe	∞	-1,000,000	-100
Don't believe	-1,000,000	∞	100

Now we have no reason to try and act in ways that will make us believe in God.

Denying that (4) follows from (2) and (3): There's some chance that your action will backfire and that by deciding to behave like a believer you'll end up less likely to believe than before.

Denying (5): Another option is to flip a coin to decide whether you will behave like a believer. This has the same infinite utility as deciding to behave like a believer.

MIT OpenCourseWare
<http://ocw.mit.edu>

24.00 Problems in Philosophy
Fall 2010

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.