

Preferences, Rationality, and Expected Utility

1 Preferences and Rationality

- Having preferences means being able to rank alternatives with respect to one another. Being able to rank them means answering the question “Which alternative do I like better?” Preference notation is intuitively just the same as mathematical relational operators that you are all familiar with.

Preference	Meaning	Math
$A \succ B$	A is strictly preferred to B	$A > B$
$A \succeq B$	A is at least as good as B	$A \geq B$
$A \sim B$	indifferent between A and B	$A = B$

- A **preference ordering** is simply one ranking of the alternatives. An actor has one preference ordering and the actor is **rational** if this ordering meets the following two conditions:
 1. **completeness**: actors can compare *all* alternatives available to them.
 2. **transitivity**: the ranking does not contradict logic. For example, you cannot claim that you strictly prefer A to B , strictly prefer B to C , and strictly prefer C to A . Suppose an actor had the following intransitive preferences: $A \succ B \succ C \succ A$. This is saying that $A \succ C$ and $C \succ A$, a contradiction. This would not be rational. It's the same as saying in math that $3 > 2 > 1 > 3$, clearly nonsense.
- AN EXAMPLE. Suppose there are four available alternatives with respect to Iraq:
 - attack Iraq unilaterally (U),
 - attack Iraq as head of global coalition (C),
 - support more UN inspections (I),
 - assassinate Saddam (A).

Examine the following preference orderings:

- (a) $U \succ C$, $I \succeq C$, $C \sim A$. These imply $U \succ C \sim A$ and $I \succeq C \sim A$. There are two ways to combine these that meet all the stated conditions, depending on where you put U and I with respect to each other. But we don't know which, and so this preference ordering is incomplete. We must be able to rank U and I .

- (b) same as above and $U \succ I$. We make a chain and get $U \succ I \succeq C \sim A$, which is consistent with the other condition $U \succ C$, and so this ordering is rational.
 - (c) same as first but $I \succ U$. This implies $I \succ U \succ C \sim A$. If this ordering is transitive, we must have $I \succ C$, which contradicts $I \succeq C$, and so the preference ordering is not rational. The problem is that $I \succeq C$ means $I \succ C$ or $I \sim C$, while $I \succ C$ excludes the second possibility.
 - (d) $C \sim U$, $C \sim I$, $I \succ A$, $I \sim U$, $C \succ A$, $U \succ A$. This implies $C \sim U \sim I \succ A$. Both complete and transitive, and so rational.
- If we assume that actors are irrational (preference orderings are either incomplete or intransitive), then anything goes. Any sort of behavior can be rationalized by some set of preferences no matter how ridiculous these preferences are. But this means we won't be able to explain anything because whenever we want to explain why some actor chose A instead of B , we would simply say that it did so because it preferred A to B . If the preferences are irrational (that is, they make no sense), then the next time it chooses B instead of A , we can just as correctly claim that it did so because it preferred B to A . There will be no contradiction here since the actor is irrational. This is the sort of "explanation" we want to avoid. Therefore, we make several assumptions about individual preferences.

2 Expected Utility

- Preference orderings are ordinal. In order to use mathematics, we want to work with cardinal numbers.
- When people say that they prefer one alternative to another, they mean that the **utility** (satisfaction) they get from the first is better than the utility they get from the second. If an actor strictly prefers A to B , it means that its utility from obtaining A is strictly higher than his utility from obtaining B . When it says that it is indifferent between B and C , it means that its utility from obtaining B is the same as the utility from obtaining C .
- The **utility function** of an actor reproduces this actor's preference ordering with real numbers. It assigns some number to each alternative in such a way that the rank ordering of the different alternatives is preserved.

2.1 An Example: Attacking Iraq

Suppose there are two ways of the United States attacking Iraq, unilaterally and as a leader of a coalition. Suppose now that you are a decision maker who is interested in the outcome of a war with Iraq. Naturally, you prefer the United States to win rather than to lose. You have two choices, attack alone, A , or attack with a coalition of other allies, O . Which one do you choose?

The answer, obviously, depends on many things. For example, you want to know the costs and benefits associated with victory given that you attack alone as well as the costs and benefits of victory if you attack with a coalition. You will want to know the costs of defeat (even if you are quite certain that the US will win either way). You also want to know the probability of success in both cases.

Let's say that the benefit of winning alone is $b_v(A) = 100$ (you can think of this number as \$100 billion, the utility from obtaining changing policies in Iraq). The cost of fighting along is $c_v(A) = 50$ (again, you can think of this in terms of billions of dollars). The corresponding numbers for a coalitional war are $b_v(O) = 70$ (because you must share in the benefits) and $c_v(O) = 35$ (because they will pay some of the cost). So, the utility of obtaining victory alone is

$$u_v(A) = b_v(A) - c_v(A) = 100 - 50 = 50,$$

and the utility of victory with a coalition is

$$u_v(O) = b_v(O) - c_v(O) = 70 - 35 = 35.$$

The utility of victory with a coalition is lower, so you might be tempted to conclude that this means you should choose go it alone instead of bothering with a coalition.

But this would be wrong because our analysis is incomplete. We also want to know the results of defeat. Let's suppose there are no benefits for the US from being defeated by Iraq, and so $b_d(A) = b_d(O) = 0$. The war is still costly, and for simplicity we shall assume that the costs are the same as in case of victory. This means that the utility of being defeated when fighting by itself is

$$u_d(A) = b_d(A) - c_d(A) = 0 - 50 = -50,$$

and the utility of being defeated as part of a coalition is

$$u_d(O) = b_d(O) - c_d(O) = 0 - 35 = -35.$$

Since the second utility is higher, you conclude that being defeated as a member of a coalition is better than being defeated while fighting alone. In this case it is better to form a coalition than fight alone.

So what do you choose? Well, you must form an estimate about your probability of prevailing if you fight along and if you fight with a coalition. Suppose the probability that the US will win if it fights alone is $p_a = .8$, and so the probability that it will lose is $1 - p_a = .2$. That is there is an 80% chance that the US will win the war if it goes it alone. Suppose further that it has a 90% chance if it goes in with a coalition (more troops and more money), that is $p_o = .9$. The probability of defeat in this case is $1 - p_o = .1$, or only 10%.

We can now calculate the **expected utility** from attacking alone:

$$EU(A) = p_a \times u_v(A) + (1 - p_a) \times u_d(A).$$

That is, the expected utility from attacking alone equals the probability of victory times the utility of victory plus the probability of defeat times to utility of defeat. Note that we used

the probabilities and utilities for fighting alone as we specified them above. Substituting the numbers from before, we get

$$EU(A) = (.8)(50) + (.2)(-50) = 40 - 10 = 30.$$

Similarly, can calculate the expected utility from attacking as a member of a coalition:

$$EU(O) = p_o \times u_v(O) + (1 - p_o) \times u_d(O).$$

Again, the expected utility from attacking as a coalition equals the probability of the coalition winning times the utility of victory plus the probability of the coalition losing times the utility of defeat. Substituting the numbers, we get

$$EU(O) = (.9)(35) + (.1)(-35) = 31.5 - 3.5 = 28.$$

Since

$$EU(A) > EU(O),$$

you expect to do better by attacking alone. Therefore, your decision *given the values of the various variables* is to attack alone and not bother with a coalition. Even though the probability of winning with a coalition is higher, and even though the utility of defeat with a coalition is also higher, it is still better to go it alone.

You could, of course, imagine a different configuration of variables that might lead to a different decision. Suppose that the costs of going it alone are even higher because the US will have to bear the opposition of its allies, and so $c_v(A) = 60$ instead of 50. Now, $u_v(A) = 100 - 60 = 40$ and $u_d(A) = -60$. We don't change any of the other variables, and so $u_v(O) = 35$ and $u_d(O) = -35$ as before. Now, however,

$$EU(A) = (.8)(40) + (.2)(-60) = 32 - 12 = 20,$$

and so

$$EU(A) < EU(O).$$

In this case, you would chose to form a coalition instead of attacking.

This nice illustration shows you how important it is to consider carefully the various factors that go into a decision. Also, a relatively innocuous change from 50 to 60 of one variable led to a complete reversal of the decision. Would you have made this choice without the little algebra? Possibly, but not probably.

What we have just done is called an **expected utility calculation** and it will be a major component of what we are going to be doing for the rest of the course. This is the extent of mathematics that you will need. If you are able to handle the multiplication and addition, you can do everything else. Again, it is not the math that's likely to cause you trouble, it getting the logic straight.

2.2 Probabilities and Critical Values

The probability of event A is the likelihood that event A will occur. A probability is a number between 0 and 1. That is, either an event does not occur for sure, in which case the probability is 0, or occurs for sure, in which case its probability is 1, or maybe occurs, in which case its probability is somewhere between 0 and 1.

You can think about probabilities as percentages if that will make it easier. Saying that event A occurs with probability .75 is equivalent to saying that the chance of event A occurring is 75%. We shall stick to the decimal notation although it may sometimes be easier to represent probabilities as fractions. For example, a probability of $\frac{1}{3}$ cannot be represented accurately in decimal form, and so we shall have to use the fractional form instead.

The first thing to remember about probabilities is that they are always real numbers between 0 and 1.

The second thing to remember about probabilities is that the probabilities of all mutually exclusive and exhaustive events must sum exactly to 1. That is, if you list all events that could occur, making sure that if one of them occurs another cannot, and you then sum the probabilities, you should get 1.

In our example, a war with Iraq could end either in victory or defeat. Suppose we assume it can end in victory, defeat, or stalemate. These three outcomes are exhaustive (there are no other possibilities) and exclusive (if one occurs, none of the other two can occur as well). We must then be careful when assigning probabilities.

Let $p_v = .8$, $p_d = .05$, and $p_s = .15$. This is a valid set of probabilities because $p_v + p_d + p_s = .8 + .05 + .15 = 1$, as required. This basically says that if we list all events, then at least one of them must occur. Also, each individual probability is between 0 and 1.

The set $p_v = .8$, $p_d = .1$, and $p_s = .15$ is not valid because while the individual probabilities are between 0 and 1, and so valid by themselves, their total is 1.05, which is not a valid probability.

Note that in the example of expected utility calculation where we had two outcomes only — victory and defeat — we only specified the probability of victory, p_a if alone and p_o if in a coalition. The probability of defeat was simply 1 minus the probability of victory. That is, we used the fact that victory and defeat are mutually exclusive and, given our assumptions, also exhaustive. That is, their probabilities must sum to 1. And so, the probability of defeat if fighting alone is $1 - p_a$, and the probability of defeat if fighting in a coalition is $1 - p_o$.

To sum it up, a valid probability is a number between 0 and 1. The sum of probabilities must also be a valid probability. If the events are exhaustive and mutually exclusive, their probabilities must sum to 1. We shall always use exhaustive and mutually exclusive events, so we must always make sure their probabilities sum up to 1.

Let's go back to our example from above. Suppose we want to know, given the utilities of the outcomes, what our chances of success must be for us to decide to attack alone. That is:

- $u_v(A) = 40$ and $u_d(A) = -60$
- $u_v(O) = 35$ and $u_d(O) = -35$, with $p_o = .9$; which means $EU(O) = 28$

We want to know what value of p_a will make it rational to attack alone. Attacking alone is the rational choice whenever the expected utility of doing so exceeds the expected utility of attacking with others:

$$EU(A) \geq EU(O).$$

We now make the appropriate substitutions:

$$\begin{aligned} p_a u_v(A) + (1 - p_a) u_d(A) &\geq EU(O) \\ p_a u_v(A) + u_d(A) - p_a u_d(A) &\geq EU(O) \\ p_a [u_v(A) - u_d(A)] &\geq EU(O) - u_d(A) \end{aligned}$$

solving for p_a , we get

$$p_a \geq \frac{EU(O) - u_d(A)}{u_v(A) - u_d(A)}$$

and, substituting the numbers, we get

$$p_a \geq \frac{28 - (-60)}{40 - (-60)} = .88$$

That is, whenever $p_a \geq .88$, the expected utility of attacking alone will exceed the expected utility of attacking with others. Remember our conclusion that the US would not attack alone when $p_a = .80$? It is obviously correct because this probability does not meet the necessary condition for lone attack. We shall call $p = .88$ the **critical value** for attacking alone because this is the value of p that separates the two decisions. Of course, this value depends on the precise values of the other variables.

2.3 Expected Utility: The General Case

How do you calculate expected utilities if you have more than two outcomes? Again, you will need to know the utility for each outcome and the probability of this outcome occurring.

Suppose you have n outcomes labeled $o_1, o_2, \dots, o_n$. You will need to know the utilities $u(o_1), u(o_2), \dots, u(o_n)$. You will also need to know the probability of each outcome actually occurring. That is, you need to know $p_1, p_2, \dots, p_n$. (We require that each p is a valid probability, and so it must be a real number between 0 and 1. We also require that the sum of all p s equals exactly 1.)

Given that you know the utilities of all outcomes and the associated probabilities, you simply multiply each utility by the corresponding probability and sum over all the products:

$$EU = u(o_1)p_1 + u(o_2)p_2 + \dots + u(o_n)p_n.$$

For example, suppose there are three outcomes, o_1, o_2 , and o_3 and the actor's utilities are $u(o_1) = 2, u(o_2) = 4$, and $u(o_3) = 7$. There are two possible actions, A and B . Action A has probabilities $p_1 = .5, p_2 = .2$, and $p_3 = .3$, while action B has probabilities $p_1 = .1$,

$p_2 = .7$, and $p_3 = .2$. (Again, these are all valid probabilities and they all sum to 1.) We then have

$$EU(A) = (.5)(2) + (.2)(4) + (.3)(7) = 1 + .8 + 2.1 = 3.9$$

$$EU(B) = (.1)(2) + (.7)(4) + (.2)(7) = .2 + 2.8 + 1.4 = 4.4$$

and so, according to the expected utility calculation, we should choose action B because it yields a higher expected utility.

Most of what we shall do in this course would involve expected utilities and similar calculations. In these examples we only had one actor making a decision. This is called **decision-theoretic** analysis. However, as I told you, we will be mostly interested in the interactions of several actors that all make expected utility calculations. That is, we will be interested in **game-theoretic** analysis. You will see, probably to your surprise, that once we start accounting for the behavior of other actors, the game-theoretic actions will be very different from the decision-theoretic ones. That is, sometimes actors will not choose the best action that a decision-theoretic model would suggest. That is, they will sometimes not choose their most preferred alternative, but instead go for the best alternative given what others are doing.