A Gentler Guide to
"Indifference, Indecision, and Coin-Flipping"

by Daniel Kian Mc Kiernan

Preferences

In the ordinary presentations of the foundations of decision theory, one starts with *preference relations*. There are two common approaches.

One starts with *two* relations, *strict preference*, and *indifference*. *<u>Strict</u> preference* corresponds to the common intuïtion of preference;

$$X_1 \succ X_2$$

means that X_1 is simply preferred to X_2.[1] The other relation, *indifference*, is typically conceptualized as one in which two things are regarded as *equally good*. This is usually written with the symbol " $\sim$ ":

$$X_1 \sim X_2$$

Then we have the relation of *<u>weak</u> preference*.

$$X_1 \succcurlyeq X_2$$

means that X_2 isn't preferred to X_1, but X_1 might-or-might-not be preferred to X_2.[2]

The other normal way of working with preference relations is to *start* with *weak preference*, then defining *strict preference* and *indifference* in terms of *weak preference*:

$$(X_1 \sim X_2) \equiv \left[(X_1 \succcurlyeq X_2) \wedge (X_2 \succcurlyeq X_1)\right]$$

says that X_1 is *indifferent* to X_2 if each is *weakly preferred* to the other. And

---

1When, in ordinary language, one says something such as "I prefer pretzels to bagels" a *context* is implied. A person who *ordinarily* prefers pretzels to bagels may really wish that she had a bagel in some contexts. The way that an economist scoops this up is by having the X_i refer to *states of the world*. So that, for example X_1 might be an otherwise ordinary state of the world in which one had a pretzel, and X_2 would be a state of the world as close as possible to the other except that one had a bagel, and there could be various other X_j in which things were less ordinary and one variously had pretzels, bagels, or neither.

2 Formally,

$$(X_1 \succcurlyeq X_2) \equiv \left[(X_1 \succ X_2) \vee (X_1 \sim X_2)\right]$$

$$\left(X_1 \succ X_2\right) \equiv \left[\left(X_1 \succcurlyeq X_2\right) \wedge \overline{X_2 \succcurlyeq X_1}\right]$$

says that X_1 is *strictly preferred* to X_2 if it's *weakly preferred* to X_2 while X_2 is not *weakly preferred* to X_1. It turns-out that this is a really *mathematically elegant* way of doing things, though the base relation (*weak preference*) is more removed from our everyday basic ideas.

Anyway, the two approaches are equivalent.

These relations have some standard properties. For example, *strict preference* is *irreflexive* (nothing is *strictly preferred* to itself), *antisymmetric* (if X_1 is *strictly preferred* to X_2 then X_2 is not *strictly preferred* to X_1), and *transitive* (if X_1 is *strictly preferred* to X_2 and X_2 is *strictly preferred* to X_3, then X_1 is *strictly preferred* to X_3). *Indifference* is *reflexive* (everything is *indifferent* to itself), *symmetrical* (if X_1 is *indifferent* to X_2 then X_2 is *indifferent* to X_1), and also *transitive*.

Notice how, with these properties, the relations *strict preference* and *indifference* seem a lot like the familiar *greater-than* and *equal-to* relations, and *weak preference* is a lot like *greater-than-or-equal-to*. Add another proposition, and the similarity increases; that proposition is that, for every pair of X_i and X_j, either *strict preference* or *indifference* obtains. Another way of putting this proposition is to say that *weak preference* is a *total ordering* of the choices.

Finding a Difference from Indifference

What would it *mean* if there were X_1 and X_2 such that neither was preferred to the other, nor were they viewed as *equally* good? A common intuïtion is that stating that two things are *equally good* is a stronger claim than that neither is preferred to the other. A person might simply not have *made up his mind*.

If we just *ask* people, then they may tell us that they haven't made up their minds about some things, and that they regard some other things as equally good, but do these *declarations* actually correspond to otherwise different behavior? Does a person who is *indecisive* really reach a different choice from that of a person who is *indifferent*? Many economists have had trouble finding an actual difference in choice.

For example, let's say that the choice is between 12-ounce cans of Coca-Cola and Pepsi Cola. If Jane thinks that these are equally good, and I say “I'll give you a Coke unless you'd rather have a Pepsi” then she has no reason to speak-up, and is given a Coke. If Richard is unable to decide which is best, and I say the same thing to him, then he likewise ends-up with a Coke. If I'd put things the other way around, then each would have ended-up with the Pepsi.

What about when the cans are placed before them, and these two are told to choose one? If Jane truly thinks that these two are equal, then she is simply held paralyzed in equal attraction between them; it is only by finding *some* difference, however tiny, that she can be pulled to one. And, so long as Richard is unable to decide which (if either) is somehow best, he is unable to decide which to choose.

Now, here many people want to say “flip a coin”. There's an important thought there, though it doesn't resolve the puzzle as presented, because we didn't offer these two a choice between a Coke, a Pepsi, and a lottery which would pay either a Coke or a Pepsi (based upon the flip of a coin).

Another problem, from the stand-point of *ordinary* economic theory, is that the coin-flip should have exactly the same value for Jane as do the Coke and the Pepsi, so that offering the coin-flip to Jane should have her paralyzed amongst *three* equally valuable options.

The only way to hold onto ordinary theory about the value of lotteries amongst indifferent alternatives is to insist that rational people will just *stay paralyzed* in the face of *indifference*; and if *indifference* must always produce the same behavior as *indecision*, then people must always just stay paralyzed whenever they have no strict preference amongst outcomes.

The Ideas and Objectives of My Paper

But, really, some times people *do* flip coins, or something like that; and some times they just fail to make any decision (even the decision to flip a coin).

My thought is that we should acknowledge *three* relations. One, *strict preference*, should be recognizable as the relation in ordinary theory. Another, which I call “*equi-indifference*” (“*equality-indifference*”), should have *mathematical* properties like those of *indifference* in ordinary theory; which properties, however, are not consonant with talk of “flipping a coin” to make a decision. The *third* relation, which I call “*undecidedness*”, is that where people *do* “flip a coin” to make a choice.

(Anthony Gamst once noted, though, that under *equi-indifference* a decision is never reached, while under what I've called “*undecidedness*” a decision is in fact made, albeït by choosing an outside mechanism of choice.)

I formally define these three relations in terms of fairly observable behavior. I also make some explicit propositions about choices. I show that the application of these propositions to the relations so defined implies properties for the relations that one would expect if *strict preference* and *indifference* didn't totally order the choices, and that are a fair fit to common intuïtions about indecision.

Choice Functions

To properly formalize and explore these relations, I had to back away from thinking and writing in terms of *preferences* to thinking and writing in terms of *choice functions*.

The idea of a choice function is really simple: given a set of possibilities, the choice function $C(\;)$ selects a subset. If only one thing is selected, this behavior corresponds to what we ordinarily interpret as a person *preferring* one option to all others, though choice functions allow for behaviors such as always choosing X_1 when the options are X_1, X_2, and X_3, but always choosing X_2 when the options are are X_1, X_2, and X_4, which really doesn't fit a simple notion of *preference*.

That example (of different choices between X_1 and X_2 depending upon what else was offered) illustrates that choice functions allow both for greater generality and for greater *observability*. To say that X_1 is *strictly preferred* to X_2 is to say that X_2 will never be chosen when X_1 is an option. That could only be observed by trying all possible combinations. On the other hand,

$$\{X_1\} = C(\{X_1, X_2, X_3\})$$

is a more modest claim, observed or falsified by testing just one set of options. (Granted that the choice function of any individual might change.)

Going back to the Coke and Pepsi example, Jane's choice function is

$$\{X_{Coke}, X_{Pepsi}\} = C(\{X_{Coke}, X_{Pepsi}\})$$

(where the X_i are the same except than in one state she has a Coke, and in the other a Pepsi) which is observed as her paralysis.

But choice functions aren't going to be perfectly observable. For example, if Richard is offered a Coca-Cola, a Pepsi Cola, and a Royal Crown Cola, and is paralyzed, one cannot tell whether this is because

$$\{X_{Coke}, X_{Pepsi}, X_{RC}\} = C(\{X_{Coke}, X_{Pepsi}, X_{RC}\})$$

or because of some paralysis between just one pair:

$$\{X_{Coke}, X_{Pepsi}\} = C(\{X_{Coke}, X_{Pepsi}, X_{RC}\})$$

unless one accepts some assumptions about his choice function and observes at least one further pair-wise decision (Coke *v.* RC or Pepsi *v.* RC).

AXIOMATA

The six “axiomata” in my paper are basic principles about choice functions. I think that I'm on pretty solid ground with each of these, though someone might be able to come-up with a more elegant set that carry the same load.

The first two are just definitional. The very first axiom just declares that a choice function selects a subset from a set. The second axiom declares that the only time that the choice function produces the empty set is when that was its input. (Some authors just rule-out empty sets; I find it cleaner to use this axiom.)

The next four axioms are rationality constraints. I want to model *indecision*, not *madness*.

Axiom (3) basically says that to figure-out what choice a person would make, one could hold a sort of tournament amongst the possibiliites, with the winner amongst each subset going-out to challenge the winners in other subsets. So, for example, if Pepsi beats RC, then we can test Pepsi against Coke, and don't have to bother testing Coke against RC to know which will be the ultimate choice.

Axiom (4) says that if the set of choices made from set B_1 is a subset of B_2 and B_2 is a subset of B_1

$$C(B_1) \subseteq B_2 \subseteq B_1$$

then the choice made from B_2 would be the choice made from B_1. So, for example, if Sally's favorite soda-pops are Coca-Cola and Pepsi Cola, then Sally's favorite colas are Coca-Cola and Pepsi.

Axiom (5) looks a lot like axiom (4), but what it says is that if the choice from set B_1 is the whole set, and if B_2 is a subset of B_1, then the choice made from B_2 is all of B_2. So, for example, if Ellen cannot rule-out anything amongst Oreos, Lorna Doones, and Nilla Wafers, then Ellen cannot decide between Oreos and Nilla Wafers.

Axiom (6) says that if two sets of options are jointly subsets of two different supersets, then it cannot be the case that just one is a subset of the choices from the first superset, and just the other is a subset of the choices from the other superset. So, for example, Peter won't chose just DC comic books from the set of DC, Marvel, and Charlton comics, and just Marvel comic books from the set of DC, Marvel, and Warren comics.

Some Definitions

At this point, I introduce some definitions.

I formally define "*paralysis*" in terms of choice functions. *Paralysis* is when the choice function produces a set with more than one option. (I don't explicitly define a *relation* of *paralysis*, though it would be trivial to do so, because I wanted *paralysis* to cover choice amongst sets with more than two members, whereäs the usual relations of decision theory are pair-wise.)

I define "*strict preference*" in a way that may seem odd. Specifically, X_1 is *strictly preferred* to X_2 when $\{X_2\}$ is not chosen from $\{X_1, X_2\}$. Under the axioms, this is equivalent to saying that $\{X_1\}$ is chosen from $\{X_1, X_2\}$, but it was easier for me to think about the math with the definition that I used. (Still, I might have to switch things around to make a referee happy.)

Finally, I define "*non-rejection*" such that

$$(X_1 \nless X_2) \equiv \left[\{X_1\} \subseteq C(\{X_1, X_2\})\right]$$

which basically says that X_1 is *not rejected* for X_2 if X_2 is not preferred to X_1. (So X_1 might be *strictly preferred* to X_2, or *paralysis* might hold.

(This is pretty much how "*weak preference*" could be defined if we represented ordinary theory in terms of choice functions. And, for a long time, I just called this relation "*weak preference*". But I became concerned that readers would confuse "*weak preference*" defined in this way with "*weak preference*" defined in terms of properties that it has in an ordinary model. So I changed the name and symbol of the relation.)

Lotteries

I define two more relations (*equi-indifference* and *undecidedness*) in terms of lotteries ("coin flips"), which requires a way of representing lotteries. I represent a lottery as a set of ordered pairs

$$\langle (X_1, p_1), (X_2, p_2), \ldots \rangle$$

Where in (X_i, p_i) , X_i is a possible outcome, and p_i is the associated probability. (The ordering is *within* the pairs; order *amongst* the pairs is unimportant.) I make the *huge* but common assumption that each p_i is a real number. (The assumption is so common that most people consider it part of the very definition of "probability".)

And I offer three equalities. These equalities claim that either side can be substituted for the other in any formula.

The first equality says that lotteries in which only one outcome is possible (*trivial* lotteries) are simply equal to that outcome.

The second equality says that pairs with the same outcome can be combined by adding the probabilities, and likewise that any pair can be split into two pairs so long as the sum of their probabilities is the same.

The third equality says that compound lotteries equal simple lotteries, by a sort of distributive property. For example, let's say that a lottery offered a 30% chance of $10, and a 70% of a second lottery; and say that the second lottery had a 40% chance of $20, and 60% of nothing. I've declared that this is equal to a lottery in which there is a 30% chance of $10, a 28% chance of $20, and a 42% chance of nothing.

Further Definitions

At this point, I formally define "*equi-indifference*" ("$\approx$"). The definition (16) is complicated. But it's just the definition, rather than some sort of principle, and it's chosen to make the relation observable. Basically, the definition says that *equi-indifference* obtains between X_1 and X_2 when

- X_1 and X_2 are both chosen from $\{X_1, X_2\}$.
- Either X_1 is X_2 or there is *paralysis* even when some "coin flip" is made an option.
- All that still holds when the choice is between the "coin-flip" and just X_1 or just X_2.

Actually, given the axioms, this defintion is equivalent to simply saying that *equi-indifference* obtains when the choice function won't rule-out anything from the set of X_1, X_2, and some "coin-flip".

The next relation is *proto-undecideness* ("¡"). It obtains when, given a choice amongst X_1, X_2, and a "coin-flip", the coin-flip is chosen. The reason that this is called "*proto-undecidedness*" is that it doesn't preclude cases where X_2 just *is* X_1, in which case the "coin flip" too just *is* X_1.

And *undecidedness* ("¿") obtains when there is *proto-undecidedness* between X_1 and X_2, and X_2 is not X_1.

Five More Propositions

The foundations supplied by the axioms and by the lottery equalities cannot answer some important questions. For example: Do the relations of *strict preference*, *equi-indifference*, and *undecidedness* exhaust all of the possibilities? If one is *undecided* between two outcomes, are *all* "coin flips" (setting aside two-headed coins and the like) preferred to the choices of certainty? If so, are all equally valued? If not, what "coin flip" is better or best?

Implications of compound lotteries and questions of precise odds rule-out some answers on practical grounds. For example, if there's exactly and only one set of odds that are preferable to the certainty of either outcome, how does one know that a coin gives exactly those odds, rather than something merely very close? Or, if all odds except certainty are equally acceptable, how does one practically distinguish certainty from *near* certainty?

It would be nice to offer a fairly complete theory of gambling, which would answer these and other questions. But I was simply aiming to make a case for there being a coherent alternative both to *strict preference* and to *indifference*. And if-and-when I offer a more complete theory, I want to dispense with that huge assumption that probabilities are always quantities.

So I just offer five propositions about how choices are made amongst lotteries, which propositions themselves seem intuïtively plausible, and which imply some appealing answers.

The first of these propositions (21) is that, when one is paralyzed between two choices, one is paralyzed between a lottery with $m:n$ odds and one with $n:m$ odds. So Jane would not be able to decided between one lottery which offered a 60% chance of Coke and a 40% chance of Pepsi, and another which offered a 40% chance of Coke and 60% chance of Pepsi. This proposition somewhat captures a sense that if one doesn't favor one alternative over another then one doesn't favor a lottery biased in favor of either alternative.

The second proposition (22) is that if one does not prefer one alternative to another, one does not prefer the certainty of the first to a lottery over the two with some chance of the other. So if Richard does not prefer 7-up to Pepsi, then Richard will not select the certainty of a 7-up over a “coin flip” to decide between it and Pepsi.

The third proposition (23) claims that, if one *does* prefer one alternative to another, then one positively prefers a lottery over the two to the certainty of the less desired outcome. So if Sally prefers Coke to 7-up, then Sally would rather have a “coin-flip” decide which she gets than have the certainty of the 7-up.

The fourth proposition (24) is that if there are three outcomes, and some lottery between two is *strictly preferred* to some other lottery across the two (and here I include *trival* lotteries, where the outcome is fixed), then there will be a lottery between one of these and the third outcome which will be *strictly preferred* to another lottery across the pairing. That claim is very weak. If we can find *something* that Peter *strictly prefers* amongst Coke, Pepsi, and various Coke-or-Pepsi “coin flips”, then we can do the same for Coke and RC and Coke-or-RC “coin-flips”, or we can do it for Pepsi and RC and Pepsi-or-RC “coin flips”; maybe both.

The fifth proposition (25) looks monstrous, but merely says that if one is *proto-undecided* between two outcomes, and paralyzed between two different lotteries across those outcomes, then one is *proto-*

undecided between the two lotteries. In the case of *undecidedness*, the implication is that deadlock between lotteries can be broken by a lottery *between* them.

Theoremata (Theorems)

Some important results follow from the axioms, equalities, definitions, and those final five propositions:

- *Strict preference*, *equi-indifference*, and *undecidedness* jointly exhaust all possibilities, and are mutually exclusive. (They cover everything and don't overlap.)
- *Strict preference* has the usual properties of *irreflexivity*, *anti-symmetr*y, and *transitivity*.
- *Equi-indifference*, like classical *indifference*, is *reflexive*, *symmetrical*, and *transitive*.
- *Undecidedness* is *irreflexive*, *symmetrical*, and *intransitive*.
- If one is *equi-indifferent* between two things, then one is *paralyzed* in attempting to choose between *any* two lotteries across those things
- If one is *undecided* between two things, then one would prefer to have the choice between them made by a non-trivial lottery, and lotteries are preferred in order of increasing “fairness”, so that the most desired “coin flip” is 50:50.
- Given X_1, X_2, and X_3, if one *strictly prefers* X_1 to X_2, and is *indifferent* or *undecided* between X_2 and X_3, then one *strictly prefers* X_1 to X_3. And if instead one *strictly prefers* X_2 to X_3, and is *indifferent* or *undecided* between X_1 and X_2, then again one *strictly prefers* X_1 to X_3.
- Given X_1, X_2, and X_3, if one is *indifferent between* X_1 *and* X_2, *and undecided* between X_2 and X_3, then one must be *undecided* between X_1 to X_3.

I believe that most or all of these properties exactly fit “common-sense” intuïtions about preference, indifference, and indecision; and that any which don't *exactly* fit will cause little or no discomfort to those intuïtions.

Discussion

Significance of the model

The real *point* here is that one *can* makes some practical sense out of some idea of *strict preference* and a notion of indifference not jointly totally ordering a set of outcomes. “I don't care” produces one set of choices, while “I don't know” produces another.

But the differences couldn't be observed without the option of (non-trivial) *lotteries*. If each specific action were uniquely associated with a specific outcome, then *equi-indifference* and *undecidedness* wouldn't be observably distinct except in what people *declared* about their internals states, and the union of *equi-indifference* and *undecidedness* would operationalize indistinguishably each from the other and jointly from the classic notion of *indifference*. This model doesn't argue that the standard model of decision-making under certainty is *wrong*, but it suggest that the standard model is less useful and relevant than has typically been presumed.

However, when modelling decision-making in the face of risk and uncertainty, ordinary theory has leaned very heavily on the assumption that *strict preference* and *indifference* totally order the options. Specifically, ordinary theory says that the value that a person uses in choosing amongst lotteries is the *expected utility* of those lotteries. The idea is that each possible outcome X_i can be assigned a quantity of desirability $u(X_1)$, and the over-all value of the lottery is then

$$\sum_i \left[p_i \cdot u(X_i) \right]$$

But the ability to assign such measures of desirability to outcomes implies a total ordering. And the notion of *undecidedness* overtly rejects the proposition that the value of a lottery is an expected utility, as a lottery is treated as somehow more desirable that each of its outcomes.

It is, therefore, when we leave the world of predetermined outcomes that I expect a lack of total ordering to be most signficant.

An Alternative: Desire for Delay

A different notion of a distinct alternative to preference and indifference might be in *desire for delay*. The idea here is that a person who hadn't made up her mind would want *time to do so*. In theory, she might be willing to *pay* for that time.

Operationalizing that notion might be harder than it initially seemed. A person who truly thought that one option were better than another might none-the-less be willing to pay for time to examine that preference. A person who were convinced that two things were of equal value might, given time, *change his mind*. And if a person refused to buy delay, it might be from indifference or it might merely be that the price of delay were too high.

If this other notion can be operationalized, there could be cases where the individual both *strictly preferred* a delay and *strictly preferred* a “coin flip” to selecting either outcome.

Attitudes toward Entropy versus Attitudes towards Risk

Ordinary economic theory does not allow for liking or disliking entropy (indeterminacy) itself. When such theory speaks of attitudes towards “risk”, it is speaking of the difference between the expected utility of the payout:

$$\sum_i \left[p_i \cdot u\left(X_i\right)\right]$$

and the utility of the expected payout:

$$u\left[\sum_i \left(p_i \cdot X_i\right)\right]$$

“Risk”, then, cannot exist unless

$$\sum_i \left[p_i \cdot u\left(X_i\right)\right] \neq u\left[\sum_i \left(p_i \cdot X_i\right)\right]$$

even when outcomes are not predetermined. In my model, choice may be affected by entropy even when there is no *risk.*

One of the propositions that I introduced to resolve questions about *equi-indifference* and about *undecidedness* rules-out cases where a person, though unable to decide between two options, would prefer *either* to a “coin-flip”. Yet we could imagine someone having such a preference. The model that I've presented would have to be changed to accept such preferences, but it's something to consider.

Areas for Future Work

The model itself is just an attempt at a *proof of concept*. It doesn't try to tell the whole story of decision-making under uncertainty, and the foundation of what it does say about such decision-making, though perhaps intuïtively plausible, is rather *ad hoc*.

One possibility would be to replace those foundations with propositions that that are less *ad hoc*, give a more complete theory of decision-making, and imply the present propositions as theorems.

Another possibility would be something more radical, but continuing to distinguish at least three mutually exclusive relations and continuing the aforementioned assumption that probabilities are *quantities*.

What I wish to pursue, though, is an integration of the idea that choices are not totally ordered with one that *plausibilities* are not totally ordered. This latter idea is that sometimes a person can neither say that one outcome is more likely than another nor that they are equally likely.

Conclusion

Although the inclusion of a relation of indecision makes *some* difference in predicted behavior, it remains to be seen whether it would have much effect on that part of economics that describes, predicts, or explains behavior without passing judgment on it. I'd need a more complete theory to have a fuller and more confident sense of the potential implications for that part of economics. And some of the *potential* implications might be empirically unreälized. The real-world effects might be mere *blips*, though I rather suspect that they will sometimes be profound.

But there's an awful lot of *normative* theory, from various quarters, that assumes that we can and should judge the merits of a policy based upon some sort of comparison and *summing-up* of costs and benefits across persons. This assumption is incoherent if costs and benefits aren't quantities. And if the foundation of cost and benefit is in some sort of preference structures, and those preference structures aren't totally ordered, then there simply aren't quantities. So the implications for *prescriptive* theory are likely to be profound even if the implications for non-normative economics are minor.