Preferences, Probabilities, and Decisions

Todd Davies, Symbolic Systems 100

May 1, 2008

1. Preferences

The development of preference theory below follows that of Kreps (1988).

DEFINITION 1.1. An *n-ary relation R* on sets $\boldsymbol{S_1,S_2,...,S_n}$, is a subset of the cross-product $\boldsymbol{S_1 \times S_2 \times ... \times S_n}$ (written as $R \subseteq \boldsymbol{S_1 \times S_2 \times ... \times S_n}$). The relation R holds between $s_1,s_2,...,s_n$ (or $R(s_1,s_2,...,s_n)$) iff the ordered pair $<s_1,s_2,...,s_n> \in R$.

- Note: If R is a *binary* relation (n=2) that holds for $<s_1,s_2>$, we may also write s_1Rs_2.

Let $\boldsymbol{X}$ be any set of *possibilities* or *outcomes*. An outcome may be thought of as a proposition which specifies an event that may happen in the future. Outcomes should be comparable to each other, and may, but need not, be mutually exclusive events.

EXAMPLE 1.2. $\boldsymbol{X}$ = {Sally receives a piece of chocolate cake tomorrow at lunch, Barack Obama is elected President of the United States in 2008, John McCain is elected President of the United States in 2008}.

Let $P_i \subseteq \boldsymbol{X \times X}$, where $<x_1,x_2> \in P_i$ denotes that for agent *i,* outcome x_1 is strictly preferred to outcome x_2. When speaking generally, we may omit the reference to an agent, and write simply: x_1Px_2 (x_1 is strictly preferred to x_2). Let $x_1 \neg Px_2$ denote that x_1 is *not* strictly preferred to x_2 .

DEFINITION 1.3. $P \subseteq \boldsymbol{X \times X}$ is a (strict) *preference relation* iff
(a) $\forall x,y\ xPy \Rightarrow y \neg Px$ (*asymmetry*),
and
(b) $\forall x,y,z\ \ x \neg Py\ \&\ \ y \neg Pz \Rightarrow x \neg Pz$ (*negative transitivity*).

THEOREM 1.4. If $P \subseteq \boldsymbol{X \times X}$ is a preference relation then $\forall x,y,z\ \ xPy\ \&\ \ yPz \Rightarrow xPz$ (*transitivity*).
Proof (due to Halpern). Suppose that xPy, yPz, and $x \neg Pz$. By asymmetry, $z \neg Py$. But by negative transitivity, $x \neg Py$, contradicting our assumption.

DEFINITION 1.5. If $P \subseteq \boldsymbol{X \times X}$ is a strict preference relation, then the relation I (*indifference*) holds between x and y iff $x \neg Py$ and $y \neg Px$.

EXERCISE 1.6. Prove that if P is a preference relation that defines I as an indifference relation, then I is transitive.

DEFINITION 1.7.. A strict preference relation P on an outcome set X is *representable in utility* iff there exists a u: $X \rightarrow \Re$ (function u that maps X into the real numbers) such that xPy iff $u(x) > u(y)$.

DEFINITION 1.8.. A set $Y \subset X$ is *order-dense* under a preference relation P iff for all x and z in X, if

xPz then there exists a y in Y such that xPy and yPz.

THEOREM 1.9. *Utility representation theorem for ordinal preferences*. A preference relation P on an outcome set X is representable in utility iff there exists a denumerable order-dense subset of X under P. For the proof, see Blume (2002).

Theorem 1.9 is called a "representation theorem" because it specifies the conditions under which data (in this case preferences) can be represented by a function. The details of the theorem are beyond the scope of this course, since it refers to the distinction between denumerable (countable) and uncountable sets, a topic requiring deeper study of set theory than we will assume. But the defining condition for a preference relation to be representable in utility is an example of the *Archimedean condition*. Applied to preferences, an implication of the Archimedean condition is that it should never be possible for an arbitrarily small difference between outcomes on one dimension to outweigh an arbitrarily large difference on another dimension. Such a possibility can occur under what are called "lexicographic preferences".

DEFINITION 1.10. Suppose that two outcomes x and y may be compared along two dimensions d_1 and d_2. The agent *lexicographically prefers* x to y iff $d_1(x) > d_1(y)$ or $[d_1(x) = d_1(y)$ & $d_2(x) > d_2(y)]$. The definition can be extended for any higher number of dimensions.

Lexicographic preference relations are not representable in utility, and are often argued to be counter-normative because they give infinite weight to infinitesimal differences.

EXERCISE 1.11. Prove that the lexicographic preference rule defined in 1.10 is a preference relation.

EXAMPLE 1.12. *Difference principle.* An example of a lexicographic preference rule that has been proposed seriously is the "difference principle" defined by John Rawls in *A Theory of Justice* (1971). The difference principle stipulates that any inequalities existing in a society "are to be of the greatest benefit to the least-advantaged members of society" (p. 303). This is also known as the *maximin* rule because it maximizes the minimum amount of resources that anyone in a society receives. The rule is lexicographic (and therefore violates the Archimedean condition) because any increase in resources for the least well-off person outweighs all changes in the resources received by others. A gain of one grain of rice by the least well-off can, under a strict interpretation of the difference principle, justify preferring that state of affairs to one in which everyone but the least well off would have vastly greater resources. This lexicographic character of the difference principle is less well known than the fact that Rawls's two main principles of justice are also lexicographically ordered (maximum equality of liberty takes lexicographic priority over economic well-being) and that the difference principle is lexicographically preferred to fair equality of opportunity.

Many studies of public opinion indicate that people hold lexicographic principles in ethical matters. For example, people often say that "human life comes before anything else" (Baron, 2000). It is difficult to determine experimentally how stringently people hold to these preferences when extreme costs on other dimensions are the result, but there is at least substantial evidence that professed policies adhere to lexicographic priority. Indeed, as shown above, philosophers have sometimes become famous for advocating lexicographic rules. Another example is Immanuel Kant, who took the view that dishonesty is always wrong, no matter what the consequences.

2. Probability

The development of probability theory below follows Kolmogorov (1933), as elaborated upon by authors such as Krantz, Luce, Suppes, and Tversky (1971) and Kreps (1988). We consider the simple case of discrete, finite sets of possibilities.

DEFINITION 2.1. Let $S = \{s_1, s_2, \ldots, s_n\}$ be a finite set of possible outcomes in a context C. S is a (finite) *sample space* for C iff exactly one outcome among the elements of S is or will be true in C.

EXAMPLE 2.2. Let C be the particular flipping of a coin. Then S = *{Heads, Tails}* is a sample space for C. Another sample space for C is S' = *{Heads is observed, Tails is observed, Cannot observe whether the coin is heads or tails}*. Yet another is S'' = *{Heads is observed and someone coughs, Heads is observed and no one coughs, Tails is observed whether someone coughs or not}*.

In what follows, we will assume the existence of a context without stating it.

DEFINITION 2.3. Let S be a sample space, and $\emptyset \neq E \subseteq 2^S$ (E is a nonempty subset of the power set of S, i.e., it is a set of subsets of S). Then E is an *event space* (or *algebra of events*) on S iff for every $A,B \in E$:
(a) $S \setminus A = A^C \in E$ (the S-complement of A is in E)
and
(b) $A \cup B \in E$ (the union of A and B is in E).
We call the elements of E consisting of single elements of S *atomic events*.

COROLLARY 2.4. If E is an event space on a sample space S, then $S \in E$.

EXERCISE 2.5. Prove 2.4.

EXAMPLE 2.6. If S = *{Heads, Tails}*, then $E = \{\emptyset, \{Heads\}, \{Tails\}, \{Heads, Tails\}\}$ is an event space on S. The atomic events are *{Heads}* and *{Tails}*.

DEFINITION 2.7. Let S be a sample space and E an event space on S. Then a function $P: E \rightarrow [0,1]$ is a (finitely additive) *probability measure* on E iff or every $A,B \in E$:
(a) $P(S) = 1$
and
(b) If $A \cap B = \emptyset$ (the intersection of A and B is empty, in which case we say that A and B are *disjoint* events), then $P(A \cup B) = P(A) + P(B)$ (*additivity*).
The triple $<S,E,P>$ is called a (finitely additive) *probability space*.

COROLLARY 2.8. If $<S,E,P>$ is a finitely additive probability space, then for all $A,B \in E$:
(a) $P(A^C) = 1 - P(A)$ (*binary complementarity)*
(b) $P(\emptyset) = 0$
(c) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
Proof.
(a) $S = A \cup A^C$ by the definition of complementarity. Therefore $P(A \cup A^C) = 1$ by 4.1.7(a). A and A^C are disjoint by the definition of complementarity, so by 2.7(b), $P(A \cup A^C) = P(A) + P(A^C)$, so $P(A) + P(A^C) = 1$ and the result follows by subtracting $P(A)$ from both sides of the equation.

(b) $S^C = S \setminus S = \varnothing$. Thus $P(\varnothing) = P(S^C) = 1 - P(S)$ by 2.8a, which by 2.7a is 1-1 = 0.
(c) From set theory, we have $A = (A \cap B) \cup (A \cap B^C)$ and $B = (B \cap A) \cup (B \cap A^C)$, $A \cup B =$ $(A \cap B) \cup (A \cap B^C) \cup (B \cap A^C) = (A \cap B) \cup [(A \cap B^C) \cup (B \cap A^C)]$, $(A \cap B) \cap (A \cap B^C) = \varnothing$, $(B \cap A) \cap (B \cap A^C) = \varnothing$, and $(A \cap B) \cap [(A \cap B^C) \cap (B \cap A^C)] = \varnothing \cap \varnothing = \varnothing$. Therefore, $P(A) = P[(A \cap B) \cup (A \cap B^C)] = P(A \cap B) + P(A \cap B^C)$ and $P(B) = P[(B \cap A) \cup (B \cap A^C)] = P(B \cap A) + P(B \cap A^C)$, and $P(A \cup B) = P\{(A \cap B) \cup [(A \cap B^C) \cup (B \cap A^C)]\} = P(A \cap B) + P(A \cap B^C) + P(B \cap A^C)$. Substituting, $P(A \cup B) =$ $P(A \cap B) + [P(A) - P(A \cap B)] + [P(B) - P(B \cap A)]$. $B \cap A = A \cap B$, so $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

DEFINITION 2.9. The *conditional probability* $P(A|B)$ of an event A given event B is defined as follows: $P(A|B) = P(A \cap B) / P(B)$.

THEOREM 2.10. *Bayes's rule.* For events A and B, $P(A|B) = [P(B|A)P(A)] / P(B)$.
Proof. By 2,9, $P(A|B) = P(A \cap B) / P(B)$. $A \cap B = B \cap A$, so $P(A \cap B) = P(B \cap A)$. From 2.9, then, we can derive $P(B \cap A) = P(B|A)P(A)$, and the theorem follows by substitution.

EXAMPLE 2.11. *Applying Bayes's rule to medical diagnosis.* Eddy (1982) gives the following example calculation of the probability that a breast lesion is cancerous based on a positive mammogram . From the literature, Eddy estimates the probability that a lesion will be detected through a mammogram as .792. Hence the test will turn up negative when cancer is actually present 20.8% of the time. When no cancer is present, the test produces a positive result 9.6% of the time (and is therefore correctly negative 90.4% of the time). The key fact is the prior probability that a patient who has a mammogram will have cancer, which is taken to be 1%. Thus, Eddy calculates the probability of cancer given as positive test as [(.792)(.01)] / [(.792)(.01) + (.096)(.99)} + .077, applying 2.11, so a patient with a positive test has less than an 8% chance of having breast cancer. Does this seem low to you?

DEFINITION 2.12. Two events A and B are *independent* iff $P(A \cap B) = P(A)P(B)$.

COROLLARY 2.13. Two events A and B satisfying $P(B)>0$ are *independent* iff $P(A|B) = P(A)$.
Proof. (a) *Only if* direction: Since A and B are independent, $P(A \cap B) = P(A)P(B)$ by 2.13. Since $P(B)>0$, $P(A|B) = P(A \cap B)/P(B) = P(A)P(B)/P(B) = P(A)$. (b) *If* direction: $P(A|B) = P(A)$, so multiplying both sides by $P(B)$, $P(A|B)P(B) = P(A)P(B) = P(A \cap B)$.

EXAMPLE 2.14. Consider two flips of a coin and let A=*{Heads on the first toss}* and B=*{Tails on the second toss}*. The probability that the coin lands on tails on the second toss is not affected by what happened on the first toss and vice versa, so $P(B|A) = P(B)$ and $P(A|B) = P(A)$. Assuming both sides of the coin have a nonzero probability of landing on top, this argument suffices to establish that the tosses are independent and that, therefore, $P(A \cap B) = P(A)P(B)$. Assuming the coin is unbiased ($P(\{Heads\}) = P(\{Tails\}) = 0.5$), this means that $P(A \cap B) = (.5)(.5) = .25$.

EXERCISE 2.15. Consider two six-sided dice (with faces varying from 1 to 6 dots) which are rolled simultaneously, and assume each roll is independent of the other. What is the probability that the sum of the two dice is 7?

3. Decisions

DEFINITION 3.1. *Preference between gambles.* If
(a) S is a *set of outcomes {x,y,z,w,...}*
(b) *x* **P** *y* means *x is preferred to y* (also known as *strict preference*)
(c) *x* **I** *y* means *x* is *indifferent to* y
(d) x **PI** y means x *is either preferred or viewed indifferently relative to* y (also known as weak preference) [note that **I** and **P** can be each be defined using **PI**]
(x,p,y) means a *gamble* (an uncertain *outcome*, or a lottery) in which *outcome* x will be received with probability p, and *outcome* y will be received with probability 1-p.

AXIOMS 3.2. Assuming that x,y,z,w ∈ S, and p,q ∈ (0,1):

1. *Closure.* (x,p,y) ∈ S.
2. *Weak ordering.*
    - x **PI** y (*Reflexivity)*
    - x **PI** y or y **PI** x *(Connectivity)*
    - x **PI** y and y **PI** z implies x **PI** z *(Transitivity)*
3. *Reducibility.* [(x,p,y),q,y] **I** (x,pq,y).
4. *Independence.* If (x,p,z) **I** (y,p,z), then (x,p,w) **I** (y,p,w).
5. *Betweenness.* If x **P** y then x **P** (x,p,y) **P** y.
6. *Solvability.* If x **P** y **P** z, then there exists p such that y **I** (x,p,z).

THEOREM 3.3. *Representation Theorem for Utility* (von Neumann & Morgenstern, 1944). If axioms 3.2.1-6 are satisfied for all outcomes in S, then there exists a real-valued utility function u defined on S, such that (1) x **P** y if and only if u(x) > u(y), and x **I** y if and only if u(x) = u(y); (2) u(x,p,y) = pu(x)+(1-p)u(y); (3) u is an interval scale, that is, if v is any other function satisfying 1 and 2, then there exist real numbers b, and a>0, such that v(x) = au(x)+b.

EXAMPLE 3.4. Paradox: (Allais, 1953) [updated version]. Compare the following two situations:

Situation 1	*Situation 2*
Choose between:	Choose between:
Gamble 1: $5000 with probability 1	Gamble 3: $5000 with probability .11
Gamble 2: $7500 with probability .10	$0 with probability .89
$5000 with probability .89	Gamble 4: $7500 with probability .10
$0 with probability .01	$0 with probability .90

Most people prefer gamble 1 to gamble 2, but prefer Gamble 4 to Gamble 3, even though this pattern is inconsistent with the independence axiom. In particular,
gamble 1 **P** gamble 2 can be rewritten as ($5000,.11,$5000) **P** [(0,1/11,$7500),.11,$5000]; and
gamble 4 **P** gamble 3 can be rewritten as [(0,1/11,$7500),.11,$0] **P** ($5000,.11,$0) (cf axiom 3.3.4).
Since expected utility theory requires an ordering consistent with the interval function of utility, this pattern of preferences cannot be accommodated. In particular, the preference for gamble 1 over gamble 2 implies that u(gamble 1) > u(gamble 2), and hence that u($5000) > .10u($7500)+.89u($5000)+.01u(0), so .11u($5000) > .10u($7500)+.01u($0). But the preference in situation 2 implies that u(gamble 4) > u(gamble 3); hence .10u($7500)+.90u($0) >

.11u($5000)+.89u($0), implying .10u($7500)+.01u($0) > .11u($5000), contradicting the inequality derived from the most common preference in situation 1.

References

Allais (1953). Le Comportement de l'Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l'Ecole Americaine. *Econometrica, 21*:503-546

Baron (2000). *Thinking and Deciding*. Cambridge University Press.

Bloom (2002). Ordinal Representations. Available online at http://www.cs.cornell.edu/courses/cs576/2002fa/blumenotes.pdf.

Eddy (1982). Probabilistic Reasoning in Clinical Medicine: Problems and Opportunities. In D. Kahneman, P. Slovic, and A. Tversky (Eds.), *Judgment Under Uncertainty: Heuristics and Biases,* pp. 249-267. Cambridge University Press.

Kolmogorov (1933). *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Springer.

Krantz, Luce, Suppes, and Tversky (1971). *Foundations of Measurement, Volume I.* Dover.

Kreps (1988). *Notes on the Theory of Choice.* Westview Press.

Rawls (1971). *A Theory of Justice.* Belknap.

von Neumann & Morgenstern (1944). *Theory of Games and Economic Behavior.* Princeton University Press.