Decision theory

Stockholm University, fall 2023

H. Orri Stefánsson
orri.stefansson@philosophy.su.se
www.orristefansson.is

The aim of this course is to introduce the basics of decision theory, and to critically discuss the philosophical foundations and implications of the theory. We start by discussing different uses of decision theory and the formulation of decision problems. After introducing the principle of expected utility maximisation, we discuss where the utilities are meant to come from and how to measure them according to the von Neumann-Morgenstern (vNM) expected utility theory. We go through the justification of the vNM axioms and consider criticism of the theory based on seemingly rational attitudes to risk that conflict with these axioms. A limitation of the vNM theory is that it assumes that the options of choice are associated with objective probability distributions over outcomes. In most realistic decision problems, however, the decision-maker has to use her own estimates of how likely an option is to lead to each of its possible outcomes (i.e., she needs to consult her subjective probabilities). Savage's decision theory is meant to accounts for this. We critically discuss Savage's axioms and his idea that a rational person's preferences determine a unique subjective probability function. We consider whether Savage's representation theorem provides a convincing argument for probabilism, the view that beliefs should be probabilistic. We conclude the course by discussing the so-called Dutch Book Argument for probabilism.

Peterson (2009) and Steele and Stefánsson (2015) are the core readings of the course. (You can also use the second, 2017, edition of Peterson. Where the relevant pages or chapters differ between the versions I list both.) Peterson contains a number of exercises, with solutions. You will definitely benefit from going through a few exercises each week, to make sure that you have understood the technical details. However, we will not go through any exercises in class, but instead focus on the conceptual issues.

Reading list—both required readings and extra readings:

- S. Bradley (2014), "Imprecise probabilities", Stanford Encyclopedia of Philosophy.
- Bradley and Stefánsson (2017), "Counterfactual desirability", British Journal for the Philosophy of Science 68 (2): 485-53.
- Briggs (2014), "Normative theories of rational choice: expected utility", Stanford Encyclopedia of Philosophy.
- Broome (1991a), *Weighing Goods*, chapter 5.
- Broome (1991b), "Utility", Economics and Philosophy 7(1): 1-12.
- Broome (1999), "Can a Humean be moderate?", in his *Ethics out of Economics*.
- Buchak (2014), "Risk and tradeoffs", Erkenntnis 79(6): 1091-1111.
- Buchak (2016), "Decision theory", in Hájek and Hitchcock (eds.) *The Oxford Handbook of Probability and Philosophy*.
- Ellsberg (1961), "Risk, Ambiguity, and the Savage Axioms", The Quarterly Journal of Economics, 75(4): 643-669.
- Eriksson and Hájek (2007), "What are degrees of belief?" Studia Logica 86(2):183-213.
- Gärdenfors and Nils-Eric Sahlin (1982), "Unreliable probabilities, risk taking, and decision making", Synthese 53(3): 361-386.
- Gustafsson (2022), *Money Pump Arguments* (available electronically).
- Hájek (2005), "Scotching Dutch books?" Philosophical Perspectives 19(1): 139-151.
- Hájek (2008), "Dutch book arguments", in Anand, Pattanaik and Puppe (eds.), *The Oxford Handbook of Rational and Social Choice*.
- Lewis (1980), "A subjectivist's guide to objective chance", in R. Jeffrey (ed.), *Studies in Inductive Logic and Probability*.
- Lewis (1994), "Humean supervenience debugged", Mind 103(412): 473-490.
- Meacham and Weisberg (2011), "Representation theorems and the foundations of decision theory", Australasian Journal of Philosophy 89(4): 641-663.
- Peterson (2009/2017), *An Introduction to Decision Theory*.
- Pettigrew (2020), *Dutch Book Arguments* (available electronmically).
- Ramsey (1926), "Truth and probability", published posthumously in 1931 in his *Foundations of Mathematics*.
- Schick (1986), "Dutch bookies and money pumps", The Journal of Philosophy 83(2): 112-119.
- Steele and Stefánsson (2015), "Decision theory", Stanford Encyclopedia of Philosophy.
- Steele and Stefánsson (2021), *Beyond Uncertainty: Reasoning with Unknown Possibilities* (available electronically).
- Stefánsson (2023), "Continuity and catastrophic risk", Economics and Philosophy 38(2): 266-274.
- Stefánsson and Bradley (2019), "What is risk aversion?" The British Journal for the Philosophy of Science, 70(1): 77-102.
- Talbott (2008), "Bayesian epistemology", Stanford Encyclopedia of Philosophy.

Course outline

Lecture 1: Introduction

There are different uses of decision theory, but the focus of this course will be on normative decision theory, as a (thin) theory of practical rationality. In this lecture we discuss the framing of decision problems and the fundamental principle of normative decision theory: you should maximize expected utility. But where do the utilities come from? And what precisely does utility maximization mean for your preferences? These questions will be discussed in lecture 2. Does rationality really require utility maximization? This question will be discussed lectures 3 and 4. What about the probabilities that go into calculating expected utilities—where do they come from? And why should one's beliefs be probabilities? These questions will be discussed in the second half of the course.

Readings:
- Peterson (2009), chapter 1 and ch. 2-2.3.

Extra readings:
- Peterson (2009), section 2.4;
- Briggs (2014), sections 1 and 2;
- Buchak (2016).

Lecture 2: Ordinal and cardinal utilities

Utility values are not primitive. Rather, a preference ordering (or comparative desire ordering) over outcomes is primitive, and the utility values simply represent this ordering. The existence of *ordinal* utilities only requires that the preference be *Transitive* and *Complete*. But ordinal utilities are not enough when making risky choices. With use of lotteries, we can however *cardinalize* utilities, if the preference satisfies certain additional constraints.

Readings:
- Stefánsson and Steele (2015), sections 2-2.2;
- Peterson (2009), ch. 5-5.2;
- Broome (1991b).

Extra reading:
- Peterson (2009), sections 5.3-5.4;
- Briggs (2014), section 3.1.

Lecture 3: von Neumann and Morgenstern's theory

When an agent satisfies the von Neumann and Morgenstern (vNM) axioms, we can be sure that her attitudes to final outcomes can be represented by cardinal utilities, and that her preference over lotteries (or risky prospects) can be represented as maximizing expected utility. We will go through the vNM axioms and discuss their plausibility as requirements of rationality. The focus will be on pragmatic arguments for Transitivity and Independence.

Readings:
- Peterson (2009), ch. 4-4.3 and ch. 8-8.4;
- Stefánsson and Steele, section 2.3.

Extra readings:
- Gustafsson (2022);
- Broome (1999a), ch. 5;
- Schick (1986), especially section II (section I will be relevant in wk. 8).

Lecture 4: vNM and risk

Many people feel that the vNM theory places implausible restrictions on attitudes to risk. One problem comes from the "phenomenology of risk aversion"—it *feels* like risk aversion has nothing to do with attitudes to outcomes (contrary to what the vNM theory entails). Another problem is that the vNM seems inconsistent with reasonable risk attitudes, such as the so-called Allais preference.

Readings:
- Stefánsson and Bradley (2019);
- Peterson (2009), sections 4.4 and 8.5. [NB: Those of you with the second (2017) edition, should replace 8.5 with 10.1 and 10.2.]

Extra readings:
- Bradley and Stefánsson (2017);
- Broome (1991a), ch. 5;
- Broome (1999b);
- Briggs (2014), section 3.2;
- Buchak (2014).

Lecture 5: Subjective and objective probabilities

The vNM framework assumes that the options we are choosing between come with predefined *objective* probabilities. In most realistic decisions, however, a decision-maker needs to consult her own *subjective* probabilities (i.e., degrees of belief or confidence) when making a choice. In lecture 6 we consider the orthodox decision theory for prospects that don't come with objective probabilities. In this lecture we first remind ourselves of the mathematics of probability, before discussing philosophical theories of objective probabilities, and how subjective probabilities should relate to objective ones.

Reading:
- Peterson (2009), chapters 6 & 7, up until p. 146 [in the second edition, up until p. 155];
- Lewis (1980).

Extra readings:
- Lewis (1994);
- Eriksson & Hájek (2007);
- Talbott (2008).

Lecture 6: Choice under uncertainty: Savage

In most real decision problems, the alternatives on offer do not come with a predetermined probability distribution over outcomes. What should you do in cases? Maximize expected utility with respect to your utility and *subjective* probability function! But can your degrees of belief really be represented by a probability function? Yes, if you satisfy Savage's rationality axioms. But does that mean that our degrees of beliefs *should* be probabilities? That is the main question of this lecture.

Readings:
- Stefánsson and Steele (2015), section 3.1;
- Peterson (2009), pp. 147-153 [in the second edition, p. 155-162].

Extra readings:
- Meacham and Weisberg (2011);
- Stefánsson and Steele (2015), section 3.2;
- Eriksson and Hájek (2007);
- Ramsey (1926).

Lecture 7: Ambiguity and imprecise probabilities

Bayesian decision theorists like Savage assume that when choosing between alternatives under uncertainty, rational people always go for the alternative with the highest expected utility, where the expectation is calculated relative to a "precise" (single) subjective probability function. Some people have argued that this ignores the fact that our evidence often only warrants "imprecise" probability judgements (or sets of probabilities), and that it can be perfectly rational to violate the assumptions of subjective expected utility theory by disliking uncertainty about probabilities (or "ambiguity"). We consider these arguments, focusing on Ellsberg's paradox and a practical version of Poppers paradox of ideal evidence.

Readings:
- Ellsberg (1961);
- Peterson (2009), section 4.5.

Extra readings:
- S. Bradley (2014);
- Gärdenfors and Sahlin (1982);
- Stefánsson and Bradley (2019);
- Steele and Stefánsson (2021).

Lecture 8: The Dutch Book Argument

Savage's axioms seem incredibly strong; arguably too strong to be requirements of rationality. But then Savage's theory does not justify *probabilism*, the view that our beliefs should be probabilistic. The Dutch book argument is meant to provide a different justification of probabilism, by showing that you risk a sure loss if and only if your degrees of belief are not probabilistic.

Readings:
- Peterson (2009), sections 7.5.2 and 7.5.3;
- Hájek (2008) [feel free to skip section 6].

Extra readings:
- Hájek (2005);
- Eriksson & Hájek (2007);
- Pettigrew (2020);
- Schick (1986);
- Ramsey (1926).