

Reading Assignment 10: Savage’s Theory

Assigned Reading Read Sections 3-3.1 of K. Steele and H. O. Stefánsson. “Decision Theory”. In: *The Stanford Encyclopedia of Philosophy*. Ed. by E. N. Zalta. Winter 2020. Metaphysics Research Lab, Stanford University, 2020.

You may stop reading Section 3.1 when you reach the sentence, “Savage’s own proof is rather complicated, but Kreps (1988) provides a useful illustration of it.” You will finish Section 3.1 in the next class. The introduction, section 1, and section 2 are optional but I recommend reading them to (1) review material from the course thus far, and (2) familiarize yourself with Steele and Stefannson’s terminology.

Words of encouragement: Some of you will find this assigned reading difficult because there is a good amount of mathematical notation. Try your best and don’t feel discouraged if you feel like you don’t understand everything. If you are the type of student who finds mathematical writing a bit discomfoting, you will likely understand way more than you appreciate if you are like my previous students.

Technical Requirements

Answer questions two, four, eight, and nine below. Together, your answers should not be longer than a single typed page.

Because the assigned reading is a webpage, please cite the *section numbers* in which questions are answered, as page numbers are not available. See the document “Reading Assignment and Quiz Guidelines” for further instructions.

1. Give examples of (a) a choices under risk and (b) a choice under uncertainty. Explain why the the choices you have described are choices under risk and uncertainty respectively.
2. What do “some of the most celebrated results in decision theory” show, and how does what they show differ from what von Neumann and Morgenstern’s theorem shows? Hint: How is probability interpreted in von Neumann and Morgenstern’s framework? How is it interpreted in Savage and Jeffrey’s frameworks?
3. According to Steele and Stefannson, what do the EU representation theorems show? To make answering this question a useful exercise, try to avoid copying Steele and Stefannson’s wording. That is, answer this question, to the best of your ability, in your own words. It might help to finish the assigned reading before trying to answering this question so that you have a better idea of how Steele and Stefannson (and others) use the following two phrases: “consistent preference attitudes” and “means to [one’s] ends.” Note: The “EU representation theorems” [of Savage and Jeffrey] are “the most celebrated results in decision theory” alluded to in the previous question. So in these opening paragraphs, Steele and Stefannson have described the implications of the theorems in two different ways. It it is worthwhile, therefore, comparing the two ways that Steele and Stefannson describe the theorems and asking yourself why the descriptions are equivalent.
4. How are “events” defined in Savage’s theory? Give an example of an event in the decision matrix below; your answer must not be a single state. What are the targets of belief in Savage’s theory? The targets of desire? What do you think Steele and Stefeannson mean by “target”?

	Rain	Snow	Sunshine
Bike	Wet ride	Cold ride	Pleasant ride
Take train	−\$3	−\$3	−\$3

5. In von Neumann and Morgenstern's theory, individuals have preferences among *lotteries*. Among what do individuals have preferences in Savage's theory? How are "acts" defined in Savage's theory? What is a "constant act"? What is "the problem" with some constant acts, according to Steele and Stefannson? Note: In an algebra class you took in the past, you graphed so-called "constant functions" like $f(x) = 2$. "Constant acts" are just constant functions in the standard mathematical sense.
6. Steele and Stefannson use the bolded capital letter \mathbb{F} to denote which of the following: (1) states, (2) acts, or (3) outcomes? Suppose f represents "Bike" in the decision matrix above and $s = \text{Rain}$. What does $f(s)$ represent?
7. Steele and Stefannson claim that, although Savage's theorem might seem surprising initially, "it is evident that some of our beliefs can be determined by examining our preferences." They then give an example to illustrate how a person's beliefs might be determined from their preferences. Give your own example to illustrate the claim.
8. Savage's definition of comparative belief (Definition 1) will likely confuse many students. Don't give up! It's an important definition, and you will understand it after class discussion. But this question is designed to force you to prepare a bit for that discussion and to stare at the definition for a bit of time.

Savage's definition is beautifully illustrated by Steele and Stefannson's example involving "a choice between two lotteries, one that results in you winning a nice prize if a coin comes up heads but getting nothing if the coin comes up tails, another that results in you winning the same prize if the coin comes up tails but getting nothing if the coin comes up heads." To understand why that example illustrates Savage's definition, suppose you interpret the variables in Definition 1 in the following way:

- Suppose E represents the event that the coin lands heads. So the event E contains precisely one state s_i , namely, the state in which the coin lands heads.
- Suppose F represents the event that the coin lands tails,
- Suppose X represents the "nice prize" that Steele and Stefannson describe.
- Suppose Y represents "getting nothing."

Given the definitions above, the expression " $E. \lesssim .F$ " denotes the claim that some decision-maker (say you!) thinks the coin is at least as likely to land on tails as it is to land on heads. The expression " $F. \lesssim .E$ " denotes the claim that some decision-maker (say you!) thinks the coin is at least as likely to land on tails as it is to land on heads.

Now here are the questions. Given those definitions of E, F, X and Y , what is represented by the variable f in the first bullet point in Savage's definition of comparative probability? What does the variable g represent? What does the expression " $Y \preceq X$ " mean? What does $f \preceq g$ mean? Finally, what does the claim " $E. \lesssim .F \Leftrightarrow f \preceq g$ mean?" Note: \Leftrightarrow means "if and only if"

9. The sure thing principle is analogous to which of von Neumann and Morgenstern's axioms?
10. What do you think distinguishes a *rationality requirement* from a *structural axiom*? This question is, at best, only partially answered in the assigned reading, and so try your best interpreting the text. To answer this question, it might help to compare and contrast Savage's last two axioms (P5 and P6) from the previous four (P1-P4).