Incomplete preferences and Knightian uncertainty

Oliver Snellman
oliver.snellman@gmail.com

October 2019

Abstract

In this essay I discuss different forms of uncertainty, how they give rise to important philosophical and practical considerations, and what is their impact on modeling the consumer behavior in Economics.

1 Introduction

The aim of Economic theory is to build coherent and useful mathematical and conceptual frameworks to improve our understanding of how humans and markets function. Uncertainty plays an important role in decision making, which poses some challenges from theoretical perspective. The reality of human beings can be deeply uncertain, due to our natural limitations and the complexity surrounding most decision making situations. Typically all of the relevant data for some decision is unavailable, even in theory, as some of it might await in the future.

There are situations, where it is impossible to quantify or even to consider all of the risks involved. These situations are said to involve Knightian uncertainty, and they violate one of the most basic axioms of decision theory; completeness of preferences. Hence, it might not be possible for a consumer to conduct a binary preference ordering of these choices.

However, not all uncertain situations exert the same degree of difficulty. It is crucial to detect, which situations allow for extensive quantitative modeling, and which might be beyond the scope of our current tools. A choice must be made in the latter case, whether to build more elaborate quantitative models, or choose a more heuristic approach.

2 Modeling of risk

Decision theory with certain outcomes starts by characterizing preferences and choices with reasonable assumptions, and building higher level theory on top of that foundation. The central assumption is, that the binary preference relation of consumer, underlying his choice behaviour, is rational. Rationality, in this context, means completeness and transitivity of preferences. [1]

Uncertainty in the form of quantifiable risk is added to this foundation with the concepts of lottery and expected utility. A simple lottery is a set of probabilities p_n summing to one, describing the objective probabilities for occurrences of different outcomes. The uncertainty can be abstracted to higher levels with compound lotteries, making it uncertain which (compound) lottery realizes. By choosing a consequentialist premise, we narrow the focus only to reduced versions of these compound lotteries - simple lotteries. The reasoning behind this choice is, that the only important thing for a decision maker is the probability distribution of final outcomes, and not the permutation of probabilities in different layers creating this end result.

We define a binary preference relation $\succsim$ on the space of simple lotteries, and assume the typical properties of continuity and rationality. Additional assumption is typically made to linearize the indifference curves on lotteries, namely independent axiom and its softer version, betweenness-axiom.

On this basis, the expected utility of a lottery can be assessed with linear (von Neumann-Morgenstern) function, weighting the Bernoulli utilities from certain outcomes with their probabilities. A rational agent should choose an outcome with highest expected utility, the payoff of an option weighted by its probability, regardless of how improbable the event is. Without some boundary conditions, this can lead to un-intuitive results. A humorous example of troubles with unbounded utilities comes from the inversion of Pascal's wager, the Pascal's mugging, where the payoff of highly improbable event grows faster than the probability diminishes [2].

Differing attitudes towards risk can be incorporated into the model by allowing Bernoulli utilities to have curvature. The risk aversiveness of an agent can then be characterized by the coefficient of relative (or absolute) risk aversion, $r_u(c) = -\frac{u''(c)}{u'(c)}$. This coefficient implies an underlying isoelastic utility function. Favourability of different lotteries can also be assessed by their expected utility and their volatility, where the latter is interpreted as a proxy for riskiness, coined First and Second order stochastic dominance, respectively.

Further, Kahneman and Tversky propose Prospect theory, which incorporates several features related to typical human behaviour with riskiness and biases of intuition, namely risk aversion, frame/state dependency and anchoring effects [3]. The analysis can also be extended to include subjective probability over the possibilities.

All of the modeling so far has been conditional on the assumption of rationality of preferences. However, the completeness assumption may not survive when we introduce the concept of Knightian uncertainty.

3 Forms of uncertainty

There are classes of randomness with highly different characteristics. Mathematician Benoit Mandelbrot characterizes these classes broadly as those exerting "mild" and "wild" randomness [4].

Standard examples of the first class are phenomena, which can be characterized by some well behaved Gaussian probability distribution, such as the distribution of heights of 20-30 year old Finnish males, or the behaviour of a molecule traveling in ideal gas. These phenomena can be completely characterized by their first (two) population moments, which can be reliably approximated from rather small representative samples, and the sample standard deviations can easily be addressed with a probability interpretation using the normality assumption.

Fat tails

By contrast, phenomena with wild randomness are hard to model reliably, as it could require larger samples than what is available, or infinitely many moments to characterise the behaviour of the distribution(s) from which the randomness comes from. The central "fat tailed" distribution exerting wild randomness is the Pareto distribution, with tails scaling in fractal manner. There even exists probability distributions which lack any finite moments, for example the Cauchy distribution. Figure 1 in the Appendix presents the walloping difference between random walks with normal and Cauchy distributed randomness. It might sometimes be hard to recognize a fat tailed distribution, as it can disguise as thin tailed distribution for some time. This also can be seen in the figure.

An example of problems when mistaking wild randomness as mild randomness comes from a dispute, whether there is reason to conclude that violence has declined in the world. Psychologist Steven Pinker presents an argument in favor of "The Long Peace", according to which the macro scale violence of wars and World Wars has permanently declined in the near past [5]. The conclusion is drawn using "thin-tailed" statistics, by noting that the percentage of people who has died in wars has, for first time in history, declined steadily during the past 100 years.

However, if a Third World War breaks lose, having potential to kill billions with modern technology, the fraction of casualties to world population could shoot off the charts. Instead of a Long Peace, it is possible that the Sword of Damocles hanging over our head has gotten larger, while hanging form fewer threds[1]. Neither conclusion can be reached in straight forward manner using Gaussian statistics.

Non-Ergodicity

Another challenge for using conventional probability estimates relates to non-Ergodic systems, where the time and ensemble probabilities have different statistical properties. The expected utility or harm is uncomputable in presence of absorbing barriers, as some of the outcomes can have infinite expected utility of harm. How much would you have to be compensated to play Russian roulette?

An example of a practical choice situation with non-Ergodicity is searching for a wife on the dating market (at least before the concept of marriage was watered down). The dating game ends with the choice, revealing only afterwards what the full payoff is. This choice also incorporates Knightian uncertainty: While choosing a wife, one also chooses a companion for older versions of himself and a mother for his offspring. How could a mere human even know what all to take into account when faced with such a hard choice?

[1] This phrasing is taken from an interview of mathematician Eric Weinstein

The unknown and unknowable

If the probabilities of all of the possible outcomes can be quantified, the risk involved can be represented with confidence or credible bands and other estimates. In contrast, with Knightian uncertainty there are no quantifiable information about some possible outcomes. Here the concepts of risk and Knightian uncertainty differ in qualitative manner, as former are something we can quantify and estimate, whereas the latter is truly unreachable. A related concept to Knightian uncertainty is a Black Swan event, which refers to highly improbable and highly influential event [6]. In presence of such events, it is not reasonable to label precaution and buffer-building "irrational". The burden of proof can be set on those introducing new option with potentially unquantifiable risks.

An interesting philosophical example about working in the limits of our knowledge spuns from the **anthropic reasoning**, which deals with observation selection effects. It is unclear, how we should reason about the probability of existence of other sentient life in the universe (or multiverse), given that we observe ourselves being sentient and alive. A naiive approach would be to calculate the odds of extraterrestrial life using for example the Drake's equation, by inserting subjective feeling about the probability, that a life permitting planet actually gives rise to life. Regardless of how probable or improbable the life originating process is, if it happens somewhere, that is where a sentient being will find herself [2] in any case. Instead, a meta-probabilistic axiom must first be acknowledged, which is influencing the actual probabilistic reasoning.

Two sensible axioms can be proposed, which in some cases lead to opposite conclusions. **The Self Sampling Argument** states, that we should regard ourselves as a random sample from a suitable reference class, say human-like creatures. Then it would be more probable, that the particular being, ME, was chosen from smaller rather than larger set of possibilities. On the other hand, **the Self Indication Assumption** states, that we should regard ourselves as a random sample from the set of all possibilities. From this point of view, we should expect there to be more instead of less of something that is being chosen at random from larger pool of possibilities. [7]

Crucial considerations

The previous example of anthropic reasoning shows, that in some cases the probability estimate is deeply dependant on the choices on meta-level. But there are also situations, where the decision can change drastically even without changes on the meta-level, by new information called *crucial considerations*[3].

For example, should I continue studies to pursue a PhD?

It takes years of time and effort, which could be used to earn money and gain experience on the real world.

- don't pursue PhD.

Studying further enables the pursuit of knowledge and truth in absence of mundane responsibilities of hectic and repetitive work.

[2]Anthropomorhpising not intended

[3]The term relates to effective altruism and derives back to Nick Bostrom

- pursue PhD.

Staying further in the academic bubble only blinds from the truth with autistic theories and models detached from the reality, which one is forced to accept to get a tenure.

- don't pursue PhD.

As a fallible human being, how do I know in advance what is reasonable or true (in some sense)? I might be able to position myself better by studying further, before starting to separate the wheat from chaff.

- pursue PhD.

...

Not only does the new information influence the choice, but it changes the policy choice completely. How can I be sure that my reasoning has exhausted all of the crucial considerations, and I should stick with the action indicated by the last step I took?

4 Dealing with uncertainty

Modeling approach

Incomplete preferences can be modeled by adding confidence levels for beliefs of agents regarding choices. Instead of assigning point estimates for the probabilities of events, the unanimity multi-prior model assigns sets of probabilities for options, making a move in the Bayesian direction.

In absence of completeness, agents can be modeled to favor the status quo. The focus can also be restricted on those options which there are enough information about. Another possibility is to incorporate all options into the analysis, regardless of how great the (quantifiable) uncertainty about them is. [8]

Some rare events can be addressed with approximate probabilities based on data. The risk can be estimated by comparing the total volume to failed events. This number can be expressed in *micromort* units, which denote one in a million chance of dying. For example, there were 48 million skydives in the US between 2000-2016, 413 of which ended in a death of the jumper. Hence, going skydiving increases the risk of death by 8 micromorts per jump.

Heuristic approach

How to function in a world which is too hard to understand or predict? We come into existence with tremendous a priori biological structure influencing the way we perceive the world, feel about social situations and conduct ourselves. On top of that, upbringing, culture and traditions can be seen as means of transmitting software on new versions of the survival machine. This software has evolved through times, and more importantly, it has been able to survive.

As fallible humans, we can't allow ourselves to rely on mere reason. From American pragmatist epistemology, time can be seen as the most appropriate meta-critic of ideas and practices - what seems to work is true enough. It is plausible, that a surviving institution has contributed to the survival and success probability of the vehicle (society/culture) it has been riding on. It is, nonetheless, not straight forward how to attach ethics to this epistemology, as many surviving institutions have lost their moral legitimacy in the light contemporary vision, for example the slavery. Naturally we also need change in order for the adaptation to new environments to be possible.

From this point of view, also ethical intuitions can be seen as evolutionary adaptations on the level of group selection, enabling us to overcome hard game theoretic control problems [9]. Our symbolic interpretations of the reality, encapsulated in mythology, can be seen as mirroring the (Jungian) archetypes of human experience [10]. Religion and tradition can hence also be viewed as collections of risk heuristics [11].

Heuristics can provide ways to robustly function on domains with Knightian uncertainty, Black Swans and non-Ergodicity. This should be kept in mind, when labelling violations of expected utility or other simplistic theories as "fallacies".

References

[1] A. Mas-Colell, Microeconomic theory, Oxford University Press, New York, 1995.

[2] N. Bostrom, Pascal's mugging, Working Paper.

[3] D. Kahneman, Thinking, fast and slow, Allen Lane, London, 2011.

[4] B. Mandelbrot, Misbehavior of Markets, Fractal view on Risk, Ruin and Reward, Vol. 21, 2006.

[5] S. Pinker, The better angels of our nature : the decline of violence in history and its causes, Allen Lane, London, 2011.

[6] N. N. Taleb, The black swan : the impact of the highly improbable, Penguin Books, London, 2008.

[7] N. Bostrom, Anthropic bias : observation selection effects in science and philosophy, Studies in philosophy (New York, N.Y.), Routledge, New York, 2002.

[8] B. Hill, Incomplete preferences and confidence, Journal of Mathematical Economics 65 (C) (2016) 83–103. doi:10.1016/j.jmateco.2016.05.

[9] J. Haidt, The righteous mind : why good people are divided by politics and religion, Pantheon Books, New York, 2012.

[10] J. B. Peterson, Maps of meaning : the architecture of belief, Routledge, New York, 1999.

[11] N. Taleb, Religion, heuristics, and intergenerational risk management, Econ Journal Watch.

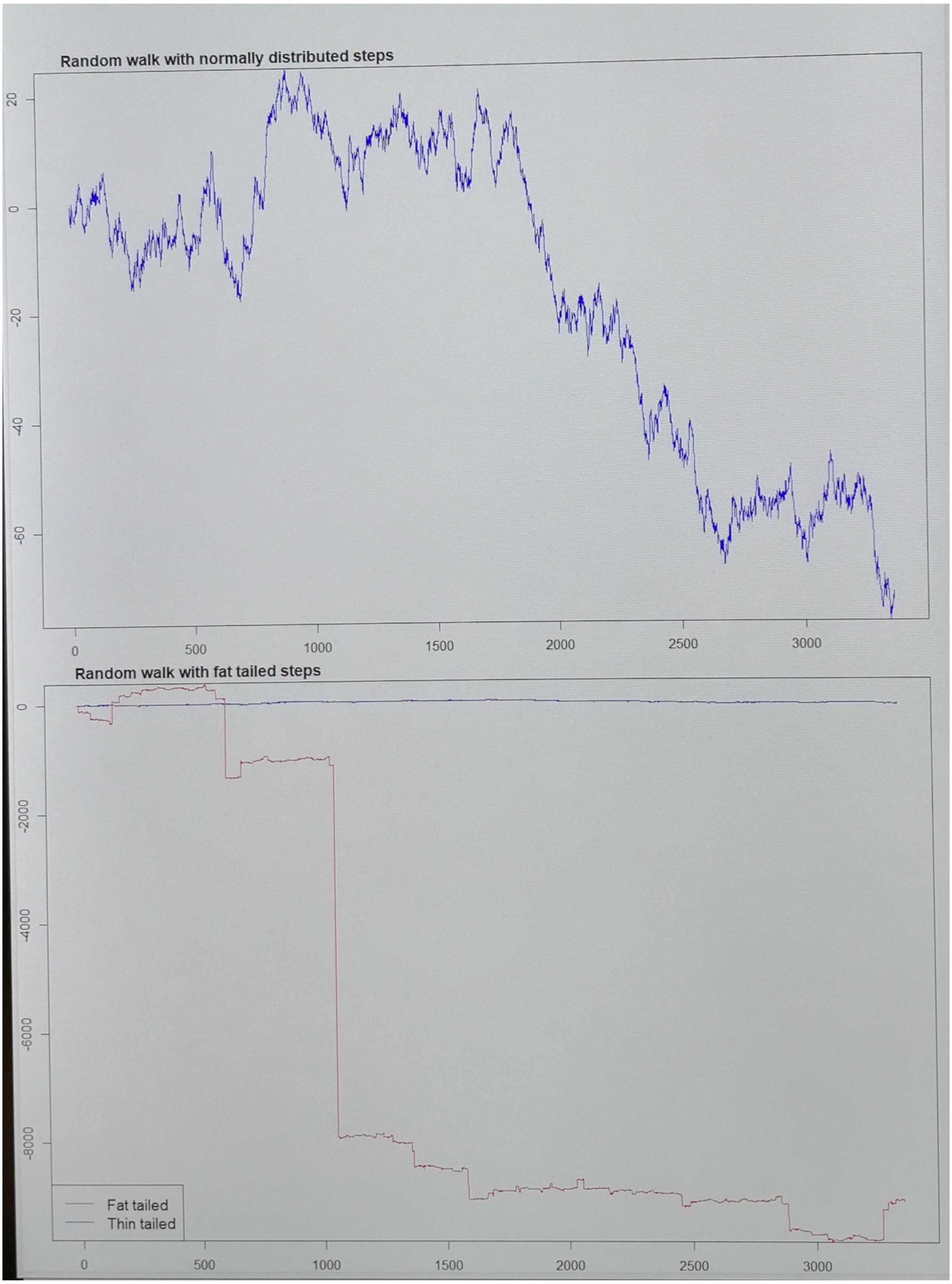

Figure 1: Difference between random walks with normal and Cauhcy-distributed (t-distribution with df=1) steps. The blue line depicting normal RW is the same in both plots.