Uncertainty, Probability and Robust Preferences

Yakov Ben-Haim
Yitzhak Moda'i Chair in Technology and Economics
Faculty of Mechanical Engineering
Technion — Israel Institute of Technology
Haifa 32000 Israel
http://www.technion.ac.il/yakov
yakov@technion.ac.il

Abstract This essay argues that uncertainty and probability are different, and that cogent reasoning about uncertainty need not rely on probabilistic premises. This is particularly important when probabilistic evidence is unavailable such as in unique or rapidly changing situations. Probability is very informative and useful, when probability models are available. But it is too highly structured and information-intensive for use in situations which are epistemically poor. Such situations abound in practical policy analysis and decision making. If cogent thought required probabilistic premises we would either have to forego cogency or assume probabilistic propositions arbitrarily (which might itself lack cogency). New formal arguments are developed in response to three specific questions: (1) Does a non-probabilistic robust preference between policy options need to assume a uniform probability distribution on the underlying uncertain events? (2) Does a robust preference assume *some* probability distribution on the uncertain events? (3) Are any probabilistic assumptions needed in order to justify robust preferences?

Keywords: Uncertainty, probability, robustness.

Can one reason cogently about uncertainty without invoking, at least implicitly, probabilistic ideas? For most people—including decision analysts, topical specialists, and philosophers—the answer to this question is fairly self evident. The curious thing is that many people answer Yes and many others answer No.

This essay will be an attempt to argue that uncertainty and probability are different, and that cogent reasoning about uncertainty need not rely on probabilistic premises. This is particularly important when probabilistic evidence is unavailable such as in unique or rapidly changing situations. Probability is very informative and useful, when probability models are available. But it is too highly structured and information-intensive for use in situations which are epistemically poor, characterized by severely deficient information. Such situations abound in practical policy analysis and decision making. In those situations, if cogent thought required probabilistic premises we would either have to forego cogency or assume probabilistic propositions arbitrarily (which might itself lack cogency).

In this essay I will address three questions: (1) *Does a non-probabilistic robust preference between policy options need to assume a uniform probability distribution on the underlying uncertain events?* It is a common mis-conception that the answer to this question is affirmative. We will dispose of this mis-understanding with a simple deductive argument showing that, while a uniform distribution would motivate robust preferences, so would many other probability distributions. This leads to the next question. (2) *Does a robust preference assume* some *probability distribution on the uncertain events?* We will find that an infinity of higher-order probability beliefs would motivate robust preferences, but belief in a specific probability distribution on the uncertain events is not required. This leads to the more generic final question. (3) *Are any probabilistic assumptions needed in order to justify robust preferences?*

The answer to all these questions will be No, non-probabilistic robust preferences do not depend upon probabilistic assumptions for their plausible justification. However, I must stress that I do not view all of my arguments as incontrovertible. I do not believe that my reasoning must convince every

[0]\papers\Uncer-Prob\up03.tex 15.6.2011.

reasonable person on all points. Some of my arguments are deductive (and thus incontrovertible), but some are arguments from pragmatic plausibility. These arguments are nonetheless relevant since we must be plausibly pragmatic in many practical decisions contexts.

Two views on uncertainty and probability. Before getting to these questions I will very briefly trot out the two main positions on the status of probability in decision making. The first position—represented here by Keynes and Carnap—views probability as fundamental to all thought about uncertainty. The second position—represented here by Knight and Wald—rejects this. There are of course many differences between the four thinkers whom I quote, and many others could be quoted. The point is that consensus on the status of probability is not available and very serious thinkers have taken both sides of the debate.

John Maynard Keynes asserts:

> Part of our knowledge we obtain direct; and part by argument. The Theory of Probability is concerned with that part which we obtain by argument, and it treats of the different degrees in which the results so obtained are conclusive or inconclusive. ...
>
> The method of this treatise has been to regard subjective probability as fundamental and to treat all other relevant conceptions as derivative from this. [8, pp.3, 281–282]

Among Rudolf Carnap's "basic conceptions" is the contention that

> all inductive reasoning, in the wide sense of nondeductive or nondemonstrative reasoning, is reasoning in terms of probability. [5, p.v]

In the other camp Frank H. Knight stressed the distinction between probabilistically measurable uncertainty, which he called risk, and unmeasurable uncertainty, which he referred to as " 'true' uncertainty" [10, p.20]. It is the latter, rather than risk, which Knight saw as underlying the greatest challenges to economic agents and analysts:

> Business decisions ... deal with situations which are far too unique, generally speaking, for any sort of statistical tabulation to have any value for guidance. The conception of an objectively measurable probability or chance is simply inapplicable. ...
>
> It is this *true uncertainty* [Knight's italics] which by preventing the theoretically perfect outworking of the tendencies of competition gives the characteristic form of 'enterprise' to economic organization as a whole and accounts for the peculiar income of the entrepreneur. [10, pp.231–232]

Taking a more extreme position, Abraham Wald wrote that

> in most of the applications not even the existence of ... an a priori probability distribution [on the class of distribution functions] ... can be postulated, and in those few cases where the existence of an a priori probability distribution ... may be assumed this distribution is usually unknown. [17, p.267]

Much has been written since the contributions of these four scholars. They nonetheless define the dispute—current to this day—with which this essay grapples.

Robust preferences. Robustness to uncertainty can be defined in many different ways, both probabilistically and non-probabilistically. In this essay I will refer to non-probabilistic robust preferences among options which are based on a set-theoretical representation of uncertainty. That is, uncertainties are characterized by a set, or family of sets, of possible contingencies, without explicitly formulating (or admitting to) any probability measure on these uncertainty sets. Specifically, the robustness might be a worst-case or min-max analysis as studied by Wald [17], Schweppe [14], Hansen and Sargent [7] or many others, or it might be an info-gap analysis [2]. Our discussion will focus on these two non-probabilistic concepts of robustness.

In a worst-case analysis the robustness of a decision is a measure of how bad the outcome would be if the worst contingency in a specified set of possibilities occurs. In an info-gap analysis a worst case is not known but a worst acceptable outcome is specified. The info-gap robustness of a decision is a measure of how large a set guarantees an acceptable decision. These two strategies are, epistemically, opposite sides of the same coin. (For further discussion of the relation between min-max and info-gap robust-satisficing see [4]).

Suppose we must choose between two options, B and C, whose robustnesses have been evaluated by some explicit procedure. Suppose that B is more robust than C. We then say that the robust preference is for B over C, denoted $B \succ_{\mathrm{r}} C$.

First question. *Does the justification of this robust preference need to assume that the probability distribution of the uncertain events is uniform?*

A uniform distribution cannot exist on the set of uncertain events if this set is of infinite extent (e.g. the entire real line or plane). In this case the answer to the first question is clearly No. If there is any justification for the robust preference, it is not a belief in the uniform distribution. We will consider the non-existence of probability distributions when we come to the third question. For now we will assume that the relevant uniform distribution exists.

In the info-gap case, the argument for this claim—that an underlying uniform distribution is necessary to justify the robust preference—is that a uniform distribution would imply that B is more likely than C to yield the required outcome. (This is because B is adequate over a larger set than C.) It is true that an underlying uniform distribution motivates a robust preference in precisely this way. The converse however—robust preference requires a uniform distribution—is not true.

The proof is that *any* probability distribution on the uncertain events for which B is more likely than C to yield the required outcome will motivate the robust preference. There are an infinity of such probability distributions. Thus, if we observe a decision maker making a robust preference—and let's suppose that the DM declares that the procedure was an info-gap robustness analysis—we cannot conclude that the DM is assuming a uniform distribution; the probabilistic assumption (if there is one) could just as well be any of an infinity of possible probability distributions.

Monty Hall's 3-door problem is an example of the arguments of this sort—and of the public as well as professional disputes which they can engender [1, 16].

In the min-max case the argument is similar. If B is more robust than C, then C leads to a worse outcome than B on the set of contingencies. Any subset of contingencies on which C is no worse than B, is smaller and thus less likely, given the uniform distribution. However, the uniform distribution is not the only probability distribution with this property. There are an infinity of such probability distributions, any of which could provide probabilistic justification for the robust preference.

In short, the answer to the first question—Is a uniform distribution necessary for justifying a robust preference?—is No.

Second question. *Is it necessary (in order to justify the robust preference) to assume the validity of one among the infinity of probability distributions which imply that B is more likely than C to yield the required outcome?*

Such an assumption would justify the robust preference, but the robust preference does not require this assumption, as I now explain. I will introduce some notation in order to enhance clarity.

u is an underlying uncertain event for which there is a set-model of uncertainty with respect to which the robustnesses of B and C are calculated.

$p(u)$ is a probability density on the u's (or a probability measure if the u's are functions, etc.). The set of all such $p(u)$'s is S. We will assume that there is some $p(u)$ in S which actually describes the probabilistic properties of u. We will refer to this $p(u)$ as the "true" distribution, and denote it by $p_{\mathrm{T}}(u)$. It is $p_{\mathrm{T}}(u)$ which some folks might claim needs to be uniform in order to justify the robust preference.

S_B is the subset of S containing all $p(u)$'s for which B is more likely to succeed than C. (The precise form of S_B is likely to be different if one is doing a min-max or an info-gap analysis, but the interpretation is the same.)

The question we are asking is: Is it necessary to assume that the true distribution, $p_{\mathrm{T}}(u)$, belongs to S_B, in order to justify the robust preference?

It is correct that if the true $p(u)$ belongs to S_B then the robust preference is justified probabilistically, which we denote:

$$p_{\mathrm{T}} \in S_B \quad \Longrightarrow \quad B \succ_{\mathrm{r}} C \tag{1}$$

The '$\Longrightarrow$' means 'justifies' or 'warrants' or 'motivates'. Thus the left hand proposition provides justification or warrant or motivation for the righthand proposition.

However, the converse does not hold. That is, acceptance of the robust preference, $B \succ_{\mathrm{r}} C$, does not require a reasonable and thoughtful decision maker to believe that $p_{\mathrm{T}} \in S_B$.

It would be sufficient for the decision maker to believe that it is *more likely* that $p_{\mathrm{T}} \in S_B$ than that $p_{\mathrm{T}} \notin S_B$, in order to probabilistically justify the robust preference. That is, a probabilistic motivation for robust preference can be described as:

$$\mathrm{Prob}(p_{\mathrm{T}} \in S_B) > \frac{1}{2} \quad \Longrightarrow \quad B \succ_{\mathrm{r}} C \tag{2}$$

"Prob" represents a probability judgment on the space of probability densities $p(u)$.

The degree of warrant of the "$\Longrightarrow$" in eq.(2) is perhaps less than the degree of warrant in eq.(1), but still relevant.

But once again, the converse of (2) need not hold. It would be sufficient for the decision maker to believe that it is more likely that $\mathrm{Prob}(p_{\mathrm{T}} \in S_B) > \frac{1}{2}$ than that $\mathrm{Prob}(p_{\mathrm{T}} \in S_B) \le \frac{1}{2}$. That is, one could claim:

$$\mathrm{Prob}\left(\mathrm{Prob}(p_{\mathrm{T}} \in S_B) > \frac{1}{2}\right) > \frac{1}{2} \quad \Longrightarrow \quad B \succ_{\mathrm{r}} C \tag{3}$$

The outer "Prob" is a probability judgment on the space of the inner probability judgments introduced in eq.(2).

This regression can go on for as long as one likes (and even longer!). There are an infinity of probability beliefs which could, to some degree, plausibly motivate the robust preference. That is, one could claim:

$$\mathrm{Prob}\left(\ldots\left[\mathrm{Prob}\left(\mathrm{Prob}(p_{\mathrm{T}} \in S_B) > \frac{1}{2}\right) > \frac{1}{2}\right]\right) > \frac{1}{2} \quad \Longrightarrow \quad B \succ_{\mathrm{r}} C \tag{4}$$

But the converse is not necessary at any step.

In short, we see that any one of an infinity of probability beliefs would justify or motivate the robust preference, at least to some extent. Higher-order probability statements might provide weaker motivation for the robust preference than lower-order statements. However, we also see that, by a regression argument, none of any finite sequence of such beliefs is necessary in order to justify the robust preference. At any step, the probabilistic belief can be deferred in favor of the next higher-order belief, which in turn, etc.

Our answer to the second question—Is eq.(1) necessary to motivate robust preferences?—is clearly negative.

Third question. *Is at least one probability belief, from among the infinite sequence of beliefs in eq.(4), necessary in order to justify the robust preference?*

I will argue that the answer is No. I will invoke a sort of pragmatism or common sense to suggest that it is more reasonable to choose the more robust over the less robust choice, even if we cannot make any probability claims at all. That is, even if I can adopt none of the probability beliefs in eq.(4), I would still adopt the robust preference. This will be an intuitive argument of plausibility, and not an indisputable apodictic deduction.

Let me begin with a quotation from Locke [12, p.57, I.i.5]:

> If we will disbelieve everything, because we cannot certainly know all things; we shall do muchwhat as wisely as he, who would not use his legs, but sit still and perish, because he had no wings to fly.

To apply this to our situation, if we disbelieve (or at least cannot justify belief in) any of the probabilistic propositions in eq.(4), we can still justify a robust preference, perhaps viewing it as an epistemic last resort. To reject even the robust preference would paralyze the decision maker, and be 'muchwhat as wise' as Locke's wingless gentleman.

It is important to note that the situation is sometimes more challenging than just being able (or not) to justify at least one of the propositions in eq.(4). There need not be *any* probability distribution on the underlying uncertain events, or any of the subsequent higher-order probabilities. We quoted Wald [17, p.267] earlier that "in most of the applications not even the existence of ... an a priori probability distribution ... can be postulated". All probability models—Bayesian, frequentist, subjective—impose a specific structure on the universe of events as defined formally by the Kolmogorov axioms [11]. Not all situations are consistent with this structure.

In situations where one holds no probabilistic beliefs, due to lack of knowledge or non-existence of the probability distributions, are we unable to justify a robust preference? The answer is No. Ignorance or non-existence of a probability distribution means that we cannot know or predict the probability (or likelihood or frequency) that B or C will succeed. This does not mean that neither B nor C will succeed, nor that there is no future probability of success. It simply means that no *a priori* prediction of that future probability can be made.

The imperative to make a decision—if we are not to perish from epistemic paralysis—remains even if we cannot make up our minds about one of the propositions in eq.(4). The imperative to decide remains also when none of the propositions are available since the probability distributions simply don't exist. So the "bottom line" is this.

- Many probabilistic beliefs justify the robust preference to some degree.

However:

- Evidence for any of these probabilistic beliefs may be unavailable.
- The underlying probabilities may, in some situations, simply not exist.

The upshot is that we need to expand our conception of uncertainty—and our ability to think about uncertainty—beyond the context of probabilistic thinking. Set-models of uncertainty, such as are used in worst-case analysis and in info-gap theory, provide a framework for doing this, as do fuzzy logic [9], p boxes [6] and imprecise probability [18] though in very different ways and with much more of a probabilistic flavor. On a more epistemological level, Shackle [15, pp.3–4, 156, 239, 401–402] and Popper [13, p.63] each advanced arguments for indeterminism, indicating the usefulness of non-probabilistic thinking about uncertainty (see also the discussion of Shackle-Popper indeterminism in [2, 3].

But still, it *is* logically possible that in either case—either one can't believe any of the probability statements in eq.(4) or the probability distributions don't exist—one must also reject the robust preference. The argument for robust preferences based on pragmatism or common sense—like Locke's story—is an intuitive rather than a deductive argument.

In the absence of a probabilistic justification should we adopt the robust preference only out of fear of epistemic paralysis? No. Probability is one reflection of the world; it is a model of, and a way of thinking about, the world. Set-models of uncertainty, such as p boxes or info-gap models, are also representations of, and ways of thinking about, the world. Set-models of uncertainty are, roughly speaking, less informative than probabilistic models. This means that we can expect less informative results if we lack probability models. We may be wrong more often if we can't probabilize. But this is not to say that we should not use what we have, even (or especially) when we are epistemically poor. A non-probabilistic set-model of uncertainty is useful because it reflects an aspect of the world's uncertainty, and usefulness is the quintessential pragmatic justification.

Acknowledgements. The author is indebted for useful comments to Mark Burgman, Yohay Carmel, Scott Cogan, Maria Demertzis, Yoshihiro Kanno, Mick McCarthy, Andrew Robinson and Barry Schwartz.

References.

1. Ben-Haim, Yakov, 1996, *Robust Reliability in the Mechanical Sciences,* Springer-Verlag, Berlin, section 7.1.
2. Ben-Haim, Yakov, 2006, *Info-Gap Decision Theory: Decisions Under Severe Uncertainty,* 2nd edition, Academic Press, London.
3. Ben-Haim, Yakov, 2007, Peirce, Haack and Info-gaps, in *Susan Haack: The Philosopher Responds to Critics,* edited by Cornelis de Waal, 2007, Prometheus Books.
4. Ben-Haim, Yakov, Clifford C. Dacso, Jonathon Carrasco and Nithin Rajan, 2009, Heterogeneous Uncertainties in Cholesterol Management, *International Journal of Approximate Reasoning,* 50: 1046–1065.
5. Carnap, Rudolf, 1962, *Logical Foundations of Probability,* 2nd ed., University of Chicago Press.
6. Ferson, Scott, 2002, *RAMAS Risk Calc 4.0 Software: Risk Assessment with Uncertain Numbers,* Boca Raton, Lewis Publishers.
7. Hansen, L.P., and Sargent, T.J., 2008. *Robustness,* Princeton University Press, Princeton and Oxford.
8. Keynes, John Maynard, 1929, *A Treatise on Probability,* Macmillan and Co., London.
9. Klir, George J. and Tina A. Folger, 1988, *Fuzzy Sets, Uncertainty, and Information,* New York, Prentice-Hall.
10. Knight, Frank H., 1921, *Risk, Uncertainty and Profit.* Houghton Mifflin Co. Re-issued by University of Chicago Press, 1971.
11. Kolmogorov, Andrei N., 1956, *Foundations of the Theory of Probability,* translation edited by N. Morrison, New York: Chelsea Publishers.
12. Locke, John, 1706, *An Essay Concerning Human Understanding,* 5th edition. Roger Woolhouse, editor. Penquin Books, 1997.
13. Popper, Karl R., 1982, *The Open Universe: An Argument for Indeterminism,* edited by W.W.Bartley III, Hutchinson & Co., London.
14. Schweppe, F.C., 1973. *Uncertain Dynamic Systems,* Prentice-Hall, Englewood Cliffs.
15. Shackle, G.L.S., 1972, *Epistemics and Economics: A Critique of Economic Doctrines,* Transaction Publishers, 1992, originally published by Cambridge University Press.
16. Terney, J., 1991, Behind Monty Hall's doors: Puzzle, debate and answer, *New York Times,* 21 July 1991, section I, p.1.
17. Wald, A., 1945. Statistical decision functions which minimize the maximum risk, *Annals of Mathematics,* 46(2), 265–280.
18. Walley, Peter, 1991, *Statistical Reasoning with Imprecise Probabilities,* London, Chapman and Hall.