Non-Bayesian Decision Theory. Beliefs and Desires as Reasons for Action, Martin Peterson. Theory and Decision Library, Springer, 2008, ix + 170 pages.

Anyone concerned in the foundations of normative decision theory and who feels puzzled by the way this theory typically proceeds will doubtless welcome Martin Peterson's *Non-Bayesian Decision Theory*. It offers a *prima facie* intriguing view of rational choice under uncertainty: a "non-Bayesian" defense of subjective expected utility theory. The view is intriguing because for most of people, both inside and outside the field, "Bayesian decision theory" and "subjective expected utility" (henceforth SEU) are taken to be interchangeable. Actually, Peterson confers a very specific meaning to the term "Bayesian". According to him, the Bayesian account of SEU consists in viewing subjective probability (on the relevant state space) and utility (on the outcomes) as *defined* in terms of (or *determined* by, or *generated* from) preferences over uncertain prospects. By contrast, a non-Bayesian account of SEU views subjective probability and utility as given antecedently to preferences over uncertain prospects, and rational choice under uncertainty as the choice of an uncertain prospect which maximizes the expected utility induced by these antecedently given probability and utility. The non-Bayesian challenge raises four basic questions:

(Q1) Why is the Bayesian view unsatisfactory? (*The case against Bayesianism*)

(Q2) Where does utility "come from"? (*The non-Bayesian analysis of utility*)

(Q3) Where does subjective probability "come from"? (*The non-Bayesian analysis of subjective probability*)

(Q4) Given some antecedently given utility and subjective probability, why maximize subjective expected utility? (*The justification of the expected utility criterion*)

Peterson's book articulates systematic answers to each of these four basic questions. The current review will deliberately focus on them, neglecting parts of the book which are less immediately linked to the non-Bayesian challenge like, for instance, Chapter 8 which is devoted to strongly risk averse decision rules. Each issue will be examined in turn, with an overview of the book's contributions and some critical comments thereafter.

The case against Bayesianism. Basic SEU models (Anscombe and Aumann, 1963 ; Savage, 1954/1972 ; Jeffrey, 1965/1983) and the so-called Bayesian perspective on SEU are reviewed and discussed in Chapter 2. According to Peterson, Bayesianism is motivated by a skeptical stance towards the introspective access to subjective probability and utility. Both have rather to be ascribed from agents' preferences over uncertain outcomes. An axiomatization for SEU is a set of properties (or axioms) on these preferences such that, if agents obey the axioms, they *can* be seen as conforming to SEU. This means that there exists a utility function and a subjective probability such that, for all prospects *a*, *b*, the agent prefers *a* to *b* iff the expected utility of *a* is greater than the expected utility of *b*. Furthermore, an axiomatization for SEU is typically accompanied by an abstract way of eliciting subjective probability and utility for people obeying the axioms. Bayesians welcome this way of having access to utility and probability from preferences over uncertain prospects as a crucial improvement on introspection. But for Peterson, this is a fatal weakness. According to him, Bayesian decision theory "puts the cart before the horse from the point of view of the deliberating decision maker" and is not "action guiding" (pp. 26-30). The reason is that, for a decision theory under uncertainty to be action guiding, it has to produce as *output* a set of preferences over uncertain prospects. But from the Bayesian point of view, these preferences are already given at the very beginning – and under appropriate conditions determine a subjective probability and a utility.

Arguably, the normative and prescriptive import of the contemporary axiomatic approach to SEU requires clarification. But it seems to me that Peterson's argument is too hasty. Firstly, he considers that Bayesians *define* utility and probability in terms of preferences over uncertain outcomes. But the axiomatic approach does not commit them to do that (see e.g., Joyce 1999). A Bayesian can *estimate* an agent's utility and probability and, given those estimates, give *reason* to choose prospect *a* rather than prospect *b*. Second, Peterson seems to consider that norms on preferences cannot by themselves be action guiding. But, for instance, transitivity can be seen as giving a reason, for an agent who prefers *a* to *b* and *b* to *c*, to prefer *a* to *c*. In view of the two preceding remarks, it seems dubious to consider that Bayesian decision theory takes as input the agent's preferences over uncertain prospects and produces as output the agent's utility and probability. It is more accurate to say that it takes as input *some* preferences over uncertain prospects and *can* produce as output some other preferences over uncertain prospects. Most of so-called *decision analysis* for risk and uncertainty proceeds roughly in this way. An analogy may be useful as well: logic seems to be belief guiding. But to produce (non-tautological) beliefs, it needs to be "fed" by others beliefs. One might perhaps claim that Bayesian decision theory requires "too many" preferences over uncertain outcomes as input. But this issue is not addressed by Peterson's case.

The non-Bayesian analysis of utility. Chapters 4 and 5 provide a non-Bayesian analysis of utility as deriving from preferences over *certain* outcomes (sometimes called "riskless utility" in the decision-theoretic literature). This account is based on a theory of indeterminate preferences (Chapter 4) that consists mainly in (i) a model of degrees of preference and (ii) an interpretative claim to the effect that for an agent *i* to have a degree of preference *p* for *a* over *b* is for *i* to believe to degree *p* that she would choose *a* rather than *b*. We may call this claim the *doxastic conception of preferences*. (i) Three properties are attributed the relation $>_p$ ("…is preferred to degree *p* to…"). They are assumed to be the counterparts of the usual properties of reflexivity, symmetry and transitivity. The less obvious one is P-transitivity according to which, if $x >_p y$ and $y >_q z$, then $x >_r z$ where $r = (p.q) / [(p.q) + (1 – p) . (1 – q)]$. P-transitivity is justified by Luce's Choice Axiom for probabilistic choice according to which for two sets of options *A*, *B*, such that $A \subset B$, the probability $p(x > \mathrm{B})$ that an option *x* is chosen from *B* is the product of the probability $p(x > A)$ that this option *x* is chosen from *A* and the probability $p(A > B)$ that *some* option from *A* is chosen when the individual chooses from *B*. P-transitivity is entailed by the Choice Axiom and the probability calculus. Peterson grants that the Choice Axiom is not universally valid (specifically, when the options are not comparable) and considers that the domain of his non-Bayesian decision theory coincides with the domain of validity of the Choice Axiom. The Choice Axiom, in line with the doxastic conception of preferences, is viewed as bearing on the agent's partial beliefs on what he would choose between pairs of options. This leads us to (ii). The construal of degrees of preference as degrees of belief is supposed to overcome the difficulties of two views on preferences: the "revealed preference" view, and "internalism" according to which preferences are "mental states that trigger choices". The end of Chapter 4 defends carefully this view against Spohn's and Levi's arguments to the effect that an agent cannot coherently assign probabilities to her own choices. Chapter 5 elaborates on the preceding account of indeterminate preferences to offer a non-Bayesian analysis of utility derived from (degrees of) preferences over certain outcomes. Peterson claims to improve on Luce's representation theorem according to which, for a finite set of options *B* if the Choice Axiom holds *and if for no pairs of option x, y, p(x > y) is 0 or 1*, there exists a utility function *u* that is unique up to multiplication by a positive constant (a ratio scale) such that for all $A \subset B$, $p(x > A) = u(x) / \Sigma_{y \in A}\, u(y)$. The weakness in this representation theorem is of course the condition that for no pairs of option *x*, *y*, $p(x > y)$ is 0 or 1. The author's proposal is to require instead that there exists a "non-perfect object" x^* such that $p(x^* > x) \neq 0, 1$ for every $x \in B$. The non-Bayesian analysis of utility is therefore that the ratio of utilities of two outcomes *x*, *y* is the ratio of the degree to which the agent believes that he will choose *x* rather than *y* on the degree to which the agent believes that he will choose *y* rather than *x*.

Some comments are in order. On the one hand, the whole book is motivated by a concern to provide to the decision maker true *reasons* to choose. On the other hand, the degree of preference for *x* over *y* is analyzed as the degree to which the agent believes that she would *choose x* rather than *y* – not the degree to which she believes that *x* is *better* than *y*. It is therefore a very peculiar kind of doxastic conception of preferences. And it is unclear to me, from a conceptual point of view, that these beliefs about one's own choices, together with one's beliefs about the environment, are sufficient to provide good reason for acting. One can see some advantages of this conception: it allows one to assume that degrees of preferences obey the probability calculus and therefore to apply Luce's representation theorem. The argument for it is mainly negative, emerging from the denial of the revealed preference view and internalism. The discussion of these views is valuable and stimulating, but internalism seems to me too quickly dismissed (and sometimes accompanied by disputable claims on mental states, e.g., the claim that the common wisdom nowadays is that "mental states are internal and readily accessible through introspection", p. 72) to give us powerful reason to see degrees of preferences as subjective beliefs. Lastly, given that one of the basic motivations of the theory of indeterminate preferences is the apparent existence of incomparable options, one may be a little worried by the fact that the author grants that "the choice axiom is invalid in decisions among incomparable objects" (p.74).

The non-Bayesian analysis of subjective probability. Chapter 6 develops a non-Bayesian analysis of subjective probability. This analysis starts from approaches based on qualitative (or comparative) probabilities. A qualitative probability is a relation > defined on an algebra of events: $X > Y$ means that the agents judges that *X* is more likely than *Y*. The author recalls a set of axioms on > proposed by DeGroot (1970) that is necessary and sufficient for the existence of a probability distribution *p* on the relevant algebra of events such that $X \geq Y$ iff $p(X) \geq p(Y)$. Peterson is not completely satisfied with this approach because he would like "a procedure for linking subjective probability to observable choice". His idea is to ask the agent for her choices among "horse-race lotteries": a horse-race lottery for event *X* is a lottery such that the agent wins some fixed amount of money if *X* is the case, and nothing if it is not. Since the amount of money is fixed across different horse-race lotteries, it is assumed that, if the agent chooses the lottery associated to event *X* rather than the one associated to event *Y*, one can safely infer that the agent judges *X* more likely than *Y*. And if one imposes on the agent's choices among horse-race lotteries the very same axioms than the one proposed by DeGroot, then one obtains a probability distribution on the relevant algebra.

The variation introduced on DeGroot's axiomatization of comparative probabilities seems to bring the proposal closer to the Bayesian way of eliciting degrees of belief since it consists mainly on a set of axioms on choices over uncertain prospects (the horse-race lotteries). (Recall that in Savage's decision theory, the elicitation of comparative probability runs as follows: pick two outcomes *x* and *y* such that the agent strictly prefers *x* (for sure) to *y* (for sure). Then ask the agent whether he would prefer the act [*x* if *X* ; *y* if not] to the act [*x* if *Y* ; *y* if not]. If this is the case, then one can infer that the agent judges *X* more likely than *Y*.) The benefit of such a variation is therefore unclear to me. Another point to stress is that since the author sees degrees of preference as degrees of belief about one's own choices, the procedure outlined in Chapter 6 for eliciting degrees of belief should also in principle be used to estimate degrees of preferences (and, ultimately, utility). One may wonder whether this cannot induce some unfortunate interferences. (It seems that I may have interest to choose [100$ if I choose *x* rather than *y* ; 0 if not] rather than [100 $ if it rains ; 0 if not] even if I am very confident that it will rain tomorrow and believe moderately than I would choose *x* rather than *y*: after all, it's up to me to choose *x* rather than *y* and therefore to win the first horse-race lottery.)

The justification of the expected utility criterion. The question of the justification of the expected utility criterion is addressed in Chapter 7. This justification is achieved by two axiomatizations of the criterion based on the investigation of choice rules made in Chapter 3. Let us recall briefly the concepts introduced in this Chapter. A formal decision problem $\pi = \langle A, S, P, U\rangle$ is a quadruple whose elements are a set *A* of acts, a set *S* of states, a set *P* of probability distributions and a set *U* of utility functions. In a formal decision problem "under risk", *P* and *U* are singletons. A *transformative* decision rule **t** on a set of decision problems Π is a function such that for all $\pi \in \Pi$, $\mathbf{t}(\pi) \in \Pi$. An *effective* decision rule **e** on Π is a function such that $\mathbf{e}(\langle A, S, P, U\rangle) \subseteq A$. The criterion of SEU can be seen as the composition of (i) a transformative rule **weigh** which transforms a decision problem under risk π in a decision problem under certainty π' with the same set of (primitive) acts and with a utility function which assigns to each act its expected utility (according to π) ; and (ii) an effective rule **max** which, in a decision problem under certainty, selects the acts with maximal utilities. The main result of Chapter 7 (Theorem 7.1) states that if a set of 7 axioms on the rationality of applications of decision rules is satisfied, then it is rational to apply the SEU criterion. (Note that this result holds only for finite set of acts and set of states.) The critical axiom is the so-called Trade-Off axiom according to which, if in a decision problem π some act generates in state *s* an outcome (*a*, *s*) whose utility is strictly greater than the utility of the outcome (*a*, *s'*) generated in another state *s'*, there exists $\delta > 0$ such that for any $\varepsilon_1 \leq \delta$ and every $p(s)$, $p(s')$, there is some ε_2 such that it is rational to transform π into π', where π' differs from π by the fact that $u(a, s)$ is deteriorated by ε_1 whereas $u(a, s')$ is increased by ε_2. Whereas Bayesian justifications of SEU (e.g., Savage) take preferences over uncertain prospects as the only primitive notion, Peterson's axiomatization considers properties of decision rules applying to decision problems, with exogenously given subjective probability and utility.

These axiomatizations of SEU constitute without doubt the most innovative part of the book for decision theorists. For this reason, it is unfortunate that there is some imprecision which may hinder the unfamiliar reader. More specifically, the four main axioms state conditions under which a decision problem may be rationally transformed in another one. But in the examples, the proofs of Theorem 7.1 and Lemma 7.2, the notion of solution-equivalence between two decision problems is invoked – two decision problems are solution-equivalent iff the acts which are rational to perform in each problem are the same. (The exposition is clearer in the paper (Peterson, 2004) from which Chapter 7 originates). One consequence of Peterson's general argument is that agents tackling a decision problem with a utility generated from preferences over *certain* outcomes should be "risk neutral" with respect to this utility (I do not refer here to the canonical notion of risk neutrality used in expected utility theory according to which an agent is risk-neutral iff he is indifferent between a monetary lottery and its expected *value*) in the sense that they should be indifferent between an act that produces for sure an outcome of 100 utils and an act that produces an outcome of 200 utils with probability ½ and an outcome of 0 utils with probability ½. Such a consequence is far from self-evident. It is of course the function of the axioms to convince one of the normative appeal of their consequences. But, reasoning backwards, one can appraise the normative appeal of the axioms on the basis of their consequences. Someone who doubts the appeal of expected utility with respect to a utility generated from *certain* outcomes may *for this very reason* doubt about the axioms. On this issue, I cannot but let the readers judge according to their normative intuitions.

Peterson's book is written in a clear and pleasant style, is well organized and shows an impressive mastery of both formal models and conceptual issues pertaining to contemporary decision theory. My critical comments were intended not to lessen these merits but to pursue the debate about the foundations of the expected utility theory that the book tackles in a stimulating *and* very constructive way. Its reading will be valuable to every scholar – philosopher or economist – interested by decision theory.

Mikael Cozic

Université Paris 12

References

Anscombe, F.J. and Aumann, R.J. (1963) “A Definition of Subjective Probability”, *Annals of Mathematical Statistics*, 34: 199-205

DeGroot, M. (1980) *Optimal Statistical Decisions,* McGraw-Hill

Jeffrey, R. (1965/1983) *The Logic of Decision*, 2nd ed., Chicago: University of Chicago Press

Joyce, J.M. (1999) *The Foundations of Causal Decision Theory*, Cambridge: Cambridge University Press

Peterson, M. (2004) “From outcomes to acts: a non-standard axiomatization of the expected utility principle”, *Journal of Philosophical Logic,* 33:361-378

Savage, L.J. (1954/1972) *The Foundations of Statistics,* 2nd ed., Dover