Extensions of Expected Utility Theory and Some Limitations of Pairwise Comparisons*

M. J. SCHERVISH
Carnegie Mellon University, USA

Of the three decision rules we discuss, perhaps the most familiar one is Γ-Maximin[2]. This rule requires that the decision maker ranks a gamble by its lower expected value, taken with respect to a closed, convex set of probabilities $\mathcal{P}$, and then to choose an option from $\mathcal{A}$ whose lower expected value is maximum. This decision rule (as simplified by the assumptions, above) was given a representation in terms of a binary preference relation over Anscombe-Aumann horse lotteries [2], has been discussed by, e.g., Section 4.7.6 of [1] and recently by [5], who defend it as a form of Robust Bayesian decision theory. The Γ-Maximin decision rule creates a preference ranking of options independent of the alternatives available in $\mathcal{A}$: it is context independent in that sense. But Γ-Maximin corresponds to a preference ranking that fails the so-called (von Neumann-Morgenstern's) "Independence" or (Savage's) "Sure-thing" postulate of SEU theory. In Section 2 of [15], we note that such theories suffer from *sequential incoherence* in particular sequential decision problems.

The second decision rule that we consider, called *E*-admissibility ('*E*' for "expectation"), was formulated in [8, 9]. *E*-admissibility constrains the decision maker's admissible choices to those gambles in $\mathcal{A}$ that are Bayes for at least one probability $P \in \mathcal{P}$. That is, given a choice set $\mathcal{A}$, the gamble f is *E*-admissible on the condition that, for at least one $P \in \mathcal{P}$, f maximizes subjective expected utility with respect to the options in $\mathcal{A}$.[3] Section 7.2 of [12][4] defends a precursor to this decision rule in connection with cooperative group decision making. *E*-admissibility does not support an ordering of options, real-valued or otherwise, so that it is inappropriate to characterize *E*-admissibility by a ranking of gambles independent of the set $\mathcal{A}$ of feasible options. However, the distinction between options that are and are not *E*-admissible does support the "Independence" postulate. For example, if neither option f nor g is *E*-admissible in a given decision problem $\mathcal{A}$, then the convex combination, the mixed option $h = \alpha f \oplus (1\text{-}\alpha)g$ $(0 \leq \alpha \leq 1)$ likewise is *E*-inadmissible when added to $\mathcal{A}$. This is evident from the basic SEU property: the expected utility of a convex combination of two gambles is the corresponding weighted average of their separate expected utilities; hence, for a given $P \in \mathcal{P}$ the expected utility of the mixture of two gambles is bounded above by the maximum of the two expected utilities. The assumption that neither of two gambles is *E*-admissible entails that their mixture has *P*-expected utility less than some *E*-admissible option in $\mathcal{A}$.

The third decision rule we consider is called *Maximality* by Walley in [17][5],

[2] When outcomes are cast in terms of a (statistical) loss function, the rule is then Γ-Minimax: rank options by their maximum expected risk and choose an option whose maximum expected risk is minimum.

[3] Levi's decision theory is lexicographic, in which the first consideration is *E*-admissibility, followed by other considerations, e.g. what he calls a Security index. Here, we attend solely to *E*-admissibility.

[4] Savage's analysis of the decision problem depicted by his Figure 1, p. 123, and his rejection of option *b*, p. 124 is the key point.

[5] There is, for our discussion here, a minor difference with Walley's formulation of Maximality

who appears to endorse it (p. 166). *Maximality* uses the strict partial order (above) to fix the admissible gambles from $\mathcal{A}$ to be those that are not strictly preferred by any other member of $\mathcal{A}$. That is, f is a *Maximal* choice from $\mathcal{A}$ provided that there is no other element $g \in \mathcal{A}$ that, for each $P \in \mathcal{P}$, carries greater expected utility than f does. *Maximality* (under different names) has been studied, for example, in [6, 8, 10, 13, 16]. Evidently, the E-admissible gambles in a decision problem are a subset of the Maximally admissible ones.

The three rules have different sets of admissible options. Here is a heuristic illustration of that difference.

Example 1 *Consider a binary-state decision problem, $\Omega = \{\omega_1, \omega_2\}$, with three feasible options. Option f yields an outcome worth 1 utile if state ω_1 obtains and an outcome worth 0 utiles if ω_2 obtains. Option g is the mirror image of f and yields an outcome worth 1 utile if ω_2 obtains and an outcomes worth 0 utiles if ω_1 obtains. Option h is constant in value, yielding an outcome worth 0.4 utiles regardless whether ω_1 or ω_2 obtains. Figure 1 graphs the expected utilities for these three acts. Let $\mathcal{P} = \{P\colon 0.25 \leqslant P(\omega_1) \leqslant 0.75\}$. The surface of Bayes solutions is highlighted. The expected utility for options f and g each has the interval of values [0.25, 0.75], whereas h of course has constant expected utility of 0.4. From the choice set of these three options $\mathcal{A} = \{f, g, h\}$ the Γ-Maximin decision rule determines that h is (uniquely) best, assigning it a value of 0.4, whereas f and g each has a Γ-Maximin value of 0.25. By contrast, under E-admissibility, only the option h is E-*inadmissible *from the trio. Either of f or g is E-admissible. And, as no option is strictly preferred to any other by expectations with respect to $\mathcal{P}$, all three gambles are admissible under* Maximality.

What normative considerations can be offered to distinguish among these three rules? For example, all three rules are immune to a Dutch Book, in the following sense:

Definition 1 *Call an option* favorable *if it is uniquely admissible in a pairwise choice against the status-quo of "no bet," which we represent as the constant 0.*

Proposition 1 *For each of the three decision rules above, no finite combination of favorable options can result in a Dutch Book, i.e., a sure loss.*

Proof. Reason indirectly. Suppose that the sum of a finite set of favorable gambles is negative in each state ω. Choose an element P from $\mathcal{P}$. The probability P is a convex combination of the extreme (0-1) probabilities, corresponding to a convex combination of the finite partition by states. The expectation of a convex

involving null-events. Walley's notion of Maximality requires, also, that an admissible gamble be classically admissible, i.e., not weakly dominated with respect to state-payoffs. This means that, e.g., our Theorem 1(i) is slightly different in content from Walley's corresponding result.

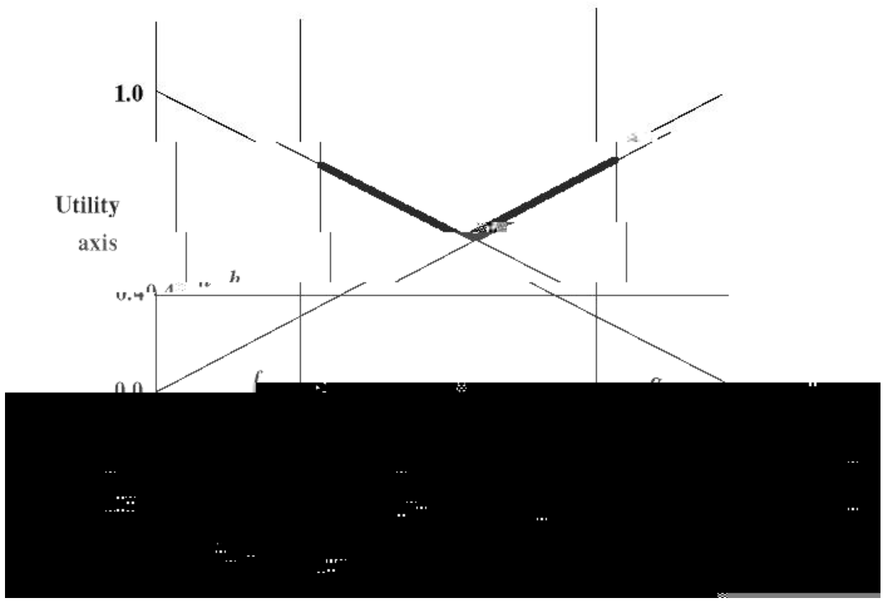


Figure 1: Expected utilities for three acts in Example 1. The thicker line indicates the surface of Bayes solutions.

combination of probabilities is the convex combination of the individual expectations. This makes the *P*-expectation of the sum of the finite set of favorable options negative. But the *P*-expectation of the sum cannot be negative unless at least one of the finitely many gambles has a negative *P*-expectation. Then that gamble cannot be favorable under any of the three decision rules. Thus, none of these three decision rules is subject to sure loss. □

In this paper, we develop an additional criterion for contrasting these decision rules. In Section 2 we address the question of what operational content the rules give to *distinguishing among different (convex) sets of probabilities*. That is, we are concerned to understand which convex sets of probabilities are treated as equivalent under a given decision rule. When do two convex sets of probabilities lead to all the same admissible options for a given decision rule? Γ-Maximin and Maximality are based solely on pairwise comparisons. Not so for *E*-admissibility. Even when the choice set $\mathcal{A}$ of feasible options is convex (e.g., closed under mixed strategies), these rules have distinct classes of admissible options.

2 Gambles and pairwise choice rules

It is evident that for Γ-Maximin generally to satisfy Criterion 1, the convex set of probabilities $\mathcal{P}$ must be closed. For an illustration why, if Example 1 is modified so that $\mathcal{P}' = \{P : 0.4 < P(\omega_1) \leqslant 0.75\}$ then, even though f and h both have the same infimum, 0.4, of expectations with respect to $\mathcal{P}'$, for each $P \in \mathcal{P}'$

Definition 2 *Let* $\mathcal{P}$ *be a convex set of p*[illegible]*ectors. We say that* $f \in \mathcal{A}$ *is* Bayes with respect to $\mathcal{P}$ *if there exists* $p \in$ [illegible] $E_p(f) \geq E_p(g)$ *for all* $g \in \mathcal{A}$.

Theorem 1 *Let* [illegible] *be the set of all* $f \in \mathcal{A}$ [illegible] *Bayes with respect to* $\mathcal{P}$. *Suppose that* $g \in \mathcal{A} \setminus$. *Assume either*

(i) that $\mathcal{P}$ *is closed, or*

(ii) that $\mathcal{A}$ *is finite and that* $\mathcal{P}$ *is open.* [illegible]

$$\{(p_1, \ldots, p_{k-} \text{[illegible]}) \in \mathcal{P}\}$$

is an open subset of $\mathbb{R}^{k-1}$.

Then there exists h in the convex hull of [illegible] $E_p(h) > E_p(g)$ *for all* $p \in \mathcal{P}$.

Corollary 1 *Assume that* $\mathcal{A}$ *is closed a*[illegible]*Let* [illegible] *be the set of all* $f \in \mathcal{A}$ *such that* f *is Bayes with respect to* $\mathcal{P}$. [illegible]*t* $g \in \mathcal{A} \setminus$ *is not Bayes with!0 T.7412 0 Td (t17/R18 9.96264 Tf*

Example 3 *This example illustrates why we assume that* $\mathcal{C}$ *is closed in Theorem 1(i). Let* Ω *consist of three states. Let*

$$\begin{aligned}\mathcal{C} &= \{(p_1,p_2,p_3) : p_2 < 2p_1 \text{ for } p_1 \le 0.2\} \\ &\cup \{(p_1,p_2,p_3) : p_2 \le 2p_1 \text{ for } 0.2 < p_1 \le 1/3\}.\end{aligned}$$

The set of acts $\mathcal{A}$ *contains only the following three acts (each expressed as a vector of its payoffs in the three states):*

$$\begin{aligned} f_1 &= (0.2, 0.2, 0.2), \\ f_2 &= (1, 0, 0), \\ g &= (-1.8, 1.2, .2). \end{aligned}$$

Notice that $E_p(f_2)$ *is the highest of the three whenever* $p_1 \ge 0.2$, $E_p(f_1)$ *is the highest whenever* $p_1 \le 0.2$, *and* $E_p(g)$ *is never the highest. So,* $= \{f_1, f_2\}$ *and* g *is not Bayes with respect to* $\mathcal{A}$*. For each* $0 \le \alpha \le 1$, *we compute*

$$\begin{aligned} E_p(\alpha f_1 + (1-\alpha) f_2) &= p_1(1-\alpha) + 0.2\alpha, \\ E_p(g) &= -2p_1 + p_2 + 0.2. \end{aligned}$$

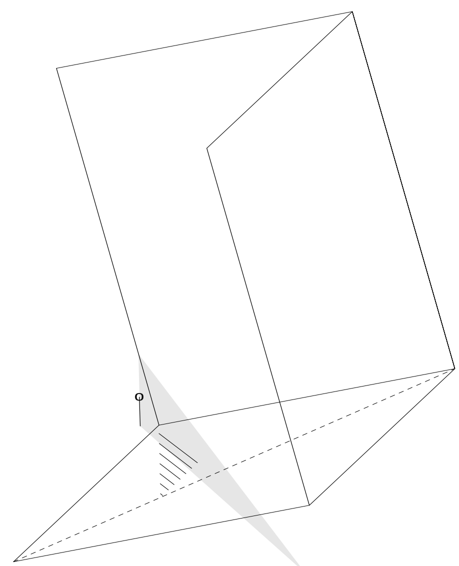
Notice that $E_p(\alpha f_1 + (1-\alpha) f_2)$ *is strictly greater than* $E_p(g)$ *if and only if* $p_2 < (3-\alpha)p_1 - 0.2(1-\alpha)$*. There is no* α *such that this inequality holds for all* $p \in \mathcal{C}$.

Remark 1 *Note that is it irrelevant to this example that* $p_2 = 0$ *for some* $p \in \mathcal{C}$.

Definition 3 *Say that two convex sets* intersect all the same supporting hyperplanes *if they have the same closure and a supporting hyperplane intersects one convex set if and only if it intersects the other.*

nnj 4131335 0 Td (aample)Tj 31.28vd (aamportan

In addition to showing that E-admissibility does not reduce to pairwise


o

either set $\mathcal{P}$ or $\mathcal{P}^*$, consider the family of decision problems defined by the three-way choices: $\mathcal{A}_{-\varepsilon} = \{f_1, f_2, g_{-\varepsilon}\}$, where $g_{-\varepsilon}$ is the act with payoffs $(1.8, 1.2-\varepsilon, 0.2)$. For each $\varepsilon > 0$, only the pair $\{f_1, f_2\}$ is E-admissible from such a three-way choice, with respect to each of the two convex sets of probabilities.

Likewise, in order to establish that the half-open line segment $(D,B]$ belongs to both sets $\mathcal{P}$ and $\mathcal{P}^*$, consider the family of decision problems defined by the three-way choices: $\mathcal{A}_{+\varepsilon} = \{f_1, f_2, g_{+\varepsilon}\}$, where $g_{+\varepsilon}$ is the act with payoffs $(1.8, 1.2+\varepsilon, .2)$. For each $\varepsilon > 0$, all three options are E-admissible with respect to each of the two convex sets of probabilities.

However, in the decision problem with options $\mathcal{A} = \{f_1, f_2, g\}$, as shown above, only the pair $\{f_1, f_2\}$ is E-admissible with respect to the convex set $\mathcal{P}$, whereas all three options are E-admissible with respect to the convex set $\mathcal{P}^*$.

By contrast, given a choice set, Maximality makes the same ruling about which options are admissible from that choice set, regardless whether convex set $\mathcal{P}$ or convex set $\mathcal{P}^*$ is used. That is, Maximality cannot distinguish between these two convex sets of probabilities in terms of admissibility of choices, as the two convex sets of probabilities intersect all the same supporting hyperplanes.

3 Summary

The discussion here contrasts three decision rules that extend Expected Utility and which apply when uncertainty is represented by a convex set of probabilities, $\mathcal{P}$, rather than when uncertainty is represented only by a single probability distribution. The decision rules are: Γ-Maximin, Maximality, and E-admissibility. We show that these decision rules have different operational content in terms of their ability to distinguish different convex sets of probabilities. When do the admissible choices differ for different convex sets of probabilities? Γ-Maximin is least sensitive among the three in this regard. We show that, even when the option set is convex, one decision rule (E-admissibility) distinguishes among more convex sets than the other two. This is because

- $C = \{x \in \mathbb{R}^k : \alpha^\top x \geq c$, *for all* $(\alpha, c) \in D\}$.
- $(\alpha, c) \in D$ *implies* $(a\alpha, ac) \in D$ *for all* $a \geq 0$,
- $(\alpha, c) \in D$ *implies* $(\alpha, c - a) \in D$ *for all* $a > 0$,

Also, for each $(\alpha, c) \in D$, $c \leq 0$.

Proof. To see that $(\alpha, c) \in D$ implies $c \leq 0$, let $\mathbf{0}$ be the origin. Then $\alpha^\top \mathbf{0} = 0 \geq c$. Define the following set

$$D_0 = \{(\alpha, c) : \alpha^\top x \geq \mathbf{c}, \text{ for all } x \in C\}. \tag{1}$$

To see that D_0 is convex, let (γ_1, d_1) and (γ_2, d_2) be in D_0 and $0 \leq \beta \leq 1$. Then, for all $x \in C$,

$$(\beta\gamma_1 + [1-\beta]\gamma_1)^\top x \geq \beta d_1 + (1-\beta) d_2.$$

This means that $\beta(\gamma_1, d_1) + [1-\beta](\gamma_2, d_2) \in D_0$, and D_0 is convex. To see that D_0 is closed, notice that $D_0 = \bigcap_{x \in C} D_x$, where $D_x = \{(\alpha, c) : \alpha^\top c \geq c\}$ and each D_x is closed. It is clear that D_0 has the last two properties in the itemized list. For the first condition, let E be the set defined in the first condition. It is clear that $C \subseteq E$. Suppose that there is $x_0 \in E$ such that $x_0 \notin C$. Then there is a hyperplane that separates $\{x_0\}$ from C. That is, there is $\gamma \in \mathbb{R}^k$ and d such that $\gamma^\top x \geq d$ for all $x \in C$ and $\gamma^\top x_0 < d$. It follows that $(\gamma, d) \in D_0$, but then $x_0 \notin E$, a contradiction.

To see that the set that satisfies the conditions 46 7.0.1534 Tf 51.0890 62.3477 Tf 317.3838 0 Td (x)Tj /R46 7 34.5322

Assume that A is nonempty. Define D to be the set of all vectors in $\mathbb{R}^{k+1}$ of the form $(a\alpha, ad-b)$ with $a,b \geq 0$ and $(\alpha,d) \in V$. Then $D = \{(\alpha,d) \in \mathbb{R}^{k+1} : \alpha^\top x \geq d$, for all $x \in A\}$.

Proof. Let $x_0 \in A$, and define

$$\begin{aligned} C &= \{x - x_0 : x \in A\}, \\ V' &= \{(\alpha, d - \alpha^\top x_0) : (\alpha,d) \in V\}. \end{aligned}$$

It follows that

$$C = \{x \in \mathbb{R}^k : \alpha^\top x \geq c, \text{ for all } (\alpha,c) \in V'\}, \tag{3}$$

and C contains the origin and is a closed convex set. Define $D_1 = \{(\alpha, d - \alpha^\top x_0) : (\alpha,d) \in D\}$. In other words, D_1 is the convex closed convex set of all vectors in $\mathbb{R}^{k+1}$ of the form $(a\alpha, ac-b)$ with $a,b \geq 0$ and $(\alpha,c) \in V'$. The definitions of D and D_1 were rigged so that D_1 satisfies all the

So, for all $x \in H_g$,

$$c_h > \alpha_h^\top x = c_h - E_p(h) + E_p(g).$$

It follows that, for all $p \in$ [illegible], $E_p(h) > E_p(g)$.

(ii) Define U, [illegible], V, A, and H_g e in a transfinite induction as follows. At each successor ordinal $\gamma+1$, find $h_{\gamma+1} \in$ [illegible] such that $E_p(h_{\gamma+1}) > E_p(h_\gamma)$ for all $p \in$ [illegible]. At a countable limit ordinal γ choose any countable sequence $\{\gamma_n\}_{n=1}^{\infty}$ of ordinals that is cofinal with γ. By the induction hypothesis, $E_p(h_{\gamma_i}) < E_p(h_{\gamma_j})$ for all $p \in$ [illegible] if $i < j$. The sequence $\{h_{\gamma_n}\}_{n=1}^{\infty}$ belongs to the closed bounded set $\mathcal{A}$, hence it has a limit h_γ and

$$E_p(h_\gamma) = \lim_{n\to\infty} E_p(h_{\gamma_n}) = \sup_n E_p(h_{\gamma_n}),$$

for all p, and hence does

[11] Levi, I. Value commitments, value conflict, and the separability of belief and value. *Philosophy of Science 66* (1999), 509–533.

[12] Savage, L. J. *The Foundations of Statistics*. Wiley, New York, 1954.

[13] Seidenfeld, T. Comment on "Probability, evidence, and judgment" by A. P. Dempster. In *Bayesian Statistics 2: Proceedings of the Second Valencia International Meeting, 1983*, e. a. J. M. Bernardo, Ed. North-Holland, Amsterdam, 1985, pp. 127–129.

[14] Seidenfeld, T., and Schervish, M. J.and Kadane, J. A representation of partially ordered preferences. *Annals of Statistics 23* (1995), 2168–2217.

[15] Seidenfeld, T., Schervish, M. J., and Kadane, J. Decisions without ordering. In *Acting and Reflecting*, W. Sieg, Ed. Kluwer Academic Publishers, Dordrecht, 1990, pp. 143–170.

[16] Sen, A. K. Social choice theory: a re-examination. *Econometrica 45* (1977), 53–89.

[17] Walley, P. *Statistical Reasoning with Imprecise Probabilities*. Chapman and Hall, London, 1990.

M. J. Schervish is with Carnegie Mellon University, USA

T. Seidenfeld is with Carnegie Mellon University, USA

J. B. Kadane is with Carnegie Mellon University, USA

I. Levi is with Columbia University, USA