A GENTLE APPROACH TO IMPRECISE PROBABILITY

GREGORY WHEELER
FRANKFURT SCHOOL OF FINANCE & MANAGEMENT
g.wheeler@fs.de

forthcoming in REFLECTIONS ON THE FOUNDATIONS OF STATISTICS: ESSAYS IN HONOR OF TEDDY SEIDENFELD

I had the good fortune of meeting Teddy Seidenfeld early in my career and learning from him over many years how to approach probability in general, and imprecise probability in particular. I first learned about Teddy's work from Henry Kyburg, Teddy's undergraduate teacher and my Ph.D. supervisor, who introduced me to Teddy's and Larry Wasserman's work on dilation, among other striking contributions that Teddy made while I was graduate student in the 1990s, many with his longtime collaborators Jay Kadane and Mark Schervish. I cannot overstate the impact Teddy's work has had on me. He will doubtlessly conclude that I have not learned from him well enough, but I am nonetheless grateful for his years of instruction.

The field of of imprecise probability has matured, in no small part because of Teddy's decades of original scholarship and essential contributions to building and sustaining the ISIPTA community. Although the basic idea behind imprecise probability is (at least) 150 years old, a mature mathematical theory has only taken full form in the last 30 years. Interest in imprecise probability during this period has also grown, but many of the ideas that the mature theory serves can be difficult to apprehend to those new to the subject. Although these fundamental ideas are common knowledge in the ISIPTA community, they are expressed, when they are expressed at all, obliquely, over the course of years with students and colleagues.

A single essay cannot convey the store of common knowledge from any research community, let alone the ISIPTA community. But, this essay nevertheless is an attempt to guide those familiar with the basic Bayesian framework to appreciate some of the elegant and powerful ideas that underpin the contemporary theory of lower previsions, which is the theory that most people associate with the term 'imprecise probability'.

---

1 BAYESIANISM IN A NUTSHELL

For a Bayesian, a probability statement represents a person's "degree of belief" about the true value of a random variable. Bayesian inference, in turn, is the practice of estimating the conditional distribution of such random quantities given observed data. While no one objects to applying Bayesian inference to instances where the prior distribution is known, or at least backed by an accepted scientific model, Bayesian statistics stands out by endorsing the practice of applying Bayesian inference without requiring any such exogenous constraints. Bayesianism instead places an internal, *coherence* condition on your degrees of belief whereby the comparable weights you assign to every event in a given class are rational just in case they satisfy the axioms of (finitely additive) probability.

There are competing explanations for what Bayesians are doing with this framework. Daniel Garber characterizes two of them as the thought police model and the learning machine model. The thought police model imagines a Bayesian scold looking over your shoulder threatening penury (de Finetti 1937) or shame (Savage 1954) should you make an incoherent assessment. The learning machine model is more aspirational. On this view the Bayesian framework describes an ideal learning machine (Carnap 1962), or the principles an ideal learning machine would follow once everything inconvenient about learning and machines is abstracted away. Garber doesn't find much of a difference between these two accounts of Bayesianism. To him, whether the thought police is coercing you to be an ideal learner or the learning machine describes an ideal citizen with a clean record "seem interchangeable" (Garber 1983, p. 101).

If however the question is why your degrees of belief ought to comport with the probability axioms, that question is answered by the *carabiniere*. De Finetti's canonical book argument is this: If your degrees of belief at a moment in time are identified with the fair prices you provide at that time for some finite number of *gambles*,[1] thereby obliging you, hypothetically, to either buy or sell units of each gamble at your named prices on demand, then you expose yourself to the possibility of sure loss if and only if your degrees of belief fail to satisfy the axioms of finitely additive probability.

For simplicity, let us restrict ourselves to propositions, which are gambles that map elements of $\mathcal{X}$ to 1 or 0, corresponding to whether a proposition is true or false, respectively. If your degree of belief is 0.7 that a proposition is true, such as the assertion that the random variable X takes the value x, then you are specifying that the certain gain of €0.70 is *equivalent*, in your judgment, to the uncertain gain of X. Thus, according to the rules of this scheme, you are *indifferent*[2] to engaging in two types of hypothetical transactions. The first transaction requires you to buy a contract for €0.70 that returns to you €1 if $X = x$ and returns to you nothing otherwise. In other words, according to the terms of this first hypothetical transaction, you are required to surrender the sure reward of 70 cents to acquire the uncertain reward of 1 euro that you receive if $X = x$ is (verifiably) true. The second transaction requires you to sell a contract for 70 cents that commits you to pay out €1 if $X = x$ and to pay nothing otherwise. In other words, the second transaction requires you to receive the sure reward of 70 cents but acquire the uncertain obligation to pay 1 euro in the event that $X = x$ is true. Your interlocutor is empowered to put together a contract consisting of any finite combination of your fair prices—obliging you to simultaneously buy this gamble, sell that gamble, et cetera. If your posted fair prices are incoherent, there exist a contract of positions he can oblige you to take such that no matter which of the possible states obtains, you will incur a loss. You are rational if and only if resolving whether in fact $X = x$ or $X \neq x$ cannot expose you to a sure loss, and you avoid sure loss if and only if the prices you give satisfy the axioms of finitely additive probability.

A degree of belief then is a judgment of uncertainty expressed in terms of your willingness to have your fortune change by an unknown outcome. The price μ at which you value a random variable X, written $P(X) = \mu$ and called a *linear prevision*, is the value you choose on the understanding that you are committed to accept any bet of the form $c(X - \mu)$, where c is a positive or negative constant chosen by your opponent, and the possible values of X are the set $\mathcal{X}$. The sign of c determines whether you are obliged to

[1] A gamble is a function from a set of possible values $\mathcal{X} = \{x_1, x_2, \ldots, x_n\}$ to $\mathbb{R}$, where the set of possible values serve as a partition against which to describe the consequences of that gamble. Strictly speaking, de Finetti formulates his theory in terms of *random quantities* that do not presume a fixed possibility space $\mathcal{X}$, where a random quantity is simply a unique real number unknown to you. Nevertheless, with a suitable translation appealing to the possible states of your random quantities and their combination (Walley 1991, §2.1.4), the two are mathematically equivalent. Gambles are discussed in more detail in Section 4.

[2] Frank Ramsey (1926) put the same argument, independently, in slightly different terms, emphasizing your *indifference* between the uncertain reward of the gamble X and the sure reward of a real number μ, your degree of belief, where the utility of the rewards are presumed to be comparable on a linear scale. In effect, μ is your "fair price" for X. The canonical Bayesian account of degrees of belief is a creole of the de Finetti (fair prices) and Ramsey (fair betting rates) approaches.

buy (positive) or sell (negative), and the magnitude of c specifies how many units of X are bought or sold. Your net outcome is the sum of the payoffs Y,

$$Y = c_1(X_1 - \mu_1) + c_1(X_2 - \mu_2) + \cdots + c_1(X_n - \mu_n)$$

and your n previsions are coherent just in case there is no linear combination of this form in which Y is negative no matter how the random variables $X_1, X_2, \ldots, X_n$ are resolved, that is, no matter which $x \in \mathcal{X}$ obtains. This is de Finetti's first definition of previsions and his first criterion for coherence.[3]

2 A Simple Extension

Every degree of belief is constructed from two hypothetical commitments, one to buy a gamble and another to sell. De Finetti's argument, and the theory of linear previsions that stands behind it, requires you to announce the same price for buying and selling a gamble. In short, this is what degrees of belief are: a collection of two-sided fair prices elicited from you about a finite collection of gambles.

A fair price of μ for a gamble expresses that you are disposed to buy the gamble for any price less than μ, and similarly disposed to sell for any price greater than μ. A simple extension of this account, first proposed by C.A.B. Smith (1961) and, independently, by Peter Williams (1975, 1976), allows for two different prices for each gamble, one for buying and another for selling. The routine for eliciting a pair of one-sided bets for each gamble is otherwise identical to the canonical procedure for eliciting a single, two-sided fair price. Committing to a maximum buying price of €0.30 for a gamble that returns to you €1 if $X = x$, and returns to you 0 otherwise, is understood colloquially as the maximum price you are willing to pay for that gamble: you are announcing your commitment to surrender a sure reward up to 30 cents for the uncertain reward of 1 euro awarded to you in the event that $X = x$ is true. Conversely, committing to a minimum selling price of €0.70 for that same gamble obliges you to surrender €1 if $X = x$, 0 otherwise, is understood as the minimum price you demand for selling that gamble: you are announcing your commitment to surrender the uncertain reward of 1 euro in the event that $X = x$ is true in exchange for the sure reward of at least 70 cents. For prices offered between 30 and 70 cents you are neither obliged to sell nor obliged to buy.

A pair of one-sided bets represents your highest buying price b and your lowest selling price s for a gamble, given the status quo alternative option of receiving no reward at all. The span (b, s) (when $b < s$) consists of prices at which you are not obligated to either buy or to sell that gamble. Note that you are not obliged to buy or to sell gambles at prices between b and s, which does not mean you are obliged to abstain come what may. Indeed, if instead of the alternative being the status quo option yielding you 0, a Smith-Williams agent were instead offered a combination of gambles that is dominated by the original gamble offered at some price between b and s, then she is not forced to abstain from this dominating choice. On the standard Bayesian two-sided fair price

[3] De Finetti's second definition of previsions and his second criterion for coherence appeals to a quadratic scoring rule. Briefly, according to the second criterion, if your price μ is scored by the amount $(\mu - 1)^2$ when $X = x$ and by μ^2 otherwise, and if the scores for different events are additive, then μ is undominated just in case μ is a finitely additive probability. The first and second criteria are equivalent in the sense that they identify the same previsions (de Finetti 1974, §3.3.7).

model, $s-b=0$, for all gambles under consideration. So, any gamble that fails to maximize expected utility may be found by pairwise comparisons between gambles under a precise probability. This pairwise comparison rule for identifying gambles as inadmissible is not valid when (non-trivial) one-sided prices are used, however.[4] More generally, since the Smith-Williams extension calls for accommodating preferences that are more general than a total preorder, a rational Smith-Williams agent may be unable to identify a preferred gamble from a given set, which is why decision making with imprecise probabilities typically involves *choice functions*.[5] The upshot is that a rational Smith-Williams agent is obligated to avoid sure loss but is nevertheless permitted to forgo a sure gain.

3 A Surprising Implication

Conceptually, the Smith-Williams extension is a straightforward generalization of the Bayesian framework. It is not, as some critics maintain, a radical break from tradition. The theory of lower previsions (Walley 1991; Troffaes and de Cooman 2014), which accommodates this simple extension and much more, includes de Finetti's theory of linear previsions as a special case. Nevertheless, the switch from linear previsions to the more general theory of lower previsions is mathematically less straightforward—and deeply interesting for that very reason.

It should come as no surprise that a mathematical generalization of the Bayesian framework will, as a matter of course, derive some properties under special circumstances that are immutable features of the original theory. No one is phased by this in other contexts. Propositional logic appears as a special case of first-order logic, but the model theory and grammar of first-order logic operate very differently than the semantics and syntax of propositional logic. The model theory for first-order logic is not a simple generalization of truth table semantics for propositional logic, as it is impossible to evaluate every formula of first-order logic by evaluating all logically exhaustive combinations of interpretations. Open formulas of first-order logic are syntactically well-formed by the rules of the grammar but fail to be semantically determinable, an impossibility for propositional logic. If semantic interpretability were deemed inseparable from syntactic well-formedness, truth tables fundamental to model theory, or decidability the *sine qua non* of logic itself, then first-order logic would be an aberration. Yet such objections have no standing in logic. They are matters resolved in the classroom, not debated in learned journals. The same unfortunately cannot be said for the theory of lower previsions.[6]

The theory of lower previsions is not constructed from the same elementary building blocks used to construct the canonical Bayesian framework. Foremost, the language of events and degrees of belief, which are fundamental to the Bayesian framework, are not foundational to lower previsions. At the heart of the Bayesian framework is a one-to-one correspondence between linear previsions and additive probability which says that a

[4]For discussion and examples, see (Seidenfeld 2004) and (Jansen, Schollmeyer, and Augustin 2021).

[5]We discuss partial preference in Section 7. See (Augustin, Coolen, de Cooman, and Troffaes 2014, Ch. 8) and (Bradley 2019) for a discussion of decision making with imprecise probability models, and the seminal work by (Seidenfeld, Schervish, and Kadane 2010) and later by (De Bock and de Cooman 2019a) involving choice functions.

[6]Some authors criticize replacing numerically determinate degrees of belief with a range of values by appealing to arguments that may be valid for precise probability models but are invalid with respect to imprecise probability models. See footnote 18 for examples.

class of additive probabilities on a domain uniquely extends to linear previsions on the set of measurable gambles defined on that domain. This equivalence between additive probability and linear previsions is the reason why additive probability, and therefore why degrees of belief expressed over the language of events, is taken as primitive in the canonical Bayesian framework. Previsions are simply expectations, defined by taking the integral with respect to a given probability.[7] The equivalence is also why de Finetti, who prioritized previsions over probability, is treated as an apostle of Bayes rather than an apostate.[8]

A similar one-to-one correspondence between lower probability and lower previsions does not hold. Unlike linear previsions, two lower previsions can have the same values on all events, and therefore express the same lower probabilities, but nevertheless have different values for some bounded gambles. In other words, coherent lower previsions defined with respect to the language of gambles is strictly more expressive than lower probabilities defined with respect to the language of events.[9] Thus, probability and degrees of belief do not play a fundamental role in imprecise probability models: instead, they are derived and appear as special cases.[10] Thus (bounded) gambles, not events, are taken as primitives in imprecise probability models. One consequence from this is that *probability* is not fundamental to imprecise probability models—as paradoxical as that sentence may sound. Yet, since linear previsions appear as a special case of lower previsions, much as propositional logic appears as a special case of first-order logic, there is a sense in which probability is not fundamental to the Bayesian framework, either.

4 Gambles and Events

Before turning to see why the theory of lower previsions is built from the language of gambles rather than the language of events, we should pause to introduce some terminology and notation. A *gamble* f is a map from the possible values of a random variable to the real numbers, representing the (possibly negative) reward to you that is associated with each of the possible outcomes of that random variable. So, if you accept the gamble f with respect to a random variable X, then when the true value of X is revealed, $X = x$, you will receive the reward $f(x)$. For example, if you were offered a gamble f on a coin flip that rewarded you 1 euro on the outcome *heads* but returned to you 0 on the outcome *tails*, your willingness to pay c for the risky reward of f would reflect your assessment of the coin turning up heads, with $c = 0$ included as a possibility to indicate that you'd take the gamble for nothing but otherwise are unwilling to exchange any sure reward for this risky one.

Let script $\mathcal{X} = \{x_1, x_2, \dots\}$ denote the set of possible values of the random variable X. A gamble f defined with respect to X will then assign a real number to each outcome in $\mathcal{X}$. The set of all gambles with respect to a set of possibilities $\mathcal{X}$ is denoted by $\mathbb{G}(\mathcal{X})$, or simply $\mathbb{G}$ when the underlying possibility space is clear. For a pair of gambles f and g,

[7] Namely, a probability p determines a unique expectation (prevision) P with respect to a random variable X with possibilities $\mathcal{X}$, if $P_p(X) := \sum_{i=1}^{n} X(x_i)p(\{x_i\})$, for finite $x \in \mathcal{X}$; and $P_p(X) := \int_{\mathcal{X}} X(x)p(dx)$, when p is a density. P_p then is the unique linear expectation, with P restricted to the events under p.

[8] Discussions and proofs of this equivalence, ordered in length from extensive to concise, are found in (de Finetti 1974, Ch. 3 and §3.8), (Walley 1991, §2.8 and §3.2), and (Troffaes and de Cooman 2014, §5.1).

[9] A running example illustrating this point begins in Section 5.

[10] This point is emphasized by Walley (1991, §2.7) and Troffaes and de Cooman (2014, §5.2).

both defined with respect to a space $\mathcal{X}$, the combination $f \circ g$ is a gamble on $\mathcal{X}$ when the arithmetical operation $\circ$ is applied point-wise to all $x \in \mathcal{X}$:

$$(f \circ g)(x) := f(x) \circ g(x) \text{ for all } x \in \mathcal{X}$$

Thus, if f and g are gambles, then so are $f - g$, $f + g$, and $f \cdot g$. Gambles may be ordered point-wise, too, such that $f \leq g$ just in case for all $x \in \mathcal{X}, f(x) \leq g(x)$. A gamble is bounded above when its supremum value is finite, bounded below when the infimum is finite, and simply *bounded* when it is both bounded above and bounded below. Although our discussion will only involve bounded gambles, and Walley's treatment is restricted to bounded gambles (Walley 1991), the general theory of lower previsions accommodates unbounded gambles (Troffaes and de Cooman 2014).

An *event* E is a subset of $\mathcal{X}$, and an indicator function I_E selects from $\mathcal{X}$ all x that are members of E:

$$I_E(x) := \begin{cases} 1 & \text{if } x \in E \\ 0 & \text{otherwise} \end{cases}$$

Indicator functions are a special type of gamble, namely bounded $\{0, 1\}$-valued gambles, which assign the value 1 if the state x is in E and the value 0 otherwise.

Alternatively, an indicator function may be viewed as the the truth conditions for a proposition q with respect to a partition $\mathcal{X}$ of states. So, rather than *choosing* the gamble q that returns to you the reward of 1 or 0 depending on which element of $\mathcal{X}$ is observed, the proposition q is true or false depending on which state in $\mathcal{X}$ is realized. There is a long tradition in philosophy, following (Jeffrey 1965), that treats propositions as fundamental. The point to emphasize is that formulations of subjective probability in terms of propositions depends on the one-to-one correspondence between linear previsions and additive probability mentioned in Section 3, and breaks down without this correspondence.

EXAMPLE 4.1 *Suppose there are two possible outcomes, $\mathcal{X} = \{x_1, x_2\}$, and let f, g, h be gambles on $\mathcal{X}$ with the associated rewards to you displayed on the left side of Figure 1. All three gambles are bounded gambles. Both f and g are $\{0, 1\}$-valued gambles and h is a constant gamble. Your payoff naturally depends on which gamble you choose and, in the case of f and g, which state, x_1 or x_2, obtains. Thus, viewed as a non-sequential decision problem, f, g, h are* acts *that you may choose from, where a* choice *refers to you selecting a single gamble from the set of gambles available to you for choice, and your* reward *is the value you receive from your choice once the true state is revealed. Your expected payoff is calculated in the usual manner, if an additive probability assessment is provided for the two states, x_1 or x_2.*

If instead a sub-additive or super-additive probability assessment is provided for the two states x_1 and x_2, then your upper and lower expected payoff, respectively, may be calculated. The dashed lines in the plot in Figure 1 represent two hypothetical probabilities that E, namely $p_1(E) = 0.3$ and $p_2(E) = 0.7$, which may be interpreted as witnesses for the lower probability and upper probability that E, respectively. The expected utility for f and g under lower and upper probabilities would then each have the interval values $[0.3, 0.7]$*, whereas the expected utility of the constant gamble h is* 0.4.

There are several proposed decision rules for imprecise probabilities, two of which are Γ-maximin and E-admissibility. According to Γ-maximin, options are ranked by

	x_1	x_2
f	0	1
g	1	0
h	0.4	0.4

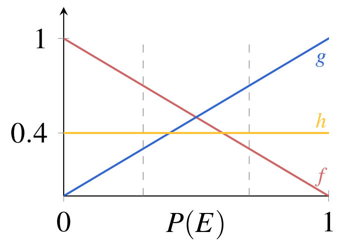


Figure 1: Left: The payoffs for gambles f, g, and h. Right: A plot of the binary-state decision problem whose expected payoff (y-axis) is a function of the probability of E (x-axis), where E is defined by the indicator $I_E(x) = 1$ if $x = x_1$ and 0 otherwise.

lower expectation and an option that is maximal under this ranking is selected (Berger 1985, §4.7.6) and (Gilboa and Schmeidler 1989). According to E-admissibility, an option is selected for choice if there is some probability in the credal set under which that option maximizes expected utility (Levi 1980). The decision problem in Example 4.1 is introduced in (Seidenfeld 2004) to show that, although h is uniquely optimal under Γ-maximin, h is uniquely excluded from choice by E-admissibility. Thus, an optimal choice under maxmin reasoning does not entail that it is an E-admissible choice. Here again we see the theoretical insights afforded by a more expressive language: the relationship between E-admissibility and Γ-maximin is not a simple logical extension of one to the other. For a discussion of decision rules for imprecise probability models, see (Troffaes 2007; Etner, Jeleva, and Tallon 2012; Bradley 2019).

5 Credal Sets and Lower Probability

What mathematical models are suited to representing coherent Smith-Williams assessments? To motivate our answer, presented over Sections 6, 7 and 8, we begin with an illustration of the language of events failing to capture all of the information contained in an imprecise probability model.

Closed convex sets of probability mass functions – called *credal sets* (Levi 1980) – are the most familiar imprecise probability model even if not, as we will see later, the most expressive or easiest with which to work. Credal sets have been used in a wide range of applications,[11] and lower and upper probabilities defined in terms of a credal set are among the most mathematically straightforward representations available for pairs of one-sided bets on $\{0,1\}$-gambles. Indeed, several topics in the foundations of imprecise probability are effectively addressed using this simple model, including dilation (Seidenfeld and Wasserman 1993; Pedersen and Wheeler 2014; Pedersen and Wheeler 2019), the value of information (Grünwald and Halpern 2004; Kadane, Seidenfeld, and

[11] A sample of examples includes models for the strength of evidence (Shafer 1976), support for an argument (Haenni, Romeijn, Wheeler, and Williamson 2011, Ch. 3), opinion pooling (Walley 1981; Stewart and Ojea Quintan 2018; Elkin and Wheeler 2018), sensitivity analysis (Walley 1991; Oberguggenberger, King, and Schmelzer 2009), 'objectively' chaotic systems (Fierens, Rêgo, and Fine 2009), probabilistic graphical models (Cozman 2000; Cozman and Seidenfeld 2009), the foundations of probability (Kadane, Schervish, and Seidenfeld 1999; de Cooman and Miranda 2007), parimutuel betting (Corsato, Pelessoni, and Vicig 2019), machine learning (Augustin, Coolen, de Cooman, and Troffaes 2014, Chs. 9, 10), and much more.

Schervish 2008; Pedersen and Wheeler 2015; Bradley and Steele 2016), different independence concepts (Cozman 2012), full conditional probabilities and indeterminacy (Wheeler and Cozman 2021), and decision making (Levi 1980; Seidenfeld 2004; Troffaes 2007; Bradley 2019). Let's turn now to consider how lower and upper probabilities are defined in terms of a credal set.

A *lower probability space*, $(\mathcal{X}, \mathscr{A}, \mathbb{P}, \underline{\mathsf{p}})$, consists of a set of possible states, $\mathcal{X}$; an algebra of events $\mathscr{A}$ over $\mathcal{X}$; a nonempty set of probability functions $\mathbb{P}$ on $\mathscr{A}$; and a *lower probability function* $\underline{\mathsf{p}}$ on $\mathscr{A}$ with respect to $\mathbb{P}$ defined by:

$$\underline{\mathsf{p}}(E) := \inf\{p(E) : p \in \mathbb{P}\} \text{ for each } E \in \mathscr{A} \tag{1}$$

When the set of probability mass functions $\mathbb{P}$ is convex and closed, $\mathbb{P}$ is a credal set. While strictly not necessary to the definition of a lower probability space, standard coherence conditions for lower previsions are defined in terms of the lower envelope of closed convex sets of probability mass functions. Specifically, every weak*-compact and convex set of linear previsions determines a unique lower prevision by constructing lower envelopes.[12] Thus, hereafter we assume that $\mathbb{P}$ is a closed convex set.

If your maximum buying price is 30 cents for a gamble that pays you one euro if E obtains and zero otherwise, your commitment can be represented by the lower probability $\underline{\mathsf{p}}(E) = 0.3$. The *upper probability function* $\overline{\mathsf{p}}$ is then defined in terms of lower probability by a conjugacy relation

$$\overline{\mathsf{p}}(E) := 1 - \underline{\mathsf{p}}(E^c) \text{ for each } E \in \mathscr{A} \tag{2}$$

So, if your minimum selling price is 70 cents for a gamble that obliges you to pay out 1 euro if E and 0 otherwise, this is equivalent to 1 minus your buying price of the gamble on the *complement* of E: $\overline{\mathsf{p}}(E) = 0.7$. Conditional lower and upper probabilities are defined as

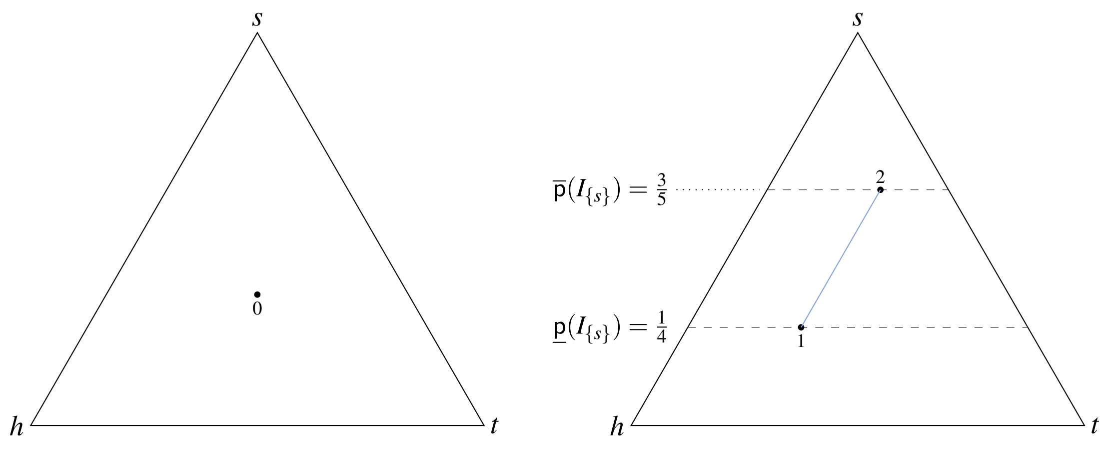
$$\underline{\mathsf{p}}(E \mid F) := \inf\{p(E \mid F) : p \in \mathbb{P}\} \quad \text{and} \quad \overline{\mathsf{p}}(E \mid F) := \sup\{p(E \mid F) : p \in \mathbb{P}\},$$

respectively. With this notation in place, we now turn to an extended example.

Imagine a soup can is to be flipped and you are interested in the outcome. The can may land topside up (*heads*), on its side (*side*), or topside down (*tails*). Suppose further that you regard these three possibilities, $\mathcal{X} = \{h, s, t\}$, to be exclusive and exhaustive for this experiment. Now suppose you wish to assess the outcome of the can landing on its side. Owing to the one-to-one correspondence between a probability mass function (p) and a linear prevision (P), when $\mathcal{X}$ is finite, $p(x) = P(I_{\{x\}}) = P(\{x\})$, for all $x \in \mathcal{X}$. So, if you were to judge the can toss to be fair, you might be disposed to assign equal probability mass to the three basic outcomes, $p_0(h) = p_0(s) = p_0(t) = \frac{1}{3}$. This equal-probability assessment is represented by the point $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$ labeled '0' in the center of the unit 3-simplex of Figure 2(a).

Each point in an n-simplex represents a probability mass function, where the dimension n corresponds to the number of possibilities in $\mathcal{X}$. So, returning to the 3-simplex

[12] That is, $\underline{\mathrm{P}}$ is the lower envelope of $\mathbb{P}$, if $\underline{\mathsf{p}}(E) = \min\{p(E) : p \in \mathbb{P}\}$ for all $E \in \mathscr{A}$. This result extends to lower previsions of gambles as well (Walley 1991, §2.7.3 and §3.6.1). Axioms for lower previsions, which include lower probability as a special case, are presented in Section 6.



(a) (b)

Figure 2: Probability 3-simplex for the three outcome can-toss experiment. (a): a single uniform probability mass function, $p_0 = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$. (b): two probability mass functions, $p_1 = \left(\frac{1}{2}, \frac{1}{4}, \frac{1}{4}\right)$ and $p_2 = \left(\frac{3}{20}, \frac{3}{5}, \frac{1}{4}\right)$. The points p_1 and p_2 and all points in the linear span between these two points is a credal set.

representation of our soup can toss in Figure 2, if you were to assign probability 1 to the event of the can landing on its side, that assessment would correspond to the probability mass function represented by the point $(0,1,0)$ at the apex. The simplex in Figure 2(b) includes two probability assessments, $p_1 = \left(\frac{1}{2}, \frac{1}{4}, \frac{1}{4}\right)$ and $p_2 = \left(\frac{3}{20}, \frac{3}{5}, \frac{1}{4}\right)$, each formatted for the outcomes (h,s,t). The probability mass function p_1 witnesses the lower probability of $\frac{1}{4}$ that the can lands on its side, whereas p_2 witnesses the upper probability $\frac{3}{5}$ that the can lands on its side. The linear span in blue and the extremal points 1 and 2 represent a credal set.

There are a continuum of lines whose endpoints witness the same upper and lower probabilities in Figure 2(b) for the event of the can landing on its side. Each of these alternative credal sets happen to realize different upper probabilities or different lower probabilities for the outcomes heads or tails. Thus, in this case, there would be no difference in credal sets without a difference in lower/upper probability. But, in general, there is a many-to-one relationship between credal sets and lower/upper probability, which we may illustrate with the six following probability assessments,

$$\begin{aligned} p_1 &= (0.50, 0.25, 0.25) & p_2 &= (0.50, 0.35, 0.15) \\ p_3 &= (0.25, 0.60, 0.15) & p_4 &= (0.15, 0.60, 0.25) \\ p_5 &= (0.15, 0.35, 0.50) & p_6 &= (0.25, 0.25, 0.50) \end{aligned}$$

which serve as the extremal points of the credal set $\mathbb{P}_1$ in Figure 3(a), and the subset $\{p_2, p_4, p_6\}$ serves as the extremal points of the credal set $\mathbb{P}_2$ in Figure 3(b).

Although the credal sets $\mathbb{P}_1$ and $\mathbb{P}_2$ are distinct, both realize the same upper and lower probabilities for all three basic outcomes of $\mathcal{X}$, namely

$$\begin{aligned} \overline{\mathsf{p}}(h) &= 0.50 & \overline{\mathsf{p}}(s) &= 0.60 & \overline{\mathsf{p}}(t) &= 0.50 \\ \underline{\mathsf{p}}(h) &= 0.15 & \underline{\mathsf{p}}(s) &= 0.25 & \underline{\mathsf{p}}(t) &= 0.15 \end{aligned}$$

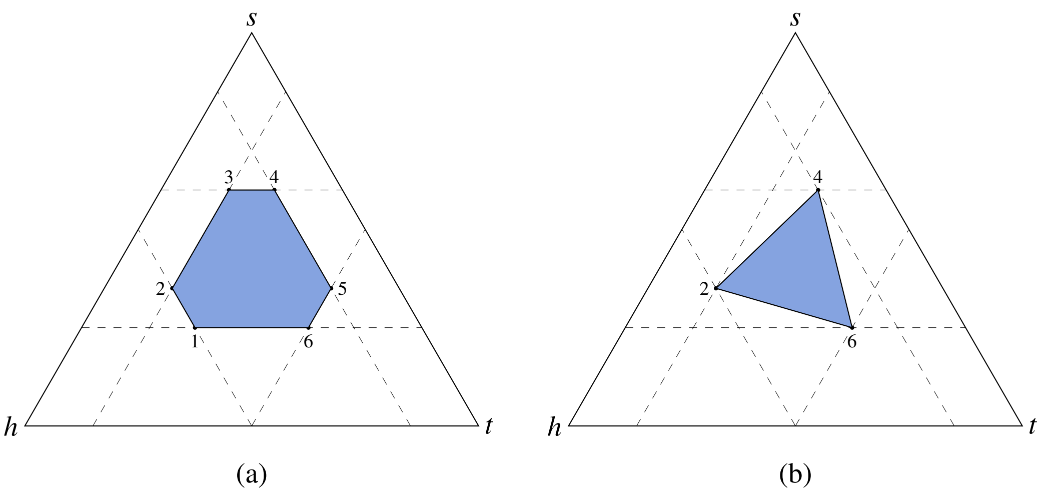


Figure 3: Two credal sets which induce the same lower and upper probabilities on $\mathcal{X}$. (a): the credal set $\mathbb{P}_1$ with extremal points $\{p_1, p_2, p_3, p_4, p_5, p_6\}$; (b): the credal set $\mathbb{P}_2$ with extremal points $\{p_2, p_4, p_6\}$.

which are represented by the three pairs of parallel dashed lines in the simplexes of Figure 3. This many-to-one relationship between credal sets and lower probability was observed by Peter Williams (1976, Theorem 2), who also remarked on the advantages afforded to the subjective interpretation of probability from moving away from the language of events. We will return to this point in Section 7.

The credal sets in Figure 3 are distinct. There are probability mass functions included in $\mathbb{P}_1$ that are excluded from $\mathbb{P}_2$, but this difference is not expressible in the language of lower/upper probability of events. A natural question to ask is whether this mismatch is merely a technical matter that has no practical bearing on probability judgments we may wish to represent and reason about. We turn in the next two sections to see that there is indeed different information encoded by different credal sets which have different behavioral consequences.

6 Gambles and Lower Previsions

The question raised in Section 5 is what mathematical models are suited to representing information encoded by credal sets that the language of lower probability on events cannot capture. In this section we introduce such a model but defer until Section 7 the issue of how your probability judgments are elicited and how the norms of the model are imagined to work, constructively, to ensure the mathematical properties necessary for coherence are invoked to evaluate such judgments. De Finetti's theory, for each of his two accounts of linear previsions, bundles together the definition of a prevision with criteria for coherence. The theory of lower previsions allows us to unpack some of these steps, but we nevertheless can begin by introducing lower previsions in a manner very similar to de Finetti's introduction of linear previsions, namely by providing the definition of a lower prevision together with conditions for coherence.

A *lower prevision model* is a triple $(\mathcal{X}, \mathbb{G}, \underline{P})$, where $\mathcal{X}$ is a set of possibilities, $\mathbb{G}$ is a linear space[13] of bounded gambles defined with respect to $\mathcal{X}$, and $\underline{P}$ is a real functional defined on $\mathbb{G}$ called a *lower prevision*. For gambles f and g in $\mathbb{G}$ and any strictly positive number λ, $\underline{P}$ is a *coherent lower prevision* if and only if the following conditions are satisfied (Walley 1991, Thm. 2.5.5):

LP1. $\underline{P}(f) \geq \inf f$ (**bounds**)

LP2. $\underline{P}(f+g) \geq \underline{P}(f) + \underline{P}(g)$ (**super-additivity**)

LP3. $\underline{P}(\lambda f) = \lambda \underline{P}(f)$ (**non-negative homogeneity**)

Lower previsions are simply lower expectations. LP1 states that the lower expectation of a gamble should never be less than the minimum of that gamble, when it is a bounded gamble, or the infimum of the gamble in general. LP2 says that lower previsions are super-additive, that is, that the lower expectation of the sum of two gambles should be never less than the sum of the lower expectations. LP3 says that if there is a non-negative constant within a lower expectation, that constant may be moved freely outside the scope of the lower expectation operator. In other words, the lower prevision of a gamble multiplied by a non-negative factor λ is the same as the lower prevision of that gamble scaled by λ. When the domain of $\underline{P}$ is restricted to only $\{0,1\}$-gambles, $\underline{P}$ is a lower probability.

Here it may be a useful to see de Finetti's coherent linear previsions expressed in these terms, namely when the domain of the prevision operator is assumed to be a linear space.[14] If P is a linear prevision on a linear space $\mathbb{G}$ of bounded gambles, then P is coherent if and only if the following two conditions are satisfied (Walley 1991, Thm. 2.8.4):

P1. $P(f) \geq \inf f$ (**bounds**)

P2. $P(f+g) = P(f) + P(g)$ (**additivity**)

The homogeneity condition, LP3 for lower previsions, follows as a consequence of P1 and P2 for linear previsions.[15]

The interpretation of lower previsions is an extension of the Smith-Williams notion of a one-sided bet, outlined in Section 2, extended to cover all (bounded) gambles and not only $\{0,1\}$-gambles. We remarked in Section 4 about the outsized role that propositions and the language of events play in some presentations of subjective probability theory. The key notion that turns lower previsions into a model of your uncertainty is the *acceptability* of a gamble, which, just as with the theory of linear previsions, is grounded in your willingness to have your fortune change in accordance with an uncertain outcome. Intuitively, the acceptability of a gamble indicates your disposition to accept it. Owing to the constructive nature of the theory, your disposition to accept might be an initial probability judgment, which may or may not be consistent, or, if consistent, may or may not be coherent. In later sections we will focus on how to check a set of initial probability judgments for consistency and how to construct a coherent model from consistent probability judgments. For now, we shall focus on the mathematical nuts and bolts of the theory and how the axioms of lower previsions are related to a coherent set of acceptable gambles.

[13] That is, $\mathbb{G}$ is closed under finite vector addition and scalar multiplication.

[14] Compare (de Finetti 1974, §3.5.1).

[15] For a proof, see (Troffaes and de Cooman 2014, p. 58) along with the observation that, for unbounded random variables, a homogeneity condition must be included as an axiom for (conditional) linear previsions as well.

Suppose the set $\mathbb{D} \subseteq \mathbb{G}(\mathcal{X})$ denotes the (non-empty) set of gambles that are acceptable to you with respect to the set of all possible gambles, $\mathbb{G}$, over the possibility space $\mathcal{X}$. Then, for any gamble $f \in \mathbb{G}$ and real number μ, the lower prevision operator $\underline{\mathrm{P}}(\cdot)$ is

$$\underline{\mathrm{P}}(f) := \sup\{\mu \in \mathbb{R} : f - \mu \in \mathbb{D}\} \tag{3}$$

which is read as the supremum price μ you are willing to pay for the gamble f. Alternatively, we say that the lower prevision of f, $\underline{\mathrm{P}}(f)$, is the supremum *acceptable* buying price for f associated with the set of gambles $\mathbb{D}$ that you find acceptable. Naturally, if $\underline{\mathrm{P}}$ is a coherent lower prevision, your announced prices are coherent. This definition offers no guidance for assessing whether your lower previsions are coherent, however. The question of assessment will be addressed in Section 7, where coherence axioms for sets of acceptable gambles are introduced. Nevertheless, if $\mathbb{D}$ is a *coherent set of acceptable gambles*, LP1-LP3 are derivable from Equation 3, allowing us to move between coherent lower previsions and coherent sets of acceptable gambles.[16]

The upper prevision operator $\overline{\mathrm{P}}(\cdot)$ is defined by

$$\overline{\mathrm{P}}(f) := \inf\{\mu \in \mathbb{R} : \mu - f \in \mathbb{D}\} \tag{4}$$

which is read as the infimum selling price μ for f that is acceptable to you with respect to $\mathbb{D}$. Just as with lower and upper probabilities, the functionals $\underline{\mathrm{P}}(\cdot)$ and $\overline{\mathrm{P}}(\cdot)$ satisfy a conjugacy condition, namely, for all possible bounded gambles $\mathbb{G}$,

$$\overline{\mathrm{P}}(f) := -\underline{\mathrm{P}}(-f) \text{ for all } f \in \mathbb{G} \tag{5}$$

which allows us to take lower previsions as primitive. Equation 5 asserts that selling the gamble f for a price μ is equivalent to buying $-f$ for $-\mu$. The reason that the lower prevision operator is typically taken as primitive is that surrendering a sure reward of μ in exchange for an uncertain reward f is regarded as more fundamental than pricing gambles to sell. For instance, the following transformations all express gambles acceptable to you as a consequence of endorsing the lower prevision $\underline{\mathrm{P}}(-f)$:

$$\begin{aligned}
\underline{\mathrm{P}}(-f) &= \quad \sup\{\mu \in \mathbb{R} : -f - \mu \in \mathbb{D}\} \\
&= -\inf\{-\mu \in \mathbb{R} : -f - \mu \in \mathbb{D}\} \\
&= -\inf\{\mu \in \mathbb{R} : -f + \mu \in \mathbb{D}\} \\
&= -\overline{\mathrm{P}}(f)
\end{aligned}$$

There are important exceptions, such as in finance (Shafer and Vovk 2019), that treat the upper prevision operator as primitive.[17]

The language of gambles affords you a larger vocabulary to express probability judgments than the language of events, even when your judgments only concern events. Let the set $\mathbb{D}$ denote to a set of gambles that are acceptable to you, and E and F be events. The following table displays some probability judgments and their translations into acceptable gambles and lower previsions.

[16]However, analogous to the many-to-one relationship between lower previsions and lower probability, there is also a many-to-one relationship between coherent sets of acceptable gambles and lower previsions. We discuss this in Section 8.

[17]See (de Cooman and Hermans 2008) for a Walley-style framework within which to embed Shafer and Vok's game-theoretic approach to probability.

Probability Judgment	**Acceptable Gamble**	**Lower Prevision**	
E is probable	$E - \frac{1}{2} \in \mathbb{D}$	$\underline{P}(E) \geq \frac{1}{2}$	
E is at least as probable as F	$E - F \in \mathbb{D}$	$\underline{P}(E - F) \geq 0$	*
E is twice as probable as F	$E - 2F \in \mathbb{D}$	$\underline{P}(E - 2F) \geq 0$	*
Conditional on G, E is at least as probable as F	$G(E - F) \in \mathbb{D}$	$\underline{P}(G(E - F) \geq 0$	*
The odds against E are at least 3 to 1	$E^c - 3E \in \mathbb{D}$	$\overline{P}(E) \leq \frac{1}{4}$	
E without F is improbable	$\frac{1}{2} - E + F \in \mathbb{D}$	$\overline{P}(E - F) \leq \frac{1}{2}$	*
Your degree of belief that E is $\frac{1}{3}$	$E - \frac{1}{3} \in \mathbb{D}$ and $\frac{1}{3} - E \in \mathbb{D}$	$\underline{P}(E) = \frac{1}{3}$ and $\overline{P}(E) = \frac{1}{3}$	

Table 1: Some probability judgments and their corresponding representations in terms of acceptable gambles and lower previsions. Even though all of the gambles are $\{0,1\}$-gambles, the comparative probability judgments marked by * cannot be modeled by lower probabilities.

As should be clear, degrees of belief are expressible and their corresponding (coherent) linear previsions are expressible in this language as well. Specifically, your degree of belief that E is recoverable when the lower prevision that E is identical to the upper prevision that E. Naturally, all events under consideration would need to be determinate valued to recover the properties of a linear prevision $P(\cdot)$. Some commentators run into trouble with examples involving some, but not all, determinate valued gambles by applying reasoning that is valid for linear previsions but invalid for lower previsions.[18]

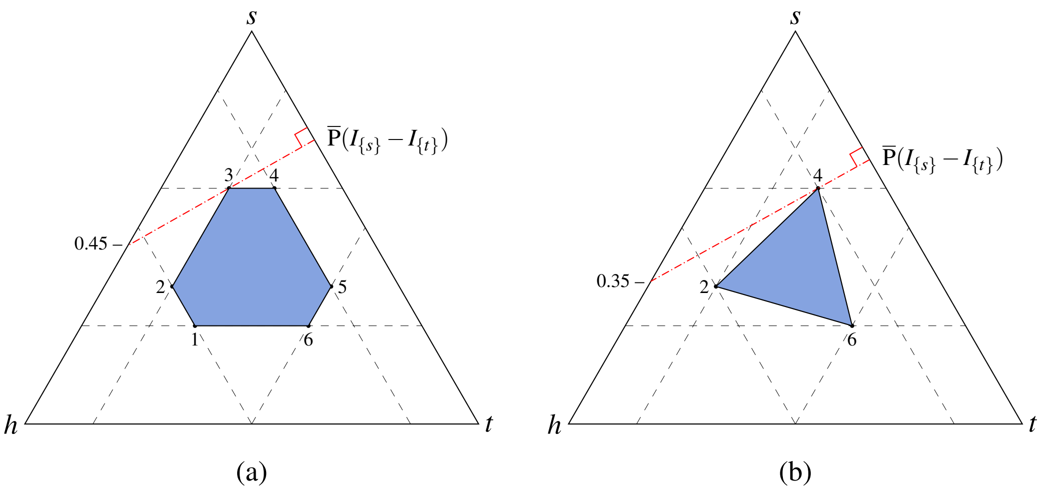


Figure 4: (a): The upper prevision $\overline{P}(I_{\{s\}} - I_{\{t\}}) = 0.45$ with respect to $\mathbb{P}_1$; (b): The upper prevision $\overline{P}(I_{\{s\}} - I_{\{t\}}) = 0.35$ with respect to $\mathbb{P}_2$.

Return now to the soup can experiment. Figure 4(a) displays the credal set $\mathbb{P}_1$ and

[18]For example, (White 2010, §5) and (Elga 2010, §4-5). See the discussions in (Pedersen and Wheeler 2014) and (Chandler 2014), respectively.

Figure 4(b) displays the credal set $\mathbb{P}_2$. Both credal sets induce the same lower and upper probabilities on $\mathcal{X}$ but two different upper previsions for the gamble $I_s - I_t$. The upper prevision $\overline{\mathrm{P}}(I_{\{s\}} - I_{\{t\}}) = 0.35$ with respect to $\mathbb{P}_2$ expresses that the gamble $0.35 - I_{\{s\}} + I_{\{t\}}$ is acceptable, which is not an acceptable gamble with respect to the upper prevision of $\overline{\mathrm{P}}(I_{\{s\}} - I_{\{t\}}) = 0.45$ defined on $\mathbb{P}_1$. Expressed in terms of probability mass functions, $\overline{\mathrm{P}}(I_{\{s\}} - I_{\{t\}}) \leq \mu$ asserts that every probability distribution p satisfying the inequality $\mu \geq p(s) - p(t)$ is acceptable, an assertion that clearly does not correspond to an event within the algebra $\mathscr{A}$ of $(\mathcal{X}, \mathscr{A}, \mathbb{P}, \underline{\mathrm{p}})$ but instead expresses a constraint on probability mass functions whereby p_3 is admissible to the set $\mathbb{P}_1$ but inadmissible to the set $\mathbb{P}_2$.

7 Acceptable Gambles and Partial Preference Orders

An acceptable gamble is one you prefer to the existing state of affairs, where the existing state of affairs is represented by a constant gamble that returns to you zero. The comparison of gambles underlying such judgments can be expressed in terms of a strict partial preference order

$$f \succ g \quad \text{iff} \quad f - g \succ 0 \quad \text{iff} \quad f - d \in \mathbb{D}. \tag{6}$$

which expresses that you strictly prefer f to g and therefore would readily exchange g for f. Alternatively, you my compare gambles under a non-strict partial preference order

$$f \succeq g \quad \text{iff} \quad f - g \succeq 0 \quad \text{iff} \quad f - d \in \mathbb{D}. \tag{7}$$

to express that you are not opposed to exchanging g for f. The structure of the set of gambles $\mathbb{D}$ in Equations 6 and 7 are different, owing to whether the preference ordering is irreflexive or reflexive, respectively. The main difference is whether the zero gamble is included among the gambles you find acceptable: under strict partial preference the zero gamble is not included in $\mathbb{D}$, under non-strict partial preference it is. The reason the rightmost equivalence in Equations 6 and 7 are the same, and do not reflect this difference, is that the expressive capacity of lower previsions cannot account for this difference and views all partial preference as non-strict partial preference. This point is discussed further in Section 8.

Coherence conditions for a set of acceptable gambles, $\mathbb{D}$, may be formulated directly in terms of axioms for partial preference. Following (Quaeghebeur 2014), we recite the axioms for coherent acceptable gambles induced by non-strict partial preference, expressed in terms of gambles f, g, h and $0 < \mu \leq 1$:

1. $f \succeq f$ (**Reflexivity**)
2. If $f \succeq g$ and $g \succeq h$, then $f \succeq h$ (**Transitivity**)
3. $f \succeq g$ iff $\mu f + (1-\mu)h \succeq \mu g + (1-\mu)h$ (**Mixture Independence**)
4. If $f > g$, then $f \succeq g$ and $g \not\succeq f$ (**Monotonicity**)

The coherence axioms for strict partial preference are identical to the axioms for non-strict partial preference, substituting $\succeq$ for $\succ$, once reflexivity is replaced by $1'$:

$1'$. $f \not\succ f$ (**Irreflexivity**)

From the axioms for non-strict preference, you can express that you are *indifferent* between f and g just in case both $f \succeq g$ and $g \succeq f$, just as you would under weak preference (von Neumann and Morgenstern 1944). However, in addition, you may also express that f and g are *incomparable* if and only if both $f \not\succeq g$ and $g \not\succeq f$.

A remark about terminology. You will increasingly find the terms 'acceptable gamble' and 'desirable gamble' used interchangeably in the literature, although Peter Walley distinguishes between three variants of desirable bounded gambles, and Teddy Seidenfeld refers to acceptable gambles as favorable gambles (Seidenfeld, Schervish, and Kadane 1990). The main issue criss-crossing this terminology is whether or not the zero gamble is included among the set of acceptable gambles, which corresponds here to whether the partial preference ordering is non-strict or strict. Following Troffaes and de Cooman (2014), we use 'acceptable' and 'desirable' interchangeably as an umbrella term, and let the specific mathematical properties specify which variety of acceptability is on order as the need arises.

8 Coherent Acceptable Gambles

Consider again the lower prevision model, $(\mathcal{X}, \mathbb{G}, \underline{\mathrm{P}})$. The key structural condition imposed on the set of bounded gambles $\mathbb{G}(\mathcal{X})$ is that it be a linear space, which a closed convex set of probability mass functions realizes in two ways. First, the convex hull of a credal set is simply the collection of all finite linear combinations of the probability mass functions in the credal set. Second, each mass function within a credal set works as a linear constraint on the lower prevision functional, $\underline{\mathrm{P}}(\cdot)$—a fact we used to show, with Figures 3 and 4, that credal sets contain more information than can be expressed by the lower and upper probabilities they support. The key information that a lower prevision uses from a credal set is the boundary of the convex hull, that is, the linear spans between the extremal points of the credal set. Lower previsions do not depend upon the probability mass functions as such. On the contrary, probability mass functions carry additional mathematical baggage that are a burden to imprecise probability models (Wheeler 2021b), especially when conditioning and updating events with probability zero (De Bock and de Cooman 2015a; Wheeler and Cozman 2021).

Credal sets also carry too little structure, a point we alluded to in Section 7. An equivalence may be drawn between non-strict partial preference of gambles, $\leq$, and lower previsions of the same domain of gambles, $\underline{\mathrm{P}}(\cdot)$: exactly the same information about your beliefs provided by one representation is provided by the other. However, strict partial preference cannot be recovered from a lower prevision. The issue concerns which gambles on the boundary should be included as acceptable and which should not, a distinction that is not supported by lower previsions (Walley 1991, §3.8.6). This additional information is nevertheless important when conditioning and updating events with probability zero. For principally this reason, and consequences that follow from it, sets of acceptable gambles are taken as theoretically primitive, which is why contemporary textbooks (Augustin, Coolen, de Cooman, and Troffaes 2014; Troffaes and de Cooman 2014) start with gambles and axioms for acceptability.

Before turning to this axiom system, there is another insight we may pull from credal sets. We remarked that the extremal points together with the linear span between those points is the necessary mathematical structure to determine a lower prevision. We nevertheless can separate these two conditions when turning to the question of how to secure such a structure. Namely, we might elicit from you a handful of judgments that in effect would yield extremal points, then construct the convex hull of those judgments through a closure operation. While far more tractable than imagining a continuum of credal judgements that one must come up with ex ante, as some commentators mistakenly presume,[19] there is nevertheless no straightforward means to elicit the specific probability mass functions which realize the extremal points of a credal set.[20]

The following four axioms are a distillation of axioms for coherent acceptable gambles proposed by (Williams 1975; Walley 2000) from (Troffaes and de Cooman 2014, §3.4). First, there are two simple rationality axioms governing gambles you should always avoid and gambles you should always accept.

A1. If $f < 0$, then $f \notin \mathbb{D}$ (**Avoid partial loss**)

A2. If $f \geq 0$, then $f \in \mathbb{D}$ (**Accept partial gain**)

Axiom A1 says that you should not accept any bounded gamble that will only yield you a negative reward, whereas axiom A2 says that you should accept any bounded gamble that will never yield you a negative reward.

The next pair of axioms construct the linearity properties that were stipulated to hold for $\mathbb{G}(\mathcal{X})$ of $(\mathcal{X}, \mathbb{G}, \underline{P})$ above.

A3. If $f \in \mathbb{D}$, then $\lambda f \in \mathbb{D}$ $(\forall \lambda \in \mathbb{R}. \lambda > 0)$ (**Positive Scale Invariance**)

A4. If $f \in \mathbb{D}$ and $g \in \mathbb{D}$, then $f + g \in \mathbb{D}$ (**Combination**)

Axioms A3 and A4 encode that we are adopting a linear utility scale to express bounded gambles. Axiom A3 says that your disposition to regard a gamble as acceptable is not changed by the introduction of a positive scale. Axiom A4 states that if you are disposed to accept f and to accept g, then you should be disposed to accept their sum. Alternatively, axioms A3 and A4 may be viewed as logical closure operations: A3 constructs a ray from the origin that passes through each gamble you accept, whereas A4 constructs a convex cone that combines all rays.

These four axioms entail a monotonicity property that underwrites dominance reasoning, namely

A5. If $f \in \mathbb{D}$ and $g \geq f$, then $g \in \mathbb{D}$ (**Monotonicity**)

A5, which follows from A2 and A4, says that any bounded gamble g that dominates an acceptable gamble f is an acceptable gamble, too.

The four axioms, A1, A2, A3, and A4, are deceptively simple, but the normative standard they encode should be familiar to any student of rational choice theory. If you

[19]See for example (Joyce 2011) and (Moss 2015), who reify the elements of credal sets and speak as if probability judgments that exercise the resources of an imprecise probability model must invoke a continuum of subjective probability judgements.

[20]Two important exceptions are group decision making, were each point is a Bayes agent (Walley 1982; Elkin and Wheeler 2018) and credal networks (Cozman 2000), although credal networks involve structural independence judgments on sets of probabilities (Cozman 2012; De Bock and de Cooman 2015b; Wheeler and Cozman 2021).

are disposed to accept a gamble that will impart to you a sure loss, you should not accept it. If you admit a collection of gambles that, on their own do not incur a sure loss but in combination do, you should not accept that combination of gambles. If a gamble cannot lose, you should accept it. Lastly, your judgments of acceptability are presumed to be invariant to positive scaling. In other words, the model presumes a linear utility scale.[21] Any set of accepted bounded gambles that satisfies A1 to A4 is called *coherent*.

The axiom system for coherent sets of acceptable gambles may also be used as a prescriptive model for rational belief and not only a normative standard for rational belief (Wheeler 2018). That is, axioms A2, A3, and A4 can be viewed as production rules that take as input a small collection of gambles you have considered and are disposed to accept. These axioms then reveal to you the rational consequences, in terms of A1-A4, that follow from holding such commitments. Hence, to speak of the acceptability of gambles and the disposition to accept a gamble, rather than of accepted gambles tout court, is to acknowledge the constructive nature of the theory. The axioms are designed to offer a prescriptive guide to better ex ante decision making, not some normative standard for meting out ex post justice. The process of *evaluating* a collection of your probability *judgments* for their suitability to construct a *coherent model* with which you are capable of making *inferences* should be familiar to students of statistical decision theory, even if the language of (bounded) gambles is perhaps unfamiliar.

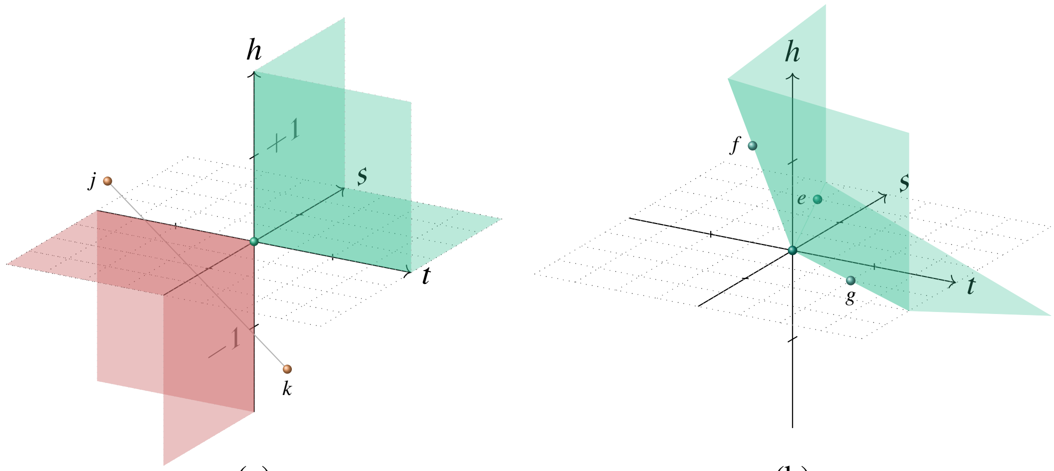
For example, suppose after deliberation you form a set of probability assessments, $\mathbb{A}$, that in your judgment are acceptable gambles. The set $\mathbb{A}$ is not a model per se, but simply a collection of assessments. The first step in evaluating the suitability of $\mathbb{A}$ for constructing a coherent model is to determine whether it is possible to extend $\mathbb{A}$ to a coherent set of acceptable gambles. In other words, the first step is to determine whether $\mathbb{A}$ is *consistent*, that is—whether your disposition to accept the gambles in $\mathbb{A}$ would, if accepted, avoid partial loss. Clearly, a set $\mathbb{A}$ of accepted gambles is consistent if it is a subset of some coherent set of acceptable gambles. But that fact does not offer much guidance. Fortunately, one can check the consistency of $\mathbb{A}$ directly by verifying there is no (non-negative) linear combination of gambles in $\mathbb{A}$ that yields a partial loss (Troffaes and de Cooman 2014, §3.4.2), namely that the following consistency condition is satisfied:

$$\sum_{j=1}^{n} \alpha_j f_j \nless 0, \text{ for bounded gambles } \{f_1, f_2, \ldots, f_n\} \text{ and non-negative } \alpha_i \in \mathbb{R} \qquad \textbf{(Con)}$$

where $f < 0$ denotes that f is nowhere positive and not zero. Given a set of assessments $\mathbb{A}$ that is consistent, a closure operation constructed from axioms A2, A3, and A4 can be applied to $\mathbb{A}$ to produce the smallest set of coherent acceptable bounded gambles, which is called the *natural extension* of $\mathbb{A}$. This inference mechanism is fundamental to the theory. As an aside, the construction of natural extensions via axioms A1-A4 subsumes classical propositional logic as a special case (de Cooman 2005, §5).

Return to the soup can experiment with outcomes $\mathcal{X} = \{h, s, t\}$. The positive orthant in Figure 5(a), the green zone whose borders are three hyperplanes, represents those gambles within the set $\mathbb{G}(\mathcal{X})$ of possible gambles that are nowhere negative. For each of

[21] This linear utility condition can be relaxed with a *discounted utility* that targets A3, the positive scale invariance condition for desirability (Wheeler 2021a)



(a) (b)

Figure 5: Coherent sets of acceptable gambles for outcomes *heads*, *side*, and *tails*. (a) The vacuous belief model with two assessments, $\mathbb{A}_a = \{j,k\}$, that are individually consistent but when combined are inconsistent. (b) A non-vacuous coherent set of acceptable gambles $\mathbb{D}$, whose boundaries are drawn in green, constructed from three assessments, $\mathbb{A}_b = \{e, f, g\}$.

these gambles you cannot lose, so they all are included in any coherent set of acceptable gambles with respect to $\mathcal{X}$ by A2. The red zone in Figure 5(a), by contrast, the negative orthant, represents all gambles that you cannot win, so must be excluded from any coherent set of acceptable gambles by A1.

Suppose you are disposed to accept the set of gambles $\mathbb{A}_a = \{j,k\}$ depicted in Figure 5(a). Each gamble in $\mathbb{A}_a$ is consistent, since each is somewhere positive: specifically, $j(h) > 0$ and $k(t) > 0$. But combinations of j and k are inconsistent, by A3 and A4. Therefore, $\mathbb{A}_a$ violates (Con). See Figure 5(a): although j and k are individually consistent, the linear span between these gambles includes gambles that are nowhere positive. Thus, $\mathbb{A}_a$ cannot be a subset of any coherent set of acceptable gambles. Yet, absent your assessments $\mathbb{A}_a$, that leaves the gambles in the green zone as your coherent set of acceptable gambles. This is the least committal yet coherent set of acceptable gambles on $\mathcal{X}$, a model that represents your disposition to refrain from having your fortune change from a toss of this can. So, the coherent set of acceptable gambles in Figure 5(a) induces a lower probability of zero for each outcome, *heads*, *side*, or *tails*. Lastly, notice the constructive nature of this example. We described a process in which you first offer an assessment, $\mathbb{A}_a$, discover from applying the axioms of the theory there is no possibility to construct a coherent model from $\mathbb{A}_a$ with which to reason, so, in light of this discovery, withdraw the assessment $\mathbb{A}_a$ and settle instead, at least for the moment, for the vacuous coherent model. Naturally, you may offer another assessment. But you are not rationally compelled to do so to construct a coherent model, nor are you prohibited, technically or otherwise, from updating a vacuous model in light of a new observation.

Figure 5(b) represents a non-vacuous coherent model generated by the assessment $\mathbb{A}_b = \{e, f, g\}$:

$\mathbb{A}_b$	*heads*	*side*	*tails*
f	$f(h) = 1.3$	$f(s) = -0.5$	$f(t) = -0.25$
e	$e(h) = -0.5$	$e(s) = 1.5$	$e(t) = 1$
g	$g(h) = 0$	$g(s) = -0.5$	$g(t) = 1$

Similar to $\mathbb{A}_a$, each gamble in $\mathbb{A}_b$ is consistent since each is somewhere positive. Unlike $\mathbb{A}_a$, $\mathbb{A}_b$ satisfies (Con). So, $\mathbb{A}_b$ is a consistent assessment. Hence A2, A3, and A4 may be applied to $\mathbb{A}_b$ to construct a coherent set of acceptable gambles, called its *natural extension*. The green hyperplanes in Figure 5(b) depict the positive hull of the natural extension of $\mathbb{A}_b$, the border of the convex cone of gambles you are, rationally, disposed to accept. As an aside, since the points in any convex cone on $\mathbb{G}(\mathcal{X})$ are gambles, not probability mass functions, this model affords much more flexibility to introduce different border structures for the positive hull than credal sets offer, as the infimum is defined whether a coherent set of acceptable gambles is an open or closed set. For principally this reason, coherent sets of acceptable gambles can represent the difference between strict and non-strict partial preference, by adjusting axioms A1 and A2 to instead compare the infimum of gambles to zero, even if the lower prevision induced by a coherent set of acceptable gambles will not encode this difference.

The vacuous coherent model of Figure 5(a) is one extreme, representing the least committal yet coherent stance you can take on the outcome of this can-toss experiment. A maximally committed stance would be one in which, for every possible gamble $f \in \mathbb{G}(\mathcal{X})$, you are either disposed to accept f, disposed to accept its negation $-f$, or disposed to accept both f and $-f$, in which case your supremum buying price for f and infimum selling price for f is identical. In other words, a maximally coherent set of desirable gambles is a special type of convex cone of gambles, namely a halfspace, whose border is a hyperplane. Since the possibility space for our gambles about the can toss are in $\mathbb{R}^3$, a maximal coherent set of acceptable gambles will be a half space whose surface is a two-dimensional plane that must pass through the origin and not bisect either the negative orthant or the positive orthant in accordance with A1 and A2, respectively. Figure 6 represents just such a maximal coherent set of desirable gambles. This hyperplane of gambles is your expectation, that is, your linear prevision over $\mathcal{X}$. With respect to this expectation, if the expectation of a gamble is positive you will accept it, and if the expectation is negative you will not. Thus, the gamble j in Figure 6(b) is not accepted. Gambles on the hyperplane are those which you assess to be fairly priced.

9 A Brief Word on Conditional Lower Previsions

Although the inference mechanism used to construct a natural extension is the most fundamental, it really only initializes a model for probabilistic reasoning.[22] But probabilistic reasoning includes updating, structural assessments, stochastic processes—all of which involve conditional lower previsions. We have mostly avoided discussing conditional

[22]Note that, when working directly with lower previsions and conditional lower previsions, an important alternative to natural extensions that is the slightly more restrictive inference mechanism for the *regular extension* (Augustin, Coolen, de Cooman, and Troffaes 2014, p. 53), introduced by Walley to handle conditioning on probability zero events. Working directly with sets of acceptable gambles removes most, if not all, the technical reasons for relying on regular extensions to handle conditioning on probability zero events.

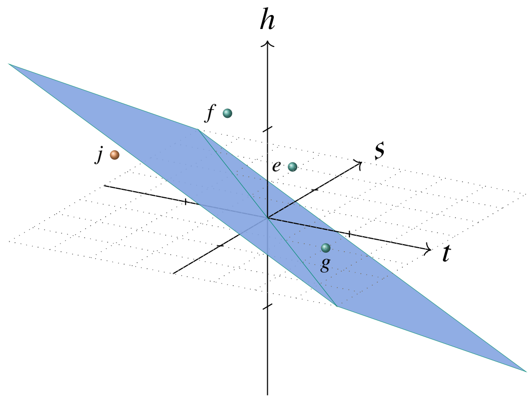


Figure 6: A maximally coherent set of acceptable gambles. With respect to the separating hyperplane in blue, gambles e, f, g are acceptable and j is not acceptable.

lower previsions, but in this section we briefly sketch how conditional lower previsions are defined with respect to a set of acceptable gambles and how updating works via a generalization of Bayes's rule.

Conditional probability judgments, whether precise or imprecise, concern your commitments at a specific time and can be interpreted in at least one of three ways.[23] First, a conditional probability may be interpreted as a contingent probability judgment, one that expresses your disposition to "call-off" a commitment to accept an unconditional gamble f unless a contingent event G occurs. This *called-off* interpretation of a conditional probability is behind the conditional probability judgment in Table 1. Since G is a $\{0,1\}$-gamble, $G = 1$ if the conditioning event G occurs, so $1 \cdot f$ represents that f contingent on G is acceptable to you. Otherwise $G = 0$, thus $0 \cdot f$, which in effect "calls-off" the gamble f and returns to you the status quo ante reward of zero if G does not occur.

For the remaining two interpretations, the conditioning event G is identified with subsets of $\mathcal{X}$ rather than merely the value, 1 or 0, of the indicator function I_G. Whereas a called-off gamble contingent on G merely depends on G occurring, hypothetical conditional judgments and temporal updating conditional judgments both depend on the restriction on the possibility space, $\mathcal{X}$, that occurs in the event that $I_G = 1$. A *hypothetical* conditional judgment is like a contingent conditional judgment in expressing your current disposition, or more generally, a disposition at a single point in time. Contingent judgments and hypothetical conditional judgments are both synchronic probability judgments. Unlike contingent conditional judgments, a hypothetical conditional judgment expresses your disposition to accept the gamble f on the hypothetical assumption that the event G occurs, where your judgment of acceptability for f is made with respect to the restriction on the set of possibilities $\mathcal{X}$ prescribed by G. A *temporal updating* conditional judgment, in turn, resembles a hypothetical conditional judgment in that both

[23]The literature on "conditionals" is enormous. But even restricting attention to probabilistic modeling of seemingly basic conditional statements in the indicative mood suggests that these three interpretations are far from exhaustive. Experimental work (Collins, Krzyżanowska, Hartmann, Wheeler, and Hahn 2020) eliciting how people's prior commitments change in light of receiving testimonial evidence in the form of a simple indicative conditional shows stable and intuitive to interpret responses that nevertheless do not admit a coherent probabilistic model.

involve the evaluation of a gamble with respect to the subset of possibilities associated with the conditioning event, G. But, unlike hypothetical conditional judgments, temporal updating conditional judgments express your future commitment to a gamble f upon observing that G obtains. In other words, temporal conditional judgments are diachronic probability judgments. Similar remarks apply to conditional value judgments and assessments of admissibility.[24]

There are several assumptions that hypothetical and temporal updated lower previsions involve, the foremost being that both forms of conditional judgments assume that the conditioning event G is identified with subsets of $\mathcal{X}$. For temporal updated previsions, procedures for making an observation and ensuring that such an observation suffices to establish that the true state is in G are also necessary, but exogenous to the theory of lower previsions. Furthermore, conditional temporal conditioning assumes that the information gathered from this observation is the only information you gather pertaining to $\mathcal{X}$.

How are these three types of conditional judgments related to one another? The short answer is they are all called-off gambles, but each imposes increasingly restrictive conditions for connecting a risky reward to a contingent event. A contingent probability judgment is the least restrictive because it only depends on G taking the value of 1 or 0, leaving unspecified the conditions under which G is assigned 1 or 0. Your yes-no assessment of G is enough. Then, by simple multiplication, the contingent gamble is called-off unless $G = 1$. Both hypothetical conditional judgments and temporal updating judgments, by contrast, explicitly identify the event G with a subset of possibilities $\mathcal{X}$, namely those $x \in \mathcal{X}$ identified by the indicator I_G. Here the supposition that $G = 1$ entails that G is an event defined with respect to a subset of $x \in \mathcal{X}$. The act of supposing that G or the act of observing that G circumscribe the space of possibilities $\mathcal{X}$, creating a *conditioned slice*, written $\mathbb{D}\rfloor_G$, of the acceptable gambles $\mathbb{D} \subseteq \mathbb{G}(\mathcal{X})$ with respect to G, For any gamble f, f is in the slice of gambles conditioned on G just in case the called off gamble $G \cdot f$ is in $\mathbb{G}$, that is

$$f \in \mathbb{D}\rfloor_G \quad \text{iff} \quad I_G \cdot f \in \mathbb{D} \qquad \textbf{(Conditioned Slice Condition)}$$

where I_G is the indicator of G. So, I_G expresses that f is called-off unless $I_G = 1$, and the conditioned slice condition connects calling off f contingent on G to the indicator I_G, and membership of f in the set of acceptable gambles $\mathbb{D}\rfloor_G$ circumscribed by $I_G = 1$.

A conditional lower prevision may be defined with respect to a coherent set of acceptable gambles $\mathbb{D}$, and may be given either a suppositional interpretation or a temporal updating interpretation. For any gamble f, $\{0,1\}$-gamble G, real number μ, and coherent set of acceptable gambles $\mathbb{D}$ defined with respect to the bounded gambles $\mathbb{G}(\mathcal{X})$, a *conditional lower prevision* is defined with respect to $\mathbb{D}$ by

$$\underline{\mathrm{P}}(f|G) := \sup\{\mu : G(f-\mu) \in \mathbb{D}\} \tag{8a}$$
$$:= \sup\{\mu : f-\mu \in \mathbb{D}\rfloor_G\}, \text{ for all } f \in \mathbb{G}(\mathcal{X}) \tag{8b}$$

[24]Related observations have been made by (Ramsey 1926; Raiffa and Schlaiffer 1961; Hacking 1967; Kyburg 1968; Levi 1980; Walley 1991; Kadane, Seidenfeld, and Schervish 2008; Pedersen and Wheeler 2015).

Equations 8a and 8b are equivalent, by the definition of a conditioned slice. However, Equation 8a corresponds directly to the suppositional interpretation of the conditional lower prevision, as the assessment of a gamble f on the supposition that G is made with respect to $\mathbb{D}$ is made at a single point in time, namely when the coherent set of acceptable gambles $\mathbb{D}$ was constructed. Although Equation 8b may be read synchronically as well, Equation 8b is the basis for the temporal updating interpretation of conditional lower previsions namely when the subset of gambles $\mathbb{D}\rfloor_G$ is treated as the "updated" set of gambles you are committed to adopting after observing G. So, the temporal updating interpretation distinguishes between the conditional judgments you make now with respect to $\mathbb{D}$ and the commitments you are prepared to make in the future after observing G, denoted by $\mathbb{D}\rfloor_G$.[25]

The salient difference between the two interpretations of Equation 8 is that a conditional lower prevision may be used either to evaluate the consequences of your conditional probability judgments, or it may be used to update your probability judgments after an observation. The theory does not mandate that you use either interpretation, nor does it mandate temporal conditionalization as a constraint on future commitments. For reasons why updated by Bayes rule should not be treated as mandatory and so-called "reflection" principles deflected, see (Levi 1987; Seidenfeld 1995).

Note that under either interpretation, Equation 8 is defined when $\underline{\mathrm{P}}(G) = 0$, and even when $\overline{\mathrm{P}}(G) = 0$. Therefore, conditioning on events of probability zero is straightforward. Furthermore, each interpretation supports the *generalized Bayes rule* (Walley 1991, Ch. 6) for conditional lower previsions, defined by

$$\underline{\mathrm{P}}\left(G[f - \underline{\mathrm{P}}(f|G)]\right) = 0 \tag{9}$$

where it is assumed that both $\underline{\mathrm{P}}(G) > 0$ and $G[f - \underline{\mathrm{P}}(f|G)] \in \mathbb{D}$. Finally, Equation 9 can be performed directly on a coherent set of acceptable gambles $\mathbb{D}$:

$$\mathbb{D}\rfloor_G = \{f \in \mathbb{G}(\mathcal{X}) : Gf \in \mathbb{D}\} \tag{10}$$

where, as before, $\mathbb{D}\rfloor_G$ may be interpreted either as the consequences from generalized Bayes rule under the supposition that G or the updated set of desirable gambles resulting from generalized Bayes updating.

10 Coherence as a Guide to Life

Coherent sets of acceptable gambles make updating, and reasoning that depends on updating, easy. The infimum of a coherent set of acceptable gambles is always defined, whether that set is open or closed, so conditioning and updating on events that have probability zero is no different than conditioning and updating on any other event. Eliciting probability judgments in terms of sets of acceptable gambles is straight forward, and affords a clear distinction between probability judgments provided as input to a model and probability judgments that follow as consequences of a coherent model. The language of

[25] Walley (1991, Ch. 6) and De Cooman and Troffaes (2014) only distinguish between contingent conditional judgments and updated conditional judgments, and De Cooman and Troffaes include additional notation for "input" conditioning events and "output" conditional lower previsions to emphasize the temporal change that updating affords.

acceptability can also accommodate non-binary choice, provided it is extended with a notion of disjunction to express that at least one of a pair of gambles is acceptable (De Bock and de Cooman 2019b). On the other hand, for many decision problems strict partial preference can offer a better representation than sets of acceptable gambles, and many examples of probabilistic reasoning are best served by a traditional probability model. But the last two remarks point to a strength of the generality of coherent sets of acceptable gambles. In addition to being a useful theory on its own terms, sets of acceptable gambles serve as an I/O bus for imprecise probability, allowing you to transform a problem posed in one uncertainty framework into the terms of another or, just as important, reveal to you conditions necessary for performing such a transformation. It is hard to fault these aims and a mathematical theory that achieves them.

Garber's view that the Bayesian framework is principally to do with punishing incoherent commitments or inspiring you to refrain from having them is shortsighted. A normative theory ought to not simply provide a normative standard by which to assess probability judgments, but also offer prescriptive guidance for how to construct a model from such judgments. People do not have probability judgments per se but construct them. Ignoring the role that elicitation plays in constructing coherent models gives free rein to speculative mathematical psychology. Contemporary disputes over the relationship between probability judgments and numerical probability mirror a 19th century debate over the relationship between value judgments and numerical utility (Stigler 1950). Comparative judgments of value afford numerical representation by a cardinal utility function, when certain conditions are met, and the elicitation of your judgment under those conditions is what both warrants and effects locating your comparative judgments on a numerical scale. Benthamite utility, with its presumption that pleasure and pain were themselves quantities, was abandoned because such quantities are impossible to measure. Similarly, one's cognitive commitments are not generated by the possession of a "credal function", or the possession of a "committee" of credal functions. Human psychology is not equipped with numerical functions. Projects to describe the "accuracy" of one's credences (Pettigrew 2013) invoke quantities that are similarly impossible to measure (Mayo-Wilson and Wheeler 2015). Isomorphisms are cheap to come by. Meaningful measurements, on the other hand, are dear.

Bishop Butler remarked that probability is the very guide to life. The ingenuity of Bayesianism is to ground this advice in coherence, bringing to heel very hard questions about what alethic probability could be, never mind how it could guide anyone, by banking on your wish to protect your interests to construct a means to reason about probabilistic judgments you might entertain. The theory of lower previsions extends this program, both in terms of extending the range of probability judgments that may be brought under management, and also improving our understanding of how to construct a coherent model from assessments one might offer. Mathematical probability is not the very guide to life. Coherence is.[26]

[26]Thanks to Thomas Augustin, Seamus Bradley, Fabio Cozman, Dominik Klein, Tim Maudlin, and Eric Schliesser for comments on earlier drafts.

References

Augustin, T., F. P. A. Coolen, G. de Cooman, and M. C. M. Troffaes (2014). *Introduction to Imprecise Probabilities*. Chichester, West Sussex: Wiley and Sons.

Berger, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis* (2nd ed.). New York: Springer.

Bradley, S. (2019). A counterexample to three imprecise decision theories. *Theoria 85*(1), 18–39.

Bradley, S. and K. Steele (2016). Can free evidence be bad? value of information for the imprecise probabilist. *Philosophy of Science 83*(1), 1–28.

Carnap, R. (1962). *The Logical Foundations of Probability* (2nd ed.). University of Chicago Press.

Chandler, J. (2014). Subjective probabilities need not be sharp. *Erkenntnis 7*(6), 1273–1286.

Collins, P. J., K. Krzyżanowska, S. Hartmann, G. Wheeler, and U. Hahn (2020). Conditionals and testimony. *Cognitive Psychology 122*, 101329.

Corsato, C., R. Pelessoni, and P. Vicig (2019). Nearly-Linear uncertainty measures. *International Journal of Approximate Reasoning 114*, 1–28.

Cozman, F. (2000). Credal networks. *Artificial Intelligence 120*(2), 199–233.

Cozman, F. (2012). Sets of probability distributions, independence, and convexity. *Synthese 186*(2), 577–600.

Cozman, F. G. and T. Seidenfeld (2009). Independence for full conditional measures and their graphoid properties. In B. Lowe, E. Pacuit, and J.-W. Romeijn (Eds.), *Reasoning about Probabilities and Probabilistic Reasoning*, Volume 16 of *Foundations of the Formal Sciences VI*, pp. 1–29. London: College Publications.

De Bock, J. and G. de Cooman (2015a). Conditioning, updating and lower probability zero. *International Journal of Approximate Reasoning 67*, 1–36.

De Bock, J. and G. de Cooman (2015b). Credal networks under epistemic irrelevance: The sets of desirable gambles approach. *International Journal of Approximate Reasoning 56, part B*, 178–207.

De Bock, J. and G. de Cooman (2019a, May). Interpreting, axiomatising and representing coherent choice functions in terms of desirability. arXiv:1903.00336.

De Bock, J. and G. de Cooman (2019b). Interpreting, axiomatising and representing coherent choice functions in terms of desirability. In J. De Bock, C. P. de Campos, G. de Cooman, E. Quaeghebeur, and G. Wheeler (Eds.), *Proceedings of the Eleventh International Symposium on Imprecise Probabilities: Theories and Applications*, Volume 103 of *Proceedings of Machine Learning Research*, Thagaste, Ghent, Belgium, pp. 125–134.

de Cooman, G. (2005). Belief models: An order-theoretic investigation. *Annals of Mathematics and Artificial Intelligence 45*(5–34).

de Cooman, G. and F. Hermans (2008). Imprecise probability trees: Bridging two theories of imprecise probability. *Artificial Intelligence 11*, 1400–1427.

de Cooman, G. and E. Miranda (2007). Symmetry of models versus models of symmetry. In W. Harper and G. Wheeler (Eds.), *Probability and Inference: Essays in Honor of Henry E. Kyburg, Jr.*, pp. 67–149. London: King's College Publications.

de Finetti, B. (1937). La prévision: ses lois logiques, ses sources subjectives. *Annales de l'Institut Henri Poincaré 7*(1), 1–68. Trans. Henry E. Kyburg, Jr., Foresight: Its logical laws, its subjective sources. In H. E. Kyburg, Jr. and H. Smokler (Eds.), *Studies in Subjective Probability*. New York: John Wiley and Sons, 1964.

de Finetti, B. (1974). *Theory of Probability: A Critical Introductory Treatment* (2017 ed.), Volume I and II. West Sussex: Wiley.

Elga, A. (2010, May). Subjective probabilities should be sharp. *Philosophers Imprint 10*(5).

Elkin, L. and G. Wheeler (2018). Resolving peer disagreements through imprecise probabilities. *Noûs 52*(2), 260–94.

Etner, J., M. Jeleva, and J.-M. Tallon (2012). Decision theory under ambiguity. *Journal of Economic Surveys 26*(2), 234–270.

Fierens, P., L. C. Rêgo, and T. L. Fine (2009). A frequentist understanding of sets of measures. *Journal of Statistical Planning and Inference 139*(6), 1879–1892.

Garber, D. (1983). Old evidence and logical omniscience in Bayesian confirmation theory. In J. Earman (Ed.), *Testing Scientific Theories*, Minneapolis, pp. 99–132. University of Minnesota.

Gilboa, I. and D. Schmeidler (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics 18*(2), 141–153.

Grünwald, P. and J. Y. Halpern (2004). When ignorance is bliss. In J. Y. Halpern (Ed.), *Proceedings of the 20th Conference on Uncertainty in Artificial Intelligence (UAI '04)*, Arlington, Virginia, pp. 226–234. AUAI Press.

Hacking, I. (1967). Slightly more realistic personal probability. *Philosophy of Science 34*(4), 311–325.

Haenni, R., J.-W. Romeijn, G. Wheeler, and J. Williamson (2011). *Probabilistic Logics and Probabilistic Networks*. Synthese Library. Dordrecht: Springer.

Jansen, C., G. Schollmeyer, and T. Augustin (2021). Quantifying degrees of E-admissibility in decision making with imprecise probabilities. In T. Augustin, F. Cozman, , and G. Wheeler (Eds.), *Essays in Honor of Teddy Seidenfeld*, Theory and Decision Library A. Springer.

Jeffrey, R. (1965). *Logic of Decision*. New York: McGraw-Hill.

Joyce, J. (2011). A defense of imprecise credences in inference and decision making. *Philosophical Perspectives 24*(1), 281–323.

Kadane, J. B., M. J. Schervish, and T. Seidenfeld (1999). *Rethinking the Foundations of Statistics*. Cambridge University Press.

Kadane, J. B., T. Seidenfeld, and M. J. Schervish (2008). Is ignorance bliss? *Journal of Philosophy 105*(1), 5–36.

Kyburg, Jr., H. E. (1968). Bets and beliefs. *American Philosophical Quarterly 1*, 54–63.

Levi, I. (1980). *The Enterprise of Knowledge*. Cambridge, MA: MIT Press.

Levi, I. (1987). The demons of decision. *Monist 70*(2), 193–211.

Mayo-Wilson, C. and G. Wheeler (2015). Epistemic decision theory's reckoning. PhilSci Archive.

Moss, S. (2015). Time-slice epistemology and action under indeterminacy. In T. S. Gendler and J. Hawthorne (Eds.), *Oxford Studies in Epistemology*, Volume 5, pp. 172–194.

Oberguggenberger, M., J. King, and B. Schmelzer (2009). Classical and imprecise probability methods for sensitivity analysis in engineering: A case study. *International Journal of Approximate Reasoning 50*(4), 680–693.

Pedersen, A. P. and G. Wheeler (2014). Demystifying dilation. *Erkenntnis 79*(6), 1305–1342.

Pedersen, A. P. and G. Wheeler (2015). Dilation, disintegrations, and delayed decisions. In *Proceedings of the 9th Symposium on Imprecise Probabilities and Their Applications (ISIPTA)*, Pescara, Italy, pp. 227–236.

Pedersen, A. P. and G. Wheeler (2019). Dilation and asymmetric relevance. In J. De Bock, C. P. Campos, G. de Cooman, E. Quaeghebeur, and G. Wheeler (Eds.), *Proceedings of Machine Learning Research*, Volume 103 of *The 11th Symposium on Imprecise Probabilities and Their Applications (ISIPTA)*, pp. 324–326.

Pettigrew, R. (2013). Epistemic utility and norms for credences. *Philosophy Compass 8*(10), 897–908.

Quaeghebeur, E. (2014). Desirability. In T. Augustin, F. P. A. Coolen, G. de Cooman, and M. C. M. Troffaes (Eds.), *Introduction to Imprecise Probabilities*, pp. 1–27. West Sussex: Wiley.

Raiffa, H. and R. Schlaiffer (1961). *Applied Statistical Decision Theory*, Volume 1 of *Studies in Managerial Economics*. Harvard Business School Publications.

Ramsey, F. P. (1926). Truth and probability. In H. E. Kyburg and H. E. Smokler (Eds.), *Studies in subjective probability* (Second (1980) ed.)., pp. 23–52. Huntington, New York: Robert E. Krieger Publishing Company.

Savage, L. J. (1954). *Foundations of Statistics*. New York: Wiley.

Seidenfeld, T. (1995). When normal and extensive form decisions differ. In D. Prawitz, B. Skyrms, and D. Westerstahl (Eds.), *Proceedings of the 9th International Congress of Logic, Methodology and Philosophy of Science*, Volume 134, pp. 451–463. Elsevier Science B. V.

Seidenfeld, T. (2004). A contrast between two decision rules for use with (convex) sets of probabilities: Γ-Maxmin versus E-admissibility. *Synthese 140*, 69–88.

Seidenfeld, T., M. J. Schervish, and J. B. Kadane (1990). Decisions without ordering. In W. Sieg (Ed.), *Acting and Reflecting*, Volume 211, pp. 143–170. Kluwer Academic Publishers.

Seidenfeld, T., M. J. Schervish, and J. B. Kadane (2010). Coherent choice functions under uncertainty. *Synthese 172*(1), 157–176.

Seidenfeld, T. and L. Wasserman (1993). Dilation for sets of probabilities. *The Annals of Statistics 21*(9), 1139–154.

Shafer, G. (1976). *A Mathematical Theory of Evidence*. Princeton, NJ: Princeton University Press.

Shafer, G. and V. Vovk (2019). *Game-Theoretic Foundations for Probability and Finance*. Hoboken, NJ: Wiley.

Smith, C. A. B. (1961). Consistency in statistical inference (with discussion). *Journal of the Royal Statistical Society 23*, 1–37.

Stewart, R. and I. Ojea Quintan (2018). Probabilistic opinion pooling with imprecise probabilities. *Journal of Philosophical Logic 47*(1), 17–47.

Stigler, G. J. (1950). The development of utility theory i. *The Journal of Political Economy 58*(4), 307–327.

Troffaes, M. C. (2007). Decision making under uncertainty using imprecise probabilities. *International Journal of Approximate Reasoning 45*, 17–29.

Troffaes, M. C. M. and G. de Cooman (2014). *Lower Previsions*. Chichester, West Sussex: Wiley and Sons.

von Neumann, J. and O. Morgenstern (1944). *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press.

Walley, P. (1981, July). Coherent lower (and upper) probabilities. Statistics Research Report 22, University of Warwick, Coventry, England.

Walley, P. (1982). The elicitation and aggregation of beliefs. Statistics Research Report 23, University of Warwick.

Walley, P. (1991). *Statistical Reasoning with Imprecise Probabilities*. London: Chapman and Hall.

Walley, P. (2000). Towards a unified theory of imprecise probability. *International Journal of Approximate Reasoning 24*, 125–148.

Wheeler, G. (2018). Bounded rationality. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2018 ed.). Metaphysics Research Lab, Stanford University.

Wheeler, G. (2021a). Discounting desirable gambles. In *Proceedings of Machine Learning Research*, The 12th Symposium on Imprecise Probabilities and Their Applications (ISIPTA).

Wheeler, G. (2021b). Moving beyond sets of probabilities. *Statistical Science 36*(2), 201–204.

Wheeler, G. and F. G. Cozman (2021). On the imprecision of full conditional probabilities. *Synthese*. https://doi.org/10.1007/s11229-020-02954-z.

White, R. (2010). Evidential symmetry and mushy credence. In T. S. Gendler and J. Hawthorne (Eds.), *Oxford Studies in Epistemology*, Volume 3, pp. 161–186. Oxford University Press.

Williams, P. M. (1975). Notes on conditional previsions. School of Mathematical and Physical Sciences, University of Sussex. Republished in *International Journal of Approximate Reasoning*, 44(3): 366–383, 2007.

Williams, P. M. (1976). Indeterminate probabilities. In M. Przełęcki, K. Szanianwski, and R. Wójcicki (Eds.), *Formal Methods in the Methodology of Empirical Sciences*, Warsaw, pp. 229–246. Ossolineum.