Accuracy Dominance on Infinite Opinion Sets

by

Mikayla Kelley

A thesis submitted in partial satisfaction of the

requirements for the degree of

Master of Arts

in

Mathematics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Lawrence Evans, Chair
Professor Steve Evans
Assistant Professor Pierre Simon
Professor Jon Wilkening

Summer 2019

Accuracy Dominance on Infinite Opinion Sets

Abstract

Accuracy Dominance on Infinite Opinion Sets

by

Mikayla Kelley

Master of Arts in Mathematics

University of California, Berkeley

Professor Lawrence Evans, Chair

There is a well-known equivalence between avoiding accuracy dominance and having probabilistically coherent credences (see, e.g., de Finetti 1974, Joyce 2009, Predd et al. 2009, Pettigrew 2016). However, this equivalence has been established in the accuracy literature only when the opinion set over which credences are defined is finite. In Chapter 1, we establish connections between accuracy dominance and coherence when credences are defined over certain classes of infinite opinion sets. One class of opinion sets for which we prove results is the class of countable *point-finite* opinion sets. In Chapter 2, we characterize the countable algebras of sets that can be generated by a point-finite collection.

Contents

Acknowledgments

Thank you to Guram Bezhanishvili, Craig Evans, Pierre Simon, James Walsh, participants at the Probability & Logic Conference 2019, and members of the Berkeley-Stanford Logic Circle for helpful comments. Thank you to Kiran Luecke for his insight and patience even in my most frustrated moments. Thank you to Wesley Holliday for his invaluable encouragement, guidance, and feedback throughout this project. Finally, thank you to my parents and sisters for being constant reminders of what is most important in life.

Chapter 1

Accuracy and Coherence: The Infinite Case

Formal epistemologists study knowledge and rationality using mathematical tools such as logic and probability theory. A central topic of study is the constraints on rational *degrees of belief*, or *credences*, in propositions. Formal epistemologists model credences using a standard mathematical framework (see, e.g., Huber 2016, Sec. 2). There is a nonempty set W whose elements represent *possible worlds*; subsets of W represent *propositions*; and a world w being an element of a proposition is interpreted to mean that the proposition is true at w. A set of propositions is called an *opinion set*; and a person's credences in the propositions in an opinion set are represented by a *credence function*, which is a map from the opinion set to the interval $[0, 1]$.

For example, consider a person's uncertainty about a sequence of n coin flips. The set of possible worlds consists of all possible ways the n coin flips could land, i.e., all sequences of "heads" and "tails" of length n. The set of sequences in which "heads" shows up in the kth entry represents the proposition "the kth coin flip lands heads." The sequence of all "heads" is an element of this proposition, which is interpreted to mean that the proposition "the kth coin flip lands heads" is true at the world in which the coin lands heads every time. The set of propositions of the form "the kth coin flip lands heads" for $k \in \{1, \ldots, n\}$ is an example of an opinion set; and an agent who believes that each of the n coin flips has a 50% chance of landing heads would have a representing credence function on this opinion set that is the constant function with value .5.

Philosophers have discussed extensively what the constraints are on a rational agent's credence function at a point in time (see, e.g., Talbott 2016, Vineberg 2016, Pettigrew 2016). For example, there seems to be something irrational about being more confident that the coin will land heads n times than that the coin will land heads at least once, since the coin lands heads at least once whenever the coin lands heads n times. Such credences are irrational according to a popular philosophical view called *probabilism*: a rational person's degrees of belief should satisfy the axioms of probability. The topic of this chapter is a collection of mathematical results that show there is reason to satisfy the constraints imposed by probabilism, based solely on a concern for having *accurate* credences.

Toward formally defining accuracy for credence functions, for $w \in W$, we let $v_w : \mathcal{F} \to \{0,1\}$ be the map given by $v_w(p) = 1$ if and only if $w \in p$. Each v_w is a possible truth assignment to the propositions in $\mathcal{F}$, namely the truth assignment which maps all and only the true propositions at world w to 1. The *inaccuracy* of a credence function c at a world $w \in W$ is the distance $\mathfrak{D}(v_w, c)$ between v_w and c, given some appropriate measure of distance $\mathfrak{D}$. There are several results dating back to de Finetti 1974 showing that for certain $\mathfrak{D}$, a credence function is extendable to a finitely additive probability function on an algebra containing $\mathcal{F}$ if and only if there is no credence function c' that *accuracy dominates* c, in the sense that $\mathfrak{D}(v_w, c) > \mathfrak{D}(v_w, c')$ for all $w \in W$. Philosophers starting with Joyce (1998) have cited these dominance results as providing an accuracy-based justification for probabilism.

However, there is a limitation to almost all of the literature on accuracy arguments for probabilism: the opinion set is assumed to be finite.[1] Indeed, de Finetti (1974), Lindley (1987), Joyce (1998, 2009), Predd et al. (2009), Leitgeb and Pettigrew (2010a,b), and Pettigrew (2016) all establish their dominance results only for finite opinion sets. In this chapter, we remove this assumption and prove dominance results that we hope to be useful in evaluating the extent to which accuracy arguments for probabilism succeed when the opinion set is infinite.

We begin in Section 1.1 by reviewing the mathematical framework and the standard dominance result for finite opinion sets. Sections 1.2-1.4 are concerned with accuracy and coherence in the infinite setting. In Sections 1.2-1.3, we explore necessary and sufficient conditions in terms of coherence for avoiding dominance on multiple domains of countable opinion sets, establishing the finite dominance result on two of them. Finally, in Section 1.4, we extend the accuracy framework to the uncountable setting and prove that coherence is necessary to avoid dominance on uncountable opinion sets.

1.1 The Finite Case

Let us first set up the framework that will be used throughout the chapter. Fix a set W (not necessarily finite) which represents the set of *possible worlds* and, for now, a finite set $\mathcal{F} \subseteq \mathcal{P}(W)$[2] that represents the set of *propositions* an agent has beliefs about.

Definition 1.1.1. An *algebra* over W is a subset $\mathcal{F}^* \subseteq \mathcal{P}(W)$ such that:

[1] Walsh (2019) gives an accuracy dominance argument in the countably infinite context, to which we return in Section 1.4. In a related but distinct area, Huttegger (2013) and Easwaran (2013) extend to the infinite setting part of the literature on using minimization of expected inaccuracy to vindicate epistemic principles. See, e.g., Greaves and Wallace (2005). Finally, Schervish et al. (2014) prove that in certain countably infinite cases, coherence is sufficient for avoiding *strong dominance*. We return to their result in Section 1.3.

[2] $\mathcal{P}(W)$ denotes the power set of W.

1. $W \in \mathcal{F}^*$;
2. if $p, p' \in \mathcal{F}^*$, then $p \cup p' \in \mathcal{F}^*$;
3. if $p \in \mathcal{F}^*$, then $W \setminus p \in \mathcal{F}^*$.

Definition 1.1.2.

i. A *credence function* on $\mathcal{F}$ is a function from $\mathcal{F}$ to $[0, 1]$.

ii. A credence function c is *coherent* if it can be extended to a finitely additive probability function on an algebra $\mathcal{F}^*$ over W containing $\mathcal{F}$. This means there is an algebra $\mathcal{F}^* \supseteq \mathcal{F}$ over W and a function $c^* : \mathcal{F}^* \to [0, 1]$ such that:

   a) $c^*(p) = c(p)$ for all $p \in \mathcal{F}$;
   b) $c^*(p \cup p') = c^*(p) + c^*(p')$ for $p, p' \in \mathcal{F}^*$ with $p \cap p' = \varnothing$;
   c) $c^*(W) = 1$.

iii. Otherwise, a credence function is *incoherent*.

Remark 1.1.3. If $\mathcal{F} = \{p_1, \ldots, p_n\}$, we identify a credence function over $\mathcal{F}$ with the vector $(c(p_1), \ldots, c(p_n)) \in [0, 1]^n$. Thus the space of all credence functions over $\mathcal{F}$ can be identified with $[0, 1]^n \subseteq \mathbb{R}^n$.

We now introduce an important subclass of the class of all credence functions, namely the credence functions that match the truth values of $\mathcal{F}$ at a world w exactly.

Definition 1.1.4. Let $|\mathcal{F}| = n$. For each $p_i \in \mathcal{F}$, let $C_{p_i} : W \to \{0, 1\}$ be defined by $C_{p_i}(w) = 1$ if and only if $w \in p_i$. Then we call $v_w = (C_{p_1}(w), \ldots, C_{p_n}(w))$ the *omniscient credence function at world* w. We let $\mathcal{V}_{\mathcal{F}}$ denote the set of all omniscient credence functions. Note that $|\mathcal{V}_{\mathcal{F}}| \leq 2^n$.

Next, we specify the inaccuracy measures we will be concerned with in this section. Fix an opinion set $\mathcal{F}$, and let $\mathcal{C}$ denote the set of credence functions on $\mathcal{F}$. We define an *inaccuracy measure* to be a function of the form

$$\mathscr{I} : \mathcal{C} \times W \to [0, \infty].$$

The class of inaccuracy measures we consider is characterized by Pettigrew (2016) and assumed in Predd et al. 2009, namely the inaccuracy measures defined in terms of an *additive Bregman divergence* (or, equivalently, an additive and continuous *strictly proper scoring rule*; see Pettigrew 2016, p. 66).

Definition 1.1.5. Suppose $\mathfrak{D} : [0, 1]^n \times [0, 1]^n \to [0, \infty]$. Then

1. $\mathfrak{D}$ is a *divergence* if $\mathfrak{D}(\mathbf{x}, \mathbf{y}) \geq 0$ for all $\mathbf{x}, \mathbf{y} \in [0, 1]^n$ with equality if and only if $\mathbf{x} = \mathbf{y}$.

2. $\mathfrak{D}$ is *additive* if there exists a function $\mathfrak{d} : [0,1]^2 \to [0,\infty]$ such that

$$\mathfrak{D}(\mathbf{x},\mathbf{y}) = \sum_{i=1}^{n} \mathfrak{d}(x_i, y_i),$$

in which case we say $\mathfrak{D}$ is *generated* by $\mathfrak{d}$.

3. $\mathfrak{D}$ is an *additive Bregman divergence* if $\mathfrak{D}$ is an additive divergence generated by $\mathfrak{d}$ and in addition there is a function $\varphi : [0,1] \to \mathbb{R}$ such that:

   a) φ is continuous, bounded, and strictly convex on $[0,1]$;

   b) φ is continuously differentiable on $(0,1)$ with the formal definition

   $$\varphi'(i) := \lim_{x\to i} \varphi'(x)$$

   for $i \in \{0,1\}$;[3]

   c) for all $x, y \in [0,1]$, we have

   $$\mathfrak{d}(x,y) = \varphi(x) - \varphi(y) - \varphi'(y)(x-y).$$

   We call such a $\mathfrak{d}$ a *one-dimensional Bregman divergence.*

See, e.g., Banerjee et al. 2005 and Gneiting and Raftery 2007 for more details on Bregman divergences as well as their connection to strictly proper scoring rules.

In line with the characterization theorem proved by Pettigrew (2016, p. 84), we introduce the following definition.

Definition 1.1.6. Let a *legitimate inaccuracy measure* be an inaccuracy measure given by

$$\mathscr{I}(c,w) = \mathfrak{D}(v_w, c)$$

where $\mathfrak{D}$ is an additive Bregman divergence.

In other words, the inaccuracy of a credence function c at a world w is the distance between c and the omniscient credence function at w, where distance is measured with an additive Bregman divergence. A popular example of a legitimate inaccuracy measure is the Brier score (see Section 12, "Homage to the Brier Score," of Joyce 2009):

$$\mathscr{I}(c,w) = \sum_{i=1}^{n} (v_w(p_i) - c(p_i))^2.$$

We now establish the dominance result connecting coherence to accuracy dominance when the opinion set is finite. It was first proved for the Brier score by de Finetti (1974, p. 87-90) and extended to any legitimate inaccuracy measure by Predd et al. (2009).

[3] We do not require $\varphi'(i) < \infty$ for $i \in \{0,1\}$.

Definition 1.1.7. For each pair of credence functions c, c^* over $\mathcal{F}$;

1. c^* *weakly dominates* c relative to an inaccuracy measure $\mathscr{I}$ if

$$\begin{cases} \mathscr{I}(c,w) \geq \mathscr{I}(c^*,w) \text{ for all } w \in W \\ \mathscr{I}(c,w) > \mathscr{I}(c^*,w) \text{ for some } w \in W; \end{cases}$$

2. c^* *strongly dominates* c relative to $\mathscr{I}$ if $\mathscr{I}(c,w) > \mathscr{I}(c^*,w)$ for all $w \in W$.

Theorem 1.1.8 (de Finetti 1974, Predd et al. 2009)**.** Let $\mathcal{F}$ be a finite opinion set, $\mathscr{I}$ a legitimate inaccuracy measure, and c a credence function on $\mathcal{F}$. Then the following are equivalent:

1. c is not strongly dominated;

2. c is not weakly dominated;

3. c is coherent.

On the basis of Theorem 1.1.8, it is concluded in the accuracy literature that incoherent credences are criticizable because there is a set of credences that do strictly better in terms of accuracy, no matter how the world turns out to be, whereas coherent credences are not criticizable in this way. Since it is the basis of the accuracy argument for probabilism in the finite case, Theorem 1.1.8 is the result we would like to extend to infinite opinion sets. We now make progress in this direction when $\mathcal{F}$ is countably infinite.

1.2 The Countable Case: Coherence is Necessary

We begin with a discussion of how to measure inaccuracy in the countably infinite setting. Fix a countably infinite opinion set $\mathcal{F}$ over a set W of worlds (of arbitrary cardinality). Let $\mathcal{C}$ be the set of credence functions over $\mathcal{F}$ which can be identified with $[0,1]^\infty$ (see Remark 1.1.3). An *inaccuracy measure* remains a map from $\mathcal{C} \times W$ into $[0, \infty]$.

Analogous to the finite case, the class of inaccuracy measures that we use are defined in terms of generalizations of additive Bregman divergences.

Definition 1.2.1. Suppose $\mathfrak{D} : [0,1]^\infty \times [0,1]^\infty \to [0,\infty]$. Then we call $\mathfrak{D}$ a *generalized additive Bregman divergence* if

$$\mathfrak{D}(\mathbf{x},\mathbf{y}) = \sum_{i=1}^{\infty} \mathfrak{d}(x_i, y_i)$$

where $\mathfrak{d}$ is a bounded one-dimensional Bregman divergence as in Definition 1.1.5.3.

Later in this section we will show that generalized additive Bregman divergences are examples of what Csiszár (1995) calls *Bregman distances*, which are generalizations of additive Bregman divergences defined on spaces of non-negative functions.

Suggestively, we make the following definition.

Definition 1.2.2. Given an enumeration of $\mathcal{F}$,[4] let a *generalized legitimate inaccuracy measure* be an inaccuracy measure $\mathscr{I} : \mathcal{C} \times W \to [0, \infty]$ given by

$$\mathscr{I}(c, w) = \mathfrak{D}(v_w, c) = \sum_{i=1}^{\infty} \mathfrak{d}(v_w(p_i), c(p_i)) \tag{1.1}$$

for $\mathfrak{D}$ a generalized additive Bregman divergence.

Notice that the Brier score extends to a generalized legitimate inaccuracy measure, namely the squared $\ell^2(\mathcal{F})$ norm

$$\mathscr{I}(c, w) = ||v_w - c||^2_{\ell^2(\mathcal{F})} = \sum_{i=1}^{\infty} (v_w(p_i) - c(p_i))^2. \tag{1.2}$$

We call (1.2) the *generalized Brier score.*

The name "generalized legitimate inaccuracy measure" is motivated by the observation that a generalized legitimate inaccuracy measure naturally restricted to the finite opinion sets is a legitimate inaccuracy measure. This is because 1) for both the generalized and finite legitimate inaccuracy measures, the score of an individual proposition is defined by a one-dimensional Bregman divergence, and 2) for both the generalized and finite legitimate inaccuracy measures, the scores of individual propositions are combined additively to give a score for the entire credence function. To use the terminology of Leitgeb and Pettigrew (2010a), in the finite and countably infinite setting, the local scores are the same and the global scores relate to the local scores in the same way. These observations support the view that, insofar as additive Bregman divergences are the appropriate functions to use for measuring inaccuracy in the finite setting, generalized additive Bregman divergences are the appropriate functions to use for measuring inaccuracy in the countably infinite setting.

We now prove that coherence is necessary to avoid accuracy dominance in the countably infinite case.

Theorem 1.2.3. Let $\mathcal{F}$ be a countably infinite opinion set, $\mathscr{I}$ a generalized legitimate inaccuracy measure and c an incoherent credence function. Then

1. c is weakly dominated relative to $\mathscr{I}$ by a coherent credence function; and
2. if $\mathscr{I}(c, w) < \infty$ for each $w \in W$, then c is strongly dominated relative to $\mathscr{I}$ by a coherent credence function.

We review the necessary background before proving Theorem 1.2.3.

[4]The choice of enumeration does not matter since the terms in the infinite sum are non-negative. Thus convergence is absolute and independent of order.

Generalized Projections

Csiszár (1995) showed that what he calls *generalized projections* onto convex sets with respect to Bregman distances exist under very general conditions. We review his relevant results here (but assume knowledge of basic measure theory).

Definition 1.2.4. Fix a σ-finite measure space $(X, \mathcal{X}, \mu)$. The *Bregman distance* of non-negative ($\mathcal{X}$-measurable) functions s and t is defined by

$$B_{\varphi,\mu}(s,t) = \int \mathfrak{d}(s(x), t(x))\mu(dx) \in [0, \infty]$$

where $\mathfrak{d}(s(x), t(x)) = \varphi(s(x)) - \varphi(t(x)) - \varphi'(t(x))(s(x) - t(x))$ for some strictly convex, differentiable function φ on $(0, \infty)$.[5] Note that $B_{\varphi,\mu}(s,t) = 0$ iff $s = t$ μ-a.e. See Csiszár (1995, p. 165) for details.

Remark 1.2.5. Notice that a generalized additive Bregman divergence $\mathfrak{D}_\varphi$ (whose generating one-dimensional Bregman divergence $\mathfrak{d}$ is given in terms of φ) has a corresponding Bregman distance $B_{\bar{\varphi},\mu}$ with

1. the measure space being $(\mathbb{N}, \mathcal{P}(\mathbb{N}), \mu)$, where μ is the counting measure on $\mathbb{N}$, and

2. $\bar{\varphi}$ on $(0, \infty)$ being a strictly convex, differentiable extension of φ on $[0, 1]$.[6]

Thus non-negative ($\mathscr{P}(\mathbb{N})$-measurable) functions are elements of $\mathbb{R}^{+\infty}$. Note, importantly, that the corresponding generalized legitimate inaccuracy measure $\mathscr{I}(c, w) = \mathfrak{D}_\varphi(v_w, c)$ is also given by the corresponding Bregman distance. That is,

$$\mathscr{I}(c, w) = B_{\bar{\varphi},\mu}(v_w, c)$$

because $\mathfrak{D}_\varphi = B_{\bar{\varphi},\mu}|_{[0,1]^\infty \times [0,1]^\infty}$.

To simplify notation, let B denote $B_{\varphi,\mu}$ a Bregman distance. Let S be the set of non-negative measurable functions. For any $E \subseteq S$ and $t \in S$, we write

$$B(E, t) = \inf_{s \in E} B(s, t).$$

If there exists $s^* \in E$ with $B(s^*, t) = B(E, t)$, then s^* is unique and is called the *B-projection of t onto E*. As Csiszár (1995) notes, these projections may not exist. However a weaker kind of projection always exists. To describe them, we need to introduce a notion of convergence called *loose in μ-measure convergence.*

---

[5] We do not need to assume $\varphi(0) = \varphi'(0) = 1$ by the remark following (1.9) in Csiszár 1995.

[6] Using that φ' exists and is finite at $x = 1$ as we assumed $\mathfrak{d}$ is bounded, we extend φ as follows: for $x \in [1, \infty)$, let $\bar{\varphi}(x) = q(x) = x^2 + bx + c$, where b and c are chosen so $\varphi(1) = q(1)$ and $\varphi'(1) = q'(1)$. Then using the fact that $\bar{\varphi}$ is differentiable at 1 by construction and a function is strictly convex if and only if its derivative is strictly increasing, it is easy to see that $\bar{\varphi}$ is differentiable and strictly convex on $(0, \infty)$.

Definition 1.2.6. We say a sequence $\{s_n\} \subseteq S$ converges *loosely in μ-measure* to t, denoted by $s_n \rightsquigarrow_\mu t$, if for every $A \in \mathcal{X}$ with $\mu(A) < \infty$, we have

$$\lim_{n\to\infty} \mu(A \cap \{|s_n - t| > \epsilon\}) = 0 \text{ for all } \epsilon > 0.$$

Definition 1.2.7.

i. Given $E \subseteq S$ and $t \in S$, we say that a sequence $\{s_n\} \subseteq E$ is a *B-minimizing* sequence if $B(s_n, t) \to B(E, t)$.

ii. If there is an $s^* \in S$ such that every B-minimizing sequence converges to s^* loosely in μ-measure, then we call s^* the *generalized B-projection of t onto E*.

The result that is integral to proving Theorem 1.2.3 is the following.

Theorem 1.2.8 (Csiszár 1995)**.** Let E be a convex subset of S and $t \in S$. If $B(E, t)$ is finite, then there exists $s^* \in S$ such that

$$B(s, t) \geq B(E, t) + B(s, s^*) \text{ for every } s \in E$$

and $B(E, t) \geq B(s^*, t)$. It follows that the generalized B-projection of t onto E exists and equals s^*.

Extending Partial Measures

We also use an extension result of Horn and Tarski (1947) in the proof of Theorem 1.2.3. Following Horn and Tarski, we introduce *partial measures* and recall that they can be extended to finitely additive probability functions. Recall the definition of a finitely additive probability function described in Definition 1.1.2 (we drop the assumption that $\mathcal{F}$ is finite), which we will abbreviate to *FA probability function*.

Remark 1.2.9. It is a simple corollary of the definition of an FA probability function c over an algebra $\mathcal{F}$ that for any $p, p' \in \mathcal{F}$: if $p \subseteq p'$, then $c(p) \leq c(p')$.

Here is another useful fact about FA probability functions.

Proposition 1.2.10. If c is an FA probability function on $\mathcal{F}$ and $a_0, \ldots, a_{m-1} \in \mathcal{F}$, then

$$\sum_{k=0}^{m-1} c(a_k) = \sum_{k=0}^{m-1} c\Big(\bigcup_{p \in S^{m,k}} \bigcap_{i \leq k} a_{p_i}\Big) \tag{1.3}$$

where $S^{m,k}$ is the set of all sequences $p = (p_0, \ldots, p_k)$ with $0 \leq p_0 < \ldots < p_k < m$.

To introduce the notion of a partial measure, we need the following definition.

Definition 1.2.11. Let $\varphi_0, \ldots, \varphi_{m-1}$ and $\psi_0, \ldots, \psi_{n-1}$ be elements of $\mathcal{F}$. Then we write

$$(\varphi_0, \ldots, \varphi_{m-1}) \subseteq (\psi_0, \ldots, \psi_{n-1})$$

to mean

$$\bigcup_{p \in S^{m,k}} \bigcap_{i \leq k} \varphi_{p_i} \subseteq \bigcup_{p \in S^{n,k}} \bigcap_{i \leq k} \psi_{p_i} \text{ for every } k < m \tag{1.4}$$

where $S^{m,k}$ is as in Proposition 1.2.10.[7]

Definition 1.2.12. A function c, defined on a subset S of an algebra $\mathcal{F}$ over W, that maps to $\mathbb{R}$ is called a *partial measure* if it satisfies the following properties:

1. $c(x) \geq 0$;
2. If $\varphi_0, \ldots, \varphi_{m-1}, \psi_0, \ldots, \psi_{n-1} \in S$ and

   $$(\varphi_0, \ldots, \varphi_{m-1}) \subseteq (\psi_0, \ldots, \psi_{n-1}),$$

   then

   $$\sum_{k=0}^{m-1} c(\varphi_k) \leq \sum_{k=0}^{n-1} c(\psi_k);$$

3. $W \in S$ and $c(W) = 1$.

The following result is the point of introducing the above definitions.

Theorem 1.2.13 (Horn and Tarski 1947)**.** Let c be a partial measure on a subset S of an algebra $\mathcal{F}$. Then there is a FA probability function c^* on $\mathcal{F}$ that extends c.

Proof of Theorem 1.2.3

We now establish the necessity of coherence to avoid dominance.

Proof of Theorem 1.2.3. Let $\mathscr{I}$ be a generalized legitimate inaccuracy measure and thus be defined by a Bregman distance $B_{\varphi,\mu}$ (see Remark 1.2.5). We write B for $B_{\varphi,\mu}$. Let S be the set of non-negative functions on $\mathcal{F}$. Let $E \subseteq S$ be the set of coherent credence functions on $\mathcal{F}$. Then clearly E is convex.

Let c be an incoherent credence function.

Case 1: $\mathscr{I}(c, w) = \infty$ for all $w \in W$. Then since $\mathscr{I}(v_w, w) = 0$ for all $w \in W$, any omniscient credence function weakly dominates c.

Case 2: $\mathscr{I}(c, w') < \infty$ for some $w' \in W$. We show that there is a coherent credence function c' such that

$$\mathscr{I}(c, w) > \mathscr{I}(c', w) \text{ for any } w \text{ such that } \mathscr{I}(c, w) < \infty.$$

[7] Note that if $m - 1 > n - 1$, this condition implies $\bigcup_{p \in S^{m,k}} \bigcap_{i \leq k} \varphi_{p_i} = \bigcup_{p \in S^{n,k}} \bigcap_{i \leq k} \psi_{p_i} = \varnothing$ for $k \geq n$.

Since $v_{w'} \in E$, we see that

$$B(E, c) \leq B(v_{w'}, c) = \mathscr{I}(c, w') < \infty.$$

Thus we can apply Theorem 1.2.8 to get a $\pi_c \in S$ such that

$$B(s, t) \geq B(E, c) + B(s, \pi_c) \text{ for every } s \in E. \tag{1.5}$$

In particular, (1.5) holds when s is the omniscient credence function at world w for any $w \in W$; and so we see that

$$\mathscr{I}(c, w) \geq B(E, c) + \mathscr{I}(\pi_c, w) \tag{1.6}$$

for all w, where all numbers in (1.6) are finite whenever $\mathscr{I}(c, w) < \infty$.

Next we show that π_c is in fact coherent. This is due to the following claim: E is closed under loose convergence in μ-measure where μ is the counting measure on $\mathcal{P}(\mathcal{F})$. To see this, let $c_n \in E$ for each n and $c \in S$. Assume $c_n \to c$ loosely in μ-measure. We show $c \in E$, i.e., c is coherent. Note c is coherent on $\mathcal{F}$ if and only if $c' : \mathcal{F} \cup \{W\} \to [0, 1]$ is coherent on $\mathcal{F} \cup \{W\}$, where $c' = c$ on $\mathcal{F}$ and $c'(W) = 1$. Thus it suffices to assume c and c_n for all n are defined on $\mathcal{F} \cup \{W\}$ with $c(W) = c_n(W) = 1$ for all n.

It is easy to see that loose convergence in the counting measure implies pointwise convergence on $\mathcal{F}$, and so

$$c(p) = \lim_{n \to \infty} c_n(p) \in [0, 1]$$

for each $p \in \mathcal{F} \cup \{W\}$. To show $c \in E$, it suffices to show c can be extended to an FA probability function on $\mathcal{P}(W)$.

We first show c is a partial measure on $\mathcal{F} \cup \{W\}$. Definitions 1.2.12.1 and 1.2.12.3 clearly hold for c so we just need to show Definition 1.2.12.2 holds. Let $\varphi_0, \ldots, \varphi_{m-1}, \psi_0, \ldots, \psi_{m'-1} \in \mathcal{F} \cup \{W\}$ and

$$\bigcup_{p \in S^{m,k}} \bigcap_{i \leq k} \varphi_{p_i} \subseteq \bigcup_{p \in S^{m',k}} \bigcap_{i \leq k} \psi_{p_i}$$

for every $k < m$. Since c_n are coherent and thus extend to FA probability functions, we have by Proposition 1.2.10 and Remark 1.2.9 that

$$\sum_{k=0}^{m-1} c_n(\varphi_k) = \sum_{k=0}^{m-1} c_n(\bigcup_{p \in S^{m,k}} \bigcap_{i \leq k} \varphi_{p_i}) \leq \sum_{k=0}^{m'-1} c_n(\bigcup_{p \in S^{m',k}} \bigcap_{i \leq k} \psi_{p_i}) = \sum_{k=0}^{m'-1} c_n(\varphi_k).$$

using that

$$\bigcup_{p \in S^{m,k}} \bigcap_{i \leq k} \psi_{p_i} = \bigcup_{p \in S^{m',k}} \bigcap_{i \leq k} \psi_{p_i} = \varnothing$$

for $k \geq m'$. Sending n to infinity and using the pointwise convergence of c_n to c on $\mathcal{F} \cup \{W\}$ we obtain that

$$\sum_{k=0}^{m-1} c(\varphi_k) \leq \sum_{k=0}^{m'-1} c(\psi_k).$$

Thus c is a partial measure on $\mathcal{F} \cup \{W\}$. By Theorem 1.2.13, it follows that there is an FA probability function c^* on $\mathcal{P}(W)$ that extends c and so $c \in E$, which concludes the proof that E is closed under loose μ-convergence.

By Theorem 1.2.8, π_c is the generalized B-projection of c onto E. Also, since

$$B(E, c) = \inf_{s \in E}(s, c) < \infty,$$

there is a sequence $\{s_n\} \subseteq E$ such that $B(s_n, c) \to B(E, c)$ by the definition of infimum. By the definition of a generalized projection, $s_n \rightsquigarrow_\mu \pi_c$. Since E is closed under loose convergence, it follows that $\pi_c \in E$. Further, we see

$$B(E, c) \geq B(\pi_c, c) > 0,$$

since $\pi_c \neq c$ (as c is incoherent) and $B(s, t) = 0 \iff s = t$ (as μ is the counting measure). So for every w such that $\mathscr{I}(c, w) < \infty$, we deduce that

$$\mathscr{I}(c, w) \geq B(E, c) + \mathscr{I}(\pi_c, w) > \mathscr{I}(\pi_c, w).$$

This proves that c is weakly dominated by π_c, and is strongly dominated by π_c if $\mathscr{I}(c, w) < \infty$ for all $w \in W$. □

1.3 The Countable Case: The Sufficiency of Coherence

Unlike in the finite case, coherent credence functions on countably infinite opinion sets can be strongly dominated.

Example 1.3.1. Let $\mathcal{F} = \{\{n \geq N : n \in \mathbb{N}\} : N \in \mathbb{N}\}$ be an opinion set over $\mathbb{N}$ (including zero). Let

$$c(\{n \geq N\}) = \frac{1}{\sqrt{N+1}}.$$

Then c is coherent, in fact countably coherent (see Definition 1.3.5), but $\mathscr{I}(c, w) = \infty$ for all $w \in W$ when $\mathscr{I}$ is the generalized Brier score. So any omniscient credence function strongly dominates c.

To deal with this problem, we restrict to certain classes of opinion sets and establish sufficient conditions—in terms of coherence and finite inaccuracy assumptions—to avoid accuracy dominance. At points, our results will only apply to the generalized Brier score. We conjecture that any such result extends to any generalized legitimate inaccuracy measure. In any case, this is the best possible restriction since the Brier score has been defended by many—including Horwich (1982), Maher (2002), Joyce (2009), and Leitgeb and Pettigrew (2010a)—as being a particularly appropriate way to measure accuracy.

Throughout this section we assume $\mathcal{F}$ is countably infinite.

Countably Discriminating Opinion Sets

We begin by proving a sufficient condition for avoiding dominance on *countably discriminating* opinion sets.

Definition 1.3.2. For $\mathcal{F} \subseteq \mathcal{P}(W)$, we define an equivalence relation $\sim$ on W such that $w \sim w'$ if and only if $\{p \in \mathcal{F} : w \in p\} = \{p \in \mathcal{F} : w' \in p\}$. We call the set of equivalence classes of W the *quotient of W relative to $\mathcal{F}$*. If the quotient of W relative to $\mathcal{F}$ is countable, then we call $\mathcal{F}$ *countably discriminating.*

The following characterization of the coherent credence functions on finite opinion sets is due to de Finetti (1974). Recall $\mathcal{V}_{\mathcal{F}}$ denotes the set of omniscient credence functions on $\mathcal{F}$, which is finite when $\mathcal{F}$ is finite.

Theorem 1.3.3 (de Finetti 1974)**.** c is a coherent credence function on a finite opinion set $\mathcal{F}$ if and only if there are $\lambda_w \in [0,1]$ with $\sum_{v_w \in \mathcal{V}_{\mathcal{F}}} \lambda_w = 1$ such that

$$c(p) = \sum_{v_w \in \mathcal{V}_{\mathcal{F}}} \lambda_w v_w(p)$$

for all $p \in \mathcal{F}$.

Theorem 1.3.3 is integral in the proof that coherence is sufficient to avoid dominance in Theorem 1.1.8. The key insight to extending this direction of Theorem 1.1.8 is that de Finetti's (1974) characterization of the coherent credence functions on finite opinion sets extends to *countably coherent* credence functions on countably discriminating opinion sets.

Definition 1.3.4. A *σ-algebra* over W is a subset $\mathcal{F}^* \subseteq \mathcal{P}(W)$ such that:

1. $W \in \mathcal{F}^*$;
2. if $\{p_i\}_{i=1}^{\infty} \subseteq \mathcal{F}^*$, then $\bigcup_{i=1}^{\infty} p_i \in \mathcal{F}^*$;
3. if $p \in \mathcal{F}^*$, then $W \setminus p \in \mathcal{F}^*$.

Definition 1.3.5. Let a credence function c be *countably coherent* if c extends to a countably additive probability function on a σ-algebra containing $\mathcal{F}$. That is, there is a σ-algebra $\mathcal{F}^* \supseteq \mathcal{F}$ and $c^* : \mathcal{F}^* \to [0,1]$ such that:

1. $c^*(p) = c(p)$ for all $p \in \mathcal{F}$;

2. $c^*(\bigcup_{i=1}^{\infty} p_i) = \sum_{i=1}^{\infty} c^*(p_i)$ for $\{p_i\}_{i=1}^{\infty} \in \mathcal{F}^*$ with $p_i \cap p_j = \varnothing$ for $i \neq j$;
3. $c^*(W) = 1$.

Otherwise, a credence function is *countably incoherent*.

We now characterize the countably coherent credence functions on countably discriminating opinion sets. The proof is almost identical to Predd et al.'s (2009) proof of Theorem 1.3.3.

Proposition 1.3.6. Let $\mathcal{F}$ be a countably discriminating opinion set. Then a credence function c is countably coherent if and only there are $\lambda_w \in [0, 1]$ with $\sum_{v_w \in \mathcal{V}_{\mathcal{F}}} \lambda_w = 1$ such that

$$c(p) = \sum_{v_w \in \mathcal{V}_{\mathcal{F}}} \lambda_w v_w(p)$$

for all $p \in \mathcal{F}$.

Proof. We adapt the proof of Proposition 1 in Predd et al. 2009. Let $\mathcal{F} = \{p_1, p_2 \ldots\}$. Let $\mathcal{X}$ be the collection of all nonempty sets of the form $\bigcap_{i=1}^{\infty} p_i^*$ where p_i^* is either p_i or its complement. Then $\mathcal{X}$ partitions W. Also, $\mathcal{X}$ is in bijection with $\mathcal{V}_{\mathcal{F}}$, the set of omniscient credence functions.

Indeed, let f map v_w to $\bigcap_{i=1}^{\infty} p_i^*$ where $p_i^* = p_i$ if $v_w(p_i) = 1$ and $p^* = p_i^c$ otherwise. Then for each w, $w \in f(v_w)$ and so $f(v_w) \in \mathcal{X}$. Note f is onto. Indeed, let $w \in \bigcap_{i=1}^{\infty} p_i^*$, where $\bigcap_{i=1}^{\infty} p_i^* \in \mathcal{X}$. Then $f(v_w) = \bigcap_{i=1}^{\infty} p_i^*$. Also, f is injective. Indeed, assume $f(v_w) = f(v_{w'})$. Then

$$f(v_w) = \bigcap_{i=1}^{\infty} p_i^1 = \bigcap_{i=1}^{\infty} p_i^2 = f(v_{w'})$$

for $p_i^j = p_i$ or its complement for all i and $j \in \{1, 2\}$. If $p_i^1 \neq p_i^2$ for some i, then without loss of generality we may assume $p_i^1 = p_i$ and $p_i^2 = p_i^c$. So $w \in p_i^1$ but $w \notin p_i^2$ and thus $w \notin \bigcap_{i=1}^{\infty} p_i^2$. But $w \in \bigcap_{i=1}^{\infty} p_i^1$ by definition of f and so $\bigcap_{i=1}^{\infty} p_i^1 \neq \bigcap_{i=1}^{\infty} p_i^2$, which is a contradiction. It follows that $p_i^1 = p_i^2$ for all i, but then by definition of f, this implies $v_w(p_i) = 1$ if and only if $v_{w'}(p_i) = 1$ for all i and so $v_w = v_{w'}$.

It is easy to see that since $\mathcal{F}$ is countably discriminating, $\mathcal{V}_{\mathcal{F}}$ is countable. It follows that $\mathcal{X}$ is countable. Enumerate the elements of $\mathcal{V}_{\mathcal{F}}$ and $\mathcal{X}$ by $v_{w_1}, v_{w_2}, \ldots$ and $e_1, e_2, \ldots$, respectively, such that $f^{-1}(e_j) = v_{w_j}$. We have that p_i is the disjoint union of e_j such that $e_j \subseteq p_i$, or equivalently the e_j where $f^{-1}(e_j)(p_i) = 1$. Note i) for any countably additive probability function μ on a σ-algebra containing $\mathcal{F}$ (and thus containing $\mathcal{X}$) and any $p_i \in \mathcal{F}$:

$$\mu(p_i) = \sum_{j=1}^{\infty} \mu(e_j) f^{-1}(e_j)(p_i).$$

Now we prove the equivalence. Assume c is countably coherent. So c extends to a countably additive probability function μ on a σ-algebra containing $\mathcal{F}$. Then by i),

$$c(p_i) = \mu(p_i) = \sum_{j=1}^{\infty} \mu(e_j) f^{-1}(e_j)(p_i)$$

for all $p_i \in \mathcal{F}$. But since $\mu(e_j)$ are non-negative and sum to 1 (since the e_j's partition W and μ is a countably additive probability function), we have that c has the form stated.

Now assume $c(p_i) = \sum_{j=1}^{\infty} \lambda_j v_{w_j}(p_i)$ for all i where $\sum_j \lambda_j = 1$. Let $\mathcal{A}$ be the smallest σ-algebra on W containing $\mathcal{F}$. Then it is easy to check that the function on $\mathcal{A}$ defined by $\bar{v}_{w_j}(p) = 1$ if and only if $w \in p$ extends v_{w_j} and is a countably additive probability function on $\mathcal{A}$. Then $\sum_{j=1}^{\infty} \lambda_j \bar{v}_{w_j}$ is a countably additive probability function on $\mathcal{A}$ since a countable sum of countably additive probability functions with coefficients that sum to 1 is a countably additive probability function. Since

$$c(p_i) = \sum_{j=1}^{\infty} \lambda_j v_{w_j}(p_i) = \sum_{j=1}^{\infty} \lambda_i \bar{v}_{w_j}(p_i)$$

for all i, it follows that c extends to a countably additive probability function on a σ-algebra containing $\mathcal{F}$. □

Recall that Example 1.3.1 shows that there are coherent credence functions on countably discriminating opinion sets that are strongly dominated without further assumptions. The following definition is used in the finite inaccuracy assumption we make in the main theorem of this section.

Definition 1.3.7. For c a countably coherent credence function on a countably discriminating opinion set $\mathcal{F}$, let $c = \sum_{v_w \in \mathcal{V}_\mathcal{F}} \lambda_w v_w$ on $\mathcal{F}$. Then for a countably additive probability function $\bar{c}$ extending c, we have

$$\mathbb{E}_{\bar{c}} \mathscr{I}(c, \cdot) = \sum_{v_w \in \mathcal{V}_\mathcal{F}} \lambda_w \mathscr{I}(c, w).$$

In light of this, we call the sum $\sum_{v_w \in \mathcal{V}_\mathcal{F}} \lambda_w \mathscr{I}(c, w)$ the *expected inaccuracy of c relative to $\mathscr{I}$*.[8]

We can now prove a sufficient condition for avoiding dominance on countably discriminating opinion sets.

Theorem 1.3.8. Let $\mathcal{F}$ be a countably discriminating opinion set and $\mathscr{I}$ a generalized legitimate inaccuracy measure. If c is a countably coherent credence function with finite expected inaccuracy relative to $\mathscr{I}$, then c is not weakly dominated relative to $\mathscr{I}$ by any credence function $d \neq c$.

[8]Note we take the convention that $0 \cdot \infty = 0$ as is typically done in defining integration of extended real-valued functions.

Proof. We adapt the proof of the second part of Theorem 1 in Predd et al. 2009. Assume d weakly dominates c. Enumerate $\mathcal{F}$ and the omniscient credence functions. Let $\mathscr{I}(c,w) = \mathfrak{D}(v_w, c)$ for a generalized additive Bregman divergence $\mathfrak{D}$ and $c = \sum_{j=1}^{\infty} \lambda_j v_{w_j}$ by Proposition 1.3.6. Then $\mathfrak{D}(v_w, c) - \mathfrak{D}(v_w, d) \geq 0$ for all w by weak dominance.

Now since

$$\mathfrak{d}(x, c(p_i)) = \varphi(x) - \varphi(c(p_i)) - \varphi'(c(p_i))(x - c(p_i))$$

for any $x \in [0,1]$, we have

$$\begin{aligned}\mathfrak{d}(x, c(p_i)) - \mathfrak{d}(x, d(p_i)) &= \varphi(d(p_i)) - \varphi(c(p_i)) + \varphi'(c(p_i))c(p_i) - \varphi'(d(p_i))d(p_i) \\ &\quad + (\varphi'(d(p_i)) - \varphi'(c(p_i))x = C_i + (\varphi'(d(p_i)) - \varphi'(c(p_i))x\end{aligned} \tag{1.7}$$

for any $x \in [0,1]$, where C_i is a constant depending on i.

Further, we make the following observation that follows from the assumption of finite expected inaccuracy and weak dominance:

$$\infty > \lambda_j(\mathfrak{D}(v_{w_j}, c) - \mathfrak{D}(v_{w_j}, d)) = \lambda_j(\sum_{i=1}^{\infty} \mathfrak{d}(v_{w_j}, c(p_i)) - \mathfrak{d}(v_{w_j}, d(p_i))). \tag{1.8}$$

for each j. In particular, $\lambda_j \mathfrak{D}(v_{w_j}, c) < \infty$ and $\lambda_j \mathfrak{D}(v_{w_j}, d) < \infty$ for each j.

Note $\mathfrak{D}(\sum_{j=1}^{\infty} \lambda_j v_{w_j}, c) = 0$ and $\mathfrak{d}(\sum_{j=1}^{\infty} \lambda_j v_{w_j}(p_i), c(p_i)) = 0$ for all i, so

$$\mathfrak{D}(\sum_{j=1}^{\infty} \lambda_j v_{w_j}, c) - \mathfrak{D}(\sum_{j=1}^{\infty} \lambda_j v_{w_j}, d) = \sum_{i=1}^{\infty} \mathfrak{d}(\sum_{j=1}^{\infty} \lambda_j v_{w_j}(p_i), c(p_i)) - \mathfrak{d}(\sum_{j=1}^{\infty} \lambda_j v_{w_j}(p_i), d(p_i)).$$

Continuing, using (1.7) and that $\sum_{i=1}^{\infty} \lambda_i x = x$ for any x, we have

$$\begin{aligned}\mathfrak{D}(\sum_{j=1}^{\infty} \lambda_j v_{w_j}, c) - \mathfrak{D}(\sum_{j=1}^{\infty} \lambda_j v_{w_j}, d) &= \sum_{i=1}^{\infty} C_i + (\varphi'(d(p_i)) - \varphi'(c(p_i))(\sum_{j=1}^{\infty} \lambda_j v_{w_j}(p_i)) \\ &= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \lambda_j [C_i + (\varphi'(d(p_i)) - \varphi'(c(p_i))v_{w_j}(p_i)] \\ &= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \lambda_j [\mathfrak{d}(v_{w_j}(p_i), c(p_i)) - \mathfrak{d}(v_{w_j}(p_i), d(p_i))].\end{aligned} \tag{1.9}$$

We claim we can switch limits in (1.9). To do so we use the monotone convergence theorem and the generalized dominance convergence theorem (see, e.g., Fabian and Hannan 1985, p. 32). Let

$$g_n(i) = \sum_{j=1}^{n} \lambda_j [\mathfrak{d}(v_{w_j}(p_i), c(p_i)) - \mathfrak{d}(v_{w_j}(p_i), d(p_i))].$$

To apply the generalized dominance convergence theorem, we need to find $h_n(i) \geq 0$ such that:

1. for each n, $\sum_{i=1}^{\infty} h_n(i) < \infty$ and h_n converges pointwise for each i;
2. $|g_n(i)| \leq h_n(i)$ for each i and n;
3. $\sum_i \lim_n h_n(i) = \lim_n \sum_i h_n(i) < \infty$.

We set

$$h_n(i) = \max\{\sum_{j=1}^{n} \lambda_j \mathfrak{d}(v_{w_j}(p_i), c(p_i)), \sum_{j=1}^{n} \lambda_j \mathfrak{d}(v_{w_j}(p_i), d(p_i))\}.$$

Then

$$\sum_{i=1}^{\infty} h_n(i) \leq \sum_{j=1}^{n} \lambda_j (\mathfrak{D}(v_w, c) + \mathfrak{D}(v_w, d)) \leq 2 \sum_{j=1}^{\infty} \lambda_j \mathfrak{D}(v_{w_j}, c) < \infty$$

where we split up the inner sum using (1.8) and establish the last bound by assumption of finite expected inaccuracy and weak dominance. Since, for each i, $\sum_{j=1}^{n} \lambda_j \mathfrak{d}(v_{w_j}(p_i), c(p_i))$ and $\sum_{j=1}^{n} \lambda_j \mathfrak{d}(v_{w_j}(p_i), d(p_i))$ are bounded and increasing in n (since $\mathfrak{d}$ is bounded and $\sum_j \lambda_j = 1$), we establish three facts: i) $h_n(i)$ is bounded and increasing in n for each i so ii) $h_n(i)$ converges pointwise for each i and iii) $\sum_i \lim h_n(i) = \lim_n \sum_i h_n(i)$ by the monotone convergence theorem. In addition, $\sum_i \lim_n h_n(i) = \lim_n \sum_i h_n(i) < \infty$ since

$$\begin{aligned}
\lim_n \sum_i h_n(i) &\leq \lim_n \sum_i \sum_{j=1}^{n} \lambda_j \mathfrak{d}(v_{w_j}(p_i), c(p_i)) + \lambda_j \mathfrak{d}(v_{w_j}(p_i), d(p_i)) \\
&= \sum_{j=1}^{\infty} \lambda_j (\sum_{i=1}^{\infty} \mathfrak{d}(v_{w_j}(p_i), c(p_i)) + \mathfrak{d}(v_{w_j}(p_i), d(p_i))) \\
&= \sum_{j=1}^{\infty} \lambda_j (\mathfrak{D}(v_{w_j}, c) + \mathfrak{D}(v_{w_j}, d)) < \infty
\end{aligned}$$

where we split up the inner sum using (1.8) and establish the last bound using the assumptions of finite expected inaccuracy and weak dominance. Moreover, it is easy to check that $|g_n(i)| \leq h_n(i)$ for each n and i.

Putting everything together, by the generalized dominance convergence theorem, we can switch the sum in j and the sum in i in (1.9). Thus we get:

$$\begin{aligned}
0 \geq \mathfrak{D}(\sum_{j=1}^{\infty} \lambda_j v_{w_j}, c) - \mathfrak{D}(\sum_{j=1}^{\infty} \lambda_j v_{w_j}, d) &= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \lambda_j [\mathfrak{d}(v_{w_j}, c(p_i)) - \mathfrak{d}(v_{w_j}, d(p_i))] \\
&= \sum_{j=1}^{\infty} \lambda_j \sum_{i=1}^{\infty} \mathfrak{d}(v_{w_j}, c(p_i)) - \mathfrak{d}(v_{w_j}, d(p_i)) \\
&= \sum_{j=1}^{\infty} \lambda_j (\mathfrak{D}(v_{w_j}, c) - \mathfrak{D}(v_{w_j}, d)) \geq 0
\end{aligned}$$

where we split up the inner sum to get the last line by (1.8) and establish the last bound by assumption of weak dominance. So $0 = \mathfrak{D}(c,c) = \mathfrak{D}(c,d)$. We can therefore conclude that $c = d$ since $\mathfrak{D}(c,d) = 0$ if and only if $c = d$. $\square$

Point-Finite and Negation-Compact Opinion Sets

In this section, we prove Theorem 1.1.8 for a certain subclass of countably discriminating opinion sets. However, for this result, we will need to restrict to the generalized Brier score (see the beginning of Section 1.3) which we denote by $\mathscr{B}$. We prove sufficient conditions for avoiding dominance on various other classes of opinion sets along the way.

I

First, we show that we can weaken finite expected inaccuracy in Theorem 1.3.8 to *somewhere finitely inaccurate* on *point-finite* opinion sets.

Definition 1.3.9. Let c be a credence function on an opinion set $\mathcal{F}$. Then we say c is *somewhere finitely inaccurate* relative to $\mathscr{I}$ if $\mathscr{I}(c,w) < \infty$ for some $w \in W$.

Definition 1.3.10. We say $\mathcal{F} \subseteq \mathcal{P}(W)$ is *point-finite* if $|\{p \in \mathcal{F} : w \in p\}| < \infty$ for each $w \in W$.

Proposition 1.3.11. Let $\mathcal{F}$ be a point-finite opinion set and c a credence function on $\mathcal{F}$. If c is countably coherent and somewhere finitely inaccurate relative to $\mathscr{B}$, then c is not weakly dominated relative to $\mathscr{B}$.

Proof. The proof is similar to the proof of Theorem 1.3.8, but we use a different criterion to switch limits. Assume d weakly dominates c. Note i) c is somewhere finitely inaccurate if and only if $\mathscr{B}(c,w) < \infty$ for all $w \in W$ if and only if $\sum_{i=1}^{\infty} c(p_i)^2 < \infty$. It follows by weak dominance that $\mathscr{B}(d,w) < \infty$ for all $w \in W$ and therefore $\sum_{i=1}^{\infty} d(p_i)^2 < \infty$. To simplify notation, let $c_i := c(p_i)$ and $d_i := d(p_i)$. Let $\mathscr{B}(c,w) = \mathfrak{D}(v_w, c)$ for $\mathfrak{D}$ a generalized additive Bregman divergence.

By a similar derivation as in Theorem 1.3.8, we have that

$$\mathfrak{D}(\sum_{j=1}^{\infty} \lambda_j v_{w_j}, c) - \mathfrak{D}(\sum_{j=1}^{\infty} \lambda_j v_{w_j}, d) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \lambda_j [(v_{w_j}(p_i) - c_i)^2 - (v_{w_j}(p_i) - d_i)^2] \tag{1.10}$$

$$= \sum_{i=1}^{\infty} (\sum_{j:w_j \notin p_i} \lambda_j)(c_i^2 - d_i^2) + (\sum_{j:w_j \in p_i} \lambda_j)((1-c_i)^2 - (1-d_i)^2)$$

$$= \sum_{i=1}^{\infty} c_i^2 - d_i^2 + 2(\sum_{j:w_j \in p_i} \lambda_j)(d_i - c_i)$$

$$= \sum_{i=1}^{\infty} -c_i^2 - d_i^2 + 2(\sum_{j:w_j \in p_i} \lambda_j) d_i$$

since $c_i = \sum_{j:w_j \in p_i} \lambda_j$. We have $\sum_{i=1}^{\infty} c_i^2 + d_i^2 < \infty$ by i). Thus

$$0 \leq \sum_{i=1}^{\infty} 2(\sum_{j:w_j \in p_i} \lambda_j) d_i < \infty \tag{1.11}$$

as

$$0 \geq \mathfrak{D}(\sum_{j=1}^{\infty} \lambda_j v_{w_j}, c) - \mathfrak{D}(\sum_{j=1}^{\infty} \lambda_j v_{w_j}, d).$$

Having established (1.11), we claim we can use the dominated convergence theorem to switch limits in (1.10). Indeed,

$$\sum_{i=1}^{\infty} \sum_{j=1}^{N} \lambda_j [(v_{w_j} - c(p_i))^2 - (v_{w_j} - d(p_i))^2] = \sum_{i=1}^{\infty} (\sum_{1 \leq j \leq N} \lambda_j)(c_i^2 - d_i^2) + 2(\sum_{\substack{j:w_j \in p_i \\ 1 \leq j \leq N}} \lambda_j)(d_i - c_i).$$

Letting

$$g_N(i) = (\sum_{1 \leq j \leq N} \lambda_j)(c_i^2 - d_i^2) + 2(\sum_{\substack{j:w_j \in p_i \\ 1 \leq j \leq N}} \lambda_j)(d_i - c_i)$$

and noting that $-(\sum_{\substack{j:w_j \in p_i \\ 1 \leq j \leq N}} \lambda_j) c_i \geq -c_i^2$ since $c_i = \sum_{j:w_j \in p_i} \lambda_j$, we see that

$$|g_N(i)| \leq 2c_i^2 + d_i^2 + 2(\sum_{j:w_j \in p_i} \lambda_j) d_i.$$

Each of c_i, d_i, and $(\sum_{j:w_j \in p_i} \lambda_j) d_i$ is summable in i. So, the dominated convergence theorem applies and we can switch limits.

Thus we have

$$\begin{aligned} 0 &\geq \mathfrak{D}(\sum_{j=1}^{\infty} \lambda_j v_{w_j}, c) - \mathfrak{D}(\sum_{j=1}^{\infty} \lambda_j v_{w_j}, d) \\ &= \sum_{j=1}^{\infty} \lambda_j \sum_{i=1}^{\infty} [(v_{w_j}(p_i) - c(p_i))^2 - (v_{w_j}(p_i) - d(p_i))^2] \\ &= \sum_{j=1}^{\infty} \lambda_j (\mathfrak{D}(v_w, c) - \mathfrak{D}(v_w, d)) \geq 0 \end{aligned}$$

where we used that c and d are both finitely inaccurate for each $w \in W$ to break up the summation in the second line. Thus we conclude that $c = d$ as $\mathfrak{D}(c, d) = 0$ if and only if $c = d$. □

II

Next, we give sufficient conditions for avoiding strong dominance on *negation-compact*[9] opinion sets and weak dominance on countably discriminating negation-compact opinion sets. Both results hold for any generalized legitimate inaccuracy measure.

Definition 1.3.12. Let $\mathcal{F} \subseteq \mathcal{P}(W)$. Let $f(n) \in \{0,1\}$ and set $p_n^{f(n)} = p_n$ if $f(n) = 0$ and $p_n^{f(n)} = p_n^c$ if $f(n) = 1$. Then we say $\mathcal{F}$ is *negation-compact* if for any choice of $\{p_n\}_{n=1}^\infty \subseteq \mathcal{F}$ and $f : \mathbb{N} \to \{0,1\}$, if $\bigcap_{n=1}^N p_n^{f(n)}$ is nonempty for every N, then $\bigcap_{n=1}^\infty p_n^{f(n)}$ is nonempty.

First, we establish that coherent credence functions (without any finite inaccuracy assumption) are not strongly dominated on negation-compact opinion sets. Unfortunately, we cannot prove a characterization result because Theorem 1.2.3 does not entail that incoherent credence functions are strongly dominated on negation-compact opinion sets. In the proof of the result, we use Konig's lemma (see, e.g., Hrbacek and Jech 1999, Sec. 12.3) and assume basic knowledge of trees.

Lemma 1.3.13 (Konig's lemma)**.** Any infinite finitely branching tree has an infinite branch.

Proposition 1.3.14. Let $\mathcal{F}$ be a negation-compact opinion set and $\mathscr{I}$ a generalized legitimate inaccuracy measure. If c is a coherent credence function on $\mathcal{F}$, then c is not strongly dominated relative to $\mathscr{I}$.

Proof. Let $\mathscr{I}(c', w) = \sum_{i=1}^\infty \mathfrak{d}(v_w(p_i), c'(p_i))$ for some enumeration of $\mathcal{F}$ and one-dimensional Bregman divergence $\mathfrak{d}$. Let $\mathscr{I}_n(c', w) = \sum_{i=1}^n \mathfrak{d}(v_w(p_i), c'(p_i))$ for each $n \in \mathbb{N}$, $w \in W$, and credence function c' on $\mathcal{F}$. Consider a credence function $d \neq c$. Define

$$T^n = \{(v_w(p_1), \ldots, v_w(p_n)) : \mathscr{I}_k(w, c) < \mathscr{I}_k(w, d) \text{ for some } k \geq n, w \in W\}$$

and $T = e \cup \bigcup_{n=1}^\infty T^n$, where e is the empty sequence. For each $s, t \in T$, we set $s < t$ if and only if s is an initial sequence of t, and we set the height of $t \in T$ to be the length of the tuple. Then T is a binary tree.

We claim T is infinite. Fix $n \in \mathbb{N}$. Then there is a $t \in T$ with height n if and only if $T_n \neq \varnothing$ if and only if $\mathscr{I}_k(c, w) < \mathscr{I}_k(d, w)$ for some $k \geq n$ and $w \in W$. Let k be the maximum of n and the smallest i such that $c(p_i) \neq d(p_i)$. Then since c restricted to any subset of $\mathcal{F}$ is coherent, by Theorem 1.1.8, $\mathscr{I}_k(c, w') < \mathscr{I}_k(d, w')$ for some $w' \in W$ and so $(v_{w'}(p_1), \ldots, v_{w'}(p_n)) \in T^n$.

By Lemma 1.3.13, there exists an infinite branch

$$\mathcal{B} = \bigcup_{n=1}^\infty \{(v_{w_n}(p_1), \ldots, v_{w_n}(p_n))\}$$

[9] This notion is introduced by Borkar et al. (2004), though they did not give a name to it.

through T, where

$$(v_{w_n}(p_1), \ldots, v_{w_n}(p_n)) < (v_{w_m}(p_1), \ldots, v_{w_m}(p_m))$$

whenever $n < m$. For each i, let $p_i^* = p_i$ if $v_{w_i}(p_i) = 1$ and $p_i^* = p_i^c$ if $v_{w_i}(p_i) = 0$. Then $w_n \in \bigcap_{i=1}^n p_i^*$ since $v_{w_i}(p_i) = 1$ if and only if $v_{w_n}(p_i) = 1$ for $i < n$ as $(v_{w_i}(p_1), \ldots, v_{w_i}(p_i)) < (v_{w_n}(p_1), \ldots, v_{w_n}(p_n))$. Thus $\bigcap_{i=1}^n p_i^* \neq \varnothing$ for each n and so by negation-compactness there is some $w \in \bigcap_{i=1}^\infty p_i^*$. Then

$$(v_w(p_1), \ldots, v_w(p_n)) = (v_{w_n}(p_1), \ldots, v_{w_n}(p_n)) \in T^n$$

for each $n \in \mathbb{N}$. By the definition of T^n, we have

$$\mathscr{I}_{k_n}(w, c) < \mathscr{I}_{k_n}(w, d)$$

for some $k_n \geq n$. Sending n to infinity, we see $\mathscr{I}(w, c) \leq \mathscr{I}(w, d)$ and thus d cannot strongly dominate c. □

Next, we use a result of Borkar et al. (2004) and Theorem 1.3.8 to show that coherence and finite expected inaccuracy are sufficient for avoiding weak dominance on countably discriminating negation-compact opinion sets.

Theorem 1.3.15 (Borkar et al. 2004)**.** Let $\mathcal{F}$ be negation-compact. If c is coherent on $\mathcal{F}$, then c is countably coherent on $\mathcal{F}$.

Corollary 1.3.16. Let $\mathcal{F}$ be a countably discriminating negation-compact opinion set and $\mathscr{I}$ a generalized legitimate inaccuracy measure. If c is a coherent credence function on $\mathcal{F}$ with finite expected inaccuracy relative to $\mathscr{I}$, then c is not weakly dominated relative to $\mathscr{I}$.

Proof. Immediate from Theorems 1.3.8 and 1.3.15. □

III

Putting together Theorem 1.2.3, Proposition 1.3.11, Proposition 1.3.14 and Theorem 1.3.15, we get the following characterization result showing that the accuracy argument for probabilism extends to point-finite negation-compact opinion sets.

Theorem 1.3.17. Let $\mathcal{F}$ be a point-finite negation-compact opinion set and c a credence function on $\mathcal{F}$. Then the following are equivalent:

1. c is not strongly dominated relative to $\mathscr{B}$;
2. c is not weakly dominated relative to $\mathscr{B}$;
3. c is coherent;
4. c is coherent and somewhere finitely inaccurate relative to $\mathscr{B}$.

Proof. Assume c is coherent and somewhere finitely inaccurate relative to $\mathscr{B}$. Then by Theorem 1.3.15, since $\mathcal{F}$ is negation-compact, c is countably coherent. Since $\mathcal{F}$ is point-finite, Proposition 1.3.11 entails that c is not weakly dominated. So (4) implies (2). By Theorem 1.2.3, (2) implies (3), and by Theorem 1.3.14 (3) implies (1).

As for the implication from (1) to (4), assume (4) does not hold. Thus c is either incoherent and somewhere finitely inaccurate or $\mathscr{B}(c, w) = \infty$ for all $w \in W$. Since $\mathcal{F}$ is point-finite, it follows that $\mathscr{B}(v_w, w') < \infty$ for all $w, w' \in W$. Thus, in the second case, c is strongly dominated relative to $\mathscr{B}$. If c is incoherent and somewhere finitely inaccurate, it is easy to check by the assumption of point-finiteness that $\mathscr{B}(c, w) < \infty$ for all $w \in W$. Thus by Theorem 1.2.3, c is strongly dominated relative to $\mathscr{B}$, proving that (1) implies (4). □

As a corollary, on point-finite negation-compact opinion sets, coherent credence functions cannot be infinitely inaccurate.

Corollary 1.3.18. Let $\mathcal{F}$ be a point-finite negation-compact opinion set and c a credence function on $\mathcal{F}$. If c is coherent, then $\mathscr{I}(c, w) < \infty$ for all $w \in W$.

Example 1.3.1 shows that Theorem 1.3.17 fails if negation-compactness is dropped, but it is open whether (1), (2), and (4) remain equivalent in this case. Further, it is open whether Theorem 1.3.17 holds when point-finiteness is dropped instead, but Proposition 1.3.14 gives us a hint.

Partitions

While partitions are not negation-compact, we can prove Theorem 1.1.8 when $\mathcal{F}$ is a countably infinite partition of W.[10] In parts of the existing literature (e.g., in Joyce 2009), credence functions are assumed to be defined on a (finite) partition of W to begin with, and so such a result might be especially relevant to extending the accuracy argument for probabilism to countably infinite opinion sets.

For this result, we assume inaccuracy is measured with the generalized Brier score $\mathscr{B}$ (see the beginning of Section 1.3). The characterization follows from two lemmas.

Lemma 1.3.19. Let $\mathcal{F}$ be a countably infinite partition and c a coherent credence function on $\mathcal{F}$. If c is strongly dominated relative to $\mathscr{B}$, then there is a finite opinion set $\mathcal{F}'$ and a coherent credence function c' on $\mathcal{F}'$ that is strongly dominated relative to $\mathscr{B}$[11].

[10] It has been noted that de Finetti's (1974) original proof of Theorem 1.1.8 assuming the Brier score extends to countably infinite opinion sets. However, the only proof we've seen is a sketch of the necessity of coherence for countably infinite partitions in Joyce (1998). Further, such a claim could not be true for arbitrary countable opinion sets as Example 1.3.1 shows.

[11] Here we mean $\mathscr{B}$ to be the Brier score which takes in credences on a finite opinion set.

Proof. Let $\mathcal{F} = \{p_1, p_2, \ldots\}$ and enumerate $\mathcal{V}_{\mathcal{F}}$ such that $v_{w_n}(p_i) = 1$ if and only if $i = n$. Let d be a credence function, and $d_i := d(p_i)$ and $c_i := c(p_i)$. Then

$$\begin{aligned}\mathscr{I}(c, w_n) - \mathscr{I}(d, w_n) &= \sum_{m \neq n} c_m^2 + (1 - c_n)^2 - \sum_{m \neq n} d_m^2 - (1 - d_n)^2 \\ &= \sum_m c_m^2 - \sum_m d_m^2 + 2(d_n - c_n).\end{aligned}$$

Thus d strongly dominates c if and only if

$$d_n > \frac{\sum_m d_m^2 - \sum_m c_m^2}{2} + c_n$$

for all n. Also, notice that c is coherent if and only if $\sum_m c_m \leq 1$, so $\sum_m c_m^2 < \infty$ if c is coherent.

Now we note three things. 1) If c is strongly dominated by some credence function d, then c is strongly dominated by a coherent credence function by Theorem 1.2.3. 2) If c is strongly dominated by a coherent credence function d, then $\sum_m d_m^2 < \sum_m c_m^2$. Indeed, if $\sum_m d_m^2 > \sum_m c_m^2$, then since $d_n \to 0$ and $c_n \to 0$, we can find a K such that

$$2|(d_n - c_n)| < |\sum_m c_m^2 - \sum_m d_m^2|$$

for $n \geq K$. Thus for $n \geq K$,

$$\sum_m c_m^2 - \sum_m d_m^2 + 2(d_n - c_n)$$

is not greater than 0 and so d does not strongly dominate c. If $\sum_m d_m^2 = \sum_m c_m^2$, then $d_n > c_n$ for all n by weak dominance, which contradicts that $\sum_m d_m^2 = \sum_m c_m^2$. 3) If c is strongly dominated by a coherent credence function d and $\sum_m d_m^2 < \sum_m c_m^2$, then c is strongly dominated by a coherent d which is 0 at all but finitely many p_i.

Indeed, assume c is strongly dominated by some coherent credence function d with $\sum_m d_m^2 < \sum_m c_m^2$. Let K be such that

$$\frac{\sum_m d_m^2 - \sum_m c_m^2}{2} + c_n < 0$$

for all $n \geq K$. We can find such a K since $\frac{\sum_m d_m^2 - \sum_m c_m^2}{2}$ is a fixed negative number and $c_n \to 0$. Then let $\overline{d_n} = d_n$ for $n < K$ and $\overline{d_n} = 0$ for $n \geq K$. Then since $\sum_m \overline{d_m}^2 < \sum d_m^2$,

$$d_n > \frac{\sum_m d_m^2 - \sum_m c_m^2}{2} + c_n$$

for all n implies

$$\overline{d_n} > \frac{\sum_m \overline{d_m}^2 - \sum_m c_m^2}{2} + c_n$$

for all n. This establishes the claim.

Assume toward a contradiction that c is strongly dominated by a credence function d. By 1), 2), and 3) above, we may assume d is coherent, $\sum_m d_m^2 < \sum_m c_m^2$, and $d(p_i) = 0$ for all but finitely many i. Let K be such that $d_n = 0$ if $n \geq K$. Then since d strongly dominates c we know

$$d_n > \frac{\sum_m d_m^2 - \sum_m c_m^2}{2} + c_n \text{ for } n < K;$$

$$0 > \frac{\sum_m d_m^2 - \sum_m c_m^2}{2} + c_n \text{ for } n \geq K.$$

Since $c_n \to 0$ and $\frac{\sum_m d_m^2 - \sum_m c_m^2}{2}$ is some fixed negative number, in fact we have that the tail terms are bounded away from zero uniformly. That is,

$$-\epsilon > \frac{\sum_m d_m^2 - \sum_m c_m^2}{2} + c_n \text{ for } n \geq K$$

for some $\epsilon > 0$.

Similarly, since there are only finitely many constraints, we can find an $\epsilon' > 0$ such that

$$d_n - \epsilon' > \frac{\sum_m d_m^2 - \sum_m c_m^2}{2} + c_n \text{ for } n < K$$

Now pick $K' > K$ such that 1) $\frac{\sum_{K'+1}^{\infty} c_n^2}{2} < \frac{\min\{\epsilon,\epsilon'\}}{2}$ and 2) $\sum_{m=1}^{K'} c_m^2 > \sum_{m=1}^{K'} d_m^2$ (which can be done since $\sum_m c_m^2 > \sum_m d_m^2$). Then d_n satisfies

$$d_n > \frac{\sum_{m=1}^{K'} d_m^2 - \sum_{m=1}^{K'} c_m^2}{2} + c_n \text{ for } n < K;$$

$$0 > \frac{\sum_{m=1}^{K'} d_m^2 - \sum_{m=1}^{K'} c_m^2}{2} + c_n \text{ for } n \geq K.$$

Now consider the finite opinion set $\mathcal{F} = \{p_1, \ldots, p_{K'}\}$ over the same set W of worlds. Then these equations show that the credence function $\overline{d} = (d_1, \ldots, d_{K'})$ strongly dominates the coherent credence function $\overline{c} = (c_1, \ldots, c_{K'})$ when restricting to $\{v_{w_1}, \ldots, v_{w_{K'}}\}$ (where the omniscient credences are also restricted to $\mathcal{F}$). For worlds w_n with $n > K'$ we note that

$$\mathscr{I}(w_n, c) - \mathscr{I}(w_n, d) = \sum_{m=1}^{K'} c_m^2 - \sum_{m=1}^{K'} d_m^2 > 0$$

by choice of K'. So $\overline{c}$ is a coherent credence function on a finite opinion set that is strongly dominated relative to $\mathscr{B}$. □

Lemma 1.3.20. Let $\mathcal{F}$ be a countably infinite partition of W and c a coherent credence function on $\mathcal{F}$. If c is weakly dominated relative to $\mathscr{B}$, then there is a coherent c' on $\mathcal{F}$ which is strongly dominated relative to $\mathscr{B}$.

Proof. Let $\mathcal{F} = \{p_1, p_2, \ldots\}$ and enumerate $\mathcal{V}_{\mathcal{F}}$ such that $v_{w_n}(p_i) = 1$ if and only if $i = n$. Let d be a credence function, and $d_i = d(p_i)$ and $c_i = c(p_i)$. Then by the same reasoning in Lemma 1.3.19, d weakly dominates c if and only if

$$d_n \geq \frac{\sum_m d_m^2 - \sum_m c_m^2}{2} + c_n$$

for all n with a strict inequality for some n. Also, notice that c is coherent if and only if $\sum_m c_m \leq 1$ so $\sum_m c_m^2 < \infty$ if c is coherent.

Now we note three things (repeating the reasoning in Lemma 1.3.19). 1) If c is weakly dominated by an incoherent credence function d, then c is strongly dominated by a coherent credence function by Theorem 1.2.3 as $\mathscr{I}(c, w) < \infty$ for all $w \in W$. Thus, by Lemma 1.3.19, c can only be weakly dominated by a coherent credence function. 2) If c is weakly dominated by a coherent credence function d, then $\sum_m d_m^2 < \sum_m c_m^2$. Indeed, if $\sum_m d_m^2 > \sum_m c_m^2$, since $d_n \to 0$ and $c_n \to 0$, we can find a K such that

$$2|(d_n - c_n)| < |\sum_m c_m^2 - \sum_m d_m^2|$$

for $n \geq K$. Thus for $n \geq K$,

$$\sum_m c_m^2 - \sum_m d_m^2 + 2(d_n - c_n)$$

is not greater than or equal to 0 and so d does not weakly dominate c. If $\sum_m d_m^2 = \sum_m c_m^2$, then $d_m \geq c_m$ for each m and $d_m > c_m$ for some m. But then $\sum_m d_m^2 \neq \sum_m c_m^2$. 3) If c is weakly dominated by a coherent credence function d with $\sum_m d_m^2 < \sum_m c_m^2$, then c is weakly dominated by a coherent d which is 0 at all but finitely many p_i. This follows by the same reasoning as in Lemma 1.3.19.

Assume toward a contradiction that c is weakly dominated by a credence function d. By 1), 2), and 3) above we may assume d is coherent, $\sum_m d_m^2 < \sum_m c_m^2$ and $d(p_i) = 0$ for all but finitely many i. Let K be such that $d(p_i) = 0$ if $i \geq K$.

Now, since d weakly dominates c we know

$$d_n \geq \frac{\sum_m d_m^2 - \sum_m c_m^2}{2} + c_n \text{ for } n < K$$

$$0 \geq \frac{\sum_m d_m^2 - \sum_m c_m^2}{2} + c_n \text{ for } n \geq K$$

with strict inequality for some n. Since $c_n \to 0$ and $\frac{\sum_m d_m^2 - \sum_m c_m^2}{2}$ is some fixed negative number, in fact we have that the tail terms are bounded away from zero if we adjust K to be larger as need be. That is,

$$-\delta > \frac{\sum_m d_m^2 - \sum_m c_m^2}{2} + c_n \text{ for } n \geq K,$$

and even with a possible adjustment of K we have that

$$d_n \geq \frac{\sum_m d_m^2 - \sum_m c_m^2}{2} + c_n \text{ for } n < K.$$

Now, let $I \subseteq \{1, \ldots, K\}$ be such that

$$d_i = \frac{\sum_m d_m^2 - \sum_m c_m^2}{2} + c_i$$

for $i \in I$. Note since

$$\frac{\sum_m d_m^2 - \sum_m c_m^2}{2} < 0,$$

this implies that $c_i \neq 0$ for each $i \in I$. Now, we will pick an $\epsilon > 0$ that satisfies the following constraints:

1. $c_i - \epsilon \geq 0$ for each $i \in I$.
2. For all $i \in I$,
$$d_i > \frac{\sum_m d_m^2 - \sum_{n \notin I} c_n^2 - \sum_{n \in I}(c_n - \epsilon)^2}{2} + c_i - \epsilon.$$
3. For all $i \in \{1, \ldots, K\} \setminus I$,
$$d_i > \frac{\sum_m d_m^2 - \sum_{n \notin I} c_n^2 - \sum_{n \in I}(c_n - \epsilon)^2}{2} + c_i.$$
4. For all $i \geq K$,
$$0 > \frac{\sum_m d_m^2 - \sum_{n \notin I} c_n^2 - \sum_{n \in I}(c_n - \epsilon)^2}{2} + c_i.$$

If we find such an $\epsilon > 0$, then by 1-4, we can set $\bar{c}_n = c_n$ for $n \notin I$ and $\bar{c}_n = c_n - \epsilon$ for $n \in I$ and get a coherent credence function (since $\sum_i \bar{c}_i \leq 1$ still) that is strongly dominated by d. Indeed, 2 ensures that d is strictly more accurate than $\bar{c}$ at worlds $i \in I$, 3 ensures that d is strictly more accurate than $\bar{c}$ at worlds $i \in \{1, 2, \ldots, K\} \setminus I$, and 4 ensures that d is strictly more accurate than $\bar{c}$ at worlds $i \geq K$.

We claim we can find an $\epsilon > 0$ since doing so requires satisfying a finite number of satisfiable constraints. To satisfy 1, we just need to pick $0 < \epsilon < c_i$ for all $i \in I$ which is possible since $c_i > 0$ and I is finite.

As for 2, since for all $i \in I$, we have

$$d_i = \frac{\sum_m d_m^2 - \sum_m c_m^2}{2} + c_i$$

the inequality in 2 holds for all i if and only if we can find an ϵ that

$$0 > (\sum_{i \in I} c_i)\epsilon - \frac{|I|\epsilon^2}{2} - \epsilon \iff 0 > \sum_{i \in I} c_i - 1 - \frac{|I|\epsilon}{2} \iff \epsilon > \frac{2(\sum_{i \in I} c_i - 1)}{|I|}.$$

Since c is coherent, we know that $\sum_{i \in I} c_i - 1 \leq 0$ and so this just amounts to satisfying $\epsilon > 0$.

As for 3, since for all $i \in \{1, 2, \ldots, K\} \setminus I$, we have

$$d_i > \frac{\sum_m d_m^2 - \sum_m c_m^2}{2} + c_i$$

and we can find a $\delta' > 0$ such that, for all $i \in I$,

$$d_i - \delta' > \frac{\sum_m d_m^2 - \sum_m c_m^2}{2} + c_i.$$

Then 3 holds if

$$(\sum_{i \in I} c_i)\epsilon - \frac{|I|\epsilon^2}{2} < \frac{\delta'}{2}.$$

Clearly we can find such an $\epsilon > 0$.

Finally, as in 3, since

$$-\delta > \frac{\sum_m a_m^2 - \sum_m c_m^2}{2} + b_n \text{ for } n \geq K$$

we can satisfy 4 if we make

$$(\sum_{i \in I} c_i)\epsilon - \frac{|I|\epsilon^2}{2} < \frac{\delta}{2}.$$

Again, clearly we can find such an $\epsilon > 0$.

Thus we see that satisfying $1-4$ requires a finite number of satisfiable constraints and so we can find such an $\epsilon > 0$ proving the lemma. □

We can now prove the generalization of Theorem 1.1.8 to countably infinite partitions.

Theorem 1.3.21. Let $\mathcal{F}$ be a countably infinite partition of W and c a credence function on $\mathcal{F}$. Then the following are equivalent:

1. c is not strongly dominated relative to $\mathscr{B}$;
2. c is not weakly dominated relative to $\mathscr{B}$;
3. c is coherent.

Proof. Theorem 1.1.8, Lemma 1.3.19 and Lemma 1.3.20 together imply that if c is coherent, then c is not weakly dominated relative to $\mathscr{B}$. Clearly if c is not weakly dominated relative to $\mathscr{B}$, then c is not strongly dominated relative to $\mathscr{B}$. Assume c is incoherent. Then either $\mathscr{B}(c, w) < \infty$ for all $w \in W$ or $\mathscr{B}(c, w) = \infty$ for all $w \in W$. In the former case, Theorem 1.2.3 implies c is strongly dominated. Since each omniscient credence function is finitely inaccurate at each world, in the latter case, c is strongly dominated by any omniscient credence function. □

	Sufficient	Necessary
Countably discriminating	finite expected inaccuracy + countably coherent	somewhere finitely inaccurate + coherent
Point-finite	somewhere finitely inaccurate + countably coherent	somewhere finitely inaccurate + coherent
Negation-compact (for strong dominance)	coherent	?
Countably discriminating & negation-compact	finite expected inaccuracy + coherent	somewhere finitely inaccurate + coherent
Point-finite & negation-compact	coherent	coherent
Partition	coherent	coherent

Figure 1.3.1: a summary of the countable case

Remark 1.3.22. Proposition 1.3.14, Theorem 1.3.17 and Theorem 1.3.21 are related to Theorem 1 of Schervish et al. 2014. However, 1) their assumptions are in some ways stronger and in some ways weaker, and 2) Schervish et al. (2014) only establish that coherence is sufficient for avoiding strong dominance—unlike Theorems 1.3.17 and 1.3.21, their results do not show that coherence is sufficient for avoiding even weak dominance and that being incoherent precludes one from avoiding dominance.

1.4 The Uncountable Case

So far we have been concerned with credences defined on countably infinite opinion sets. We now consider what can be said in favor of probabilism when credences are defined on uncountable opinion sets. When extending from the finite to the countably infinite setting, we kept the additivity requirement for legitimate inaccuracy measures. Similarly, we suggest that the natural generalization of the additivity requirement to the uncountable case is defining, when possible, inaccuracy by integration against a uniform measure. Here we prove a general result concerning inaccuracy defined by integration against any finite measure.

Due to the measure theoretic construction of the inaccuracy measures we consider, we restrict our attention to measurable credence functions and equate credence functions that are equal almost everywhere. In some measure spaces, like the counting measure space underlying generalized legitimate inaccuracy measures, we lose nothing since every credence function is measurable and only the empty set is measure zero. However, in other cases, these assumptions are worth evaluating. We begin by extending the accuracy framework to the measure theoretic setting.

Definition 1.4.1. Let $(\mathcal{F}, \mathcal{A}, \mu)$ be a measure space and $c : \mathcal{F} \to \mathbb{R}^+$. If c is $\mathcal{A}$-measurable and $\mu(\{p : c(p) \notin [0,1]\}) = 0$, we call c a *μ-credence function.* We say a μ-credence function c is *μ-coherent* if there is a coherent (in the usual sense) credence function c' on $\mathcal{F}$ with $c = c'$ μ-a.e. We say a μ-credence function is *μ-incoherent* if there is no coherent credence function c' such that $c = c'$ μ-a.e.

Definition 1.4.2. Let $\mathcal{F}$ be an opinion set (of arbitrary cardinality) over a set W of worlds. Let $(\mathcal{F}, \mathcal{A}, \mu)$ be a σ-finite measure space over the opinion set $\mathcal{F}$. Let $\mathcal{C}$ be the space of all μ-credence functions. Assume $\mathscr{I} : \mathcal{C} \times W \to [0, \infty]$ is such that, for all $(c, w) \in \mathcal{C} \times W$, we have

$$\mathscr{I}(c, w) = B_{\varphi,\mu}(v_w, c)$$

where $B_{\varphi,\mu}$ is a Bregman distance relative to φ[12] and $(\mathcal{F}, \mathcal{A}, \mu)$. In particular, each v_w is a μ-credence function. Then we call $\mathscr{I}$ an *integral inaccuracy measure on* $(\mathcal{F}, \mathcal{A}, \mu)$.

We now prove a dominance result about integral inaccuracy measures. The proof is essentially a measure-theoretic version of the proof of Theorem 1.2.3.

Theorem 1.4.3. Let $\mathscr{I}$ be an integral inaccuracy measure on a finite[13] measure space $(\mathcal{F}, \mathcal{A}, \mu)$. Then for every μ-credence function c, if c is μ-incoherent, then there is a μ-coherent μ-credence function c' that strongly dominates c relative to $\mathscr{I}$.

Proof. Let $\mathscr{I}(c, w) = B_{\varphi,\mu}(v_w, c)$. We write B for $B_{\varphi,\mu}$. Let S be the set of non-negative $\mathcal{A}$-measurable functions on $\mathcal{F}$. Let $E \subseteq S$ be the set of μ-coherent μ-credence functions over $\mathcal{F}$. Then clearly E is convex. Let c be a μ-incoherent μ-credence function. Because μ is finite and $\mathfrak{d}$ is bounded,

$$B(E, c) < \infty.$$

Thus we can apply Theorem 1.2.8 to get a $\pi_c \in S$ such that

$$B(s, t) \geq B(E, c) + B(s, \pi_c) \text{ for every } s \in E. \tag{1.12}$$

In particular, (1.12) holds when s is an omniscient credence function at world w for each w, so we obtain

$$\mathscr{I}(w, c) \geq B(E, c) + \mathscr{I}(w, \pi_c) \tag{1.13}$$

for all w, where all numbers in (1.13) are finite. We show that π_c is in fact a μ-coherent μ-credence function. It suffices to show that π_c is μ-a.e. equal to a coherent

[12] Again, we assume the one-dimensional Bregman divergence $\mathfrak{d}$ generated by φ is bounded.

[13] We may replace finite with σ-finite if in addition i) the μ-coherent μ-credence functions are closed under loose convergence in μ-measure and ii) we weaken the theorem to a two-part theorem as in Theorem 1.2.3. μ being finite is sufficient for i) and implies $\mathscr{I}(c, w) < \infty$ for all $c \in \mathcal{C}$ and $w \in W$.

credence function on $\mathcal{F}$ (since $\pi_c \in \mathcal{S}$, it is $\mathcal{A}$-measurable). To do so, we prove the following claim: E is closed under loose-convergence in μ-measure.

To see this, let $c_n \in E$ for each n and $c \in S$. Assume $c_n \to c$ loosely in μ-measure. The first thing to notice is, since μ is finite, loose μ-convergence implies μ-a.e. convergence on a subsequence $\{a_n\} \subseteq \{n\}$, so that

$$c(p) = \lim_{n\to\infty} c_{a_n}(p) \in [0,1]$$

for each $p \in \mathcal{G}$ with $\mu(\mathcal{G}^c) = 0$. Since the c_{a_n} are μ-coherent, we can change each c_{a_n} on a (measurable) measure zero set $\mathcal{X}_n$ to get coherent μ-credence functions c_{a_n}. Further, we replace $\mathcal{G}$ with $\mathcal{G} \setminus (\cup_{n=1}^{\infty} \mathcal{X}_n)$. Assuming these adjustments have been made, we have that $c_{a_n} \to c$ on $\mathcal{G}$ with $\mu(\mathcal{G}^c) = 0$ and each c_{a_n} is coherent. We now show $c \in E$ by showing it is equal to a coherent credence function on $\mathcal{F}$ when restricting to $\mathcal{G}$.

First, we extend c (resp. c_{a_n}) to $\overline{c}$ (resp. $\overline{c_{a_n}}$), where $\overline{c}$ (resp. $\overline{c_{a_n}}$) is a credence function on $\mathcal{G} \cup \{W\}$ such that $c = \overline{c}$ (resp. $c_{a_n} = \overline{c_{a_n}}$) on $\mathcal{G}$ and $\overline{c}(W) = 1$ (resp. $\overline{c_{a_n}}(W) = 1$). Then notice that c (resp. c_{a_n}) is coherent on $\mathcal{G}$ if and only if $\overline{c}$ (resp. $\overline{c_{a_n}}$) is coherent on $\mathcal{G} \cup \{W\}$. Thus we work with $\overline{c}$ and $\overline{c_{a_n}}$ instead noting that $\overline{c} = \lim_n \overline{c_{a_n}}$ on $\mathcal{G} \cup \{W\}$. To show $c \in E$, we first show $\overline{c}$ is a partial measure on $\mathcal{G} \cup \{W\}$.

Definitions 1.2.12.1 and 1.2.12.3 clearly hold for $\overline{c}$ so we just need to show Definition 1.2.12.2 holds. Let $\varphi_0, \ldots, \varphi_{m-1}, \psi_0, \ldots, \psi_{m'-1} \in \mathcal{G} \cup \{W\}$ and

$$\bigcup_{p \in S^{m,k}} \bigcap_{i \leq k} \varphi_{p_i} \subseteq \bigcup_{p \in S^{m',k}} \bigcap_{i \leq k} \psi_{p_i}$$

for every $k < m$. Since $\overline{c_{a_n}}$ are coherent on $\mathcal{G} \cup \{W\}$ and thus extend to measures on an algebra containing $\mathcal{G} \cup \{W\}$, we have by Corollary 1.2.10 that

$$\sum_{k=0}^{m-1} \overline{c_{a_n}}(\varphi_k) = \sum_{k=0}^{m-1} \overline{c_{a_n}}(\bigcup_{p \in S^{m,k}} \bigcap_{i \leq k} \varphi_{p_i}) \leq \sum_{k=0}^{m'-1} \overline{c_{a_n}}(\bigcup_{p \in S^{m',k}} \bigcap_{i \leq k} \psi_{p_i}) = \sum_{k=0}^{m'-1} \overline{c_{a_n}}(\varphi_k)$$

using that

$$\bigcup_{p \in S^{m,k}} \bigcap_{i \leq k} \psi_{p_i} = \bigcup_{p \in S^{m',k}} \bigcap_{i \leq k} \psi_{p_i} = \varnothing$$

for $k \geq m'$. Sending n to infinity and using the pointwise convergence of $\overline{c_{a_n}}$ to $\overline{c}$ on $\mathcal{G} \cup \{W\}$ we conclude that

$$\sum_{k=0}^{m-1} \overline{c}(\varphi_k) \leq \sum_{k=0}^{m'-1} \overline{c}(\psi_k).$$

Thus $\overline{c}$ is a partial measure on $\mathcal{G} \cup \{W\}$. By Theorem 1.2.13, it follows that there is an FA probability function c^* on $\mathcal{A}(\mathcal{F})$ (the smallest algebra containing $\mathcal{F}$) such that $c^* = \overline{c}$ on $\mathcal{G} \cup \{W\}$. Thus $c^*|_{\mathcal{F}}$ is a coherent credence function on $\mathcal{F}$ and

$$c = \overline{c} = c^*|_{\mathcal{F}}$$

μ-a.e. (specifically off $\mathcal{G}^c$). Further, we already assumed c is $\mathcal{A}$-measurable and $\{p : c(p) \in [0,1]\} \subseteq \mathcal{G}$. Thus c is a μ-coherent μ-credence function.

The proof is finished just as in the proof of Theorem 1.2.3. By Theorem 1.2.8, π_c is the *generalized projection* of c onto E. Also, since

$$B(E, c) = \inf_{s \in E} (s, c) < \infty$$

there is a sequence $\{s_n\} \subseteq E$ such that $B(s_n, c) \to B(E, c)$ by the definition of infimum. By the definition of a generalized projection, $s_n \rightsquigarrow_\mu \pi_c$. Since E is closed under loose convergence, it follows that $\pi_c \in E$. Further, since c is μ-incoherent we know $c \neq \pi_c$ (up to μ-a.e. equivalence) so we see $B(E, c) \geq B(\pi_c, c) > 0$ since $B(s, t) = 0$ if and only if $s = t$ μ-a.e. Since $\mathscr{I}(w, c) < \infty$ for all w, we deduce that

$$\mathscr{I}(w, c) \geq B(E, c) + \mathscr{I}(w, \pi_c) > \mathscr{I}(w, \pi_c)$$

for all $w \in W$. This proves that c is strongly dominated by π_c, and we are done. $\square$

We now show one direction of the accuracy dominance result in Walsh 2019 follows from Theorem 1.4.3. We first recall his result.

Theorem 1.4.4 (Walsh 2019)**.** Let $\mathcal{F}$ be a countably infinite opinion set. Let

$$\mathscr{I}(w, c) = \sum_{i=1}^{\infty} 2^{-i} (v_w(p_i) - c(p_i))^2. \tag{1.14}$$

Then:

1. if c is incoherent, it is strongly dominated relative to $\mathscr{I}$ by a coherent credence function;

2. if c is coherent, it is not weakly dominated relative to $\mathscr{I}$ by any credence function $d \neq c$.

As a corollary to Theorem 1.4.3, we prove the first part of Walsh's (2019) result. We also allow the coefficients in (1.14) to be any $a_n \in (0, 1]$ with $\sum_n a_n < \infty$.

Corollary 1.4.5. Let $\mathcal{F}$ be a countably infinite opinion set. Let

$$\mathscr{I}(c, w) = \sum_{i=1}^{\infty} a_n (v_w(p_i) - c(p_i))^2$$

with $a_n \in (0, 1]$ and $\sum_n a_n < \infty$. Then if c is incoherent, it is strongly dominated relative to $\mathscr{I}$ by a coherent credence function.

Proof. We take the measure space $(\mathcal{F}, \mathcal{P}(\mathcal{F}), \mu)$ over $\mathcal{F}$ where

$$\mu(A) = \sum_{\{i : p_i \in A\}} a_i < \infty$$

for each $A \in \mathcal{P}(\mathcal{F})$. Note $(x-y)^2 = x^2 + y^2 - 2xy$ so letting $\varphi(x) = x^2$

$$(x-y)^2 = \varphi(x) - \varphi(y) - \varphi'(y)(x-y).$$

Also note φ is strictly convex and differentiable on $(0, \infty)$. Thus we have that for all $(c, w) \in \mathcal{C} \times W$,

$$\mathscr{I}(c, w) = B_{\varphi,\mu}(c, w)$$

for $B_{\varphi,\mu}$ a Bregman distance relative to φ and the finite measure space $(\mathcal{F}, \mathcal{P}(\mathcal{F}), \mu)$. Thus the conditions in Theorem 1.4.3 hold. The last thing to notice is that, as with the counting measure, two credence functions are equal μ-a.e. if and only if they are equal everywhere. □

Remark 1.4.6. Corollary 1.4.5 can be generalized to allow any bounded one-dimensional Bregman divergence in place of $\mathfrak{d}(x, y) = (x-y)^2$. The proof is essentially the same.

Here is an example of how Theorem 1.4.3 can be used to give an accuracy argument in a concrete uncountable setting. Assume we have a coin with unknown bias $\theta \in [0, 1]$ and a set of propositions of the form $a \leq \theta \leq b$ for each $a, b \in [0, 1]$. Then a credence function on this uncountable opinion set can be represented by a function

$$c : X \to [0, 1]$$

where $X = \{(a, b) : 0 \leq a \leq b \leq 1\} \subseteq [0, 1]^2$. We put the Lebesgue measure λ on X to generalize the additive constraint used in the countable setting. We let

$$\mathscr{I}(c, w) = \int \mathfrak{d}(v_w(\mathbf{x}), c(\mathbf{x}))\lambda(d\mathbf{x})$$

for a bounded one-dimensional divergence $\mathfrak{d}$. Then the assumptions of Theorem 1.4.3 hold so we get the following dominance result: for any λ-credence function c, if c is a λ-incoherent, then there is a λ-coherent λ-credence function that strongly dominates c.

Chapter 2

Point-Finitely Generated Algebras of Sets

In the general topology literature, characterizing the topologies that have a *point-countable* base (each point is in at most countably many sets from the base) or a *σ-point-finite* base (a countable union of point-finite collections) has been a topic of interest (see, e.g., Heath 1967, Aull 1971, Collins et al. 1990). We ask a related question in the case of generating collections for algebras of sets: which countable algebras of sets can be generated by[1] a point-finite collection?

In answering this question, we link two classes of collections of sets considered in Section 1.3: point-finite collections and countably discriminating algebras (see Definitions 1.3.10 and 1.3.2). We prove that the countable algebras that can be generated by a point-finite collection are exactly the countably discriminating algebras.

Theorem 2.0.1. Let $\mathcal{F}$ be a countable algebra over a set X. Then $\mathcal{F}$ is generated by a point-finite collection if and only if $\mathcal{F}$ is countably discriminating.

We begin the proof by showing that any countable algebra over a countable set can be generated by a point-finite collection. This will be a consequence of a more general result showing that certain properties can be passed from an algebra of sets representing the countable free Boolean algebra to any countable algebra of sets, assuming both are over countable sets.

Definition 2.0.2. We say a class C of algebras of sets has the *point-based transfer property* (PTP) if: given any pair of countable algebras $\mathcal{F}, \mathcal{F}'$ over countable sets X, X' respectively and onto homomorphism $\Psi : \mathcal{F} \to \mathcal{F}'$ such that there is an $f : X' \to X$ (not necessarily onto) where

$$f(\Psi(A)) \subseteq A$$

for all $A \in \mathcal{F}$, if $\mathcal{F} \in \mathsf{C}$, then $\mathcal{F}' \in \mathsf{C}$.

[1] $\mathcal{G} \subseteq \mathcal{P}(X)$ *generates* $\mathcal{F}$ if the smallest algebra containing $\mathcal{G}$ is $\mathcal{F}$.

Remark 2.0.3. Since we assume Ψ is an onto homomorphism in Definition 2.0.2, the condition that $f(\Psi(A)) \subseteq A$ for all $A \in \mathcal{F}$ is equivalent to

$$x' \in \Psi(A) \leftrightarrow f(x') \in A$$

for all $x' \in X'$ and $A \in \mathcal{F}$.

Speaking informally, a sufficient condition for a class having the PTP is that the class is defined by a universal property concerning all points in the underlying set of the algebra of sets. For example, each point being contained in exactly one set and each point being contained in at most N sets (for any N) define classes with the PTP.

We now prove that a representation of the countable free Boolean algebra being in a class with the PTP is in some sense "contagious." In the proof, we assume basic knowledge of the Stone representation theorem for Boolean algebras (see, e.g., Sikorski 1969, p. 23) but often provide citations with page numbers. Also, we will need the following definition in the proof (see, e.g., Bennett 1972).

Definition 2.0.4. A space X is *countable dense homogeneous* if it is separable and for any two countable dense subsets A and B there is a homeomorphism h such that $h(A) = B$.

Theorem 2.0.5. Let C have the point-based transfer property. If there is some realization of the countable free Boolean algebra over a countable set in C, then C contains every countable algebra over a countable set.

Proof. Let $\mathcal{F}$ be a countably infinite algebra of sets over a countable set X and $\mathbb{B}$ the countable free Boolean algebra with $f : \mathcal{F} \to \mathbb{B}$ an isomorphism and where $\mathcal{F} \in \mathsf{C}$. Let $S(\mathbb{B})$ be the Stone space of $\mathbb{B}$. Then by the Stone representation theorem, there is an isomorphism $\Psi : \mathbb{B} \to \text{Clop}(S(\mathbb{B}))$, where $\text{Clop}(S(\mathbb{B}))$ is the set of clopen subsets of $S(\mathbb{B})$ and $\Psi(b) := \{x \in S(\mathbb{B}) : b \in x\}$. Let $\varphi : X \to S(\mathbb{B})$ be given by

$$x \mapsto \{f(A) : x \in A, A \in \mathcal{F}\}.$$

It is easy to see that φ outputs an element of $S(\mathbb{B})$. We claim 1) the image of φ is dense in $S(\mathbb{B})$ and 2) $r^{-1} := f^{-1} \circ \Psi^{-1} : \text{Clop}(S(\mathbb{B})) \to \mathcal{F}$ is induced by φ (that is, $r^{-1}(U) = \{x \in X : \varphi(x) \in U\}$ for all $U \in \text{Clop}(S(\mathbb{B}))$).

First, let V be a nonempty open set in $S(\mathbb{B})$. Then V contains some nonempty clopen set $\{x \in S(\mathbb{B}) : b \in x\}$ for some $b \in \mathbb{B}$. Since f is an isomorphism, let $f(A) = b$ for some nonempty $A \in \mathcal{F}$. Then for any $x' \in A$, we have $b \in \varphi(x')$ and so $\varphi(x') \in \{x \in S(\mathbb{B}) : b \in x\} \subseteq V$. Thus the image of φ is dense in $S(\mathbb{B})$.

Second, let $U = \{x \in S(\mathbb{B}) : b \in x\}$ for $b \in \mathbb{B}$. We need to show

$$r^{-1}(U) = \{x \in X : \varphi(x) \in U\}.$$

Let $x' \in r^{-1}(U)$. Then $x' \in f^{-1}(\Psi^{-1}(U)) = f^{-1}(b)$. But if $x' \in f^{-1}(b)$, then $b \in \varphi(x')$ and so $\varphi(x') \in U$. Thus we conclude $r^{-1}(U) \subseteq \{x \in X : \varphi(x) \in U\}$. Now let $x' \in \{x \in X : \varphi(x) \in U\}$. So $b \in \varphi(x')$ from which it follows $f(A) = b$ for some $A \in \mathcal{F}$ with $x' \in A$. But notice

$$r^{-1}(U) = f^{-1}(\Psi^{-1}(U)) = f^{-1}(b) = A$$

and so $x' \in r^{-1}(U)$, concluding the proof that r^{-1} is induced by φ.

Let $\mathcal{F}'$ be a realization of another countable Boolean algebra $\mathbb{B}'$ over countably infinite X' with $f' : \mathcal{F}' \to \mathbb{B}'$ an isomorphism. Using the diagram in Figure 2.0.1 as a road map, we are going to build an onto homomorphism from $\mathcal{F}$ to $\mathcal{F}'$ that satisfies the constraint in Definition 2.0.2.

$$\begin{array}{ccccccc} & & \overset{h}{\circlearrowright} & & & & \\ \mathcal{F} & \xrightarrow{r} & \mathrm{Clop}(S(\mathbb{B})) & \xrightarrow{g} & \mathrm{Clop}(S(\mathbb{B}')) & \xrightarrow{r'} & \mathcal{F}' \\ & & & & & & \\ X & \overset{\varphi}{\dashrightarrow} & S(\mathbb{B}) & \overset{\chi}{\dashleftarrow} & S(\mathbb{B}') & \overset{\varphi'}{\dashleftarrow} & X' \end{array}$$

Figure 2.0.1

The argument establishing 1) and 2) above did not depend on choosing $\mathbb{B}$ to be the countable free Boolean algebra. Thus, we can run the whole argument again to get an $r' : \mathrm{Clop}(S(\mathbb{B}')) \to \mathcal{F}'$ that is induced by φ' defined similarly as above. Then the image of φ' is dense in $S(\mathbb{B}')$.

Since every countable Boolean algebra is the homomorphic image of the countable free Boolean algebra (Monk et al. 1989, p. 132), there is an onto homomorphism $g : \mathrm{Clop}(S(\mathbb{B})) \to \mathrm{Clop}(S(\mathbb{B}'))$, and since $\mathrm{Clop}(S(\mathbb{B}))$ is a perfect algebra of sets (Sikorski 1969, p. 20), g is induced by a point map $\chi : S(\mathbb{B}') \to S(\mathbb{B})$ (Sikorski 1969, p. 33). Now consider $\chi \circ \varphi' : X' \to S(\mathbb{B})$. The image of $\chi \circ \varphi'$ will be a countable set since the image of φ' is countable in $S(\mathbb{B}')$, so χ takes the image to an at most countable set. Note the image of $\chi \circ \varphi'$ is contained in a countable dense subset D of $S(\mathbb{B})$. Using the fact that the Stone space of the free countable Boolean algebra is homeomorphic to the classical Cantor set (see, e.g., Monk et al. 1989, p. 104) and that the Cantor set is countable dense homogeneous (see, e.g., Hernandez-Gutierrez et al. 2018), there is a homeomorphism h of $S(\mathbb{B})$ sending the image of φ onto D.

Now let $\Gamma = r' \circ g \circ h \circ r$. Then since r, r' are isomorphisms, h is a homeomorphism, and g is an onto homomorphism it follows that Γ is an onto homomorphism. To see that Γ satisfies Definition 2.0.2, we notice the following:

1. Since r^{-1} is induced by φ, for each $A \in \mathcal{F}$

$$x \in A \leftrightarrow \varphi(x) \in r(A).$$

2. For each $A \in \mathrm{Clop}(S(\mathbb{B}))$ and $x \in S(\mathbb{B})$,

$$x \in A \leftrightarrow h(x) \in h(A).$$

3. $\mathrm{Im}\,\chi \circ \varphi' \subseteq \mathrm{Im}\, h|_{\mathrm{Im}\,\varphi}$.

4. Since g is given by the point-map χ, for each $x \in S(\mathbb{B}')$ and $A \in \mathrm{Clop}(S(\mathbb{B}))$,

$$\chi(x) \in A \leftrightarrow x \in g(A).$$

5. Since r' is induced by φ', for each $x \in X'$ and $A \in \mathrm{Clop}(S(\mathbb{B}'))$,

$$\varphi'(x) \in A \leftrightarrow x \in r'(A).$$

Let $f : X' \to X$ map $x' \in X$ to any $x \in X$ such that $h(\varphi(x)) = \chi(\varphi'(x'))$, which we know to exist by 3. We show $f(\Gamma(A)) \subseteq A$ for each $A \in \mathcal{F}$. Consider $f(x')$ for $x' \in \Gamma(A)$ where $\Gamma(A) = r' \circ g \circ h \circ r(A)$. Then by 5, $\varphi'(x') \in g \circ h \circ r(A)$. By 4, $\chi(\varphi'(x')) \in h(r(A))$ and so by definition of f we have $h(\varphi(f(x'))) \in h(r(A))$. By 2, it follows that $\varphi(f(x')) \in r(A)$ and finally by 1 that $f(x') \in A$. Thus, indeed $f(\Gamma(A)) \subseteq A$ proving that Γ is an onto homomorphism as needed in Definition 2.0.2. It follows that $\mathcal{F}' \in \mathsf{C}$ as C has the PTP, completing the proof. □

We use Theorem 2.0.5 to establish that every countable algebra over a countable set can be generated by a point-finite collection, and so we must first show a realization of the countable free Boolean algebra over a countable set can be generated by a point-finite collection. Note that the countable free Boolean algebra can be realized as the periodic subsets of $\mathbb{N}$ with period a power of 2 (see, e.g., Givant and Halmos 2010, p. 140.).

Lemma 2.0.6. The algebra over $\mathbb{N}$ consisting of all periodic sets with period a power of 2 can be generated by a point-finite collection.

Proof. Let

$$\mathcal{A} = \{\{n \in \mathbb{N} : n \equiv j_1, \ldots, j_k \mod 2^k\} : 1 \leq j_1 < \ldots < j_k \leq 2^k, k \in \mathbb{N}\}$$

be the periodic sets of $\mathbb{N}$ with period a power of 2. It is clear that $\mathcal{A}$ is generated by

$$\mathcal{F}' = \{\{n \in \mathbb{N} : n \equiv j \mod 2^k\} : 1 \leq j \leq 2^k, k \in \mathbb{N}\}.$$

We claim the following subset of $\mathcal{A}$ is point-finite and generates $\mathcal{F}'$:

$$\mathcal{F} = \{\{x : x \equiv j \mod 2^k\} : k+1 \leq j \leq 2^k, k \in \mathbb{N}\}.$$

It is easy to check that $\mathcal{F}$ is point-finite—for every $n \in \mathbb{N}$ and some k sufficiently large, n is no longer in the sets of period greater than 2^k. To prove $\mathcal{F}$ generates $\mathcal{F}'$, we induct on k to show $\mathcal{F}$ generates all sets in $\mathcal{F}'$ of the form $\{x : x \equiv j \mod 2^k\}$.

There are two elements in $\mathcal{F}'$ of period 2, namely $2 \mod 2$ and $1 \mod 2$, both of which are in $\mathcal{F}$ so the base case holds. Assume $\{x : x \equiv j \mod 2^k\}$ for $1 \leq j \leq 2^k$ can be generated by elements in $\mathcal{F}'$. We note the following simple but key claim: for each $k, j \in \mathbb{N}$,

$$\{x : x \equiv j \mod 2k\} = \{x : x \equiv j \mod k\} \setminus \{x : x \equiv j + k \mod 2k\}. \tag{2.1}$$

To see (2.1), let $x \equiv j \mod 2k$. Then $x = j + 2kl$ for some $l \in \mathbb{N}$. So clearly $x \equiv j \mod k$. Also notice that $x - j$ is an even multiple of k. But if $x \equiv j + k \mod 2k$, then

$$x = j + k + 2kl' = j + (2l' + 1)k$$

for some $l' \in \mathbb{N}$ and so $x - j$ is an odd multiple of k, which is a contradiction. Thus $x \in \{x' : x' \equiv j + k \mod 2k\}^c$, proving the left-to-right containment in (2.1). Let $x \in \{x' : x \equiv j \mod k\} \setminus \{x' : x' \equiv j + k \mod 2k\}$. Then $x = j + kl$ for $l \in \mathbb{N}$. Since $x \notin \{x' : x' \equiv j+k \mod 2k\}$ we know l must not be odd and so must be even. Thus $x = j + 2kl'$ for some $l' \in \mathbb{N}$ and so $x \equiv j \mod 2k$, proving the right-to-left containment in (2.1).

Now let $1 \leq j \leq k$. By the lemma,

$$\{x : x \equiv j \mod 2^{k+1}\} = \{x : x \equiv j \mod 2^k\} \setminus \{x : x \equiv j + k \mod 2^{k+1}\}.$$

We know $\{x : x \equiv j \mod 2^k\}$ can be generated by elements in $\mathcal{F}$ by induction and $\{x : x \equiv j+k \mod 2^{k+1}\} \in \mathcal{F}$ since $j + k \geq k + 1$. Thus $\{x : x \equiv j \mod 2^{k+1}\}$ can be generated by elements in $\mathcal{F}'$ for all $1 \leq j \leq k$. Since $\{x : x \equiv j \mod 2^{k+1}\} \in \mathcal{F}$ for $k + 1 \leq j \leq 2^{k+1}$, we have proved the claim. □

That every countable algebra over a countable set can be generated by a point-finite collection is an easy corollary to Theorem 2.0.5 and Lemma 2.0.6.[2]

Corollary 2.0.7. Every countable algebra over a countable set can be generated by a point-finite collection.

Proof. We show that the class C of algebras of sets that can be point-finitely generated has the PTP. Let $\mathcal{F}, \mathcal{F}'$ be countable algebras over countable sets X, X' respectively. Let $\Psi : \mathcal{F} \to \mathcal{F}'$ be onto with a map $f : X' \to X$ such that

$$f(\Psi(A)) \subseteq A$$

for all $A \in \mathcal{F}$. Assume $\mathcal{F} \in \mathsf{C}$ and G is a point-finite generating collection for $\mathcal{F}$. Let

$$G' = \{\Psi(A) : A \in G\}.$$

[2]Thanks to Pierre Simon for pointing out that there is a direct way to prove Corollary 2.0.7. Here is a sketch. Let $\mathcal{F} = \{A_i\}_{i=1}^{\infty}$ be a countable algebra over $X = \{x_i\}_{i=1}^{\infty}$. Since $(A \Delta B) \Delta B = A$, for any i, j, replacing A_i with $A_i \Delta A_j$ does not change the algebra generated by the resulting collection. Let i_1 be the minimal i such that $x_1 \in A_i$. For $i > i_1$ with $x_1 \in A_i$, replace A_i with $A_i \Delta A_{i_1}$ so that x_1 is no longer in any A_i for $i > i_1$. Continue for each x_j making sure that $i_j > i_{j-1}$ so that A_{i_j} will not contain $x_{j'}$ for any $j' < j$.

Then G' generates $\mathcal{F}'$ since Ψ is an onto homomorphism. G' is also point-finite since for each $x' \in X'$ and $\Psi(A) \in G'$ (so $A \in G$),

$$x' \in \Psi(A) \leftrightarrow f(x') \in A.$$

So if x' were in infinitely many elements of G', it would follow that $f(x')$ is in infinitely many elements of G, which is a contradiction. Thus $\mathcal{F}' \in \mathsf{C}$ and so C has the PTP. $\square$

We now consider what happens when the underlying space of the countable algebra is uncountable.

Lemma 2.0.8. Let $\mathcal{F}$ be a countable algebra over an uncountable set X. Then $\mathcal{F}$ is generated by a point-finite collection if and only if $\mathcal{F}$ is countably discriminating.

Proof. Consider a countable algebra $\mathcal{F}$ over an uncountably infinite set X. Let $\overline{X} = \{[x] : x \in X\}$ where $[x]$ is the $\sim$ equivalence class of x (see Definition 1.3.2). For each $A \in \mathcal{F}$, let

$$\overline{A} = \{[x] : x \in A\}$$

and

$$\overline{\mathcal{F}} = \{\overline{A} : A \in \mathcal{F}\}.$$

We claim $\mathcal{F}$ is a point-finitely generated algebra if and only if $\overline{\mathcal{F}}$ is. In particular, we claim $\mathcal{G}$ generates $\mathcal{F}$ if and only if $\overline{\mathcal{G}} = \{\overline{B} : B \in \mathcal{G}\}$ generates $\overline{F}$. This follows from two facts. First, it is easy to check that $A = \bigcap_{i=1}^{n} \bigcup_{j=1}^{m_n} B_{ij}$, where $B_{ij} \in \mathcal{F}$ or $B_{ij}^c \in \mathcal{F}$, if and only if $\overline{A} = \bigcap_{i=1}^{n} \bigcup_{j=1}^{m_n} \overline{B_{ij}}$. Second, since $x \in A$ if and only if $[x] \in \overline{A}$, we have that

$$|\{A : x \in A\}| = |\{\overline{A} : [x] \in \overline{A}\}$$

and so each x is in finitely many elements of $\mathcal{G}$ if and only if $[x]$ is in finitely many elements of $\overline{\mathcal{G}}$

Now, we show $\overline{\mathcal{F}}$ is point-finitely generated if and only if $\overline{X}$ is countable, which will prove the lemma. Let $\mathcal{G}$ generate $\overline{\mathcal{F}}$ and be point-finite. Then take $u([x])$ to be the ultrafilter generated by $[x]$. Since $\mathcal{G}$ is point-finite, we have

$$u([x]) \cap \mathcal{G} < \infty.$$

But note that $u([x])$ is completely determined by $u([x]) \cap \mathcal{G}$. So if

$$u([x]) \cap \mathcal{G} = u([y]) \cap \mathcal{G},$$

then $u([x]) = u([y])$ which implies $[x] = [y]$ by the definition of $\sim$. In other words, the map from $\overline{X}$ to finite subsets of $\mathcal{G}$ given by

$$[x] \mapsto u([x]) \cap \mathcal{G}$$

is an injection. This implies $\overline{X}$ is countable. Also if $\overline{X}$ is countable, then $\overline{\mathcal{F}}$ is point-finitely generated by Theorem 2.0.7, and so we are done. $\square$

The main theorem of this chapter is a simple consequence of Corollary 2.0.7 and Lemma 2.0.8.

Theorem 2.0.1. Let $\mathcal{F}$ be a countable algebra over a set X. Then $\mathcal{F}$ is generated by a point-finite collection if and only if $\mathcal{F}$ is countably discriminating.

Proof. If X is countable, then $\mathcal{F}$ is countably discriminating and, by Corollary 2.0.7, generated by a point-finite collection, proving the equivalence. If X is uncountable, the equivalence is given by Lemma 2.0.8. □

To conclude, consider the following generalization of the notion of point-finiteness.

Definition 2.0.9. For $\mathcal{F}$ an algebra of sets, we define the *agglomerativity* of $\mathcal{F}$ to be the least cardinal κ such that there is a generating collection for $\mathcal{F}$ in which every point belongs to fewer than κ sets from the generating collection.[3]

For example, the agglomerativity of the finite-cofinite algebra over a countable set is 2 since it can be generated by the set of singletons. By Theorem 2.0.7, the agglomerativity of any countable algebra over a countable set is at most $\aleph_0$. It is also easy to show that the agglomerativity of an algebra of sets over a countable set representing the countable free Boolean algebra is $\aleph_0$.[4]

More generally, can we give a theory of the agglomerativity of algebras of sets in terms of properties of the Boolean algebras they represent? While the remarks just made are a first step toward such a theory of agglomerativity, we leave the rest for future work.

[3] Thanks to Wesley Holliday for suggesting this formulation of the notion of agglomerativity.

[4] The proof comes down to showing that if $\mathcal{A}$ has finite agglomerativity then $\mathcal{A}$ has an atom.

Bibliography

C. E. Aull. Topological spaces with a σ-point finite base. *Proceedings of the American Mathematical Society*, 29(2):411–416, 1971.

A. Banerjee, Xin Guo, and Hui Wang. On the optimality of conditional expectation as a Bregman predictor. *IEEE Transactions on Information Theory*, 51(7):2664–2669, July 2005.

R. Bennett. Countable dense homogeneous spaces. *Fundamenta Mathematicae*, 74 (3):189–194, 1972.

Vivek Borkar, Vijay Konda, Goldman Sachs, and Sanjoy Mitter. On de Finetti coherence and kolmogorov probability. *Statistics & Probability Letters*, 2004.

P.J. Collins, G.M. Reed, and A.W. Roscoe. Open problems in topology. chapter The Point-Countable Base Problem. North Holland, 1990.

I. Csiszár. Generalized projections for non-negative functions. *Acta Mathematica Hungarica*, 68(1-2):161–186, 3 1995.

Bruno de Finetti. *Theory of Probability.* New York: John Wiley, 1974.

Kenny Easwaran. Expected accuracy supports conditionalization—and conglomerability and reflection. *Philosophy of Science*, 80(1):119–142, 2013.

Vaclav Fabian and James Hannan. *Introduction to Probability and Mathematical Statistics*. John Wiley Sons Inc, 1985.

Steven Givant and Paul Halmos. Introduction to Boolean algebras. undergraduate texts in mathematics. *Bulletin of Symbolic Logic*, 16:281–282, 2010.

Tilmann Gneiting and Adrian E Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477):359–378, 2007.

Hilary Greaves and David Wallace. Justifying conditionalization: conditionalization maximizes expected epistemic utility. *Mind*, 115(459):607–632, 2005.

Roy W. Heath. On certain first countable spaces. 1967.

Rodrigo Hernandez-Gutierrez, Michale Hrusak, and Jan van Mill. Countable dense homogeneity and λ-sets. *Fundamenta Mathematicae*, 2018.

Alfred Horn and Alfred Tarski. Measures in Boolean algebras. *Transactions of the American Mathematical Society*, 1947.

Paul Horwich. *Probability and Evidence*. Cambridge University Press, 1982.

Karel Hrbacek and Thomas Jech. *Introduction to Set Theory*. Marcel Dekker, 1999.

Franz Huber. Formal representations of belief. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, spring 2016 edition, 2016.

Simon M. Huttegger. In defense of reflection. *Philosophy of Science*, 80(3):413–433, 2013.

James Joyce. Accuracy and coherence: Prospects for an alethic epistemology of partial belief. In Franz Huber and Christoph Schmidt-Petri, editors, *Degrees of Belief*, pages 263–297. Synthese, 2009.

James M. Joyce. A nonpragmatic vindication of probabilism. *Philosophy of Science*, 65(4):575–603, 1998.

Hannes Leitgeb and Richard Pettigrew. An objective justification of Bayesianism I: Measuring inaccuracy. *Philosophy of Science*, 77(2):201–235, 2010a.

Hannes Leitgeb and Richard Pettigrew. An objective justification of Bayesianism II: The consequences of minimizing inaccuracy. *Philosophy of Science*, 77(2):236–272, 2010b.

D. V. Lindley. System design for human interaction. chapter Scoring Rules and the Inevitability of Probability, pages 182–208. IEEE Press, Piscataway, NJ, USA, 1987.

Patrick Maher. Joyce's argument for probabilism. *Philosophy of Science*, 69(1): 73–81, 2002.

James Donald Monk, Sabine Koppelberg, and Robert Bonnet. *Handbook of Boolean Algebras*. North-Holland, 1989.

Richard Pettigrew. *Accuracy and the Laws of Credence*. Oxford University Press UK, 2016.

J. B. Predd, R. Seiringer, E. H. Lieb, D. N. Osherson, H. V. Poor, and S. R. Kulkarni. Probabilistic coherence and proper scoring rules. *IEEE Transactions on Information Theory*, 55(10):4786–4792, Oct 2009.

M. J. Schervish, Teddy Seidenfeld, and J. B. Kadane. Dominating countably many forecasts. *Ann. Statist.*, 42(2):728–756, 04 2014. doi: 10.1214/14-AOS1203.

Roman Sikorski. *Boolean Algebras*. Springer-Verlag Berlin Heidelberg, 1969.

William Talbott. Bayesian epistemology. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, winter 2016 edition, 2016.

Susan Vineberg. Dutch book arguments. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, spring 2016 edition, 2016.

Sean Walsh. Probabilism in infinite dimensions. 2019. Unpublished manuscript.