Philosophical Foundations of Imprecise Probability

MUNICH CENTER FOR MATHEMATICAL PHILOSOPHY

LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

Gregory Wheeler

ISIPTA 2015 Tutorial

The history of philosophy is to a great extent that of a certain clash of human temperaments.

The history of philosophy is to a great extent that of a certain clash of human temperaments. … [Temperament] loads the evidence … one way or the other, making for a more sentimental or a more hard-hearted view of the universe, just as this fact or that principle would (James 1907, 8–9).

Zermelo-Fraenkel

$0 = \{\}$

$1 = \{0\} = \{\{\}\}$

$2 = \{1\} = \{\{\{\}\}\}$

$3 = \{2\} = \{\{\{\{\}\}\}\}$

Von Neumann

$0 = \{\}$

$1 = \{0\} = \{\{\}\}$

$2 = \{0, 1\} = \{\{\}, \{\{\}\}\}$

$3 = \{0, 1, 2\} = \{\{\}, \{\{\}\}, \{\{\}, \{\{\}\}\}\}$

Subjective Interpretation

Ramsey (1926), de Finetti (1937), Savage (1954), Anscombe-Aumann (1963)
Jeffreys (1939), Fisher (1936)

Logical Interpretation

Carnap (1945, 1952), Paris & Vencovská (2015), Kyburg (1961, 2001)

Frequency / Propensity Interpretation

Reichenbach[1] and Popper (1959)

[1] See (Glymour and Eberhardt 2012).

What is probability?

Any response should answer at least three questions (Salmon 1967):

1. Why should probability have particular **mathematical properties**?
2. How do are probabilities **determined or measured**?
3. Why and when is probability **useful**?

3. *Why is logical probability useful*?

   Measures the strength of evidential support.

3. *Why is logical probability useful?*

   Measures the strength of evidential support.

2. *How are logical probabilities measured?*

Carnap	Kyburg
?	From statistical data

3. *Why is logical probability useful?*

   Measures the strength of evidential support.

2. *How are logical probabilities measured?*

Carnap	Kyburg
?	From statistical data

1. *Why does logical probability satisfy the axioms?*

Carnap	Kyburg
Analytic	?

3. *Why is subjective probability useful?*

Measures the strength of partial belief.

Allows us to calculate expected utility calculations.

3. *Why is subjective probability useful?*

   Measures the strength of partial belief.

   Allows us to calculate expected utility calculations.

2. *How is subjective probability measured?*

   Betting behavior

   Accurate forecasting

   Preferences among compound lotteries

   Preferences among acts

1. *Why does subjective probability satisfy the axioms?*

Measurement Procedure	Rationality Criteria
Betting behavior	Avoiding sure loss
Accurate forecasting	Minimizing squared-error loss
Qualitative verbal comparisons	Qualitative probability axioms
Preferences among lotteries	VNM & Anscombe-Aumann axioms
Preference among acts	Savage axioms

1. *Why does subjective probability satisfy the axioms?*

Measurement Procedure	**Rationality Criteria**
Betting behavior	Avoiding sure loss
Accurate forecasting	Minimizing squared-error loss
Qualitative verbal comparisons	Qualitative probability axioms
Preferences among lotteries	VNM & Anscombe-Aumann axioms
Preference among acts	Savage axioms

Theorem: If probability is elicited via the *measurement procedure*, then for the corresponding *rationality criteria*:

$$\textit{Rationality criteria} \Leftrightarrow \textit{Satisfy Probability Axioms.}$$

- Traditional dF-style use of (strictly) proper scoring rules:

Measurement Procedure	Rationality Criteria
Accurate forecasting	Minimizing squared-error loss

- Purely epistemic interpretation of (strictly) proper scoring rules:[2]

Alethic Property	Rationality Criteria
Gradational (in)accuracy	'Distance' from the truth

[2]See (Joyce 1998; Joyce 2009; Leitgeb and Pettigrew 2010; Pettigrew 2013).

Two philosophical temperaments:

Tender-minded: cling to the belief that facts should be related to values and that values seen as predominant.

Tough-minded: want facts to be dissociated from values and left to themselves.

Joyce's Commitments (1998, 2009)

- Credal commitments (belief) modeled by IP &
- Purely epistemic interpretation of (strictly) proper scoring rules:

Theorem (Seidenfeld et al. (2012)[3])

Admissibility, Imprecision, Continuity, Quantifiability, Extensionality, and Strict Immodesty are jointly inconsistent.

[3] A mild mathematical generalization is in Mayo-Wilson and Wheeler, forthcoming.

For the purposes of this talk,

a *scoring rule* $I(b, \omega)$ denotes the ‘inaccuracy’ of the belief b about a proposition φ when the truth-value of φ is $\omega \in \{0, 1\}$.

Claim: *there is no strictly proper IP scoring rule.*

Plan: give 6 necessary postulates that cannot all be satisfied.

Admissibility Let b, c, and d be three (not necessarily distinct) belief states, and suppose that d is at least as accurate as c whatever the truth.

If your belief state is b and the set of rational belief states R_b from your perspective contains c, then it also contains d.

Imprecision: A belief state is a set of real numbers between 0 and 1.

Quantifiability: Degrees of inaccuracy are represented by non-negative real numbers.

Extensionality: For every truth-value ω and every belief state b, there is a single degree of inaccuracy $I(b, \omega)$ representing how inaccurate belief b is.

Moreover, this degree depends only upon b and the truth-value ω of the proposition φ of interest.

Strict Immodesty: If your belief state is b, then the set of rational belief states R_b from your perspective is $\{b\}$.

Problem: It is unclear how to represent the distance between arbitrary sets of numbers between 0 and 1.

How "close" are the beliefs that

(i) a flipped coin lands heads in the interval $[\frac{1}{4}, \frac{3}{4}]$, and
(ii) a flipped coin lands heads in the interval $[\frac{1}{4}, \frac{3}{4}]$ other than $\frac{4}{7}$?

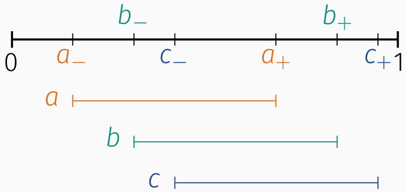
PARETO CONSTRAINT

Suppose that belief states a, b, c are such that

$$a_- \leq b_- \leq c_- \text{ or } a_- \geq b_- \geq c_-, \text{ and}$$
$$a_+ \leq b_+ \leq c_+ \text{ or } a_+ \geq b_+ \geq c_+.$$

Constraint P The distance between the belief states a and c ought to be at least as great as the distance between the belief states a and b.

Example:



Continuity Sufficiently similar belief states are similarly inaccurate. More precisely, for all ω, the function $I(b, \omega)$ restricted to the set of interval beliefs b is continuous with respect to the parameter b, where the metric on beliefs satisfies Constraint P.

Theorem (Seidenfeld et al. (2012)[4])

Admissibility, Imprecision, Continuity, Quantifiability, Extensionality, and Strict Immodesty are jointly inconsistent.

[4] A mild mathematical generalization is in Mayo-Wilson and Wheeler, forthcoming.

One to many propositions: Although formulated for a single proposition, our result extends to finitely many propositions with additional mathematical machinery to ensure the topological invariance of dimension.[5]

[5] Thanks here to Catrin Campbell-Moore.

One to many propositions: Although formulated for a single proposition, our result extends to finitely many propositions with additional mathematical machinery to ensure the topological invariance of dimension.[5]

Other Uncertainty Models: The theorem applies to *Dempster-Shafer Belief functions* and *Ranking functions.*

[5] Thanks here to Catrin Campbell-Moore.

THE OPTIONS

- Admissibility
- Extensionality

Central to accuracy-first epistemology

- Admissibility
- Extensionality

Central to accuracy-first epistemology

- Imprecision

Central to IP

THE OPTIONS

- Admissibility
- Extensionality

Central to accuracy-first epistemology

- Imprecision

Central to IP

- Continuity
- Quantifiability
- Strict Immodesty

Remaining options

Drop Continuity: Perhaps discontinuities for extreme belief states are okay, such as assigning probability zero to a true proposition.

Drop Continuity: Perhaps discontinuities for extreme belief states are okay, such as assigning probability zero to a true proposition.

Our proof shows the stronger result that a measure of inaccuracy ***must be*** *discontinuous almost everywhere if it is to satisfy the other 5 axioms.*

So, continuity stays.

Drop Quantifiability: Perhaps the extended reals would work, such as giving the score ∞ to the vacuous belief state $[0, 1]$.

Drop Quantifiability: Perhaps the extended reals would work, such as giving the score ∞ to the vacuous belief state $[0, 1]$.

*Our proof holds for the extended reals, too; one cannot weaken **Quantifiability** by a small trick.*

SSK Lexicographic probabilities

Joyce Inaccuracy can be measured by a single real number **only** when degrees of belief are; Sturgeon calls this principle *Character Matching*[6]

[6] See (Wheeler 2014) for a reply.

Character Matching: If your degrees of belief are indeterminate, then the distance between your degrees of belief and the truth is likewise indeterminate.

Character Matching: If your degrees of belief are indeterminate, then the distance between your degrees of belief and the truth is likewise indeterminate.

So, perhaps inaccuracy should be represented by **a set** of real numbers rather than **a single** one.

Example:

Suppose p is a precise credence and $I(p, \omega)$ its inaccuracy if $\omega = 1$

Example:

Suppose p is a precise credence and $I(p, \omega)$ its inaccuracy if $\omega = 1$

Suppose p is replaced by $[\frac{1}{3}, \frac{2}{3}]$.

Example:

Suppose p is a precise credence and $I(p, \omega)$ its inaccuracy if $\omega = 1$

Suppose p is replaced by $[\frac{1}{3}, \frac{2}{3}]$.

A natural idea is then for inaccuracy to be $\{I(q, \omega) : q \in [\frac{1}{3}, \frac{2}{3}]\}$.

But if this argument for Imprecision is convincing, then one should abandon the idea that inaccuracy is numerically quantifiable at all.

Why?

But if this argument for Imprecision is convincing, then one should abandon the idea that inaccuracy is numerically quantifiable at all.

Why?

Just as an **indeterminate** credal state may admit an **indeterminate** degree of inaccuracy with respect to a single proposition,

But if this argument for Imprecision is convincing, then one should abandon the idea that inaccuracy is numerically quantifiable at all.

Why?

Just as an **indeterminate** credal state may admit an **indeterminate** degree of inaccuracy with respect to a single proposition, so too can a **precise** credal state admit **indeterminate** degrees of inaccuracy with respect to multiple propositions.

- Admissibility
- Extensionality
- Quantifiability

Quantifiability and **Pure Epistemic Loss**, including “accuracy,” are incompatible.

- Admissibility
- Extensionality
- Quantifiability

Quantifiability and **Pure Epistemic Loss**, including "accuracy," are incompatible.

THE OPTIONS

- Admissibility
- Extensionality

Central to accuracy-first epistemology

- Imprecision

Central to IP

- Continuity
- Quantifiability
- Strict Immodesty

MILDLY PROPER IP SCORING RULES

PLAN

Claim: *there are ~~strictly~~ mildly proper IP scoring rules.*

PLAN

Claim: *there are ~~strictly~~ mildly proper IP scoring rules.*

Plan: give new postulates to replace Strict Immodesty.

~~Strict Immodesty~~ If an agent's belief state is b, then the set of rational belief states R_b from her perspective is equal to the singleton $\{b\}$.

~~Strict Immodesty~~ If an agent's belief state is b, then the set of rational belief states R_b from her perspective is equal to the singleton $\{b\}$.

Mild Immodesty If an agent's belief state is b, then the set of rational belief states R_b from her perspective includes b.

Problem: There are lots of mildly immodest scoring rules.

Here is one:

Problem: There are lots of mildly immodest scoring rules.

Here is one:

> **Lucky 7:** Score every belief by your lucky number, $I(b, \omega) = 7$.

Problem: There are lots of mildly immodest scoring rules.

Here is one:

> **Lucky 7:** Score every belief by your lucky number, $I(b, \omega) = 7$.

Lucky 7 satisfies *Imprecision*, *Continuity*, *Quantifiability*, *Extensionality*, *Admissibility*, and (non-strict) *Mild Immodesty*.

The remaining postulates aim to pick out *reasonable* mildly immodest scoring rules.

Truth-Directedness Let $b, c \in B$ be any two beliefs. If $|p - \omega| < |q - \omega|$ for all precise credences $p \in b$ and $q \in c$, then $I(b, \omega) < I(c, \omega)$.

Truth directedness rules out vacuous rules like **Lucky 7**.

Adding a bad egg to the pan cannot improve the omelet.

Adding a bad egg to the pan cannot improve the omelet.

Example:

If Hans believes that *Miami is south of Munich* to degree .99 and Klaus believes it only to degree .9,

Adding a bad egg to the pan cannot improve the omelet.

Example:

If Hans believes that *Miami is south of Munich* to degree .99 and Klaus believes it only to degree .9, Hans cannot be more accurate weakening his belief from .99 to [.9, .99].

Adding a bad egg to the pan cannot improve the omelet.

SOL Let $b, c \in B$ be any two beliefs such that $b \subseteq c$ and $|q - \omega| > |p - \omega|$ for all $q \in c \setminus b$ and $p \in b$. Then $I(b, \omega) < I(c, \omega)$.

Adding accurate credences cannot make a belief state less accurate.

Adding accurate credences cannot make a belief state less accurate.

If Hans's belief that *Miami is south of Munich* are represented by [.9, .99],

Adding accurate credences cannot make a belief state less accurate.

If Hans's belief that *Miami is south of Munich* are represented by $[.9, .99]$, then his belief cannot be made less accurate by weakening to $[.9, 1]$.

Adding accurate credences cannot make a belief state less accurate.

Monotonicity Let $b, c \in B$ be any two beliefs such that $b \subseteq c$ and $|q - \omega| \leq |p - \omega|$ for all $q \in c \setminus b$ and $p \in b$. Then $I(b, \omega) \geq I(c, \omega)$.

Let b, c, and d be three (not necessarily distinct) belief states, and suppose that d is at least as accurate as c whatever the truth.

~~Admissibility~~ If your belief state is b and the set of rational belief states R_b from your perspective contains c, then it also contains d.

Let b, c, and d be three (not necessarily distinct) belief states, and suppose that d is at least as accurate as c whatever the truth.

~~Admissibility~~ If your belief state is b and the set of rational belief states R_b from your perspective contains c, then it also contains d.

Dominance If d is strictly less accurate than c whatever the truth, then d is not a rational belief state.

Regardless of one's belief state b, the set of rational beliefs R_b does not contain d.

Theorem ((Mayo-Wilson and Wheeler 2015))

Dominance, Imprecision, Quantifiability, Extensionality, Mild Immodesty, Continuity, Truth-Directedness, SOL, and Monotonicity

[7] For a *single* proposition

Theorem ((Mayo-Wilson and Wheeler 2015))

Dominance, Imprecision, Quantifiability, Extensionality, Mild Immodesty, Continuity, Truth-Directedness, SOL, and Monotonicity entail there is a function $f : B \rightarrow [0, 1]$ *such that, for any belief b:*

[7] For a *single* proposition

Theorem ((Mayo-Wilson and Wheeler 2015))

Dominance, Imprecision, Quantifiability, Extensionality, Mild Immodesty, Continuity, Truth-Directedness, SOL, and Monotonicity entail there is a function $f : B \to [0, 1]$ *such that, for any belief b:*

- $f(p) \in [b_-, b_+]$, *and*
- $I(b, \omega) = I(f(p), \omega)$ *for all* ω[7]

[7] For a *single* proposition

THEOREM

Any mildly immodest method of measuring inaccuracy of an imprecise belief *b* must reduce to measuring the inaccuracy of *exactly one precise credence.*

CONVERSE?

We would like to say that any measure of inaccuracy $I_{\mathcal{A}}$[8], together with

- the definition of R_b, and
- having the functional form $I(b,\omega) = I(f(p),\omega)$ for all ω, then:

$I_{\mathcal{A}}$ satisfies Dominance, Imprecision, Quantifiability, Extensionality, Mild Immodesty, Continuity, Truth-Directedness, SOL, and Monotonicity

[8] Again, for a *single* proposition

Suppose $I_{\mathcal{A}}$ satisfies Extensionality, Continuity, and Truth-Directedness, and

for all $b, c \in \mathcal{B}$

1. $f(b) \in [b_-, b_+]$,
2. If $p < q$ for all $p \in b$ and $q \in c$, then $f(b) < f(c)$,
3. If $b \subseteq c$ and $p < q$ for all $p \in b$ and $q \in c \setminus b$, then $f(b) \leq f(c)$, and
4. If $b \subseteq c$ and $q < p$ for all $p \in b$ and $q \in c \setminus b$, then $f(c) \leq f(b)$.

then $I_{\mathcal{A}}$ satisfies Dominance, Imprecision, Quantifiability, Mild Immodesty, SOL, and Monotonicity.

How to use this measure to score an imprecise belief state?

Mid-point scoring Measure inaccuracy of b by scoring its midpoint

There are a wide range of ways to satisfy the axioms.

Midpoint scoring is one way to formalize "average" inaccuracy of an interval-valued belief state.

Limitations to mid-point scoring

- The midpoint of $\frac{1}{2}$ and $[0, 1]$ are the same.
- Useless for elicitation

Recall that the original motivation for studying strictly proper scoring rules was for elicitation. But if $\frac{1}{2}$ scores the same as $[0, 1]$, then a rational agent has no accuracy-related incentive to report one credal state over the other.

Joyce His arguments for IP, like most IP theorists, do not appeal to accuracy.

Instead, imprecision is thought to reflect the quality of *evidence*.

Arguments for imprecision **should not** appeal to accuracy.

Arguments for imprecision **should not** appeal to accuracy.

Example: Alice is a US history scholar and knows that Lincoln wore a stovetop hat.

Bill thinks every 19th Century US President wore a stovetop hat.

Arguments for imprecision **should not** appeal to accuracy.

Example: Alice is a US history scholar and knows that Lincoln wore a stovetop hat.

Bill thinks every 19th Century US President wore a stovetop hat.

Alice and Bill are just as accurate about Lincoln.

Moral If imprecision is determined by strength of evidence, then precision and accuracy may come apart.

Moral If imprecision is determined by strength of evidence, then precision and accuracy may come apart.

There may be many belief states of different precision that are equally accurate.

If so, then considerations of accuracy will *generally* fail to narrow the set of rational beliefs to a single state as Strict Immodesty requires.

3 Ideas

3 ideas

Interpretation: Interpretation of probability is entangled with use.

Temperament: An understandable desire for objectivity.

Interpretation: Interpretation of probability is entangled with use.

Temperament: An understandable desire for objectivity.

Mildly Proper IP Scoring Rules: How to manage the math and meaning of such a thing.
To reconcile 'accuracy' and 'imprecision', two options

1. Drop Quantifiability [Joyce, Seidenfeld / SSK]
2. Replace Strict Immodesty by Mild Immodesty [us]

Dropping strict immodesty can be motivated by evidential considerations, like imprecision.

References

Anscombe, F. J. and R. J. Aumann (1963).
A definition of subjective probability.
Annals of Mathematical Statistics 34, 199–205.

Carnap, R. (1945).
On inductive logic.
Philosophy of Science 12, 72–97.

Carnap, R. (1952).
The Continuum of Inductive Methods.
University of Chicago Press.

de Finetti, B. (1937).
La prévision: ses lois logiques, ses sources subjectives.
Annales de l'Institut Henri Poincaré 7(1), 1–68.

Fisher, R. A. (1936).
Uncertain inference.
Proceedings of the American Academy of Arts and Sciences 71, 245–258.

Glymour, C. and F. Eberhardt (2012).
Hans Reichenbach.
In E. N. Zalta (Ed.), *Stanford Enclyclopedia of Philosophy* (Winter 2012 ed.). CSLI Publications.

James, W. (1907).
The present dilemma in philosophy.
In G. Gunn (Ed.), *Pragmatism and Other Writings*, 2000. London: Penguin Books.

Jeffreys, H. (1939).
Theory of Probability.
Oxford: Clarendon Press.

Joyce, J. (1998).
A nonpragmatic vindication of probabilism.
Philosophy of Science 65(4), 575–603.

Joyce, J. M. (2009).
Accuracy and coherence: Prospects for an alethic epistemology of partial belief.
In *Degrees of belief*, pp. 263–297. Springer.

Kyburg, Jr., H. E. (1961).
Probability and the Logic of Rational Belief.
Middletown, CT: Wesleyan University Press.

Kyburg, Jr., H. E. and C. M. Teng (2001).
Uncertain Inference.
Cambridge: Cambridge University Press.

Leitgeb, H. and R. Pettigrew (2010).
An objective justification of Bayesianism I: Measuring accuracy.
Philosophy of Science 77(2), 201–235.

Mayo-Wilson, C. and G. Wheeler (2015).
Scoring imprecise credences: A mildly immodest proposal.
Philosophy and Phenomenological Research.

Paris, J. and A. Vencovská (2015).
Pure Inductive Logic.
ASL Perspectives in Mathematical Logic. Cambridge: Cambridge University Press.

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

Popper, K. R. (1959).
The Logic of Scientific Discovery.
Routledge.

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

Salmon, W. (1967).
The Foundations of Scientific Inference, Volume 28.
Pittsburgh, PA: University of Pittsburgh Press.

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

Seidenfeld, T., M. J. Schervish, and J. B. Kadane (2012).
Forecasting with imprecise probabilities.
International Journal of Approximate Reasoning 53(8), 1248–1261.

Wheeler, G. (2014).
Character matching and the envelope of belief.
In F. Lihoreau and M. Rebuschi (Eds.), *Epistemology, Context, and Formalism*, Synthese Library, pp. 185–94. Springer.