Imprecise Probabilistic Machine Learning:

Integral Imprecise Probability Metrics & Conformal Uncertainty Quantification

Siu Lun Chau

Epistemic Intelligence & Computation Lab,
College of Computing & Data Science,
Nanyang Technological University, Singapore

February 2026

A bit of me

1. (2014-2018) MMATH in Mathematics and Statistics, Oxford
2. (2018-2023) DPhil in Statistics (OxWaSP CDT), Oxford
   - Supervisors: Dino Sejdinovic (main), Xiaowen Dong, Mihai Cucuringu
3. (2023-2025) Postdoc at CISPA Helmholtz Center for Information Security, Germany
   - Supervisor: Krikamol Muandet
4. (2025-onwards) Assistant Professor of Statistical Machine Learning at NTU Singapore

Epistemic Intelligence & Computation Lab

@College of Computing & Data Science, NTU Singapore

Siu Lun Chau

PI

Kaizheng Wang

Postdoc Fellow

Clayton Chong

PhD Student

Yuqi Zhang

PhD student (co-supervised with Edwin Fong@HKU)

Xiaotao Liu

Research Assistant

Epistemic Intelligence & Computation Lab

@College of Computing & Data Science, NTU Singapore

Siu Lun Chau

PI

Kaizheng Wang

Postdoc Fellow

Clayton Chong

PhD Student

Yuqi Zhang

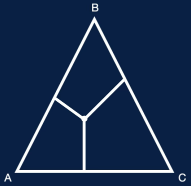
PhD student
(co-supervised with
Edwin Fong@HKU)

Xiaotao Liu

Research
Assistant

Mathematical Foundations

How should uncertainty, ambiguity, imprecision, be represented mathematically?



Epistemic Intelligence & Computation Lab

@College of Computing & Data Science, NTU Singapore

Siu Lun Chau

PI

Kaizheng Wang

Postdoc Fellow

Clayton Chong

PhD Student

Yuqi Zhang

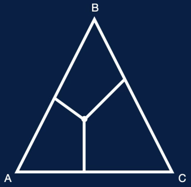
PhD student (co-supervised with Edwin Fong@HKU)

Xiaotao Liu

Research Assistant

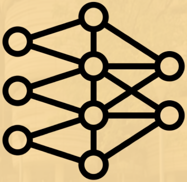
Mathematical Foundations

How should uncertainty, ambiguity, imprecision, be represented mathematically?



Algorithmic Development

How to build learning machines that are aware of their uncertainty?



Epistemic Intelligence & Computation Lab

@College of Computing & Data Science, NTU Singapore

Siu Lun Chau

PI

Kaizheng Wang

Postdoc Fellow

Clayton Chong

PhD Student

Yuqi Zhang

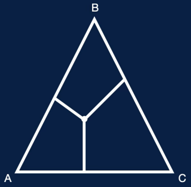
PhD student (co-supervised with Edwin Fong@HKU)

Xiaotao Liu

Research Assistant

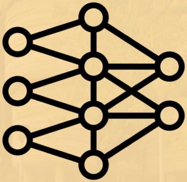
Mathematical Foundations

How should uncertainty, ambiguity, imprecision, be represented mathematically?



Algorithmic Development

How to build learning machines that are aware of their uncertainty?



Translational Research

How to develop uncertainty-aware AI systems, regulations, and other applications?

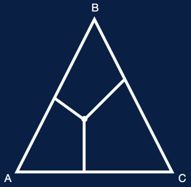
Epistemic Intelligence & Computation Lab

@College of Computing & Data Science, NTU Singapore

Siu Lun Chau	Kaizheng Wang	Clayton Chong	Yuqi Zhang	Xiaotao Liu
				
PI	Postdoc Fellow	PhD Student	PhD student (co-supervised with Edwin Fong@HKU)	Research Assistant

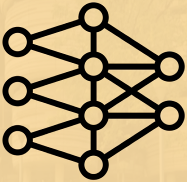
Mathematical Foundations

How should uncertainty, ambiguity, imprecision, be represented mathematically?



Algorithmic Development

How to build learning machines that are aware of their uncertainty?



Translational Research

How to develop uncertainty-aware AI systems, regulations, and other applications?

Check out epiclab-sg.github.io for more!

Today's talk

By the end of the talk, you will learn a little bit more about:

Today's talk

By the end of the talk, you will learn a little bit more about:

1. Generalisation of Classical Probability Theory, known as Imprecise Probabilities,

Today's talk

By the end of the talk, you will learn a little bit more about:

1. Generalisation of Classical Probability Theory, known as Imprecise Probabilities,
2. their applications to Statistical Machine Learning, and

Today's talk

By the end of the talk, you will learn a little bit more about:

1. Generalisation of Classical Probability Theory, known as Imprecise Probabilities,
2. their applications to Statistical Machine Learning, and
3. Something about imprecise probability metrics and conformal uncertainty quantification.

Table of Contents

Historical Motivation of Imprecise Probabilities

Flexible Uncertainty Representation

Historical Motivation of Imprecise Probabilities

Flexible Uncertainty Representation

- $\mathbb{P}_{\text{Alan}}(\{RainTomorrow\}) = 1$

Historical Motivation of Imprecise Probabilities

Flexible Uncertainty Representation

- $\mathbb{P}_{\text{Alan}}(\{\mathit{RainTomorrow}\}) = 1$
- $\mathbb{P}_{\text{Dino}}(\{\mathit{RainTomorrow}\}) > 0.5$

Historical Motivation of Imprecise Probabilities

Flexible Uncertainty Representation

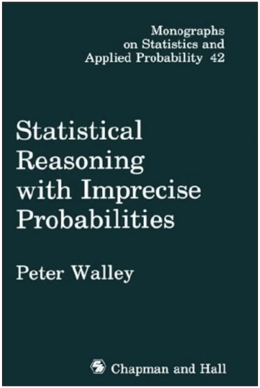
- $\mathbb{P}_{\text{Alan}}(\{\mathit{RainTomorrow}\}) = 1$
- $\mathbb{P}_{\text{Dino}}(\{\mathit{RainTomorrow}\}) > 0.5$
- $\mathbb{P}_{\text{Krik}}(\{\mathit{RainTomorrow}\}) \in [0, 1]$

Historical Motivation of Imprecise Probabilities

Flexible Uncertainty Representation

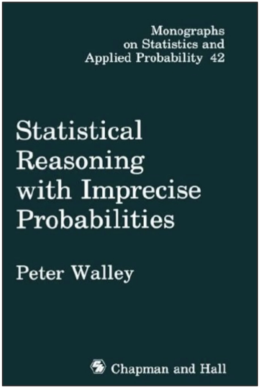
- $\mathbb{P}_{\text{Alan}}(\{\textit{RainTomorrow}\}) = 1$
- $\mathbb{P}_{\text{Dino}}(\{\textit{RainTomorrow}\}) > 0.5$
- $\mathbb{P}_{\text{Krik}}(\{\textit{RainTomorrow}\}) \in [0, 1]$

- This kicked off the studies of Imprecise Probabilities (IPs): Keynes [1921], Kyburg [1961], Levi [1980], Walley [1991], Troffaes and De Cooman [2014], Augustin et al. [2014a].



Walley [1991, Chapter 1.1.4]

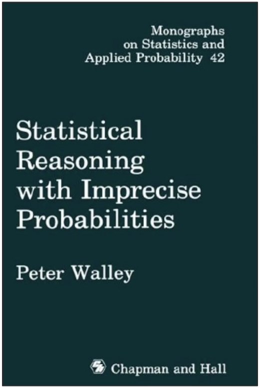
What are Imprecise Probabilities about?



Walley [1991, Chapter 1.1.4]

- Reasoning begins with the recognition and acknowledgement of uncertainty and ignorance.

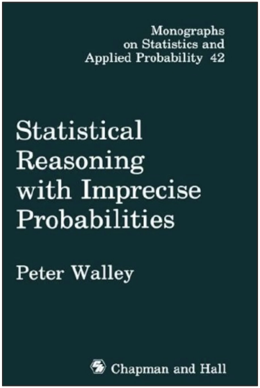
What are Imprecise Probabilities about?



Walley [1991, Chapter 1.1.4]

- Reasoning begins with the recognition and acknowledgement of uncertainty and ignorance.
- Uncertainty: How likely facts or events are to happen
  - $\approx$ Risk, Known Unknown, first-order uncertainty, Statistical variation, Aleatoric Uncertainty

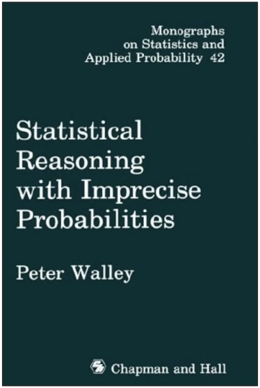
What are Imprecise Probabilities about?



Walley [1991, Chapter 1.1.4]

- Reasoning begins with the recognition and acknowledgement of uncertainty and ignorance.
- Uncertainty: How likely facts or events are to happen
  - ≈ Risk, Known Unknown, first-order uncertainty, Statistical variation, Aleatoric Uncertainty
  - Modelled by Probabilities

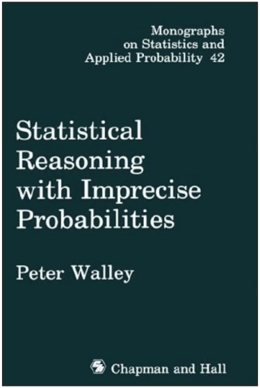
What are Imprecise Probabilities about?



Walley [1991, Chapter 1.1.4]

- Reasoning begins with the recognition and acknowledgement of uncertainty and ignorance.
- Uncertainty: How likely facts or events are to happen
  - $\approx$ Risk, Known Unknown, first-order uncertainty, Statistical variation, Aleatoric Uncertainty
  - Modelled by Probabilities
- Ignorance: How ignorant are we about facts or events
  - $\approx$ Indeterminacy, Unknown Unknown, higher-order uncertainty, Knightian uncertainty [Knight, 1921], Epistemic Uncertainty

What are Imprecise Probabilities about?



Walley [1991, Chapter 1.1.4]

- Reasoning begins with the recognition and acknowledgement of uncertainty and ignorance.
- Uncertainty: How likely facts or events are to happen
  - $\approx$ Risk, Known Unknown, first-order uncertainty, Statistical variation, Aleatoric Uncertainty
  - Modelled by Probabilities
- Ignorance: How ignorant are we about facts or events
  - $\approx$ Indeterminacy, Unknown Unknown, higher-order uncertainty, Knightian uncertainty [Knight, 1921], Epistemic Uncertainty
  - Modelled by Imprecise Probabilities

So What are Imprecise Probabilities?

So What are Imprecise Probabilities?

Imprecise Probabilities

Mathematical models that allow to conduct analyses in the presence of both (partial) ignorance and uncertainty.

So What are Imprecise Probabilities?

Imprecise Probabilities

Mathematical models that allow to conduct analyses in the presence of both (partial) ignorance and uncertainty.

Research questions that people study include

- Uncertainty representation and quantification
- Update of imprecise beliefs upon evidence
- Imprecise beliefs elicitation
- Rational decision-making under imprecision
- ...

So What are Imprecise Probabilities?

Imprecise Probabilities

Mathematical models that allow to conduct analyses in the presence of both (partial) ignorance and uncertainty.

Research questions that people study include

- Uncertainty representation and quantification
- Update of imprecise beliefs upon evidence
- Imprecise beliefs elicitation
- Rational decision-making under imprecision
- ...

IP has strong ties to Robust Frequentist Statistics [Huber, 1964] and Bayesian Sensitivity Analysis [Berger, 1984]!

Historical Motivations for Imprecise Probabilities

Other Approaches to Statistical Inference

- Fiducial Inference Fisher [1935], Hannig [2009]
- Dempster-Shafer Theory of Evidence Dempster [1967], Shafer [1976], Gong and Meng [2021]
- Probability Kinematics Jeffrey [1965], Levi [1967], Diaconis and Zabell [1982], Marchetti and Antonucci [2018]; Caprio and Gong [2023]
- Inferential models Martin [2019]

Historical Motivations for Imprecise Probabilities
An Elephant in the Room: What about Bayesianism?

Clarification:

- Bayesianism is (mostly) about coherent update of subject beliefs encoded as probability distributions (over parameters)
- Imprecise probabilities studies suitable representations of partial ignorance and related operations.
- No conflict!

Clarification:

- Bayesianism is (mostly) about coherent update of subject beliefs encoded as probability distributions (over parameters)
- Imprecise probabilities studies suitable representations of partial ignorance and related operations.
- No conflict!

See

- Walley [1991]'s Statistical Reasoning with Imprecise Probabilities
- Williamson [2010]'s "In Defence of Objective Bayesianism"
- Hájek [2019]'s "Interpretation of Probability"

- Economics and Decision Theory Amarante and Maccheroni [2006], Cerreia-Vioglio et al. [2023], Maccheroni et al. [2006], Gilboa et al. [2010]
- Game Theory Battigalli et al. [2015, 2019]
- Finance Vicig [2008], Cerreia-Vioglio et al. [2022]
- Mathematics De Cooman [1997], Coletti and Scozzafava [2002], Marinacci and Montrucchio [2004], Cerreia-Vioglio et al. [2015], Cuzzolin [2020], Maccheroni and Marinacci [2005]
- Physics Benavoli et al. [2021], De Vos et al. [2023]
- Engineering Ferson et al. [2003], Louis J. M. Aslett [2022], Vasile [2021]; [Augustin et al., 2014a, Chapter 13]
- Computer Science Halpern [2003], Casanova et al. [2022]
- Logic Saffiotti [1992], Gerla [1994]
- ...

Table of Contents

Acknowledging Uncertainties in Statistics and ML

Statistical Inference

Observing samples $X_1, \dots, X_n \overset{iid}{\sim} \mathbb{P}_{\theta,\mathcal{X}}$, how to test hypothesis about θ?

Acknowledging Uncertainties in Statistics and ML

Statistical Inference

Observing samples $X_1, \ldots, X_n \stackrel{iid}{\sim} \mathbb{P}_{\theta,\mathcal{X}}$, how to test hypothesis about θ?

Supervised Learning

Observing samples $\{X_i, Y_i\}_{i=1}^n \stackrel{iid}{\sim} \mathbb{P}_{\theta,\mathcal{X}\times\mathcal{Y}}$, how to learn an optimal mapping $f : \mathcal{X} \to \mathcal{Y}$?

Acknowledging Uncertainties in Statistics and ML

Statistical Inference

Observing samples $X_1, \dots, X_n \overset{iid}{\sim} \mathbb{P}_{\theta,\mathcal{X}}$, how to test hypothesis about θ?

Supervised Learning

Observing samples $\{X_i, Y_i\}_{i=1}^n \overset{iid}{\sim} \mathbb{P}_{\theta,\mathcal{X}\times\mathcal{Y}}$, how to learn an optimal mapping $f : \mathcal{X} \to \mathcal{Y}$?

Reasonable sources of imprecision and uncertainties

- What if data not iid? data contaminated?

Acknowledging Uncertainties in Statistics and ML

Statistical Inference

Observing samples $X_1, \ldots, X_n \overset{iid}{\sim} \mathbb{P}_{\theta,\mathcal{X}}$, how to test hypothesis about θ?

Supervised Learning

Observing samples $\{X_i, Y_i\}_{i=1}^n \overset{iid}{\sim} \mathbb{P}_{\theta,\mathcal{X}\times\mathcal{Y}}$, how to learn an optimal mapping $f : \mathcal{X} \to \mathcal{Y}$?

Reasonable sources of imprecision and uncertainties

- What if data not iid? data contaminated?
- What if a precise hypothesis cannot be formulated?

Acknowledging Uncertainties in Statistics and ML

Statistical Inference

Observing samples $X_1, \dots, X_n \overset{iid}{\sim} \mathbb{P}_{\theta,\mathcal{X}}$, how to test hypothesis about θ?

Supervised Learning

Observing samples $\{X_i, Y_i\}_{i=1}^n \overset{iid}{\sim} \mathbb{P}_{\theta,\mathcal{X}\times\mathcal{Y}}$, how to learn an optimal mapping $f : \mathcal{X} \to \mathcal{Y}$?

Reasonable sources of imprecision and uncertainties

- What if data not iid? data contaminated?
- What if a precise hypothesis cannot be formulated?
- What if distribution family of $\mathbb{P}_\theta$ unknown? Distribution shifts? Adversarial attacks? Performative effects?

Acknowledging Uncertainties in Statistics and ML

Statistical Inference

Observing samples $X_1, \ldots, X_n \overset{iid}{\sim} \mathbb{P}_{\theta,\mathcal{X}}$, how to test hypothesis about θ?

Supervised Learning

Observing samples $\{X_i, Y_i\}_{i=1}^n \overset{iid}{\sim} \mathbb{P}_{\theta,\mathcal{X}\times\mathcal{Y}}$, how to learn an optimal mapping $f : \mathcal{X} \to \mathcal{Y}$?

Reasonable sources of imprecision and uncertainties

- What if data not iid? data contaminated?
- What if a precise hypothesis cannot be formulated?
- What if distribution family of $\mathbb{P}_\theta$ unknown? Distribution shifts? Adversarial attacks? Performative effects?
- What if inductive biases (e.g. model class, causal assumptions, optimisation objectives) are no longer warranted?

Acknowledging Uncertainties in Statistics and ML

Statistical Inference

Observing samples $X_1, \ldots, X_n \overset{iid}{\sim} \mathbb{P}_{\theta,\mathcal{X}}$, how to test hypothesis about θ?

Supervised Learning

Observing samples $\{X_i, Y_i\}_{i=1}^n \overset{iid}{\sim} \mathbb{P}_{\theta,\mathcal{X}\times\mathcal{Y}}$, how to learn an optimal mapping $f : \mathcal{X} \to \mathcal{Y}$?

Reasonable sources of imprecision and uncertainties

- What if data not iid? data contaminated?
- What if a precise hypothesis cannot be formulated?
- What if distribution family of $\mathbb{P}_\theta$ unknown? Distribution shifts? Adversarial attacks? Performative effects?
- What if inductive biases (e.g. model class, causal assumptions, optimisation objectives) are no longer warranted?

Can we be faithful and precise about imprecision and still **do something useful**?

First Applications of IP to ML and AI

- Classification problems
  - Denoeux [2000]: Evidential Neural Networks (ENNs)
  - Zaffalon [2002]: Naive Credal Classifier (NCC)

First Applications of IP to ML and AI

- Classification problems
  - Denoeux [2000]: Evidential Neural Networks (ENNs)
  - Zaffalon [2002]: Naive Credal Classifier (NCC)
- Causal Inference
  - Cozman [2000]: Credal Networks (CNs)

First Applications of IP to ML and AI

- Classification problems
  - Denoeux [2000]: Evidential Neural Networks (ENNs)
  - Zaffalon [2002]: Naive Credal Classifier (NCC)
- Causal Inference
  - Cozman [2000]: Credal Networks (CNs)
- Vision
  - Cuzzolin and Frezza [1999]: Evidential Object Tracking (EOT)

First Applications of IP to ML and AI

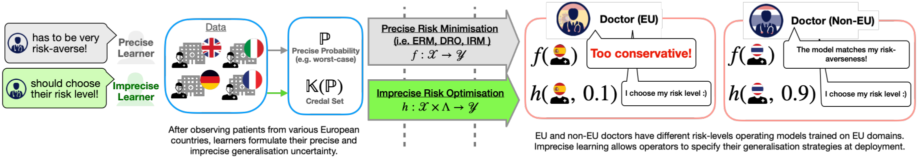
- Classification problems
  - Denoeux [2000]: Evidential Neural Networks (ENNs)
  - Zaffalon [2002]: Naive Credal Classifier (NCC)
- Causal Inference
  - Cozman [2000]: Credal Networks (CNs)
- Vision
  - Cuzzolin and Frezza [1999]: Evidential Object Tracking (EOT)
- Dynamical Systems
  - Hüllermeier [1997]: Qualitative and Model-Based Reasoning

First Applications of IP to ML and AI

- Classification problems
  - Denoeux [2000]: Evidential Neural Networks (ENNs)
  - Zaffalon [2002]: Naive Credal Classifier (NCC)
- Causal Inference
  - Cozman [2000]: Credal Networks (CNs)
- Vision
  - Cuzzolin and Frezza [1999]: Evidential Object Tracking (EOT)
- Dynamical Systems
  - Hüllermeier [1997]: Qualitative and Model-Based Reasoning
- Some more examples...

Domain Generalisation via Imprecise Learning

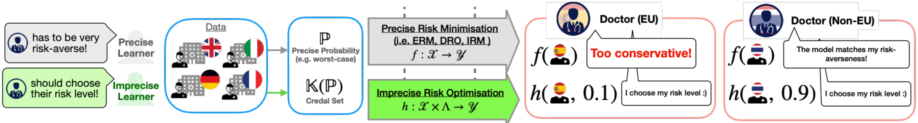
Singh, **Chau**, Bouabid, Muandet [ICML2024 Spotlight]



- **Research Question:** In domain generalisation, can distributional uncertainty be propagated to inference time, enabling downstream users — rather than model designers — to resolve residual ambiguity?

Domain Generalisation via Imprecise Learning

Singh, **Chau**, Bouabid, Muandet [ICML2024 Spotlight]



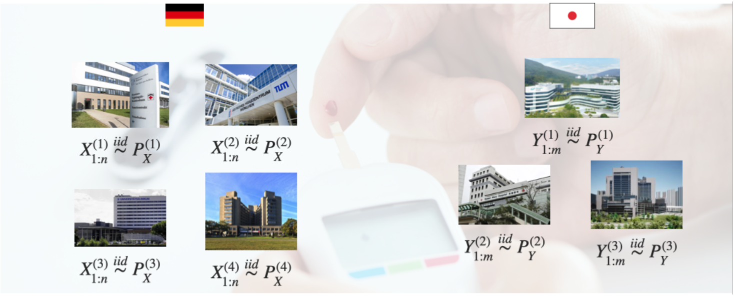
After observing patients from various European countries, learners formulate their precise and imprecise generalisation uncertainty.

EU and non-EU doctors have different risk-levels operating models trained on EU domains. Imprecise learning allows operators to specify their generalisation strategies at deployment.

- **Research Question:** In domain generalisation, can distributional uncertainty be propagated to inference time, enabling downstream users — rather than model designers — to resolve residual ambiguity?
- **Challenges:** How to represent distributional uncertainty? How to learn an infinite collection of models?

Credal Two-sample Tests of Epistemic Uncertainties

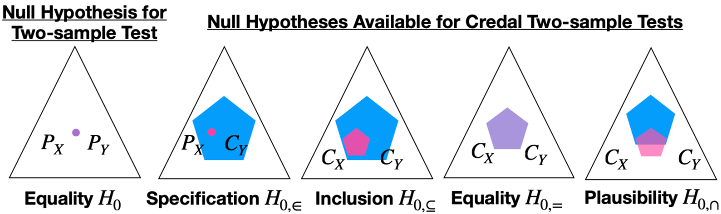
Chau, Schrab, Gretton, Sejdinovic, Muandet [AISTATS 2025]



- **Research Question:** Without direct access to the distribution of interest, and relying only on proxies, can we compare distributional uncertainties statistically?
- **Challenges:** You never have access from the true distribution of interests. A new kind of null "imprecise" hypotheses.

Credal Two-sample Tests of Epistemic Uncertainties

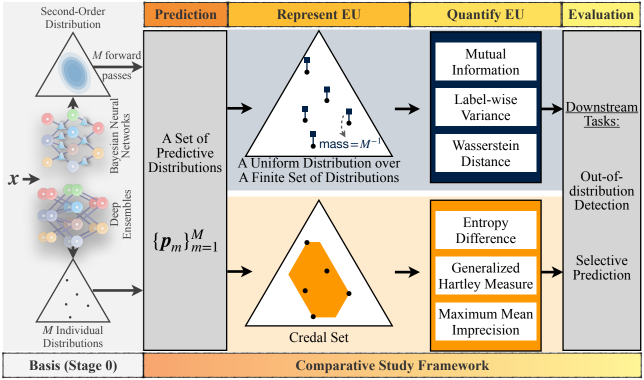
Chau, Schrab, Gretton, Sejdinovic, Muandet [AISTATS 2025]



- **Research Question:** Without direct access to the distribution of interest, and relying only on proxies, can we compare distributional uncertainties statistically?
- **Challenges:** You never have access from the true distribution of interests. A new kind of null "imprecise" hypotheses.

Set-based v.s. Distribution-based Representations of Epistemic Uncertainty: A Comparative Study

Wang, Wang, Cuzzolin, Moens, Hallez, **Chau** [In Submission 2026]



Research Question: If we remove the confounding effect of base learners, what can we say about set-based v.s. distribution-based representations of epistemic uncertainty through a comparative study?

Modern Research in IPML: Methodological

- Active Learning Oala et al. [2020], Nguyen et al. [2022]; Dutta et al. [2025]
- Reinforcement Learning Oren et al. [2022], Zanger et al. [2023], Kaufmann et al. [2023]
- Continual Learning Lu et al. [2025]
- Bayesian Learning Marquardt et al. [2023]; Caprio et al. [2024a]; Singh et al. [2025b]
- Evidential Learning Benavoli et al. [2009], Denœux [2023], Juergens et al. [2024]; Caprio et al. [2025b]
- Classification Caprio et al. [2024b]; Nguyen et al. [2025], Wang et al. [2025b]
- Conformal Prediction Stutz et al. [2023], Javanmardi et al. [2024], Huang et al. [2025]; Caprio et al. [2025d], Wu et al. [2025]

Modern Research in IPML: Methodological

- Causal Inference Maua et al. [2014], Zaffalon et al. [2020, 2023]
- Machine Vision Teeti et al. [2022], Giunchiglia et al. [2023]
- Medical Applications Bansback et al. [2016]
- Domain Generalization **Singh et al. [2024]**
- Credal Set Calibration Acharya et al. [2015], Gao et al. [2018], Mortier et al. [2023], Liu and Briol [2024] **Chau et al. [2024]**
- Proper scoring rule **Singh et al. [2025a]**
- Support Vector Machines Faccini et al. [2022], Maggioni and Spinelli [2024]
- Large Language Models (LLMs) Mubashar et al. [2025], Ji et al. [2025]
- ...

Modern Research in IPML: Theoretical

- Uncertainty Measures and Representation Abellán et al. [2006], Bengs et al. [2022], Hüllermeier and Waegeman [2021a], Wimmer et al. [2023], Destercke et al. [2008a], Troffaes and Destercke [2011]; Sale et al. [2023, 2024]; Bülte et al. [2025]
- Imprecise Probability Theory Campagner [2023], Derr and Williamson [2023], Fröhlich et al. [2023]; Caprio and Seidenfeld [2023]; Caprio and Gong [2023]; Caprio and Mukherjee [2023]; Caprio et al. [2024c]; sipta.org; IJAR
- Probability Metrics Hofman et al. [2024]; **Chau et al. [2025]**
- Fixed Point Theory **Caprio et al. [2025a]**

Table of Contents

This part of the presentation is based on the following paper:

Integral Imprecise Probability Metrics

Siu Lun Chau[1], **Michele Caprio**[2,3], and **Krikamol Muandet**[4]

[1]College of Computing & Data Science, Nanyang Technological University, Singapore
[2]Department of Computer Science, The University of Manchester, United Kingdom
[3]Manchester Centre for AI Fundamentals, Manchester, United Kingdom
[4]Rational Intelligence Lab, CISPA Helmholtz Center for Information Security, Germany

Abstract

Quantifying differences between probability distributions is fundamental to statistics and machine learning, primarily for comparing statistical uncertainty. In contrast, epistemic uncertainty—due to incomplete knowledge—requires richer representations than those offered by classical probability. Imprecise probability (IP) theory offers such models, capturing ambiguity and partial belief. This has

This part of the presentation is based on the following paper:

Integral Imprecise Probability Metrics

Siu Lun Chau[1], **Michele Caprio**[2,3], and **Krikamol Muandet**[4]

[1]College of Computing & Data Science, Nanyang Technological University, Singapore
[2]Department of Computer Science, The University of Manchester, United Kingdom
[3]Manchester Centre for AI Fundamentals, Manchester, United Kingdom
[4]Rational Intelligence Lab, CISPA Helmholtz Center for Information Security, Germany

Abstract

Quantifying differences between probability distributions is fundamental to statistics and machine learning, primarily for comparing statistical uncertainty. In contrast, epistemic uncertainty—due to incomplete knowledge—requires richer representations than those offered by classical probability. Imprecise probability (IP) theory offers such models, capturing ambiguity and partial belief. This has

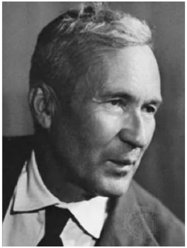
- Research Question: Can we devise an integral-based metric for imprecise probabilities? What can we get out of it?

How should we represent uncertainty?

Probability theory is the gold standard for modelling stochasticity in random phenomena.

How should we represent uncertainty?

Probability theory is the gold standard for modelling stochasticity in random phenomena.



Probability [Kolmogorov, 1933]

A probability measure $\mathbb{P} \in \mathcal{P}(\mathcal{X})$ on a measurable space $(\mathcal{X}, \Sigma_{\mathcal{X}})$ is a set function $\mathbb{P} : \Sigma_{\mathcal{X}} \to [0, 1]$ such that

1. $\mathbb{P}(A) \geq 0,$ for all event $A \in \Sigma_{\mathcal{X}}$
2. $\mathbb{P}(\mathcal{X}) = 1,$
3. for any sequence of disjoint sets $\{A_i\}_{i\geq 1}$ $\mathbb{P}(\bigcup_{i\geq 1} A_i) = \sum_{i\geq 1} \mathbb{P}(A_i)$

How should we represent uncertainty?

But there are limitations to standard probability when it comes to ignorance

1. Standard probability (e.g. real-valued random variables) lack formal framework to handle imprecise observations:

[1] therefore probability shouldn't be used as confidence measure but people use it regardlessly.

How should we represent uncertainty?

But there are limitations to standard probability when it comes to ignorance

1. Standard probability (e.g. real-valued random variables) lack formal framework to handle imprecise observations:
   - Set-valued observations (random set, ontic v.s. ontological), missing observations, qualitative data, probability bounds

[1] therefore probability shouldn't be used as confidence measure but people use it regardlessly.

How should we represent uncertainty?

But there are limitations to standard probability when it comes to ignorance

1. Standard probability (e.g. real-valued random variables) lack formal framework to handle imprecise observations:
   - Set-valued observations (random set, ontic v.s. ontological), missing observations, qualitative data, probability bounds
2. Probability fail to distinguish between indifference and ignorance.

[1] therefore probability shouldn't be used as confidence measure but people use it regardlessly.

How should we represent uncertainty?

But there are limitations to standard probability when it comes to ignorance

1. Standard probability (e.g. real-valued random variables) lack formal framework to handle imprecise observations:
   - Set-valued observations (random set, ontic v.s. ontological), missing observations, qualitative data, probability bounds
2. Probability fail to distinguish between indifference and ignorance.
   - $\mathbb{P}(\{\text{Coin flip} = \text{head}\}) = 0.5$. Precise equal likelihood belief or pick 0.5 because you don't know anything?

[1] therefore probability shouldn't be used as confidence measure but people use it regardlessly.

How should we represent uncertainty?

But there are limitations to standard probability when it comes to ignorance

1. Standard probability (e.g. real-valued random variables) lack formal framework to handle imprecise observations:
   - Set-valued observations (random set, ontic v.s. ontological), missing observations, qualitative data, probability bounds
2. Probability fail to distinguish between indifference and ignorance.
   - $\mathbb{P}(\{\text{Coin flip} = \text{head}\}) = 0.5$. Precise equal likelihood belief or pick 0.5 because you don't know anything?
3. Probability cannot encode absolute strength of belief.
   - If I do not know anything about event A, I want to encode $\mathbb{P}(A) = 0$, but that implies I know everything about A^c since now $\mathbb{P}(A^c) = 1$.[1]

[1] therefore probability shouldn't be used as confidence measure but people use it regardlessly.

How should we represent uncertainty?

But there are limitations to standard probability when it comes to ignorance

1. Standard probability (e.g. real-valued random variables) lack formal framework to handle imprecise observations:
   - Set-valued observations (random set, ontic v.s. ontological), missing observations, qualitative data, probability bounds
2. Probability fail to distinguish between indifference and ignorance.
   - $\mathbb{P}(\{\text{Coin flip} = \text{head}\}) = 0.5$. Precise equal likelihood belief or pick 0.5 because you don't know anything?
3. Probability cannot encode absolute strength of belief.
   - If I do not know anything about event A, I want to encode $\mathbb{P}(A) = 0$, but that implies I know everything about A^c since now $\mathbb{P}(A^c) = 1$.[1]
4. Probability cannot represent ignorance anyway...
   - Uniform distribution is not closed under reparametrisation.

[1] therefore probability shouldn't be used as confidence measure but people use it regardlessly.

How should we represent ignorance about uncertainty?

Candidate models from IP and their recent applications

Some candidates to model and represent degree of belief of truthfulness to events:

How should we represent ignorance about uncertainty?

Candidate models from IP and their recent applications

Some candidates to model and represent degree of belief of truthfulness to events:

1. Lower and upper probabilities [Wang et al., 2024, 2025a]

$$\underline{\mathbb{P}}(A) = \inf_{P \in \mathcal{C}} P(A) \qquad \overline{\mathbb{P}}(A) = \sup_{P \in \mathcal{C}} P(A)$$

How should we represent ignorance about uncertainty?

Candidate models from IP and their recent applications

Some candidates to model and represent degree of belief of truthfulness to events:

1. Lower and upper probabilities [Wang et al., 2024, 2025a]

$$\underline{\mathbb{P}}(A) = \inf_{P \in \mathcal{C}} P(A) \qquad \overline{\mathbb{P}}(A) = \sup_{P \in \mathcal{C}} P(A)$$

2. Belief Functions/Random sets [Manchingal and Cuzzolin, 2022]

$$\mathrm{Bel}(A) = \sum_{B \subseteq A} m(A)$$

How should we represent ignorance about uncertainty?

Candidate models from IP and their recent applications

Some candidates to model and represent degree of belief of truthfulness to events:

1. Lower and upper probabilities [Wang et al., 2024, 2025a]

$$\underline{\mathbb{P}}(A) = \inf_{P \in \mathcal{C}} P(A) \qquad \overline{\mathbb{P}}(A) = \sup_{P \in \mathcal{C}} P(A)$$

2. Belief Functions/Random sets [Manchingal and Cuzzolin, 2022]

$$\mathrm{Bel}(A) = \sum_{B \subseteq A} m(A)$$

3. Interval Probability (construct plausible candidates) [Wang et al., 2025c]

$$\{P \in \mathcal{P}(\mathcal{X} : \ell_i \leq P(\{x_i\}) \leq u_i : \sum_i P(\{x_i\}) = 1\}$$

How should we represent ignorance about uncertainty?

Candidate models from IP and their recent applications

Some candidates to model and represent degree of belief of truthfulness to events:

1. Lower and upper probabilities [Wang et al., 2024, 2025a]

$$\underline{\mathbb{P}}(A) = \inf_{P \in \mathcal{C}} P(A) \qquad \overline{\mathbb{P}}(A) = \sup_{P \in \mathcal{C}} P(A)$$

2. Belief Functions/Random sets [Manchingal and Cuzzolin, 2022]

$$\mathrm{Bel}(A) = \sum_{B \subseteq A} m(A)$$

3. Interval Probability (construct plausible candidates) [Wang et al., 2025c]

$$\{P \in \mathcal{P}(\mathcal{X} : \ell_i \leq P(\{x_i\}) \leq u_i : \sum_i P(\{x_i\}) = 1\}$$

4. Possibilistic measure [Hieu et al., 2025]

$$\overline{\mathbb{P}}(A) = \sup_{x \in A} \pi(x)$$

How should we represent ignorance about uncertainty?

Underlying Principle of IPML

Work with probabilistic information that are compatible to current evidence.

How should we represent ignorance about uncertainty?

Underlying Principle of IPML

Work with probabilistic information that are compatible to current evidence.

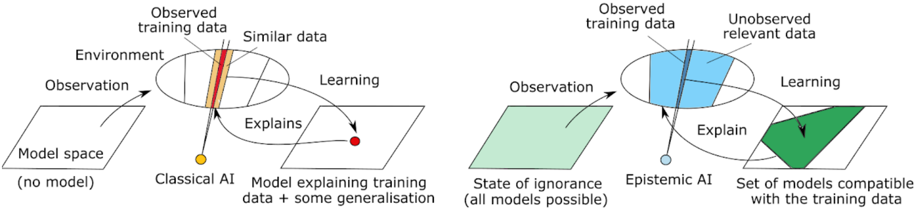
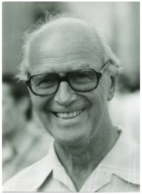


Figure: An illustrative picture from the Epistemic AI project; funded by the European Union's Horizon 2020 research and innovation programme under grant agreement No. 964505 (E-pi).

How should we represent ignorance about uncertainty?

Generalisations of probability



Capacities [Choquet, 1954]

A capacity $\nu \in \mathcal{V}(\mathcal{X})$ on a measurable space $(\mathcal{X}, \Sigma_{\mathcal{X}})$ is a set function $\nu : \Sigma_{\mathcal{X}} \to [0, 1]$ such that

- $\nu(\emptyset) = 0, \quad \nu(\mathcal{X}) = 1$
- for any $A, B \in \Sigma_{\mathcal{X}}$, $A \subseteq B \implies \nu(A) \leq \nu(B)$.

- non-additive representation of credence!

How should we represent ignorance about uncertainty?

Capacities generalises many IP and P models

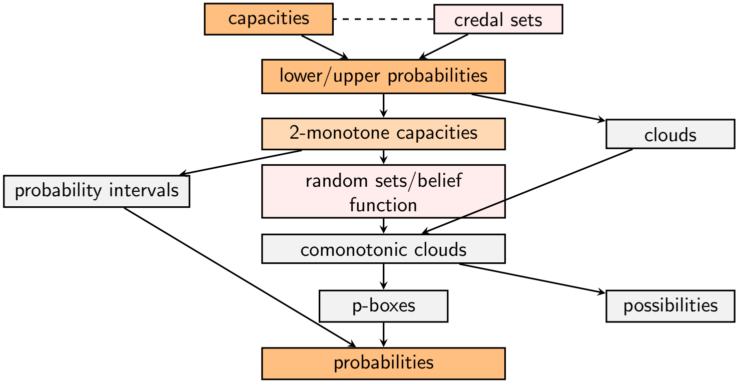


Figure: A skeleton demonstrating the connection between various uncertainty calculi. "A $\rightarrow$ B" means A generalises B, meaning that B is a specific instance of A. The figure is adopted from Destercke et al. [2008b] and Hüllermeier and Waegeman [2021b]. Most of these frameworks generalise classical probability theory.

Some more properties of capacities

- $\nu(A) \leftarrow$ credence to the truthfulness of event A.
- $\nu(A^c) \leftarrow$ credence to the truthfulness of complementary event A^c

Some more properties of capacities

- $\nu(A) \leftarrow$ credence to the truthfulness of event A.
- $\nu(A^c) \leftarrow$ credence to the truthfulness of complementary event A^c
- In general, $\nu(A) \neq 1 - \nu(A^c)$!

Some more properties of capacities

- $\nu(A) \leftarrow$ credence to the truthfulness of event A.
- $\nu(A^c) \leftarrow$ credence to the truthfulness of complementary event A^c
- In general, $\nu(A) \neq 1 - \nu(A^c)$!
- Unlike precise probability.

Some more properties of capacities

- $\nu(A) \leftarrow$ credence to the truthfulness of event A.
- $\nu(A^c) \leftarrow$ credence to the truthfulness of complementary event A^c
- In general, $\nu(A) \neq 1 - \nu(A^c)$!
- Unlike precise probability.

Conjugate capacity

Given a capacity ν, the conjugate capacity $\nu^\star : \Sigma_{\mathcal{X}} \to [0, 1]$ is given by

$$\nu^\star(A) = 1 - \nu(A^c)$$

Some more properties of capacities

- $\nu(A) \leftarrow$ credence to the truthfulness of event A.
- $\nu(A^c) \leftarrow$ credence to the truthfulness of complementary event A^c
- In general, $\nu(A) \neq 1 - \nu(A^c)$!
- Unlike precise probability.

Conjugate capacity

Given a capacity ν, the conjugate capacity $\nu^\star : \Sigma_{\mathcal{X}} \to [0, 1]$ is given by

$$\nu^\star(A) = 1 - \nu(A^c)$$

- Credence for A is now given by the none-truthfulness of event A^c.

Some more properties of capacities

- $\nu(A) \leftarrow$ credence to the truthfulness of event A.
- $\nu(A^c) \leftarrow$ credence to the truthfulness of complementary event A^c
- In general, $\nu(A) \neq 1 - \nu(A^c)$!
- Unlike precise probability.

Conjugate capacity

Given a capacity ν, the conjugate capacity $\nu^\star : \Sigma_{\mathcal{X}} \to [0, 1]$ is given by

$$\nu^\star(A) = 1 - \nu(A^c)$$

- Credence for A is now given by the none-truthfulness of event A^c.
- Example: $\sup_{\mathbb{P} \in C} P(A) = \overline{P}(A) = 1 - \underline{P}(A^c) = 1 - \inf_{P \in C} P(A^c)$

Integral-based Metric for Imprecise Probabilities

How is it done for precise probability first?

Integral Probability Metrics [Müller, 1997]

Given a set of continuous bounded real-valued measurable functions $\mathcal{F} \subseteq C_b(\mathcal{X})$ and probability measures $\mathbb{P}, \mathbb{Q} \in \mathcal{P}(\mathcal{X})$, the IPM associated to $\mathcal{F}$ is

$$\text{IPM}_{\mathcal{F}}(\mathbb{P}, \mathbb{Q}) = \sup_{f \in \mathcal{F}} \left| \int f d\mathbb{P} - \int f d\mathbb{Q} \right|$$

Integral-based Metric for Imprecise Probabilities

How is it done for precise probability first?

Integral Probability Metrics [Müller, 1997]

Given a set of continuous bounded real-valued measurable functions $\mathcal{F} \subseteq C_b(\mathcal{X})$ and probability measures $\mathbb{P}, \mathbb{Q} \in \mathcal{P}(\mathcal{X})$, the IPM associated to $\mathcal{F}$ is

$$\mathrm{IPM}_{\mathcal{F}}(\mathbb{P}, \mathbb{Q}) = \sup_{f \in \mathcal{F}} \left| \int f d\mathbb{P} - \int f d\mathbb{Q} \right|$$

- Popular examples include the total variation distance, MMD, and the Wasserstein distance.

Integral-based Metric for Imprecise Probabilities

How is it done for precise probability first?

Integral Probability Metrics [Müller, 1997]

Given a set of continuous bounded real-valued measurable functions $\mathcal{F} \subseteq C_b(\mathcal{X})$ and probability measures $\mathbb{P}, \mathbb{Q} \in \mathcal{P}(\mathcal{X})$, the IPM associated to $\mathcal{F}$ is

$$\mathrm{IPM}_{\mathcal{F}}(\mathbb{P}, \mathbb{Q}) = \sup_{f \in \mathcal{F}} \left| \int f d\mathbb{P} - \int f d\mathbb{Q} \right|$$

- Popular examples include the total variation distance, MMD, and the Wasserstein distance.
- Can we extend to capacities trivially?

Integral-based Metric for Imprecise Probabilities

How is it done for precise probability first?

Integral Probability Metrics [Müller, 1997]

Given a set of continuous bounded real-valued measurable functions $\mathcal{F} \subseteq C_b(\mathcal{X})$ and probability measures $\mathbb{P}, \mathbb{Q} \in \mathcal{P}(\mathcal{X})$, the IPM associated to $\mathcal{F}$ is

$$\mathrm{IPM}_{\mathcal{F}}(\mathbb{P}, \mathbb{Q}) = \sup_{f \in \mathcal{F}} \left| \int f d\mathbb{P} - \int f d\mathbb{Q} \right|$$

- Popular examples include the total variation distance, MMD, and the Wasserstein distance.
- Can we extend to capacities trivially? No! Lebesgue integral is not designed for non-additive "measures"!

Choquet Integrals [Choquet, 1954]

Given a capacity $\nu \in \mathcal{V}(\mathcal{X})$ and a real-valued function f on $\mathcal{X}$, the Choquet integral of f with respect to ν is

$$\oint f d\nu = \int_0^\infty \nu(\{x : \max(0, f(x)) \geq t\})dt - \int_0^\infty \nu(\{x : -\min(0, f(x)) \geq t\})dt$$

provided the difference is well defined. For bounded f, we have

$$\oint f d\nu = \underline{f} + \int_{\underline{f}}^{\overline{f}} \nu(\{x : f(x) \geq t\})dt \tag{1}$$

where $\underline{f} = \inf_x f(x)$ and $\overline{f} = \sup_x f(x)$.

IIPM [Chau et al., 2025]

For function class $\mathcal{F} \subseteq C_b(\mathcal{X})$ and capacities $\nu, \mu \in \mathcal{V}(\mathcal{X})$, the integral imprecise probability metric associated with $\mathcal{F}$ between ν, μ is

$$\text{IIPM}_{\mathcal{F}}(\nu, \mu) = \sup_{f \in \mathcal{F}} \left| ∯ f d\nu - ∯ f d\mu \right|$$

Integral-based Metric for Imprecise Probabilities

Integral Imprecise Probability Metric (IIPM)

IIPM [Chau et al., 2025]

For function class $\mathcal{F} \subseteq C_b(\mathcal{X})$ and capacities $\nu, \mu \in \mathcal{V}(\mathcal{X})$, the integral imprecise probability metric associated with $\mathcal{F}$ between ν, μ is

$$\mathrm{IIPM}_{\mathcal{F}}(\nu, \mu) = \sup_{f \in \mathcal{F}} \left| \oint f d\nu - \oint f d\mu \right|$$

Properties:

- IIPM is a pseudo-metric in general and a metric, e.g., when $\mathcal{F}$ is dense in $C_b(\mathcal{X})$.

IIPM [Chau et al., 2025]

For function class $\mathcal{F} \subseteq C_b(\mathcal{X})$ and capacities $\nu, \mu \in \mathcal{V}(\mathcal{X})$, the integral imprecise probability metric associated with $\mathcal{F}$ between ν, μ is

$$\mathrm{IIPM}_{\mathcal{F}}(\nu, \mu) = \sup_{f \in \mathcal{F}} \left| ⨏ f d\nu - ⨏ f d\mu \right|$$

Properties:

- IIPM is a pseudo-metric in general and a metric, e.g., when $\mathcal{F}$ is dense in $C_b(\mathcal{X})$.
- IIPM recovers IPM when ν, μ is a precise probability.

IIPM [Chau et al., 2025]

For function class $\mathcal{F} \subseteq C_b(\mathcal{X})$ and capacities $\nu, \mu \in \mathcal{V}(\mathcal{X})$, the integral imprecise probability metric associated with $\mathcal{F}$ between ν, μ is

$$\text{IIPM}_{\mathcal{F}}(\nu, \mu) = \sup_{f \in \mathcal{F}} \left| \oint f d\nu - \oint f d\mu \right|$$

Properties:

- IIPM is a pseudo-metric in general and a metric, e.g., when $\mathcal{F}$ is dense in $C_b(\mathcal{X})$.
- IIPM recovers IPM when ν, μ is a precise probability.
- For $\mathcal{F} \subseteq C_b(\mathcal{X})$ that are dense, $\text{IIPM}_{\mathcal{F}}$ metrises the (Choquet) weak convergence of $\mathcal{V}(\mathcal{X})$.

Example use cases of the IIPM framework

- Derive new distances for imprecise probabilities: e.g. lower Dudley metric and lower total variation.
- Recover results in optimal transport with ϵ contamination sets
- Derive a kernel MMD for comparing ϵ contamination sets.
- more to come...

Quantifying Epistemic Uncertainty with IIPM

The Maximum Mean Imprecision (MMI)

The IIPM framework extends beyond its theoretical appeal!

The IIPM framework extends beyond its theoretical appeal!
We can now measure discrepancies between conjugate representations!

The IIPM framework extends beyond its theoretical appeal!
We can now measure discrepancies between conjugate representations!

The Maximum Mean Imprecision (MMI)

Let $\mathcal{F} \subseteq C_b(\mathcal{X})$. The **Maximum Mean Imprecision** with respect to $\mathcal{F}$ is

$$\mathrm{MMI}_{\mathcal{F}}(\underline{\mathbb{P}}) = \mathrm{IIPM}_{\mathcal{F}}(\underline{\mathbb{P}}, \overline{\mathbb{P}}) = \sup_{f \in \mathcal{F}} \left\{ \oint f d\overline{\mathbb{P}} - \oint f d\underline{\mathbb{P}} \right\}$$

The IIPM framework extends beyond its theoretical appeal!
We can now measure discrepancies between conjugate representations!

The Maximum Mean Imprecision (MMI)

Let $\mathcal{F} \subseteq C_b(\mathcal{X})$. The **Maximum Mean Imprecision** with respect to $\mathcal{F}$ is

$$\mathrm{MMI}_{\mathcal{F}}(\underline{\mathbb{P}}) = \mathrm{IIPM}_{\mathcal{F}}(\underline{\mathbb{P}}, \overline{\mathbb{P}}) = \sup_{f \in \mathcal{F}} \left\{ \oint f d\overline{\mathbb{P}} - \oint f d\underline{\mathbb{P}} \right\}$$

Is this a sensible epistemic uncertainty quantification measure?

Quantifying Epistemic Uncertainty with IIPM

Desirable Properties for epistemic UQ "measures"

Let $\mathbb{U} : \underline{\mathbb{P}}(\mathcal{X}) \to \mathbb{R}$ be a credal uncertainty measure. Here are some commonly accepted desirable properties:

1. Non-negative and boundedness

Let $\mathbb{U} : \underline{\mathbb{P}}(\mathcal{X}) \to \mathbb{R}$ be a credal uncertainty measure. Here are some commonly accepted desirable properties:

1. Non-negative and boundedness
2. Continuity: $\mathbb{U}$ is a continuos functional

Let $\mathbb{U} : \underline{\mathbb{P}}(\mathcal{X}) \to \mathbb{R}$ be a credal uncertainty measure. Here are some commonly accepted desirable properties:

1. Non-negative and boundedness
2. Continuity: $\mathbb{U}$ is a continuos functional
3. Monotonicity: for $\underline{\mathbb{P}}, \underline{\mathbb{Q}}$, if $\underline{\mathbb{P}}(A) \leq \underline{\mathbb{Q}}(A)$ for all $A \in \Sigma_{\mathcal{X}}$, then $\mathbb{U}(\underline{\mathbb{P}}) \geq \mathbb{U}(\underline{\mathbb{Q}})$.

Quantifying Epistemic Uncertainty with IIPM

Desirable Properties for epistemic UQ "measures"

Let $\mathbb{U} : \underline{\mathbb{P}}(\mathcal{X}) \to \mathbb{R}$ be a credal uncertainty measure. Here are some commonly accepted desirable properties:

1. Non-negative and boundedness
2. Continuity: $\mathbb{U}$ is a continuos functional
3. Monotonicity: for $\underline{\mathbb{P}}, \underline{\mathbb{Q}}$, if $\underline{\mathbb{P}}(A) \leq \underline{\mathbb{Q}}(A)$ for all $A \in \Sigma_{\mathcal{X}}$, then $\mathbb{U}(\underline{\mathbb{P}}) \geq \mathbb{U}(\underline{\mathbb{Q}})$.
4. Probability consistency: for all precise $\mathbb{P} \in \mathcal{P}(\mathcal{X})$ we have $\mathbb{U}(\mathbb{P}) = 0$.

Let $\mathbb{U} : \underline{\mathbb{P}}(\mathcal{X}) \to \mathbb{R}$ be a credal uncertainty measure. Here are some commonly accepted desirable properties:

1. Non-negative and boundedness
2. Continuity: $\mathbb{U}$ is a continuos functional
3. Monotonicity: for $\underline{\mathbb{P}}, \underline{\mathbb{Q}}$, if $\underline{\mathbb{P}}(A) \leq \underline{\mathbb{Q}}(A)$ for all $A \in \Sigma_{\mathcal{X}}$, then $\mathbb{U}(\underline{\mathbb{P}}) \geq \mathbb{U}(\underline{\mathbb{Q}})$.
4. Probability consistency: for all precise $\mathbb{P} \in \mathcal{P}(\mathcal{X})$ we have $\mathbb{U}(\mathbb{P}) = 0$.
5. Sub-additivity (not relevant for predictive uncertainty)
6. Additivity (not relevant for predictive uncertainty)

Quantifying Epistemic Uncertainty with IIPM

Desirable Properties for epistemic UQ "measures"

Let $\mathbb{U} : \underline{\mathbb{P}}(\mathcal{X}) \to \mathbb{R}$ be a credal uncertainty measure. Here are some commonly accepted desirable properties:

1. Non-negative and boundedness
2. Continuity: $\mathbb{U}$ is a continuos functional
3. Monotonicity: for $\underline{\mathbb{P}}, \underline{\mathbb{Q}}$, if $\underline{\mathbb{P}}(A) \leq \underline{\mathbb{Q}}(A)$ for all $A \in \Sigma_{\mathcal{X}}$, then $\mathbb{U}(\underline{\mathbb{P}}) \geq \mathbb{U}(\underline{\mathbb{Q}})$.
4. Probability consistency: for all precise $\mathbb{P} \in \mathcal{P}(\mathcal{X})$ we have $\mathbb{U}(\mathbb{P}) = 0$.
5. Sub-additivity (not relevant for predictive uncertainty)
6. Additivity (not relevant for predictive uncertainty)

MMI satisfy top 4 for any $\mathcal{F}$, and the rest when $\underline{P}$ satisfies 2-monotonicity.

Maximum Mean Imprecision in Action

A simple instantiation for classifiers

Epistemic uncertainty-aware models either model second-order uncertainty via

Maximum Mean Imprecision in Action

A simple instantiation for classifiers

Epistemic uncertainty-aware models either model second-order uncertainty via

1. distributions (posterior parameter distribution in Bayesian settings, or conjugate Dirichlet prior in evidential neural networks)

Maximum Mean Imprecision in Action

A simple instantiation for classifiers

Epistemic uncertainty-aware models either model second-order uncertainty via

1. distributions (posterior parameter distribution in Bayesian settings, or conjugate Dirichlet prior in evidential neural networks)
2. sets (e.g. probability bounds, intervals, convex hull)

Maximum Mean Imprecision in Action

A simple instantiation for classifiers

Epistemic uncertainty-aware models either model second-order uncertainty via

1. distributions (posterior parameter distribution in Bayesian settings, or conjugate Dirichlet prior in evidential neural networks)
2. sets (e.g. probability bounds, intervals, convex hull)
    - ensemble based approaches [Wang et al., 2024, 2025a,b]
    - relative likelihood [Löhr et al., 2025]
    - random set theory [Manchingal et al., 2025, Mubashar et al., 2025]
    - probability bounds and outer measure [Hüllermeier, 2003, Hao et al., 2022]

Maximum Mean Imprecision in Action

A simple instantiation for classifiers

Epistemic uncertainty-aware models either model second-order uncertainty via

1. distributions (posterior parameter distribution in Bayesian settings, or conjugate Dirichlet prior in evidential neural networks)
2. sets (e.g. probability bounds, intervals, convex hull)
   - ensemble based approaches [Wang et al., 2024, 2025a,b]
   - relative likelihood [Löhr et al., 2025]
   - random set theory [Manchingal et al., 2025, Mubashar et al., 2025]
   - probability bounds and outer measure [Hüllermeier, 2003, Hao et al., 2022]

and will obtain a lower probability $\underline{\mathbb{P}}$ as the second-order uncertainty representation.

Maximum Mean Imprecision in Action

Analytical form

For classification problems $(\mathcal{Y}, 2^{\mathcal{Y}})$, pick $\mathcal{H}_{TV} = \{\mathbf{1}_A : A \in 2^{\mathcal{Y}}\}$, leading to

$$\mathrm{MMI}_{\mathcal{H}_{TV}}(\underline{\mathbb{P}}) = \sup_{A \subseteq \mathcal{Y}} |\overline{\mathbb{P}}(A) - \underline{\mathbb{P}}(A)|$$

Maximum Mean Imprecision in Action

Analytical form

For classification problems $(\mathcal{Y}, 2^{\mathcal{Y}})$, pick $\mathcal{H}_{TV} = \{\mathbf{1}_A : A \in 2^{\mathcal{Y}}\}$, leading to

$$\mathrm{MMI}_{\mathcal{H}_{TV}}(\underline{\mathbb{P}}) = \sup_{A \subseteq \mathcal{Y}} |\overline{\mathbb{P}}(A) - \underline{\mathbb{P}}(A)|$$

Linear-time computable upper bound

$$\mathrm{MMI}_{\mathcal{H}_{TV}}(\underline{\mathbb{P}}) \leq 1 - \sum_{y \in \mathcal{Y}} \underline{\mathbb{P}}(\{y\})$$

Maximum Mean Imprecision in Action

Epistemic Uncertainty model evaluation

How should we evaluate the quantified EU empirically?

Maximum Mean Imprecision in Action

Epistemic Uncertainty model evaluation

How should we evaluate the quantified EU empirically?
This is not a supervised learning problem, no ground-truth epistemic uncertainty!

Maximum Mean Imprecision in Action

Epistemic Uncertainty model evaluation

How should we evaluate the quantified EU empirically?
This is not a supervised learning problem, no ground-truth epistemic uncertainty!

Standard practice: compare EU based on their informativeness for decision-making: e.g. active learning, Bayesian optimisation, learning-to-defer, OOD detection, selective classification.

Maximum Mean Imprecision in Action

Selective Classification

Consider this abstraction, for each $i \in 1, \ldots, n_{test}$,

$$x_i \mapsto (\ \underbrace{\hat{y}_i}_{\text{prediction}}\ ,\ \underbrace{\mathrm{EU}(x_i)}_{\text{predictive EU}}\ )$$

Maximum Mean Imprecision in Action

Selective Classification

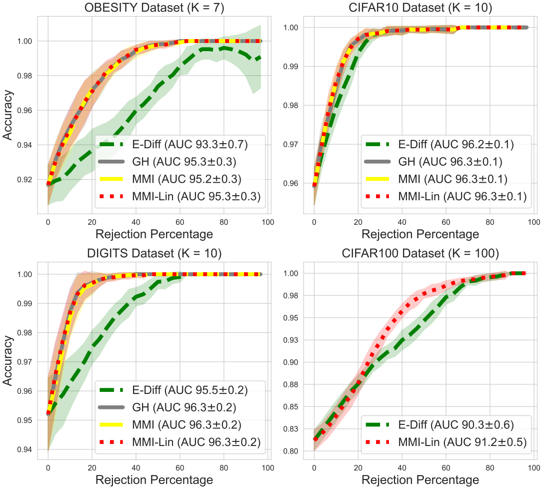
Consider this abstraction, for each $i \in 1, \ldots, n_{test}$,

$$x_i \mapsto \Big(\underbrace{\hat{y}_i}_{\text{prediction}} , \underbrace{\mathrm{EU}(x_i)}_{\text{predictive EU}} \Big)$$

1. Sort $x_1, \ldots, x_{n_{test}}$ in ascending order based on their estimated predictive confidence (inverse EU).
2. Remove the top $p\%$ of least confident inputs, predict the rest, record performance.
3. Plot performance against threshold $p\%$

Maximum Mean Imprecision in Action

Experiment Results



1. Imprecise probabilities offer flexible uncertainty representation frameworks through probability sets, intervals, and bounds, etc.

Integral Imprecise Probability Metrics

Takeaways

1. Imprecise probabilities offer flexible uncertainty representation frameworks through probability sets, intervals, and bounds, etc.
2. IIPM generalises IPM for IP.

Integral Imprecise Probability Metrics

Takeaways

1. Imprecise probabilities offer flexible uncertainty representation frameworks through probability sets, intervals, and bounds, etc.
2. IIPM generalises IPM for IP.
3. MMI measures the maximal imprecision in the Choquet expectation, allowing us to quantify "epistemic uncertainty" of IP-based predictors.

Table of Contents

This part of the presentation is based on the following paper:

Quantifying Epistemic Predictive Uncertainty in Conformal Prediction

Siu Lun Chau[1], **Soroush H. Zargarbashi**[2], **Yusuf Sale**[3,4], and **Michele Caprio**[5, 6]

[1]Epistemic Intelligence & Computation Lab, College of Computing & Data Science, Nanyang Technological University, Singapore
[2]CISPA Helmholtz Center for Information Security, Germany
[3]Munich Center for Machine Learning; Institute of Informatics, LMU Munich, Germany
[4]Munich Center for Machine Learning, Germany
[5]Manchester Centre for AI Fundamentals, Manchester, United Kingdom
[6]Department of Computer Science, The University of Manchester, United Kingdom

February 3, 2026

This part of the presentation is based on the following paper:

Quantifying Epistemic Predictive Uncertainty in Conformal Prediction

Siu Lun Chau[1], **Soroush H. Zargarbashi**[2], **Yusuf Sale**[3,4], and **Michele Caprio**[5, 6]

[1]Epistemic Intelligence & Computation Lab, College of Computing & Data Science, Nanyang Technological University, Singapore
[2]CISPA Helmholtz Center for Information Security, Germany
[3]Munich Center for Machine Learning; Institute of Informatics, LMU Munich, Germany
[4]Munich Center for Machine Learning, Germany
[5]Manchester Centre for AI Fundamentals, Manchester, United Kingdom
[6]Department of Computer Science, The University of Manchester, United Kingdom

February 3, 2026

- Research Question: Can we build the link from split Conformal Prediction to Imprecise Probabilities and quantify epistemic predictive uncertainty therein?

Split Conformal Prediction
A Brief Recap

Given calibration set D_{cal}, nonconformity score $s(x,y)$, error tolerance $\alpha \in [0,1]$:

$$\mathcal{C}_\alpha(x) := \{y \in \mathcal{Y} : s(x,y) \le \widehat{q}_{1-\alpha}\}.$$

Equivalent via conformal transducer / p-values:

$$\pi_x(y) = \frac{1 + |\{i \in D_{cal} : s_i \ge s(x,y)\}|}{1 + n_{\mathrm{cal}}} \quad \Rightarrow \quad \mathcal{C}_\alpha(x) = \{y : \pi_x(y) > \alpha\}.$$

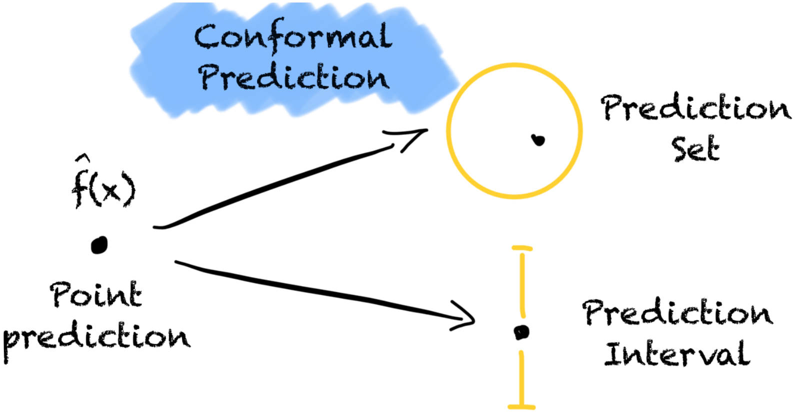
With marginal coverage

$$\mathbb{P}(Y_{n+1} \in \mathcal{C}_\alpha(X_{n+1})) \ge 1 - \alpha$$

- $\pi_x(y)$ can be interpreted as a p-value for exchangeability of $D_{\mathrm{cal}} \cup \{(x,y)\}$.
- We will use the *whole function* $\pi_x(\cdot)$ to expose model multiplicity.

Split Conformal Prediction

A brief recap



Split Conformal Prediction

How to perform Uncertainty Quantification for CP?

Set sizes!

Split Conformal Prediction

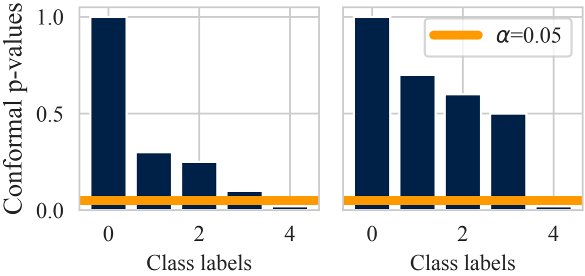
How to perform Uncertainty Quantification for CP?

Set sizes! but...

Split Conformal Prediction

How to perform Uncertainty Quantification for CP?

Set sizes! but...



Same set size, but very different confidence profile? What does the uncertainty really mean?

Split Conformal Prediction

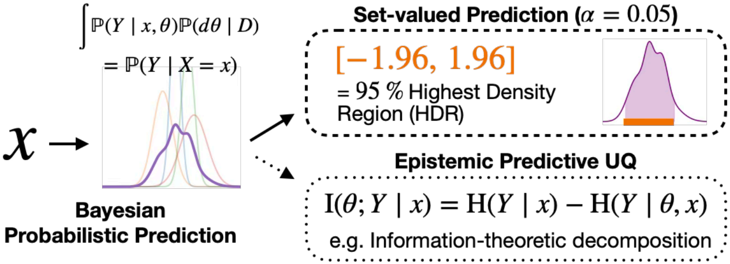
Predictive Uncertainty v.s. Uncertainty of the prediction.

Take Bayesian probabilistic Prediction as an example,

Split Conformal Prediction

Predictive Uncertainty v.s. Uncertainty of the prediction.

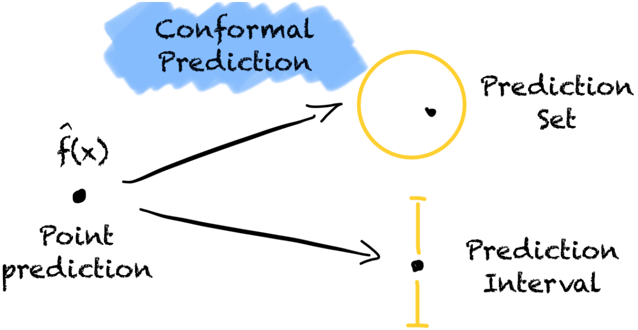
Take Bayesian probabilistic Prediction as an example,



Split Conformal Prediction

Epistemic Predictive Uncertainty v.s. Uncertainty of the prediction.

Set sizes only reflect uncertainty therein the (set-valued) prediction.



Epistemic Predictive Uncertainty

"The difficulty of making predictions about outcome Y given an observation $X = x$ due to model uncertainty, i.e., when multiple plausible predictive models exist."

Epistemic Predictive Uncertainty

"The difficulty of making predictions about outcome Y given an observation $X = x$ due to model uncertainty, i.e., when multiple plausible predictive models exist."

For instance....

- In Bayesian models, $\mathbb{P}(\theta \mid D)$ defines plausible models $\mathbb{P}(Y \mid X = x,\ \theta)$.

Epistemic Predictive Uncertainty

"The difficulty of making predictions about outcome Y given an observation $X = x$ due to model uncertainty, i.e., when multiple plausible predictive models exist."

For instance....

- In Bayesian models, $\mathbb{P}(\theta \mid D)$ defines plausible models $\mathbb{P}(Y \mid X = x,\ \theta)$.
- Mutual information then measures the expected reduction in predictive entropy under the posterior predictive relative to alternative plausible models!

Epistemic Predictive Uncertainty

"The difficulty of making predictions about outcome Y given an observation $X = x$ due to model uncertainty, i.e., when multiple plausible predictive models exist."

For instance....

- In Bayesian models, $\mathbb{P}(\theta \mid D)$ defines plausible models $\mathbb{P}(Y \mid X = x,\ \theta)$.
- Mutual information then measures the expected reduction in predictive entropy under the posterior predictive relative to alternative plausible models!
- Does something like this exist for Conformal prediction?

Recent results by Cella and Martin [2022], Caprio et al. [2025c], Caprio [2025] have shown that (full) CP and IP shares deep connection!

Conformal Prediction and Imprecise Probabilities

CP induces implicit credal set

Recent results by Cella and Martin [2022], Caprio et al. [2025c], Caprio [2025] have shown that (full) CP and IP shares deep connection!

High-level Intuition

- Under the consonance assumption ($\sup_{y \in \mathcal{Y}} \pi_x(y) = 1$), every full CP procedure induces a closed and convex set of predictive probability distributions (credal set). [Cella and Martin, 2022]

Recent results by Cella and Martin [2022], Caprio et al. [2025c], Caprio [2025] have shown that (full) CP and IP shares deep connection!

High-level Intuition

- Under the consonance assumption ($\sup_{y \in \mathcal{Y}} \pi_x(y) = 1$), every full CP procedure induces a closed and convex set of predictive probability distributions (credal set). [Cella and Martin, 2022]
- The pessimistic highest density region of these distributions gives exactly the conformal prediction set for full CP. [Caprio, 2025, Caprio et al., 2025c]

Conformal Prediction and Imprecise Probabilities

CP induces implicit credal set

Recent results by Cella and Martin [2022], Caprio et al. [2025c], Caprio [2025] have shown that (full) CP and IP shares deep connection!

High-level Intuition

- Under the consonance assumption ($\sup_{y \in \mathcal{Y}} \pi_x(y) = 1$), every full CP procedure induces a closed and convex set of predictive probability distributions (credal set). [Cella and Martin, 2022]
- The pessimistic highest density region of these distributions gives exactly the conformal prediction set for full CP. [Caprio, 2025, Caprio et al., 2025c]

- Let me show you an alternative way to perform CP!

An alternative way to perform Conformal Prediction

Step 1: the consonance condition (bridge CP $\leftrightarrow$ IP)

Consonance

A transducer $\pi_x(\cdot)$ is **consonant** if $\sup_{y \in \mathcal{Y}} \pi_x(y) = 1$.

- Holds naturally in regression for many scores (e.g. absolute residual at $y = \hat{f}(x)$).
- In classification, may fail; can be enforced by stretching the largest p-value to 1:

$$\pi_x(y) \leftarrow \begin{cases} 1, & y = y_{\sigma(1)} \\ \pi_x(y), & \text{otherwise.} \end{cases}$$

An alternative way to perform Conformal Prediction

Step 1: the consonance condition (bridge CP $\leftrightarrow$ IP)

Consonance

A transducer $\pi_x(\cdot)$ is **consonant** if $\sup_{y \in \mathcal{Y}} \pi_x(y) = 1$.

- Holds naturally in regression for many scores (e.g. absolute residual at $y = \hat{f}(x)$).
- In classification, may fail; can be enforced by stretching the largest p-value to 1:

$$\pi_x(y) \leftarrow \begin{cases} 1, & y = y_{\sigma(1)} \\ \pi_x(y), & \text{otherwise.} \end{cases}$$

- Effect: only prevents empty prediction sets; preserves marginal coverage and upgrades a stronger reliability notion (Type II validity).

Define the set function (a plausibility / maxitive upper probability):

$$\overline{\mathbb{P}}(A) := \sup_{y \in A} \pi_x(y), \qquad \overline{\mathbb{P}}(\emptyset) = 0.$$

- Under consonance, $\overline{\mathbb{P}}$ is an **upper probability** and defines a **core credal set** $\mathcal{M}(\overline{\mathbb{P}})$ as follows:

$$\mathcal{M}(\overline{P}) := \{P \in \mathcal{P}(\mathcal{Y}) : P(A) \leq \overline{P}(A)\ \forall A\}.$$

- Interpretation: split CP induces a *set of plausible predictive distributions* (model multiplicity) based on the conformal p-values π_x.

(1-α) Imprecise Highest Density Region (IHDR))

The $(1-\alpha)$ IHDR of a lower probability $\underline{\mathbb{P}}$ is the set $\mathrm{IR}_\alpha \subseteq \mathcal{Y}$ such that

$$\underline{\mathbb{P}}(Y \in \mathrm{IR}_\alpha) = 1 - \alpha$$

and the size $|\,\mathrm{IR}_\alpha\,|$ is minimal.

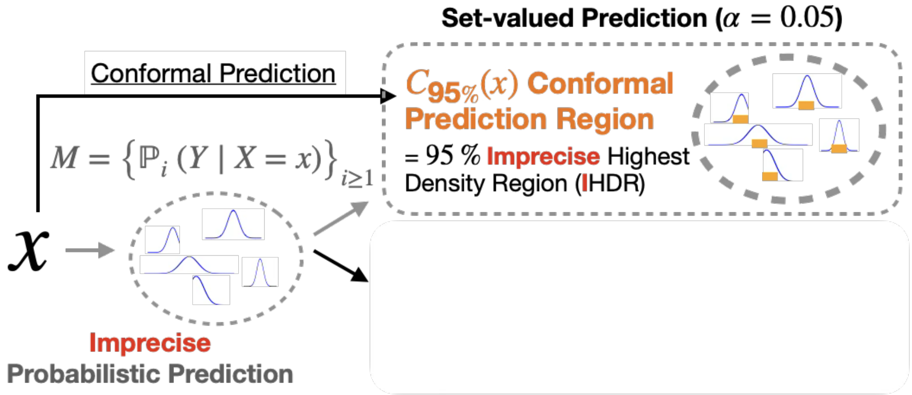
Now we can finally obtain our conformal prediction

Conformal prediction done differently

Denote the $(1-\alpha)$ IHDR of $\overline{\mathbb{P}_x}$ as IR_α, then for any $\alpha \in [0,1]$ we have $\mathrm{IR}_\alpha = \mathcal{C}_\alpha(x)$

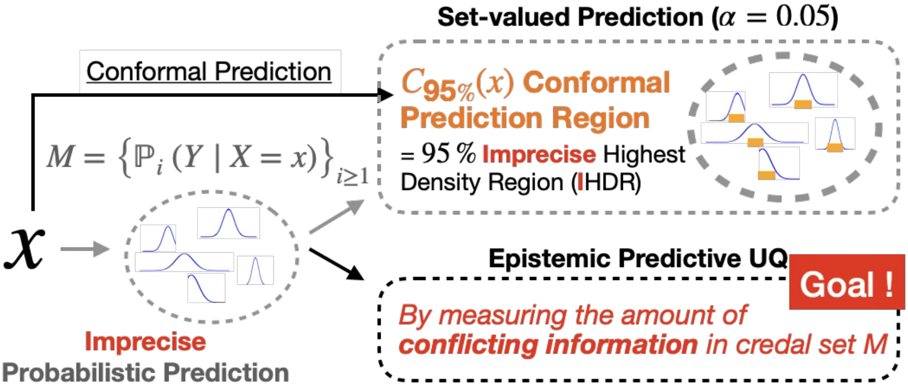
An alternative way to perform Conformal Prediction

A summary



An alternative way to perform Conformal Prediction

A summary



Quantifying EPU: two MMI-CP instantiations

We quantify EPU as *conflict* encoded in the induced credal set via MMI on $\overline{P}_x$.

(A) Total-variation test functions: $\mathrm{MMI}_{\mathrm{TV}}$

Let $\mathcal{H}_{\mathrm{TV}} = \{\mathbf{1}_A : A \subseteq \mathcal{Y}\}$:

$$\mathrm{MMI}_{\mathcal{H}_{\mathrm{TV}}}(\overline{P}_x) = \sup_{A \subseteq \mathcal{Y}} |\overline{P}_x(A) - \underline{P}_x(A)| = \pi_{\sigma(2)}.$$

Quantifying EPU: two MMI-CP instantiations

We quantify EPU as *conflict* encoded in the induced credal set via MMI on $\overline{P}_x$.

(A) Total-variation test functions: $\mathrm{MMI}_{\mathrm{TV}}$

Let $\mathcal{H}_{\mathrm{TV}} = \{\mathbf{1}_A : A \subseteq \mathcal{Y}\}$:

$$\mathrm{MMI}_{\mathcal{H}_{\mathrm{TV}}}(\overline{P}_x) = \sup_{A\subseteq\mathcal{Y}} |\overline{P}_x(A) - \underline{P}_x(A)| = \pi_{\sigma(2)}.$$

- Extremely simple: in classification, EPU is the **second-largest** conformal p-value.
- Intuition: the largest p-value is forced to 1 under consonance; what matters is the "best competing alternative."
- Coincides with existing notion of "confidence" for CP.

Quantifying EPU: portfolio-aware MMI using π_x

A more global measure accounts for the entire p-value portfolio.

(B) Use π_x as the test function

$$\mathrm{MMI}_{\{\pi_x\}}(\overline{P}_x) = \int_0^1 \sup_{y \notin \mathcal{C}_\alpha(x)} \pi_x(y)\, d\alpha.$$

Quantifying EPU: portfolio-aware MMI using π_x

A more global measure accounts for the entire *p*-value portfolio.

(B) Use π_x as the test function

$$\mathrm{MMI}_{\{\pi_x\}}(\overline{P}_x) = \int_0^1 \sup_{y \notin \mathcal{C}_\alpha(x)} \pi_x(y)\, d\alpha.$$

- Aggregates across confidence levels: at each α, look at the most plausible label *excluded* from $\mathcal{C}_\alpha(x)$.
- Naturally depends on both calibration data and score design (as CP does).

Closed form in conformal classification

For K-class classification with sorted p-values $\pi_{\sigma(1)} \geq \pi_{\sigma(2)} \geq \cdots \geq \pi_{\sigma(K)}$ and $\pi_{\sigma(K+1)} := 0$:

$$\mathrm{MMI}_{\{\pi_x\}}(\overline{P}_x) = \sum_{k=2}^{K+1} \left(\pi_{\sigma(k-1)} - \pi_{\sigma(k)}\right) \pi_{\sigma(k)}.$$

- Not just a sum: weights by discrete "gradients" of the sorted p-value profile.

Conformal regression: a fundamental limitation

For many standard regression scores satisfying $\inf_{y \notin \mathcal{C}_\alpha(x)} s(x, y) = \widehat{q}_{1-\alpha}$,

$$\mathrm{MMI}_{\{\pi_x\}}(\overline{P}_x) = 1 + \int_0^1 \frac{1 - \lceil (n_{\mathrm{cal}} + 1)(1 - \alpha) \rceil}{n_{\mathrm{cal}} + 1} \, d\alpha,$$

which is **constant across test instances**.

- Explains why standard conformal regression intervals often fail to rank instances by difficulty: the conformal component is dominated by a global calibration quantile.
- Takeaway: instance-adaptivity typically comes from the base model / score choice, not CP itself.

Experiments: evaluation via downstream decision-making

No ground-truth EPU exists $\Rightarrow$ evaluate informativeness via tasks where good EPU helps.

Compared methods

- Baselines: prediction-set size $|\mathcal{C}_\alpha(x)|$ at $\alpha \in \{0.01, 0.05, 0.1, 0.2, 0.3\}$.
- Proposed:
  - **MMI-CP**$_{\mathrm{TV}}$: $\pi_{\sigma(2)}$.
  - **MMI-CP**$_\pi$: portfolio-aware $\mathrm{MMI}_{\{\pi_x\}}(\overline{P}_x)$.

- Active learning: choose point with highest EPU (uncertainty sampling).
- Selective classification: rank by EPU to abstain on hardest instances; evaluate via ARC / AUC.

Results

Table: Experiment results: †/⋆ denote statistically significantly worse performance than MMI-CP$_{\pi_x}$/MMI-CP$_{\mathrm{TV}}$, respectively (Wilcoxon signed-rank test, 5%). Results are averaged over 10 seeds; one standard error is reported.

(a) (Active learning) Accuracies achieved at the final step.

	Digits v1	Digits v2	Digits v3	Letters
# classes	10	10	10	26
MMI-CP$_{\mathrm{TV}}$	$95.30_{\pm 0.29}{}^{\dagger}$	$97.69_{\pm 0.29}$	$95.75_{\pm 0.47}$	$\mathbf{82.00_{\pm 0.39}}$
MMI-CP$_{\pi_x}$	$\mathbf{95.80_{\pm 0.40}}$	$\mathbf{97.83_{\pm 0.42}}$	$\mathbf{95.90_{\pm 0.41}}$	$81.40_{\pm 0.61}$
$\lvert C_{0.01}(\cdot)\rvert$	$94.00_{\pm 0.22}{}^{\dagger\star}$	$94.91_{\pm 0.14}{}^{\dagger\star}$	$94.30_{\pm 0.40}{}^{\dagger\star}$	$81.32_{\pm 0.33}{}^{\dagger\star}$
$\lvert C_{0.05}(\cdot)\rvert$	$95.30_{\pm 0.53}{}^{\dagger}$	$97.01_{\pm 0.14}{}^{\dagger\star}$	$94.90_{\pm 0.51}{}^{\dagger}$	$80.86_{\pm 0.17}{}^{\dagger\star}$
$\lvert C_{0.1}(\cdot)\rvert$	$95.05_{\pm 0.53}{}^{\dagger}$	$96.74_{\pm 0.40}{}^{\dagger\star}$	$95.85_{\pm 0.34}{}^{\dagger}$	$80.75_{\pm 0.85}$
$\lvert C_{0.2}(\cdot)\rvert$	$95.40_{\pm 0.25}{}^{\dagger}$	$97.01_{\pm 0.28}{}^{\dagger\star}$	$95.80_{\pm 0.19}$	$81.10_{\pm 0.41}$
$\lvert C_{0.3}(\cdot)\rvert$	$95.40_{\pm 0.64}{}^{\dagger}$	$96.90_{\pm 0.23}{}^{\dagger\star}$	$95.55_{\pm 0.80}$	$81.05_{\pm 0.27}$

(b) (Selective classification) Area under ARC.

	Cifar10	Cifar100	Caltech	FMNIST
# classes	10	100	100	10
MMI-CP$_{\mathrm{TV}}$	$97.34_{\pm 1.42}$	$88.16_{\pm 0.87}$	$\mathbf{98.53_{\pm 0.03}}$	$\mathbf{97.54_{\pm 0.27}}$
MMI-CP$_{\pi_x}$	$\mathbf{97.68_{\pm 0.46}}$	$\mathbf{88.28_{\pm 0.91}}$	$\mathbf{98.53_{\pm 0.03}}$	$97.52_{\pm 0.29}{}^{\star}$
$\lvert C_{0.01}(\cdot)\rvert$	$97.37_{\pm 0.43}{}^{\dagger}$	$85.92_{\pm 0.94}{}^{\dagger\star}$	$98.43_{\pm 0.14}{}^{\dagger\star}$	$97.29_{\pm 0.29}$
$\lvert C_{0.05}(\cdot)\rvert$	$95.66_{\pm 0.62}{}^{\dagger\star}$	$86.81_{\pm 1.01}{}^{\dagger\star}$	$98.43_{\pm 0.08}{}^{\dagger\star}$	$94.39_{\pm 0.76}{}^{\dagger\star}$
$\lvert C_{0.1}(\cdot)\rvert$	$95.15_{\pm 0.52}{}^{\dagger\star}$	$86.36_{\pm 1.16}{}^{\dagger\star}$	$98.30_{\pm 0.14}{}^{\dagger\star}$	$93.59_{\pm 0.86}{}^{\star}$
$\lvert C_{0.2}(\cdot)\rvert$	$94.30_{\pm 0.83}{}^{\dagger\star}$	$83.19_{\pm 2.11}{}^{\dagger\star}$	$98.09_{\pm 0.22}{}^{\dagger\star}$	$93.00_{\pm 0.58}{}^{\star}$
$\lvert C_{0.3}(\cdot)\rvert$	$93.73_{\pm 1.74}{}^{\dagger\star}$	$81.53_{\pm 1.97}{}^{\dagger\star}$	$97.95_{\pm 0.24}{}^{\dagger\star}$	$92.69_{\pm 1.09}{}^{\dagger\star}$

Experiments: Active Learning

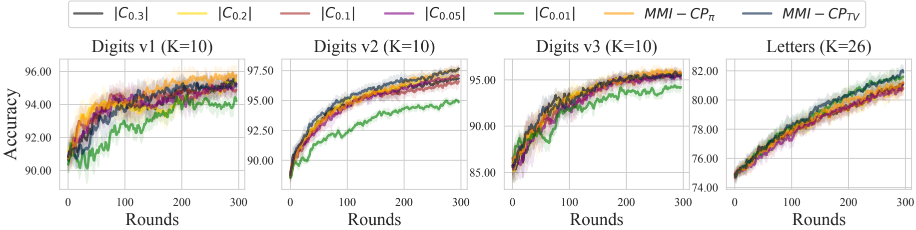


Figure: Active Learning

Experiments: Selective Classification

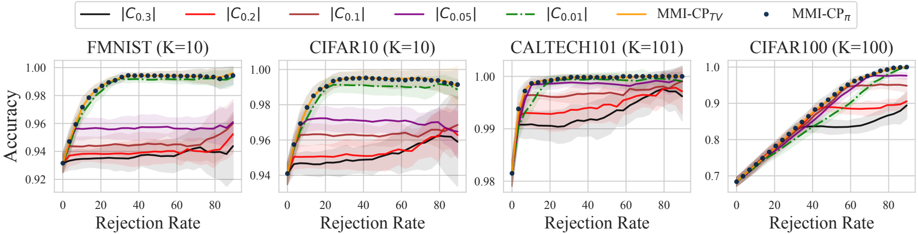


Figure: Selective Classification

Takeaways

1. Split CP implicitly defines a **predictive credal set** via a plausibility upper probability.

Takeaways

1. Split CP implicitly defines a **predictive credal set** via a plausibility upper probability.
2. CPRs are **imprecise HDRs**; EPU should be about *model multiplicity*, not set size.

Takeaways

1. Split CP implicitly defines a **predictive credal set** via a plausibility upper probability.
2. CPRs are **imprecise HDRs**; EPU should be about *model multiplicity*, not set size.
3. **MMI-CP** provides efficient, analytically tractable EPU measures:
   - $\mathrm{MMI}_{\mathrm{TV}} = \pi_{\sigma(2)}$ (cheap, strong baseline).
   - Portfolio-aware $\mathrm{MMI}_{\{\pi_x\}}$ (uses all p-values).

Takeaways

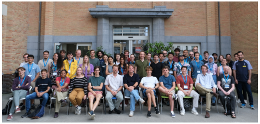
1. Split CP implicitly defines a **predictive credal set** via a plausibility upper probability.
2. CPRs are **imprecise HDRs**; EPU should be about *model multiplicity*, not set size.
3. **MMI-CP** provides efficient, analytically tractable EPU measures:
   - $\mathrm{MMI}_{\mathrm{TV}} = \pi_{\sigma(2)}$ (cheap, strong baseline).
   - Portfolio-aware $\mathrm{MMI}_{\{\pi_x\}}$ (uses all p-values).
4. Standard conformal regression offers a limited instance-dependent EPU signal.

Takeaways

1. Split CP implicitly defines a **predictive credal set** via a plausibility upper probability.
2. CPRs are **imprecise HDRs**; EPU should be about *model multiplicity*, not set size.
3. **MMI-CP** provides efficient, analytically tractable EPU measures:
   - $\mathrm{MMI}_{\mathrm{TV}} = \pi_{\sigma(2)}$ (cheap, strong baseline).
   - Portfolio-aware $\mathrm{MMI}_{\{\pi_x\}}$ (uses all p-values).
4. Standard conformal regression offers a limited instance-dependent EPU signal.
5. We now have a more fine-grained uncertainty signal for conformal prediction through connections to imprecise probabilities.

Curious to learn more about IP?

- Society of Imprecise Probability: Theory and Application (SIPTA) (`https://sipta.org/`)
- SIPTA Summer School
- EIML Workshops (EurIPS 2025)
- "Introductory materials": Walley [1991], Augustin et al. [2014b], Troffaes and De Cooman [2014]



Thank you! Any questions?

References I

Joaquín Abellán, George Jiří Klir, and Serafín Moral. Disaggregated total uncertainty measure for credal sets. *International Journal of General Systems*, 1 (35):29–44, 2006.

Jayadev Acharya, Constantinos Daskalakis, and Gautam Kamath. Optimal testing for properties of distributions. *Advances in Neural Information Processing Systems*, 28, 2015.

Massimiliano Amarante and Fabio Maccheroni. When an event makes a difference. *Theory and Decision*, 60:119–126, 2006.

Thomas Augustin, Frank P. A. Coolen, Gert de Cooman, and Matthias C. M. Troffaes. *Introduction to imprecise probabilities*. John Wiley & Sons,, West Sussex, England, 2014a.

Thomas Augustin, Frank P. A. Coolen, Gert De Cooman, and Matthias C. M. Troffaes, editors. *Introduction to imprecise probabilities*. Wiley series in probability and statistics. Wiley, Hoboken, NJ, 2014b. ISBN 978-0-470-97381-3.

Nick Bansback, Mark Harrison, and Carlo Marra. Does introducing imprecision around probabilities for benefit and harm influence the way people value treatments? *Medical Decision Making*, 36(4):490–502, 2016.

References II

Pierpaolo Battigalli, Simone Cerreia-Vioglio, Fabio Maccheroni, and Massimo Marinacci. Self-confirming equilibrium and model uncertainty. *The American Economic Review*, 105(2):646–677, 2015.

Pierpaolo Battigalli, Roberto Corrao, and Martin Dufwenberg. Incorporating belief-dependent motivation in games. *Journal of Economic Behavior & Organization*, 167:185–218, 2019.

A. Benavoli, B. Ristic, A. Farina, M. Oxenham, and L. Chisci. An application of evidential networks to threat assessment. *IEEE Transactions on Aerospace and Electronic Systems*, 45(2):620–639, 2009.

Alessio Benavoli, Alessandro Facchini, and Marco Zaffalon. The weirdness theorem and the origin of quantum paradoxes. *Foundations of Physics*, 51, 2021. URL https://api.semanticscholar.org/CorpusID:244228312.

Viktor Bengs, Eyke Hüllermeier, and Willem Waegeman. Pitfalls of epistemic uncertainty quantification through loss minimisation. In Alice H. Oh, Alekh Agarwal, Danielle Belgrave, and Kyunghyun Cho, editors, *Advances in Neural Information Processing Systems*, 2022. URL https://openreview.net/forum?id=epjxT_ARZW5.

References III

James O. Berger. The robust Bayesian viewpoint. In Joseph B. Kadane, editor, *Robustness of Bayesian Analyses*. Amsterdam : North-Holland, 1984.

Christopher Bülte, Yusuf Sale, Gitta Kutyniok, and Eyke Hüllermeier. Uncertainty quantification for regression: A unified framework based on kernel scores, 2025. URL https://arxiv.org/abs/2510.25599.

Andrea Campagner. *Credal Learning: Weakly Supervised Learning from Credal Sets*, pages 327–334. IOS Press, 09 2023. ISBN 9781643684369. doi: 10.3233/FAIA230287.

Michele Caprio. The joys of categorical conformal prediction, 2025. URL https://arxiv.org/abs/2507.04441.

Michele Caprio and Ruobin Gong. Dynamic precise and imprecise probability kinematics. In Enrique Miranda, Ignacio Montes, Erik Quaeghebeur, and Barbara Vantaggi, editors, *Proceedings of the Thirteenth International Symposium on Imprecise Probability: Theories and Applications*, volume 215 of *Proceedings of Machine Learning Research*, pages 72–83. PMLR, 11–14 Jul 2023.

References IV

Michele Caprio and Sayan Mukherjee. Ergodic theorems for dynamic imprecise probability kinematics. *International Journal of Approximate Reasoning*, 152: 325–343, 2023.

Michele Caprio and Teddy Seidenfeld. Constriction for sets of probabilities. In Enrique Miranda, Ignacio Montes, Erik Quaeghebeur, and Barbara Vantaggi, editors, *Proceedings of the Thirteenth International Symposium on Imprecise Probability: Theories and Applications*, volume 215 of *Proceedings of Machine Learning Research*, pages 84–95. PMLR, 11–14 Jul 2023. URL https://proceedings.mlr.press/v215/caprio23b.html.

Michele Caprio, Souradeep Dutta, Kuk Jin Jang, Vivian Lin, Radoslav Ivanov, Oleg Sokolsky, and Insup Lee. Credal bayesian deep learning. *Transactions on Machine Learning Research*, 2024a. ISSN 2835-8856. URL https://openreview.net/forum?id=4NHF9AC5ui.

Michele Caprio, Kuk Jin Jang, Souradeep Dutta, Shireen K. Manchingal, Fabio Cuzzolin, Oleg Sokolsky, and Insup Lee. Credal and interval deep evidential classifications. Presented at WUML 2024, 2024b.

References V

Michele Caprio, Yusuf Sale, Eyke Hüllermeier, and Insup Lee. A Novel Bayes' Theorem for Upper Probabilities. In Fabio Cuzzolin and Maryam Sultana, editors, *Epistemic Uncertainty in Artificial Intelligence*, pages 1–12, Cham, 2024c. Springer Nature Switzerland.

Michele Caprio, Siu Lun Chau, and Krikamol Muandet. When do credal sets stabilize? fixed-point theorems for credal set updates, 2025a. URL https://arxiv.org/abs/2510.04769.

Michele Caprio, Shireen K. Manchingal, and Fabio Cuzzolin. Credal and interval deep evidential classifications, 2025b. URL https://arxiv.org/abs/2512.05526.

Michele Caprio, Yusuf Sale, and Eyke Hüllermeier. Conformal prediction regions are imprecise highest density regions. In Sébastien Destercke, Alexander Erreygers, Max Nendel, Frank Riedel, and Matthias C. M. Troffaes, editors, *Proceedings of the Fourteenth International Symposium on Imprecise Probabilities: Theories and Applications*, volume 290 of *Proceedings of Machine Learning Research*, pages 47–59. PMLR, 15–18 Jul 2025c. URL https://proceedings.mlr.press/v290/caprio25b.html.

References VI

Michele Caprio, David Stutz, Shuo Li, and Arnaud Doucet. Conformalized credal regions for classification with ambiguous ground truth. *Transactions on Machine Learning Research*, 2025d. ISSN 2835-8856. URL https://openreview.net/forum?id=L7sQ8CW2FY.

Arianna Casanova, Juerg Kohlas, and Marco Zaffalon. Information algebras in the theory of imprecise probabilities, an extension. *International Journal of Approximate Reasoning*, 150:311–336, 2022. ISSN 0888-613X. doi: https://doi.org/10.1016/j.ijar.2022.09.003. URL https://www.sciencedirect.com/science/article/pii/S0888613X22001396.

Leonardo Cella and Ryan Martin. Validity, consonant plausibility measures, and conformal prediction. *International Journal of Approximate Reasoning*, 141: 110–130, 2022. Probability and Statistics: Foundations and History. In honor of Glenn Shafer.

Simone Cerreia-Vioglio, Fabio Maccheroni, and Massimo Marinacci. Ergodic theorems for lower probabilities. *Proceedings of the American Mathematical Society*, 144:3381–3396, 2015.

References VII

Simone Cerreia-Vioglio, Roberto Corrao, and Giacomo Lanzani. Nonlinear fixed points and stationarity: Economic applications (cid:3). In *Accepted for presentation at DTEA 22 and ESSET 23*, 2022. URL https://api.semanticscholar.org/CorpusID:248914283.

Simone Cerreia-Vioglio, Roberto Corrao, and Giacomo Lanzani. Dynamic opinion aggregation: Long-run stability and disagreement. *The Review of Economic Studies*, page rdad072, 2023.

Siu Lun Chau, Antonin Schrab, Arthur Gretton, Dino Sejdinovic, and Krikamol Muandet. Credal two-sample tests of epistemic ignorance. *arXiv preprint arXiv:2410.12921*, 2024.

Siu Lun Chau, Michele Caprio, and Krikamol Muandet. Integral imprecise probability metrics. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025. URL https://openreview.net/forum?id=KM2XzHq2Rm.

Gustave Choquet. Theory of capacities. *Annales de l'Institut Fourier*, 5:131–295, 1954.

References VIII

Giulianella Coletti and Romano Scozzafava. *Probabilistic Logic in a Coherent Setting*. Trends in Logic. Dordrecht : Springer, 2002.

Fabio Gagliardi Cozman. Credal networks. *Artificial Intelligence*, 120:199–233, 2000.

Fabio Cuzzolin. *The Geometry of Uncertainty*. Artificial Intelligence: Foundations, Theory, and Algorithms. Cham : Springer, 2020.

Fabio Cuzzolin and Ruggero Frezza. Evidential reasoning framework for object tracking. In Matthew R. Stein, editor, *Telemanipulator and Telepresence Technologies VI*, volume 3840, pages 13 – 24. International Society for Optics and Photonics, SPIE, 1999.

Gert De Cooman. Possibility theory I: the measure-and integral-theoretic groundwork. *International Journal of General Systems*, 25(4):291–323, 1997.

References IX

Keano De Vos, Gert de Cooman, and Jasper De Bock. Indistinguishability through exchangeability in quantum mechanics? In Enrique Miranda, Ignacio Montes, Erik Quaeghebeur, and Barbara Vantaggi, editors, *Proceedings of the Thirteenth International Symposium on Imprecise Probability: Theories and Applications*, volume 215 of *Proceedings of Machine Learning Research*, pages 177–188. PMLR, 11–14 Jul 2023. URL https://proceedings.mlr.press/v215/de-vos23a.html.

Arthur Pentland Dempster. Upper and lower probabilities induced by a multivalued mapping. *The Annals of Mathematical Statistics*, 38(2):325–339, 1967.

Thierry Denoeux. A neural network classifier based on Dempster-Shafer theory. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 30(2):131–150, 2000.

Thierry Denœux. Quantifying prediction uncertainty in regression using random fuzzy sets: the ENNreg model. *IEEE Transactions on Fuzzy Systems*, pages 1–10, 2023.

References X

Rabanus Derr and Robert C. Williamson. Systems of precision: Coherent probabilities on pre-dynkin systems and coherent previsions on linear subspaces. *Entropy*, 25(9), 2023.

S. Destercke, D. Dubois, and E. Chojnacki. Unifying practical uncertainty representations – i: Generalized p-boxes. *International Journal of Approximate Reasoning*, 49(3):649–663, 2008a. ISSN 0888-613X. doi: https://doi.org/10.1016/j.ijar.2008.07.003. URL https://www.sciencedirect.com/science/article/pii/S0888613X0800114X.

Sébastien Destercke, Didier Dubois, and Eric Chojnacki. Unifying practical uncertainty representations–i: Generalized p-boxes. *International Journal of Approximate Reasoning*, 49(3):649–663, 2008b.

Persi Diaconis and Sandy L. Zabell. Updating subjective probability. *Journal of the American Statistical Association*, 77(380):822–830, 1982.

References XI

Souradeep Dutta, Michele Caprio, Vivian Lin, Matthew Cleaveland, Kuk Jin Jang, Ivan Ruchkin, Oleg Sokolsky, and Insup Lee. Distributionally robust statistical verification with imprecise neural networks. In *Proceedings of the 28th ACM International Conference on Hybrid Systems: Computation and Control*, HSCC '25, New York, NY, USA, 2025. Association for Computing Machinery. ISBN 9798400715044. doi: 10.1145/3716863.3718040. URL https://doi.org/10.1145/3716863.3718040.

Daniel Faccini, Francesca Maggioni, and Florian A. Potra. Robust and distributionally robust optimization models for linear support vector machine. *Computers & Operations Research*, 147:105930, 2022. ISSN 0305-0548. doi: https://doi.org/10.1016/j.cor.2022.105930. URL https://www.sciencedirect.com/science/article/pii/S0305054822001861.

Scott Ferson, Vladik Kreinovick, Lev Ginzburg, Davis S. Myers, and Kari Sentz. Constructing probability boxes and Dempster-Shafer structures. Technical Report SAND2002-4015, Sandia National Laboratiories, 2003.

Ronald Aylmer Fisher. The fiducial argument in statistical inference. *Annals of Eugenics*, 5(4):391–398, 1935.

References XII

Christian Fröhlich, Rabanus Derr, and Robert C Williamson. Strictly frequentist imprecise probability. *Available at arxiv:2302.03520*, 2023.

Rui Gao, Liyan Xie, Yao Xie, and Huan Xu. Robust hypothesis testing using wasserstein uncertainty sets. *Advances in Neural Information Processing Systems*, 31, 2018.

Giangiacomo Gerla. Inferences in probability logic. *Artificial Intelligence*, 70(1): 33–52, 1994. ISSN 0004-3702. doi: https://doi.org/10.1016/0004-3702(94)90102-3. URL https://www.sciencedirect.com/science/article/pii/0004370294901023.

Itzhak Gilboa, Fabio Maccheroni, Massimo Marinacci, and David Schmeidler. Objective and subjective rationality in a multiple prior model. *Econometrica*, 78:755–770, 2010.

Eleonora Giunchiglia, Mihaela Cătălina Stoian, Salman Khan, Fabio Cuzzolin, and Thomas Lukasiewicz. ROAD-R: the autonomous driving dataset with logical requirements. *Machine Learning*, 1(112):3261–3291, 2023.

References XIII

Ruobin Gong and Xiao-Li Meng. Judicious judgment meets unsettling updating: dilation, sure loss, and Simpson's paradox. *Statistical Science*, 36(2):169–190, 2021.

Joseph Y. Halpern. *Reasoning About Uncertainty*. MIT Press, 2003.

Jan Hannig. On generalized fiducial inference. *Statistica Sinica*, 19(2):491–544, 2009.

Pei-Yi Hao, Jung-Hsien Chiang, and Yu-De Chen. Possibilistic classification by support vector networks. *Neural Networks*, 149:40–56, May 2022. ISSN 08936080. doi: 10.1016/j.neunet.2022.02.007. URL https://linkinghub.elsevier.com/retrieve/pii/S0893608022000405.

Nong Minh Hieu, Jeremie Houssineau, Neil K Chada, and Emmanuel Delande. Decoupling epistemic and aleatoric uncertainties with possibility theory. In *The 28th International Conference on Artificial Intelligence and Statistics*, pages 2899–2907, 2025.

Paul Hofman, Yusuf Sale, and Eyke Hüllermeier. Quantifying aleatoric and epistemic uncertainty with proper scoring rules, 2024. URL https://arxiv.org/abs/2404.12215.

References XIV

Jiayi Huang, Sangwoo Park, Nicola Paoletti, and Osvaldo Simeone. Distilling calibration via conformalized credal inference, 2025. URL https://arxiv.org/abs/2501.06066.

Peter J Huber. Robust estimation of a location parameter. *Ann. Math. Statist.*, 35(4):73–101, 1964.

Alan Hájek. Interpretations of Probability. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, fall 2019 edition, 2019. URL https://plato.stanford.edu/archives/fall2019/entries/probability-interpret/.

Eyke Hüllermeier. An approach to modelling and simulation of uncertain dynamical systems. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 5(2):117–137, 1997.

Eyke Hüllermeier. Possibilistic instance-based learning. *Artificial Intelligence*, 148 (1-2):335–383, August 2003. ISSN 00043702. doi: 10.1016/S0004-3702(03)00019-5. URL https://linkinghub.elsevier.com/retrieve/pii/S0004370203000195.

References XV

Eyke Hüllermeier and Willem Waegeman. Aleatoric and epistemic uncertainty in machine learning: an introduction to concepts and methods. *Machine Learning*, 3(110):457–506, 2021a.

Eyke Hüllermeier and Willem Waegeman. Aleatoric and Epistemic Uncertainty in Machine Learning: An Introduction to Concepts and Methods. *Machine Learning*, 110(3):457–506, March 2021b. ISSN 0885-6125, 1573-0565. doi: 10.1007/s10994-021-05946-3. arXiv:1910.09457 [cs, stat].

Alireza Javanmardi, David Stutz, and Eyke Hüllermeier. Conformalized credal set predictors. *arXiv preprint arXiv:2402.10723*, 2024.

Richard C. Jeffrey. *The Logic of Decision*. University of Chicago Press, New York, NY, USA, 1965.

Shihao Ji, Zihui Song, and Jiajie Huang. Credal transformer: A principled approach for quantifying and mitigating hallucinations in large language models, 2025. URL https://arxiv.org/abs/2510.12137.

References XVI

Mira Juergens, Nis Meinert, Viktor Bengs, Eyke Hüllermeier, and Willem Waegeman. Is epistemic uncertainty faithfully represented by evidential deep learning methods? In *Forty-first International Conference on Machine Learning*, 2024. URL https://openreview.net/forum?id=mxjB0LIgpT.

Timo Kaufmann, Paul Weng, Viktor Bengs, and Eyke Hüllermeier. A survey of reinforcement learning from human feedback. *Available at arxiv:2312.14925*, 2023.

John Maynard Keynes. *A treatise on probability*. 1921.

Frank Hyneman Knight. *Risk, uncertainty and profit*, volume 31. Houghton Mifflin, 1921.

Kolmogorov. Sulla determinazione empirica di una legge didistribuzione. *Giorn Dell'inst Ital Degli Att*, 4:89–91, 1933.

Henry Ely Kyburg. *Probability and the Logic of Rational Belief*. Wesleyan University Press, Middletown, CT, USA, 1961.

Isaac Levi. Probability Kinematics. *Philosophy of Science*, 18(3):197–209, 1967.

Isaac Levi. *The Enterprise of Knowledge*. London, UK : MIT Press, 1980.

References XVII

Xing Liu and François-Xavier Briol. On the robustness of kernel goodness-of-fit tests. *arXiv preprint arXiv:2408.05854*, 2024.

Timo Löhr, Paul Hofman, Felix Mohr, and Eyke Hüllermeier. Credal prediction based on relative likelihood. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025.

Jasper De Bock Louis J. M. Aslett, Frank P. A. Coolen, editor. *Uncertainty in Engineering: Introduction to Methods and Applications*, volume 7 of *SpringerBriefs in Statistics*. Springer Cham, 2022.

Pengyuan Lu, Michele Caprio, Eric Eaton, and Insup Lee. IBCL: Zero-shot Model Generation for Task Trade-offs in Continual Learning. *arXiv preprint arXiv:2305.14782*, 2025.

Fabio Maccheroni and Massimo Marinacci. A strong law of large numbers for capacities. *Annals of Probability*, 33(3):1171–1178, 2005.

Fabio Maccheroni, Massimo Marinacci, and Aldo Rustichini. Ambiguity aversion, robustness, and the variational representation of preferences. *Econometrica*, 74: 1447–1498, 2006.

References XVIII

Francesca Maggioni and Andrea Spinelli. A robust optimization model for nonlinear support vector machine, 2024. URL https://arxiv.org/abs/2306.06223.

Shireen Kudukkil Manchingal and Fabio Cuzzolin. Epistemic deep learning. *Available at arxiv:2206.07609*, 2022.

Shireen Kudukkil Manchingal, Muhammad Mubashar, Kaizheng Wang, Keivan Shariatmadar, and Fabio Cuzzolin. Random-set neural networks. In *The Thirteenth International Conference on Learning Representations*, 2025.

Sabina Marchetti and Alessandro Antonucci. Imaginary kinematics. In Amir Globerson and Ricardo Silva, editors, *UAI 2018: Proceedings of the Thirty-Fourth Conference on Uncertainty in Artificial Intelligence*, pages 104–113, Monterey, California, USA, 2018. AUAI Press. URL http://auai.org/uai2018/proceedings/papers/42.pdf.

Massimo Marinacci and Luigi Montrucchio. Introduction to the mathematics of ambiguity. In Itzhak Gilboa, editor, *Uncertainty in economic theory: a collection of essays in honor of David Schmeidler's 65th birthday*. London : Routledge, 2004.

References XIX

Alexander Marquardt, Julian Rodemann, and Thomas Augustin. An empirical study of prior-data conflicts in Bayesian neural networks. *Poster presented at 13th International Symposium on Imprecise Probabilities: Theories and Applications (ISIPTA) 2023*, 2023.

Ryan Martin. False confidence, non-additive beliefs, and valid statistical inference. *International Journal of Approximate Reasoning*, 113:39–73, October 2019. ISSN 0888613X. doi: 10.1016/j.ijar.2019.06.005. URL http://arxiv.org/abs/1607.05051. arXiv:1607.05051 [math, stat].

D. D. Maua, C. P. de Campos, A. Benavoli, and A. Antonucci. Probabilistic inference in credal networks: New complexity results. *Journal of Artificial Intelligence Research*, 50(1):603–637, 2014.

Thomas Mortier, Viktor Bengs, Eyke Hüllermeier, Stijn Luca, and Willem Waegeman. On the calibration of probabilistic classifier sets. In *International Conference on Artificial Intelligence and Statistics*, pages 8857–8870. PMLR, 2023.

References XX

Muhammad Mubashar, Shireen Kudukkil Manchingal, and Fabio Cuzzolin. Random-set large language models, 2025. URL https://arxiv.org/abs/2504.18085.

Alfred Müller. Integral probability metrics and their generating classes of functions. *Advances in applied probability*, 29(2):429–443, 1997.

Vu-Linh Nguyen, Mohammad Hossein Shaker, and Eyke Hüllermeier. How to measure uncertainty in uncertainty sampling for active learning. *Machine Learning*, 111:89–122, 2022.

Vu-Linh Nguyen, Haifei Zhang, and Sébastien Destercke. Credal ensembling in multi-class classification. *Machine Learning*, 114(19):62, 2025.

Luis Oala, Cosmas Heiß, Jan Macdonald, Maximilian März, Wojciech Samek, and Gitta Kutyniok. Interval neural networks: Uncertainty scores. *arXiv preprint arXiv:2003.11566*, 2020.

Yaniv Oren, Matthijs T. J. Spaan, and Wendelin Böhmer. E-MCTS: Deep Exploration in Model-Based Reinforcement Learning by Planning with Epistemic Uncertainty. *Available at arXiv:2210.13455*, 2022.

References XXI

Alessandro Saffiotti. A belief-function logic. In *AAAI*, pages 642–647, 1992.

Yusuf Sale, Michele Caprio, and Eyke Hüllermeier. Is the volume of a credal set a good measure for epistemic uncertainty? In Robin J. Evans and Ilya Shpitser, editors, *Proceedings of the Thirty-Ninth Conference on Uncertainty in Artificial Intelligence*, volume 216 of *Proceedings of Machine Learning Research*, pages 1795–1804. PMLR, 31 Jul–04 Aug 2023.

Yusuf Sale, Viktor Bengs, Michele Caprio, and Eyke Hüllermeier. Second-order uncertainty quantification: A distance-based approach. In Ruslan Salakhutdinov, Zico Kolter, Katherine Heller, Adrian Weller, Nuria Oliver, Jonathan Scarlett, and Felix Berkenkamp, editors, *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pages 43060–43076. PMLR, 21–27 Jul 2024. URL https://proceedings.mlr.press/v235/sale24a.html.

Glenn Shafer. *A Mathematical Theory of Evidence*. Princeton : Princeton University Press, 1976.

References XXII

Anurag Singh, Siu Lun Chau, Shahine Bouabid, and Krikamol Muandet. Domain generalisation via imprecise learning. *arXiv preprint arXiv:2404.04669, accepted to ICML 2024*, 2024.

Anurag Singh, Siu Lun Chau, and Krikamol Muandet. Truthful elicitation of imprecise forecasts. In *The 41st Conference on Uncertainty in Artificial Intelligence*, 2025a.

Jasraj Singh, Shelvia Wongso, Jeremie Houssineau, and Badr-Eddine Chérief-Abdellatif. Maxitive donsker-varadhan formulation for possibilistic variational inference, 2025b. URL https://arxiv.org/abs/2511.21223.

David Stutz, Abhijit Guha Roy, Tatiana Matejovicova, Patricia Strachan, Ali Taylan Cemgil, and Arnaud Doucet. Conformal prediction under ambiguous ground truth. *Transactions on Machine Learning Research*, 2023.

Izzeddin Teeti, Valentina Musat, Salman Khan, Alexander Rast, Fabio Cuzzolin, and Andrew Bradley. Vision in adverse weather: Augmentation using CycleGANs with various object detectors for robust perception in autonomous racing. *Available at arXiv:2201.03246*, 2022.

References XXIII

Matthias Troffaes and Sebastien Destercke. Probability boxes on totally preordered spaces for multivariate modelling. *International Journal of Approximate Reasoning*, 52(6):767–791, 2011. ISSN 0888-613X. doi: https://doi.org/10.1016/j.ijar.2011.02.001. URL https://www.sciencedirect.com/science/article/pii/S0888613X1100048X.

Matthias CM Troffaes and Gert De Cooman. *Lower previsions*. John Wiley & Sons, 2014.

Massimiliano Vasile, editor. *Optimization Under Uncertainty with Applications to Aerospace Engineering*, volume 6 of *Physics and Astronomy*. Springer Cham, 2021.

Paolo Vicig. Financial risk measurement with imprecise probabilities. *International Journal of Approximate Reasoning*, 49(1):159–174, 2008. Special Section on Logical Approaches to Imprecise Probabilities and Special Section on Imprecise Probabilities in Finance and Economics.

Peter Walley. *Statistical Reasoning with Imprecise Probabilities*, volume 42 of *Monographs on Statistics and Applied Probability*. London : Chapman and Hall, 1991.

References XXIV

Kaizheng Wang, Fabio Cuzzolin, Keivan Shariatmadar, David Moens, Hans Hallez, et al. Credal deep ensembles for uncertainty quantification. *Advances in Neural Information Processing Systems*, 37:79540–79572, 2024.

Kaizheng Wang, Fabio Cuzzolin, Keivan Shariatmadar, David Moens, and Hans Hallez. Credal wrapper of model averaging for uncertainty estimation in classification. In *International Conference on Learning Representations*, 2025a.

Kaizheng Wang, Keivan Shariatmadar, Shireen Kudukkil Manchingal, Fabio Cuzzolin, David Moens, and Hans Hallez. Creinns: Credal-set interval neural networks for uncertainty estimation in classification tasks. *Neural Networks*, 185:107198, 2025b. ISSN 0893-6080. doi: https://doi.org/10.1016/j.neunet.2025.107198. URL https://www.sciencedirect.com/science/article/pii/S0893608025000772.

Kaizheng Wang, Keivan Shariatmadar, Shireen Kudukkil Manchingal, Fabio Cuzzolin, David Moens, and Hans Hallez. Creinns: Credal-set interval neural networks for uncertainty estimation in classification tasks. *Neural Networks*, page 107198, 2025c.

References XXV

Jon Williamson. *In Defence of Objective Bayesianism*. Oxford University Press, April 2010. ISBN 978-0-19-922800-3. doi: 10.1093/acprof:oso/9780199228003.001.0001.

Lisa Wimmer, Yusuf Sale, Paul Hofman, Bernd Bischl, and Eyke Hüllermeier. Quantifying aleatoric and epistemic uncertainty in machine learning: Are conditional entropy and mutual information appropriate measures? In Robin J. Evans and Ilya Shpitser, editors, *Proceedings of the Thirty-Ninth Conference on Uncertainty in Artificial Intelligence*, volume 216 of *Proceedings of Machine Learning Research*, pages 2282–2292. PMLR, 31 Jul–04 Aug 2023. URL https://proceedings.mlr.press/v216/wimmer23a.html.

Fanyi Wu, Veronika Lomanova, and Michele Caprio. Bayesian conformal prediction as a decision risk problem using bayesian quadrature. Technical report, The University of Manchester, 2025.

Marco Zaffalon. The naive credal classifier. *Journal of Statistical Planning and Inference*, 105(1):5–21, 2002. Imprecise Probability Models and their Applications.

References XXVI

Marco Zaffalon, Alessandro Antonucci, and Rafael Cabañas. Structural causal models are (solvable by) credal networks. In Manfred Jaeger and Thomas Dyhre Nielsen, editors, *Proceedings of the 10th International Conference on Probabilistic Graphical Models*, volume 138 of *Proceedings of Machine Learning Research*, pages 581–592. PMLR, 23–25 Sep 2020.

Marco Zaffalon, Alessandro Antonucci, Rafael Cabañas, and David Huber. Approximating counterfactual bounds while fusing observational, biased and randomised data sources. *International Journal of Approximate Reasoning*, 162: 109023, 2023.

Moritz A. Zanger, Wendelin Böhmer, and Matthijs T. J. Spaan. Diverse projection ensembles for distributional reinforcement learning. *Available at arXiv:2306.07124*, 2023.