Conformal Prediction and Its Uncertainty Quantification Capability via Credal Sets

Michele Caprio

Department of Computer Science, The University of Manchester
Manchester Centre for AI Fundamentals

SIPTA.org/seminars
Society for Imprecise Probabilities: Theories and Applications

SIPTA

MANCHESTER 1824
The University of Manchester

Snapshot of the Paper

The Joys of Categorical Conformal Prediction

Michele Caprio

Conformal prediction (CP) is an Uncertainty Representation technique that delivers finite-sample calibrated prediction regions for any underlying Machine Learning model. Its status as an Uncertainty Quantification (UQ) tool, though, has remained conceptually opaque: While Conformal Prediction Regions (CPRs) give an ordinal representation of uncertainty (larger regions typically indicate higher uncertainty), they lack the capability to cardinally quantify it (twice as large regions do not imply twice the uncertainty). We adopt a category-theoretic approach to CP -- framing it as a morphism, embedded in a commuting diagram, of two newly-defined categories -- that brings us three joys. First, we show that -- under minimal assumptions -- CP is intrinsically a UQ mechanism, that is, its cardinal UQ capabilities are a structural feature of the method. Second, we demonstrate that CP bridges (and perhaps subsumes) the Bayesian, frequentist, and imprecise probabilistic approaches to predictive statistical reasoning. Finally, we show that a CPR is the image of a covariant functor. This observation is relevant to AI privacy: It implies that privacy noise added locally does not break the global coverage guarantee.

Motivation

- Conformal Prediction (CP) quantifies uncertainty in an *ordinal* way

Motivation

- Conformal Prediction (CP) quantifies uncertainty in an *ordinal* way

Research Question

Does is do so also **cardinally**?

Motivation

- Conformal Prediction (CP) quantifies uncertainty in an *ordinal* way

Research Question

Does is do so also **cardinally**?

- The journey to answer this question led to other interesting discoveries

(Full) Conformal Prediction

- Consider an exchangeable process $\mathcal{Y}_1, \mathcal{Y}_2, \ldots$ with distribution $\mathfrak{P}$
  - Each $\mathcal{Y}_i$ is a random element taking values in the m.s. (Y, Σ_Y)

(Full) Conformal Prediction

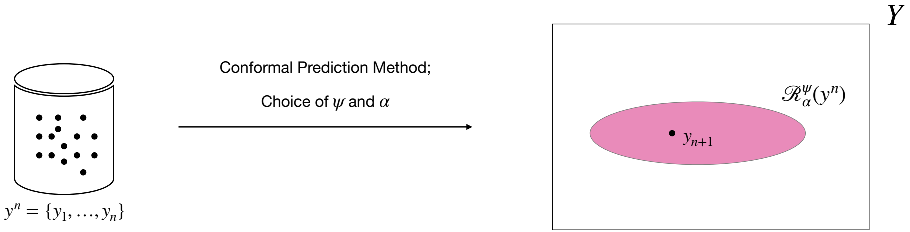
- Consider an exchangeable process $\mathcal{Y}_1, \mathcal{Y}_2, \ldots$ with distribution $\mathfrak{P}$
  - Each $\mathcal{Y}_i$ is a random element taking values in the m.s. (Y, Σ_Y)
  - A sequence is exchangeable if, for any $k \in \mathbb{N}$, $(\mathcal{Y}_1, \ldots, \mathcal{Y}_k)$ and $(\mathcal{Y}_{\text{perm}(1)}, \ldots, \mathcal{Y}_{\text{perm}(k)})$ have the same joint distribution

(Full) Conformal Prediction

- Consider an exchangeable process $\mathcal{Y}_1, \mathcal{Y}_2, \dots$ with distribution $\mathfrak{P}$
  - Each $\mathcal{Y}_i$ is a random element taking values in the m.s. (Y, Σ_Y)
  - A sequence is exchangeable if, for any $k \in \mathbb{N}$, $(\mathcal{Y}_1, \dots, \mathcal{Y}_k)$ and $(\mathcal{Y}_{\text{perm}(1)}, \dots, \mathcal{Y}_{\text{perm}(k)})$ have the same joint distribution
- We observe the first n terms of the process $\mathcal{Y}^n = (\mathcal{Y}_1, \dots, \mathcal{Y}_n)$

(Full) Conformal Prediction

- Consider an exchangeable process $\mathcal{Y}_1, \mathcal{Y}_2, \dots$ with distribution $\mathfrak{P}$
  - Each $\mathcal{Y}_i$ is a random element taking values in the m.s. (Y, Σ_Y)
  - A sequence is exchangeable if, for any $k \in \mathbb{N}$, $(\mathcal{Y}_1, \dots, \mathcal{Y}_k)$ and $(\mathcal{Y}_{\text{perm}(1)}, \dots, \mathcal{Y}_{\text{perm}(k)})$ have the same joint distribution
- We observe the first n terms of the process $\mathcal{Y}^n = (\mathcal{Y}_1, \dots, \mathcal{Y}_n)$
- With this data + exchangeability, goal: predict $\mathcal{Y}_{n+1}$ using a method that is valid/reliable



(Full) Conformal Prediction, cont'd

- Let $\mathcal{Y}^{n+1} = (\mathcal{Y}^n, \mathcal{Y}_{n+1})$

(Full) Conformal Prediction, cont'd

- Let $\mathcal{Y}^{n+1} = (\mathcal{Y}^n, \mathcal{Y}_{n+1})$
- Consider the transform

$$\mathcal{Y}^{n+1} \to T^{n+1} = (T_1, \dots, T_{n+1})$$

(Full) Conformal Prediction, cont'd

- Let $\mathcal{Y}^{n+1} = (\mathcal{Y}^n, \mathcal{Y}_{n+1})$
- Consider the transform

$$\mathcal{Y}^{n+1} \to T^{n+1} = (T_1, \ldots, T_{n+1})$$

defined by the rule

$$T_i \coloneqq \psi_i\left(\mathcal{Y}^{n+1}\right) \equiv \psi\left(y_{-i}^{n+1}, y_i\right), \quad i \in \{1, \ldots, n+1\}$$

  - $y_{-i}^{n+1} = y^{n+1} \setminus \{y_i\}$,

(Full) Conformal Prediction, cont'd

- Let $\mathcal{Y}^{n+1} = (\mathcal{Y}^n, \mathcal{Y}_{n+1})$
- Consider the transform

$$\mathcal{Y}^{n+1} \to T^{n+1} = (T_1, \dots, T_{n+1})$$

defined by the rule

$$T_i := \psi_i \left(\mathcal{Y}^{n+1}\right) \equiv \psi \left(y_{-i}^{n+1}, y_i\right), \quad i \in \{1, \dots, n+1\}$$

  - $y_{-i}^{n+1} = y^{n+1} \setminus \{y_i\}$,
  - $\psi : Y^n \times Y \to \mathbb{R}$ is a fixed function, invariant to permutations in its first vector argument. E.g. $\psi_i(y^{n+1}) = |\text{mean}(y_{-i}^{n+1}) - y_i|$

(Full) Conformal Prediction, cont'd

- Let $\mathcal{Y}^{n+1} = (\mathcal{Y}^n, \mathcal{Y}_{n+1})$
- Consider the transform

$$\mathcal{Y}^{n+1} \to T^{n+1} = (T_1, \ldots, T_{n+1})$$

defined by the rule

$$T_i := \psi_i\left(\mathcal{Y}^{n+1}\right) \equiv \psi\left(y_{-i}^{n+1}, y_i\right), \quad i \in \{1, \ldots, n+1\}$$

  - $y_{-i}^{n+1} = y^{n+1} \setminus \{y_i\}$,
  - $\psi : Y^n \times Y \to \mathbb{R}$ is a fixed function, invariant to permutations in its first vector argument. E.g. $\psi_i(y^{n+1}) = |\text{mean}(y_{-i}^{n+1}) - y_i|$
  - Function ψ: non-conformity measure; $\psi_i(y^{n+1})$ is small if and only if y_i agrees with – i.e. is "close to" – a prediction based on the data y_{-i}^{n+1}

(Full) Conformal Prediction, cont'd

- The value $\mathcal{Y}_{n+1}$ has not yet been observed; it is the prediction target

(Full) Conformal Prediction, cont'd

- The value $\mathcal{Y}_{n+1}$ has not yet been observed; it is the prediction target
- Exchangeability-preserving properties of the transformations T_i give us a procedure to rank candidate values $\tilde{y}$ of $\mathcal{Y}_{n+1}$
  - Based on the observed $\mathcal{Y}^n = y^n$

Algorithm 1 Full Conformal prediction (CP)

Initialize: data y^n, non-conformity measure ψ, grid of $\tilde{y}$ values
for each $\tilde{y}$ value in the grid **do**
  set $y_{n+1} = \tilde{y}$ and write $y^{n+1} = y^n \cup \{y_{n+1}\}$;
  define $T_i = \psi_i(y^{n+1})$, for all $i \in \{1, \ldots, n+1\}$;
  evaluate $\pi(\tilde{y}, y^n) = (n+1)^{-1} \sum_{i=1}^{n+1} \mathbb{1}[T_i \geq T_{n+1}]$;
end for
return $\pi(\tilde{y}, y^n)$ for each $\tilde{y}$ on the grid.

(Full) Conformal Prediction, cont'd

- The value $\mathcal{Y}_{n+1}$ has not yet been observed; it is the prediction target
- Exchangeability-preserving properties of the transformations T_i give us a procedure to rank candidate values $\tilde{y}$ of $\mathcal{Y}_{n+1}$
  - Based on the observed $\mathcal{Y}^n = y^n$

Algorithm 1 Full Conformal prediction (CP)

Initialize: data y^n, non-conformity measure ψ, grid of $\tilde{y}$ values
for each $\tilde{y}$ value in the grid **do**
  set $y_{n+1} = \tilde{y}$ and write $y^{n+1} = y^n \cup \{y_{n+1}\}$;
  define $T_i = \psi_i(y^{n+1})$, for all $i \in \{1, \ldots, n+1\}$;
  evaluate $\pi(\tilde{y}, y^n) = (n+1)^{-1} \sum_{i=1}^{n+1} \mathbb{1}[T_i \geq T_{n+1}]$;
end for
return $\pi(\tilde{y}, y^n)$ for each $\tilde{y}$ on the grid.

- Data-dependent function $\tilde{y} \mapsto \pi(\tilde{y}, y^n) \in [0,1]$ can be interpreted as a measure of plausibility of the assertion that $\mathcal{Y}_{n+1} = \tilde{y}$, given data y^n
- Vovk et al. (2005) refer to the function π as conformal transducer

(Full) Conformal Prediction, cont'd

- For any $\alpha \in [0,1]$, the α-level Conformal Prediction Region (CPR) is defined as (Vovk, 2013, Equation (2))

$$\mathscr{R}_\alpha^\psi(y^n) \coloneqq \{y_{n+1} \in Y : \pi(y_{n+1}, y^n) > \alpha\}$$

(Full) Conformal Prediction, cont'd

- For any $\alpha \in [0, 1]$, the α-level Conformal Prediction Region (CPR) is defined as (Vovk, 2013, Equation (2))

$$\mathscr{R}_{\alpha}^{\psi}(y^n) := \{y_{n+1} \in Y : \pi(y_{n+1}, y^n) > \alpha\}$$

- It satisfies

$$P\left[\mathcal{Y}_{n+1} \in \mathscr{R}_{\alpha}^{\psi}(y^n)\right] \geq 1 - \alpha, \tag{1}$$

uniformly in n and in P (Vovk et al., 2005). That is, (1) is satisfied for all $n \in \mathbb{N}$ and all exchangeable distributions P

(Full) Conformal Prediction as a Correspondence

- Caprio et al. (2025, Section 5) point out how Full CP can be written as a correspondence

$$\kappa : [0,1] \times Y^n \times \mathscr{F} \rightrightarrows Y, \quad (\alpha, y^n, \psi) \mapsto \kappa(\alpha, y^n, \psi),$$

(Full) Conformal Prediction as a Correspondence

- Caprio et al. (2025, Section 5) point out how Full CP can be written as a correspondence

$$\kappa : [0,1] \times Y^n \times \mathscr{F} \rightrightarrows Y, \quad (\alpha, y^n, \psi) \mapsto \kappa(\alpha, y^n, \psi),$$

where $\mathscr{F}$ is the set

$$\{\psi : Y^n \times Y \to \mathbb{R},\ \psi \text{ is invariant to permutations in its first argument}\},$$

(Full) Conformal Prediction as a Correspondence

- Caprio et al. (2025, Section 5) point out how Full CP can be written as a correspondence

$$\kappa : [0,1] \times Y^n \times \mathscr{F} \rightrightarrows Y, \quad (\alpha, y^n, \psi) \mapsto \kappa(\alpha, y^n, \psi),$$

where $\mathscr{F}$ is the set

$$\{\psi : Y^n \times Y \to \mathbb{R},\ \psi \text{ is invariant to permutations in its first argument}\},$$

and the image $\kappa(\alpha, y^n, \psi)$ of κ is defined as

$$\left\{ y_{n+1} \in Y : \underbrace{\frac{1}{n+1} \sum_{i=1}^{n+1} \mathbb{1} \left[\psi\left(y_{-i}^{n+1}, y_i\right) \geq \psi\left(y_{-(n+1)}^{n+1}, y_{n+1}\right) \right]}_{=:\pi(y_{n+1}, y^n)} > \alpha \right\}$$

(Full) Conformal Prediction as a Correspondence

- Caprio et al. (2025, Section 5) point out how Full CP can be written as a correspondence

$$\kappa : [0,1] \times Y^n \times \mathscr{F} \rightrightarrows Y, \quad (\alpha, y^n, \psi) \mapsto \kappa(\alpha, y^n, \psi),$$

where $\mathscr{F}$ is the set

$$\{\psi : Y^n \times Y \to \mathbb{R},\ \psi \text{ is invariant to permutations in its first argument}\},$$

and the image $\kappa(\alpha, y^n, \psi)$ of κ is defined as

$$\left\{ y_{n+1} \in Y : \underbrace{\frac{1}{n+1} \sum_{i=1}^{n+1} \mathbb{1}\left[\psi\left(y_{-i}^{n+1}, y_i\right) \geq \psi\left(y_{-(n+1)}^{n+1}, y_{n+1}\right)\right]}_{=:\pi(y_{n+1}, y^n)} > \alpha \right\}$$

- It is easy to see, then, that $\kappa(\alpha, y^n, \psi) \equiv \mathscr{R}_\alpha^\psi(y^n)$

Research Question, Explained

- CP Regions depend only on the *ranks* of the non-conformity scores: for any strictly increasing transform f, $\kappa_\alpha(y^n, \psi) = \kappa_\alpha(y^n, f \circ \psi)$
- $\text{diam}(\mathscr{R})$ is numerically well-defined once a metric on Y is fixed, but: no canonical cardinal scale of "amount of uncertainty" attached to $\mathscr{R}$,

$$\text{diam}(\mathscr{R}_1) = 3\text{diam}(\mathscr{R}_2) \not\Rightarrow \text{“}\mathscr{R}_1 \text{ is three times more uncertain than } \mathscr{R}_2\text{”}$$

Research Question, Explained

- CP Regions depend only on the *ranks* of the non-conformity scores: for any strictly increasing transform f, $\kappa_\alpha(y^n, \psi) = \kappa_\alpha(y^n, f \circ \psi)$
- $\mathrm{diam}(\mathscr{R})$ is numerically well-defined once a metric on Y is fixed, but: no canonical cardinal scale of "amount of uncertainty" attached to $\mathscr{R}$,

$$\mathrm{diam}(\mathscr{R}_1) = 3\mathrm{diam}(\mathscr{R}_2) \not\Rightarrow \text{"}\mathscr{R}_1 \text{ is three times more uncertain than } \mathscr{R}_2\text{"}$$

- CP natively supports *ordinal* comparisons (larger/smaller regions), not cardinal arithmetic

The Conformal Transducer Induces a Credal Set

- Call Δ_Y the space of finitely additive probabilities on Y, endowed with the weak* topology
- A credal set $\mathcal{M}$ is a (nonempty) weak*-closed and convex subset of Δ_Y (Levi, 1980)

The Conformal Transducer Induces a Credal Set

- Call Δ_Y the space of finitely additive probabilities on Y, endowed with the weak* topology
- A credal set $\mathcal{M}$ is a (nonempty) weak*-closed and convex subset of Δ_Y (Levi, 1980)
- Cella and Martin (2022) show that, under the consonance assumption, i.e. $\sup_{y_{n+1} \in Y} \pi(y_{n+1}, y^n) = 1$, then conformal transducer π induces

$$\overline{\Pi}(A) := \sup_{y_{n+1} \in A} \pi(y_{n+1}, y^n), \quad \forall A \in \Sigma_Y,$$

and in turn a credal set

$$\mathcal{M}(\overline{\Pi}) := \{P : P(A) \leq \overline{\Pi}(A), \forall A \in \Sigma_Y\}$$

Credal Set Derivation as a Correspondence

- Credal set $\mathcal{M}(\overline{\Pi})$ can be seen as the image of the correspondence $\text{CRED} : Y^n \times \mathscr{F} \rightrightarrows \Delta_Y$, $(y^n, \psi) \mapsto \text{CRED}(y^n, \psi) :=$

$$\left\{ P \in \Delta_Y : P(A) \leq \underbrace{\sup_{y_{n+1} \in A} \left(\frac{1}{n+1} \sum_{i=1}^{n+1} \mathbb{1} \left[\psi \left(y_{-i}^{n+1}, y_i \right) \geq \psi \left(y_{-(n+1)}^{n+1}, y_{n+1} \right) \right] \right)}_{\equiv \sup_{y_{n+1} \in A} \pi(y_{n+1}, y^n) =: \overline{\Pi}(A)}, \forall A \in \Sigma_Y \right\}.$$

Credal Set Derivation as a Correspondence

- Credal set $\mathcal{M}(\overline{\Pi})$ can be seen as the image of the correspondence $\text{CRED} : Y^n \times \mathscr{F} \rightrightarrows \Delta_Y$, $(y^n, \psi) \mapsto \text{CRED}(y^n, \psi) :=$

$$\left\{ P \in \Delta_Y : P(A) \leq \underbrace{\sup_{y_{n+1} \in A} \left(\frac{1}{n+1} \sum_{i=1}^{n+1} \mathbb{1}\left[\psi\left(y_{-i}^{n+1}, y_i\right) \geq \psi\left(y_{-(n+1)}^{n+1}, y_{n+1}\right)\right]\right)}_{\equiv \sup_{y_{n+1} \in A} \pi(y_{n+1}, y^n) =: \overline{\Pi}(A)}, \forall A \in \Sigma_Y \right\}.$$

- The elements of $\text{CRED}(y^n, \psi)$ are characterized in Martin (2025, Theorem 1).

Imprecise Highest Density Region (IHDR)

- We define (a version of) a confidence interval for $\mathcal{M}(\overline{\Pi})$

Imprecise Highest Density Region (IHDR)

- We define (a version of) a confidence interval for $\mathcal{M}(\overline{\Pi})$
- Conjugate Lower Probability: $\underline{\Pi}(A) = 1 - \overline{\Pi}(A^c)$, for all $A \in \Sigma_Y$

Imprecise Highest Density Region (IHDR)

- We define (a version of) a confidence interval for $\mathcal{M}(\overline{\Pi})$
- Conjugate Lower Probability: $\underline{\Pi}(A) = 1 - \overline{\Pi}(A^c)$, for all $A \in \Sigma_Y$
- Fix a credibility level $\alpha \in [0, 1]$. The α-level IHDR of credal set $\mathcal{M}(\overline{\Pi})$ is the set $B^\alpha_\mathcal{M} \subset Y$ such that $\underline{\Pi}(B^\alpha_\mathcal{M}) = 1 - \alpha$, and its "size is minimal" (Coolen, 1992; Caprio et al., 2025),

$$B^\alpha_\mathcal{M} = \bigcap \{A \in \Sigma_Y : \underline{\Pi}(A) \geq 1 - \alpha\}$$

Imprecise Highest Density Region (IHDR)

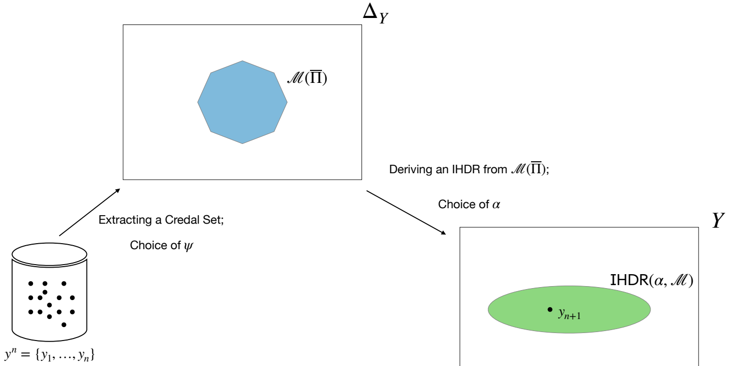
- We define (a version of) a confidence interval for $\mathcal{M}(\overline{\Pi})$
- Conjugate Lower Probability: $\underline{\Pi}(A) = 1 - \overline{\Pi}(A^c)$, for all $A \in \Sigma_Y$
- Fix a credibility level $\alpha \in [0, 1]$. The α-level IHDR of credal set $\mathcal{M}(\overline{\Pi})$ is the set $B^\alpha_\mathcal{M} \subset Y$ such that $\underline{\Pi}(B^\alpha_\mathcal{M}) = 1 - \alpha$, and its "size is minimal" (Coolen, 1992; Caprio et al., 2025),

$$B^\alpha_\mathcal{M} = \bigcap \{A \in \Sigma_Y : \underline{\Pi}(A) \geq 1 - \alpha\}$$

- The IHDR $B^\alpha_\mathcal{M}$ can be written as the image of the following correspondence $\text{IHDR} : [0, 1] \times \mathscr{C} \rightrightarrows Y$,

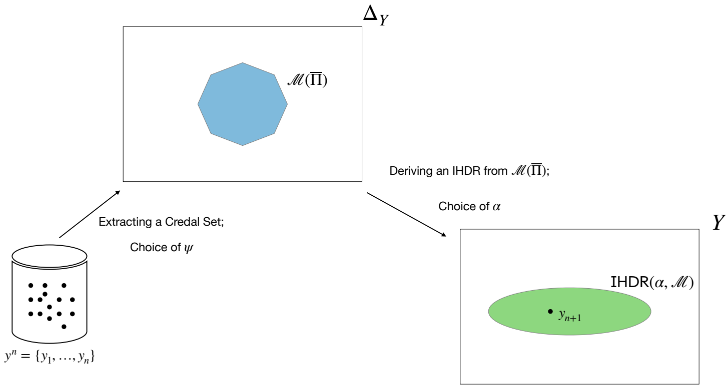
$$(\alpha, \mathcal{M}(\overline{\Pi})) \mapsto \text{IHDR}(\alpha, \mathcal{M}(\overline{\Pi})) := \bigcap \{A \in \Sigma_Y : \underline{\Pi}(A) \geq 1 - \alpha\}$$

IHDR is Equivalent to CPR



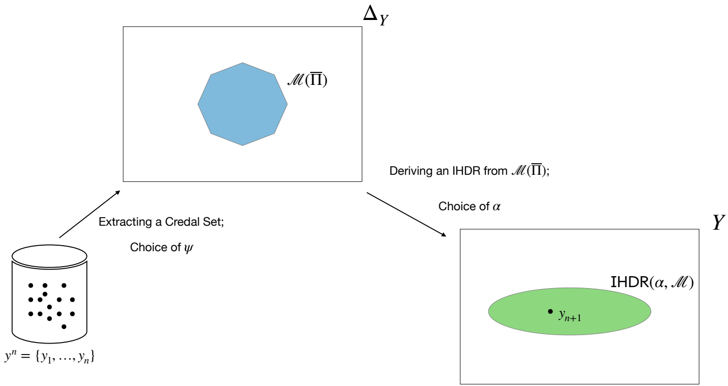
- Caprio et al. (2025, Propositions 5 and 6): the α-level IHDR associated with $\mathcal{M}(\overline{\Pi})$ corresponds to the Conformal Prediction Region $\mathscr{R}_\alpha^\psi(y^n)$
- It enjoys the same probabilistic guarantee as the CPR

Addressing the Research Question



- Is this "indirect route" to obtaining a prediction region the one that gives us cardinal UQ capabilities for Conformal Prediction?

Addressing the Research Question



- Is this "indirect route" to obtaining a prediction region the one that gives us cardinal UQ capabilities for Conformal Prediction?
  - We use a Category-Theoretic approach to answer this question

Conformal Prediction as a Morphism

- Category Theory (CT) can be thought of as a general theory of mathematical objects and their relations (Perrone, 2024)

Structure **UHCont**

Define a structure **UHCont** as follows,

- **Objects:** Topological spaces (X, τ_X).
- **Morphisms:** Upper hemicontinuous correspondences $\Phi : X \rightrightarrows Y$.

Conformal Prediction as a Morphism

- Category Theory (CT) can be thought of as a general theory of mathematical objects and their relations (Perrone, 2024)

Structure UHCont

Define a structure **UHCont** as follows,

- **Objects:** Topological spaces (X, τ_X).
- **Morphisms:** Upper hemicontinuous correspondences $\Phi : X \rightrightarrows Y$.

- **UHCont** is a well-defined category (Caprio, 2025, Theorem 2)
  - That is, a structure satisfying desirable properties

Conformal Prediction as a Morphism

- Category Theory (CT) can be thought of as a general theory of mathematical objects and their relations (Perrone, 2024)

Structure **UHCont**

Define a structure **UHCont** as follows,

- **Objects:** Topological spaces (X, τ_X).
- **Morphisms:** Upper hemicontinuous correspondences $\Phi : X \rightrightarrows Y$.

- **UHCont** is a well-defined category (Caprio, 2025, Theorem 2)
  - That is, a structure satisfying desirable properties
- Under minimal assumptions, the CP correspondence κ is a morphism of **UHCont**,

Conformal Prediction as a Morphism

- Category Theory (CT) can be thought of as a general theory of mathematical objects and their relations (Perrone, 2024)

Structure UHCont

Define a structure **UHCont** as follows,

- **Objects:** Topological spaces (X, τ_X).
- **Morphisms:** Upper hemicontinuous correspondences $\Phi : X \rightrightarrows Y$.

- **UHCont** is a well-defined category (Caprio, 2025, Theorem 2)
  - That is, a structure satisfying desirable properties
- Under minimal assumptions, the CP correspondence κ is a morphism of **UHCont**, and it is part of a commutative diagram involving the IHDR and the credal set correspondences

Conformal Prediction as a Morphism

(i) Y is compact Hausdorff and $\Sigma_Y = \mathcal{B}(Y)$ is its Borel σ-field

Conformal Prediction as a Morphism

(i) Y is compact Hausdorff and $\Sigma_Y = \mathcal{B}(Y)$ is its Borel σ-field

(ii) Each $\psi \in \mathscr{F}$ is jointly continuous on $Y^n \times Y$, i.e. $\mathscr{F} \subset C(Y^n \times Y)$, and is endowed with the uniform topology

Conformal Prediction as a Morphism

(i) Y is compact Hausdorff and $\Sigma_Y = \mathcal{B}(Y)$ is its Borel σ-field

(ii) Each $\psi \in \mathscr{F}$ is jointly continuous on $Y^n \times Y$, i.e. $\mathscr{F} \subset C(Y^n \times Y)$, and is endowed with the uniform topology

(iii) *No-tie parameter:* The credibility level α satisfies $\alpha \notin S_{n+1} := \{0, \frac{1}{n+1}, \frac{2}{n+1}, \dots, \frac{n}{n+1}, 1\}$

Conformal Prediction as a Morphism

(i) Y is compact Hausdorff and $\Sigma_Y = \mathcal{B}(Y)$ is its Borel σ-field

(ii) Each $\psi \in \mathscr{F}$ is jointly continuous on $Y^n \times Y$, i.e. $\mathscr{F} \subset C(Y^n \times Y)$, and is endowed with the uniform topology

(iii) *No-tie parameter:* The credibility level α satisfies $\alpha \notin S_{n+1} := \{0, \frac{1}{n+1}, \frac{2}{n+1}, \dots, \frac{n}{n+1}, 1\}$

(iv) Δ_Y is endowed with the weak* topology, and $\mathscr{C} \subset 2^{\Delta_Y}$ is endowed with the Vietoris topology

Conformal Prediction as a Morphism

(i) Y is compact Hausdorff and $\Sigma_Y = \mathcal{B}(Y)$ is its Borel σ-field

(ii) Each $\psi \in \mathscr{F}$ is jointly continuous on $Y^n \times Y$, i.e. $\mathscr{F} \subset C(Y^n \times Y)$, and is endowed with the uniform topology

(iii) *No-tie parameter:* The credibility level α satisfies $\alpha \notin S_{n+1} := \{0, \frac{1}{n+1}, \frac{2}{n+1}, \dots, \frac{n}{n+1}, 1\}$

(iv) Δ_Y is endowed with the weak* topology, and $\mathscr{C} \subset 2^{\Delta_Y}$ is endowed with the Vietoris topology

The Conformal Prediction Diagram Commutes, UHCont

Assume that (i)-(iv) hold, and that the conformal transducer π is *consonant*.

Conformal Prediction as a Morphism

(i) Y is compact Hausdorff and $\Sigma_Y = \mathcal{B}(Y)$ is its Borel σ-field

(ii) Each $\psi \in \mathscr{F}$ is jointly continuous on $Y^n \times Y$, i.e. $\mathscr{F} \subset C(Y^n \times Y)$, and is endowed with the uniform topology

(iii) *No-tie parameter:* The credibility level α satisfies $\alpha \notin S_{n+1} := \{0, \frac{1}{n+1}, \frac{2}{n+1}, \ldots, \frac{n}{n+1}, 1\}$

(iv) Δ_Y is endowed with the weak* topology, and $\mathscr{C} \subset 2^{\Delta_Y}$ is endowed with the Vietoris topology

The Conformal Prediction Diagram Commutes, UHCont

Assume that (i)-(iv) hold, and that the conformal transducer π is *consonant*. Consider the restrictions κ_α and IHDR_α to any $\alpha \in [0,1]$ that satisfies (iii).

Conformal Prediction as a Morphism

(i) Y is compact Hausdorff and $\Sigma_Y = \mathcal{B}(Y)$ is its Borel σ-field

(ii) Each $\psi \in \mathscr{F}$ is jointly continuous on $Y^n \times Y$, i.e. $\mathscr{F} \subset C(Y^n \times Y)$, and is endowed with the uniform topology

(iii) *No-tie parameter:* The credibility level α satisfies $\alpha \notin S_{n+1} := \{0, \frac{1}{n+1}, \frac{2}{n+1}, \ldots, \frac{n}{n+1}, 1\}$

(iv) Δ_Y is endowed with the weak* topology, and $\mathscr{C} \subset 2^{\Delta_Y}$ is endowed with the Vietoris topology

The Conformal Prediction Diagram Commutes, UHCont

Assume that (i)-(iv) hold, and that the conformal transducer π is *consonant*. Consider the restrictions κ_α and IHDR_α to any $\alpha \in [0,1]$ that satisfies (iii). Then, $\text{IHDR}_\alpha \circ \text{CRED} = \kappa_\alpha$. That is, the *Full CP Diagram* commutes in **UHCont**.

Conformal Prediction as a Morphism

(i) Y is compact Hausdorff and $\Sigma_Y = \mathcal{B}(Y)$ is its Borel σ-field

(ii) Each $\psi \in \mathscr{F}$ is jointly continuous on $Y^n \times Y$, i.e. $\mathscr{F} \subset C(Y^n \times Y)$, and is endowed with the uniform topology

(iii) *No-tie parameter:* The credibility level α satisfies $\alpha \notin S_{n+1} := \{0, \frac{1}{n+1}, \frac{2}{n+1}, \dots, \frac{n}{n+1}, 1\}$

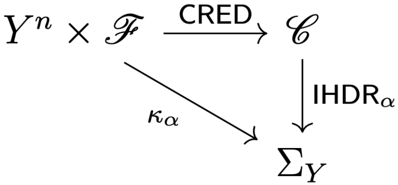
(iv) Δ_Y is endowed with the weak* topology, and $\mathscr{C} \subset 2^{\Delta_Y}$ is endowed with the Vietoris topology

The Conformal Prediction Diagram Commutes, UHCont

Assume that (i)-(iv) hold, and that the conformal transducer π is *consonant*. Consider the restrictions κ_α and IHDR_α to any $\alpha \in [0,1]$ that satisfies (iii). Then, $\text{IHDR}_\alpha \circ \text{CRED} = \kappa_\alpha$. That is, the *Full CP Diagram* commutes in **UHCont**.

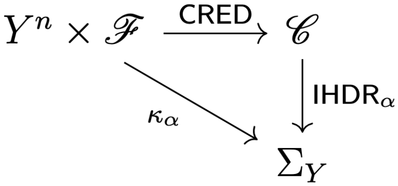
- Conditions (i)-(iv) + consonance are truly minimal: without any of them, the result does not hold

Answering the Research Question



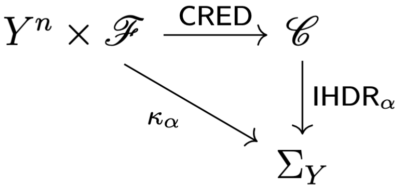
- Commuting factorization $\kappa_\alpha = \text{IHDR}_\alpha \circ \text{CRED}$ reveals hidden layer
- Computing a CP Region is equivalent to (i) forming the credal set $\mathcal{M}(\overline{\Pi})$ and (ii) extracting its IHDR

Answering the Research Question



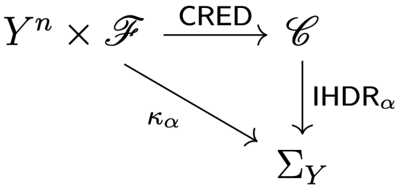
- Commuting factorization $\kappa_\alpha = \mathrm{IHDR}_\alpha \circ \mathrm{CRED}$ reveals hidden layer
- Computing a CP Region is equivalent to (i) forming the credal set $\mathcal{M}(\overline{\Pi})$ and (ii) extracting its IHDR
  - The result in Caprio et al. (2025) that CPR $=$ IHDR was not enough to conclude the intrinsic equality between the two methods of obtaining the prediction region

Answering the Research Question



- Commuting factorization $\kappa_\alpha = \text{IHDR}_\alpha \circ \text{CRED}$ reveals hidden layer
- Computing a CP Region is equivalent to (i) forming the credal set $\mathcal{M}(\overline{\Pi})$ and (ii) extracting its IHDR
  - The result in Caprio et al. (2025) that CPR = IHDR was not enough to conclude the intrinsic equality between the two methods of obtaining the prediction region
- Because $\mathcal{M}(\overline{\Pi})$ is a credal set, existing metrics (Abellán et al., 2006; Hüllermeier and Waegeman, 2021; Chau et al., 2025) endow it with a cardinal scale

Answering the Research Question



- Commuting factorization $\kappa_\alpha = \mathrm{IHDR}_\alpha \circ \mathrm{CRED}$ reveals hidden layer
- Computing a CP Region is equivalent to (i) forming the credal set $\mathcal{M}(\overline{\Pi})$ and (ii) extracting its IHDR
  - The result in Caprio et al. (2025) that CPR = IHDR was not enough to conclude the intrinsic equality between the two methods of obtaining the prediction region
- Because $\mathcal{M}(\overline{\Pi})$ is a credal set, existing metrics (Abellán et al., 2006; Hüllermeier and Waegeman, 2021; Chau et al., 2025) endow it with a cardinal scale that is able to quantify (different types of) uncertainty, in particular, reducible (epistemic) and irreducible (aleatoric) (Javanmardi et al., 2025; Cabezas et al., 2025; Stutts et al., 2024)

Answering the Research Question

- The uncertainty will be "marginal",[1] thus reflecting the marginal guarantee of the Conformal Prediction Region

[1] That is, averaged over the possible realization y^n.

Answering the Research Question

- The uncertainty will be "marginal",[1] thus reflecting the marginal guarantee of the Conformal Prediction Region
- Conjecture: The analyst-controlled choices – α and $\psi \in \mathscr{F}$ – are primarily associated with epistemic uncertainty
- Conjecture (cont'd): aleatoric uncertainty stems from the intrinsic variability of the exchangeable process that we study, encoded in $\mathfrak{P}$

[1] That is, averaged over the possible realization y^n.

An ML-Inspired Example

- Let $Y = [0,1]^d$ with its Euclidean topology (compact and Hausdorff), and let $\varphi_\theta \colon Y \to \mathbb{R}^m$ be a feed-forward neural network whose activation functions are continuous (e.g. ReLU)
  - Hence, φ_θ is continuous in both its parameters θ and input y

An ML-Inspired Example

- Let $Y = [0,1]^d$ with its Euclidean topology (compact and Hausdorff), and let $\varphi_\theta \colon Y \to \mathbb{R}^m$ be a feed-forward neural network whose activation functions are continuous (e.g. ReLU)
  - Hence, φ_θ is continuous in both its parameters θ and input y
- For a training set $y^n = (y_1, \ldots, y_n) \in Y^n$ and a candidate point $\tilde{y} \in Y$, define

$$\psi_\theta(y^n, \tilde{y}) = -\left\| \varphi_\theta(\tilde{y}) - \frac{1}{n} \sum_{i=1}^{n} \varphi_\theta(y_i) \right\|_2^2$$

An ML-Inspired Example

- Let $Y = [0,1]^d$ with its Euclidean topology (compact and Hausdorff), and let $\varphi_\theta \colon Y \to \mathbb{R}^m$ be a feed-forward neural network whose activation functions are continuous (e.g. ReLU)
  - Hence, φ_θ is continuous in both its parameters θ and input y
- For a training set $y^n = (y_1, \dots, y_n) \in Y^n$ and a candidate point $\tilde{y} \in Y$, define

$$\psi_\theta(y^n, \tilde{y}) = -\left\| \varphi_\theta(\tilde{y}) - \frac{1}{n} \sum_{i=1}^{n} \varphi_\theta(y_i) \right\|_2^2$$

- *Permutation invariance:* The sample mean is symmetric in $y_1, \dots, y_n$
- *Joint continuity:* Composition, finite sums, Euclidean norm, and squaring preserve continuity; Thus $(y^n, \tilde{y}) \mapsto \psi_\theta(y^n, \tilde{y})$ is jointly continuous on $Y^n \times Y$. Hence, $\psi_\theta \in \mathscr{F}$

An ML-Inspired Example (cont'd)

- Consonance does not automatically hold in this example. To enforce it, it is enough to consider the normalized conformal transducer $\pi'(\tilde{y}, y^n) := \frac{\pi(\tilde{y}, y^n)}{\sup_{y \in Y} \pi(y, y^n)}$ (Cella and Martin, 2021, Section 7)

An ML-Inspired Example (cont'd)

- Consonance does not automatically hold in this example. To enforce it, it is enough to consider the normalized conformal transducer $\pi'(\tilde{y}, y^n) := \frac{\pi(\tilde{y}, y^n)}{\sup_{y \in Y} \pi(y, y^n)}$ (Cella and Martin, 2021, Section 7)
- This example is interesting because of the following implications for Machine Learning
  - *Few shot/prototype learning:* $\psi_\theta(y^n, y)$ is the squared Euclidean distance to the class prototype mean $\varphi_\theta(y_i)$

An ML-Inspired Example (cont'd)

- Consonance does not automatically hold in this example. To enforce it, it is enough to consider the normalized conformal transducer $\pi'(\tilde{y}, y^n) := \frac{\pi(\tilde{y}, y^n)}{\sup_{y \in Y} \pi(y, y^n)}$ (Cella and Martin, 2021, Section 7)
- This example is interesting because of the following implications for Machine Learning
  - *Few shot/prototype learning:* $\psi_\theta(y^n, y)$ is the squared Euclidean distance to the class prototype mean $\varphi_\theta(y_i)$
  - *Out-Of-Distribution (OOD) detection:* The same distance acts as a learned Mahalanobis-style OOD score

An ML-Inspired Example (cont'd)

- Consonance does not automatically hold in this example. To enforce it, it is enough to consider the normalized conformal transducer $\pi'(\tilde{y}, y^n) := \frac{\pi(\tilde{y}, y^n)}{\sup_{y \in Y} \pi(y, y^n)}$ (Cella and Martin, 2021, Section 7)
- This example is interesting because of the following implications for Machine Learning
  - *Few shot/prototype learning:* $\psi_\theta(y^n, y)$ is the squared Euclidean distance to the class prototype mean $\varphi_\theta(y_i)$
  - *Out-Of-Distribution (OOD) detection:* The same distance acts as a learned Mahalanobis-style OOD score
  - *Heteroscedastic conformal regression:* If φ_θ is the penultimate layer of a regression network, $\psi_\theta(y^n, y)$ yields locally adaptive, yet still valid, prediction sets

Full CP Diagram commuting in a Different Category

- The following is a category (Caprio, 2025, Theorem 23)

Structure $\mathbf{WMeas}_{\text{uc}}$

Define a structure $\mathbf{WMeas}_{\text{uc}}$ as follows,

- **Objects:** Compact Polish spaces (X, Σ_X), with $\Sigma_X = \mathcal{B}(X)$, the Borel σ-algebra.
- **Morphisms:** Weakly measurable, uniformly compact-valued correspondences $\Phi : X \rightrightarrows Y$.

Full CP Diagram commuting in a Different Category

- The following is a category (Caprio, 2025, Theorem 23)

Structure $\mathbf{WMeas}_{\text{uc}}$

Define a structure $\mathbf{WMeas}_{\text{uc}}$ as follows,

- **Objects:** Compact Polish spaces (X, Σ_X), with $\Sigma_X = \mathcal{B}(X)$, the Borel σ-algebra.
- **Morphisms:** Weakly measurable, uniformly compact-valued correspondences $\Phi : X \rightrightarrows Y$.

- The Full CP Diagram commutes in $\mathbf{WMeas}_{\text{uc}}$ as well, under essentially the same conditions (i)-(iv) + consonance (Caprio, 2025, Theorem 24)

Full CP Diagram commuting in a Different Category

- The following is a category (Caprio, 2025, Theorem 23)

Structure $\mathbf{WMeas}_{\text{uc}}$

Define a structure $\mathbf{WMeas}_{\text{uc}}$ as follows,

- **Objects:** Compact Polish spaces (X, Σ_X), with $\Sigma_X = \mathcal{B}(X)$, the Borel σ-algebra.
- **Morphisms:** Weakly measurable, uniformly compact-valued correspondences $\Phi : X \rightrightarrows Y$.

- The Full CP Diagram commutes in $\mathbf{WMeas}_{\text{uc}}$ as well, under essentially the same conditions (i)-(iv) + consonance (Caprio, 2025, Theorem 24)
- Under our assumptions, cardinal uncertainty quantification is an intrinsic feature of Conformal Prediction methods
  - The Full CP diagram commutes when we focus on both measurability and continuity aspects of the CP correspondence κ

An Unexpected Byproduct

- Define

$$\text{BCP} : Y^n \times \Delta_\Theta^{\text{dens}} \to Y^n \times C(Y^n \times Y), \quad \text{BCP}\big(y^n, P\big) := \big(y^n, \psi_{(y^n,P)}\big),$$

where

$$\psi_{(y^n,P)}(y_{n+1}) := -p_P(y_{n+1} \mid y^n)$$

An Unexpected Byproduct

- Define

$$\text{BCP}: Y^n \times \Delta_\Theta^{\text{dens}} \to Y^n \times C(Y^n \times Y), \quad \text{BCP}\big(y^n, P\big) := \big(y^n, \psi_{(y^n,P)}\big),$$

where

$$\psi_{(y^n,P)}(y_{n+1}) := -p_P(y_{n+1} \mid y^n)$$

- We write BCP: an acronym for Bayesian Conformal Prediction (Fong and Holmes, 2021)

An Unexpected Byproduct

- Call QUANT : $(0,1) \times Y^n \times \Delta_{\Theta}^{\text{dens}} \rightrightarrows Y$ the set-valued map that extracts the α-level set of the posterior predictive distribution,

$$(\alpha, y^n, P) \mapsto \text{QUANT}(\alpha, y^n, P) \coloneqq H\big(c(\alpha, y^n, P), y^n, P\big),$$

where

$$H : \mathbb{R} \times Y^n \times \Delta_{\Theta}^{\text{dens}} \rightrightarrows Y, \qquad H\big(c, y^n, P\big) \coloneqq \{y \in Y : p_P(y \mid y^n) \geq c\}$$

An Unexpected Byproduct

- Assume that (i)-(iv),

An Unexpected Byproduct

- Assume that (i)-(iv), consonance,

An Unexpected Byproduct

- Assume that (i)-(iv), consonance, and the usual minimal technical machinery to ensure the posterior predictive is well defined and depends continuously on both the data y^n and the prior P, hold (Caprio, 2025, Conditions (v)-(x))

An Unexpected Byproduct

- Assume that (i)-(iv), consonance, and the usual minimal technical machinery to ensure the posterior predictive is well defined and depends continuously on both the data y^n and the prior P, hold (Caprio, 2025, Conditions (v)-(x))

Unifying Bayes, Conformal, and Imprecise Prediction in **UHCont**

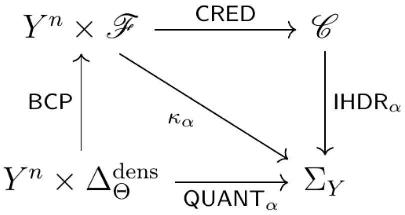
Fix any $\alpha \in (0,1)$ and define $\beta_n(\alpha) := \min\{k/(n+1) : k/(n+1) > \alpha\}$. Then,

$$d_H\left(\kappa_\alpha(y^n, \psi_{(y^n,P)}), \mathrm{QUANT}_\alpha(y^n, P)\right) \xrightarrow{\mathfrak{P}} 0.$$

Consequently, since $\mathrm{IHDR}_\alpha \circ \mathrm{CRED} = \kappa_\alpha$ under consonance,

$$d_H\left(\mathrm{IHDR}_\alpha\left(\mathrm{CRED}(y^n, \psi_{(y^n,P)})\right), \mathrm{QUANT}_\alpha(y^n, P)\right) \xrightarrow{\mathfrak{P}} 0.$$

An Unexpected Byproduct



- Equivalently, the diagram above *commutes asymptotically in probability* in **UHCont** (in the Hausdorff distance d_H)
  - If Y is not compact metric, a similar result can be proved with convergence taken in the Vietoris topology on closed subsets of Y

Discussing the Result

- Theorem above is the central payoff of our categorical treatment
- It confirms intuition in Martin (2022) and Caprio et al. (2025, Section 5.2): CP is a bridge between the Bayesian, frequentist, and imprecise approaches to predictive statistical reasoning

Discussing the Result

- Theorem above is the central payoff of our categorical treatment
- It confirms intuition in Martin (2022) and Caprio et al. (2025, Section 5.2): CP is a bridge between the Bayesian, frequentist, and imprecise approaches to predictive statistical reasoning
- When seen as u.h.c. correspondences,

Discussing the Result

- Theorem above is the central payoff of our categorical treatment
- It confirms intuition in Martin (2022) and Caprio et al. (2025, Section 5.2): CP is a bridge between the Bayesian, frequentist, and imprecise approaches to predictive statistical reasoning
- When seen as u.h.c. correspondences, the **model-based** Bayesian and imprecise methods,

Discussing the Result

- Theorem above is the central payoff of our categorical treatment
- It confirms intuition in Martin (2022) and Caprio et al. (2025, Section 5.2): CP is a bridge between the Bayesian, frequentist, and imprecise approaches to predictive statistical reasoning
- When seen as u.h.c. correspondences, the **model-based** Bayesian and imprecise methods, and the **model-free** conformal construction, yield (asymptotically) the same α-level prediction region

Discussing the Result

- Theorem above is the central payoff of our categorical treatment
- It confirms intuition in Martin (2022) and Caprio et al. (2025, Section 5.2): CP is a bridge between the Bayesian, frequentist, and imprecise approaches to predictive statistical reasoning
- When seen as u.h.c. correspondences, the **model-based** Bayesian and imprecise methods, and the **model-free** conformal construction, yield (asymptotically) the same α-level prediction region
  - A very profound connections between these three only apparently far apart prediction mechanisms

Further Results from the Paper

- Under a minimal condition, an upper bound to the Generalized Bayes' Upper Posterior is an e-posterior in the sense of Grünwald (2023)

Further Results from the Paper

- Under a minimal condition, an upper bound to the Generalized Bayes' Upper Posterior is an e-posterior in the sense of Grünwald (2023)
- We can write the IHDR, and in turn the CPR, as a fuctor between two well-defined categories, $\mathscr{C}$ and Σ_Y with inclusion morphisms "$\subseteq$"
    - This has interesting AI privacy implications (Caprio, 2025, Section 5)

Further Results from the Paper

- Under a minimal condition, an upper bound to the Generalized Bayes' Upper Posterior is an e-posterior in the sense of Grünwald (2023)
- We can write the IHDR, and in turn the CPR, as a fuctor between two well-defined categories, $\mathscr{C}$ and Σ_Y with inclusion morphisms "$\subseteq$"
  - This has interesting AI privacy implications (Caprio, 2025, Section 5)
- We sudy the properties (monads, functors) of **UHCont** and $\textbf{WMeas}_{\text{uc}}$ (Caprio, 2025, Section 6)

Conclusion

- In this paper, we took a Category Theory route to show that Full Conformal Prediction exhibits intrinsic cardinal UQ capabilities
- As byproduct, we were able to prove that CP methods bridge Bayesian, frequentist, and imprecise approaches to statistical prediction

Conclusion

- In this paper, we took a Category Theory route to show that Full Conformal Prediction exhibits intrinsic cardinal UQ capabilities
- As byproduct, we were able to prove that CP methods bridge Bayesian, frequentist, and imprecise approaches to statistical prediction
- We open up the study of Categorical Conformal Prediction; many unsolved problems
    - How to leverage the structure of **UHCont** and $\mathbf{WMeas}_{\mathrm{uc}}$ to learn further aspects of the CP methodology?

Conclusion

- In this paper, we took a Category Theory route to show that Full Conformal Prediction exhibits intrinsic cardinal UQ capabilities
- As byproduct, we were able to prove that CP methods bridge Bayesian, frequentist, and imprecise approaches to statistical prediction
- We open up the study of Categorical Conformal Prediction; many unsolved problems
  - How to leverage the structure of **UHCont** and $\mathbf{WMeas}_{\text{uc}}$ to learn further aspects of the CP methodology?
  - The morphisms of these categories are correspondences:

Conclusion

- In this paper, we took a Category Theory route to show that Full Conformal Prediction exhibits intrinsic cardinal UQ capabilities
- As byproduct, we were able to prove that CP methods bridge Bayesian, frequentist, and imprecise approaches to statistical prediction
- We open up the study of Categorical Conformal Prediction; many unsolved problems
  - How to leverage the structure of **UHCont** and $\mathbf{WMeas}_{uc}$ to learn further aspects of the CP methodology?
  - The morphisms of these categories are correspondences: what can the Kleisli categories (Kleisli, 1962, 1965), (Perrone, 2024, Section 5.1) arising from the subcategories of **UHCont** and $\mathbf{WMeas}_{uc}$ admitting a monad, tell us about Conformal Prediction?

Conclusion

- In this paper, we took a Category Theory route to show that Full Conformal Prediction exhibits intrinsic cardinal UQ capabilities
- As byproduct, we were able to prove that CP methods bridge Bayesian, frequentist, and imprecise approaches to statistical prediction
- We open up the study of Categorical Conformal Prediction; many unsolved problems
  - How to leverage the structure of **UHCont** and $\mathbf{WMeas}_{\text{uc}}$ to learn further aspects of the CP methodology?
  - The morphisms of these categories are correspondences: what can the Kleisli categories (Kleisli, 1962, 1965), (Perrone, 2024, Section 5.1) arising from the subcategories of **UHCont** and $\mathbf{WMeas}_{\text{uc}}$ admitting a monad, tell us about Conformal Prediction?
- CP as a morphism of different categories so to remove the need for assumptions altogether, in particular compact state space Y

Carthago Delenda Est

THANK YOU FOR YOUR ATTENTION!

EIML | EPISTEMIC INTELLIGENCE IN MACHINE LEARNING

Still think Epistemic Uncertainty is just error bars?

*Join us at the EIML workshop @ EurIPS 2025; where we bring together researchers to explore **foundational**, **methodological**, and **practical** questions around Epistemic Uncertainty in machine learning!*

brought to you by:


Michele Caprio (Manchester UK)


Siu Lun Chau (NTU Singapore)


Ruobin Gong (Rutgers US)


Shireen Manchingal (Oxford Brookes UK)


Krikamol Muandet (CISPA Germany)


Bob Williamson (Tubingen Germany)

AI Privacy Implications of IHDR Functoriality

- Suppose that we have many sources producing credal sets $\mathcal{M}_j$

AI Privacy Implications of IHDR Functoriality

- Suppose that we have many sources producing credal sets $\mathcal{M}_j$
- Differentially-private transformations applied to each $\mathcal{M}_j$ propagates through the functor without violating the coverage guarantee
  - Statistical validity is preserved "for free"

AI Privacy Implications of IHDR Functoriality

- Suppose that we have many sources producing credal sets $\mathcal{M}_j$
- Differentially-private transformations applied to each $\mathcal{M}_j$ propagates through the functor without violating the coverage guarantee
  - Statistical validity is preserved "for free"
- Collaboration can proceed in a strictly federated manner
  - Agents needs to transmit only (possibly privatized) summary object $\mathcal{M}_j$, never the underlying data

AI Privacy Implications of IHDR Functoriality

- Suppose that we have many sources producing credal sets $\mathcal{M}_j$
- Differentially-private transformations applied to each $\mathcal{M}_j$ propagates through the functor without violating the coverage guarantee
  - Statistical validity is preserved "for free"
- Collaboration can proceed in a strictly federated manner
  - Agents needs to transmit only (possibly privatized) summary object $\mathcal{M}_j$, never the underlying data
  - Minimizing exposure while still enabling the consortium to construct globally valid prediction region

AI Privacy Implications of IHDR Functoriality: Caveat

- The functorial guarantee is monotone in the credal set

AI Privacy Implications of IHDR Functoriality: Caveat

- The functorial guarantee is monotone in the credal set
- To preserve coverage after privatization, each site should output a *superset* $\widetilde{\mathcal{M}}_j \supseteq \mathcal{M}_j$

AI Privacy Implications of IHDR Functoriality: Caveat

- The functorial guarantee is monotone in the credal set
- To preserve coverage after privatization, each site should output a *superset* $\widetilde{\mathcal{M}}_j \supseteq \mathcal{M}_j$
- For federation, aggregation should be performed via the closed convex hull $\overline{\mathrm{co}}(\cdot)$,

$$\mathcal{M}_{\mathrm{agg}} = \overline{\mathrm{co}}\Big(\bigcup_j \widetilde{\mathcal{M}}_j\Big),$$

which ensures $B^{\alpha}_{\widetilde{\mathcal{M}}_j} \subseteq B^{\alpha}_{\mathcal{M}_{\mathrm{agg}}}$, and thus global validity whenever the true law lies in some $\widetilde{\mathcal{M}}_j$

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.

Luben M. C. Cabezas, Vagner S. Santos, Thiago R. Ramos, and Rafael Izbicki. Epistemic uncertainty in conformal scores: A unified approach, 2025. URL `https://arxiv.org/abs/2502.06995`.

Michele Caprio. The joys of categorical conformal prediction. *ArXiV*, 2025.

Michele Caprio, Yusuf Sale, and Eyke Hüllermeier. Conformal Prediction Regions are Imprecise Highest Density Regions. *Accepted to ISIPTA 2025*, 2025. URL `https://arxiv.org/abs/2502.06331`.

Leonardo Cella and Ryan Martin. Valid inferential models for prediction in supervised learning problems. In *International Symposium on Imprecise Probability: Theories and Applications*, pages 72–82. PMLR, 2021.

References II

Leonardo Cella and Ryan Martin. Validity, consonant plausibility measures, and conformal prediction. *International Journal of Approximate Reasoning*, 141:110–130, 2022.

Siu Lun Chau, Michele Caprio, and Krikamol Muandet. Integral imprecise probability metrics, 2025. URL https://arxiv.org/abs/2505.16156.

Franciscus Petrus Antonius Coolen. Imprecise highest density regions related to intervals of measures. *Memorandum COSOR; Volume 9254*, 1992.

Edwin Fong and Chris C Holmes. Conformal bayesian computation. In M. Ranzato, A. Beygelzimer, Y. Dauphin, P.S. Liang, and J. Wortman Vaughan, editors, *Advances in Neural Information Processing Systems*, volume 34, pages 18268–18279. Curran Associates, Inc., 2021. URL https://proceedings.neurips.cc/paper_files/paper/2021/file/97785e0500ad16c18574c64189ccf4b4-Paper.pdf.

References III

Peter D Grünwald. The e-posterior. *Philosophical Transactions of the Royal Society A*, 381(2247):20220146, 2023.

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 2021. doi: 10.1007/s10994-021-05946-3. URL https://link.springer.com/article/10.1007/s10994-021-05946-3.

Alireza Javanmardi, Soroush H. Zargarbashi, Santo M. A. R. Thies, Willem Waegeman, Aleksandar Bojchevski, and Eyke Hüllermeier. Optimal conformal prediction under epistemic uncertainty, 2025. URL https://arxiv.org/abs/2505.19033.

Heinrich Kleisli. Homotopy theory in abelian categories. *Canadian Journal of Mathematics*, 14:139–169, 1962. doi: 10.4153/CJM-1962-011-x.

References IV

Heinrich Kleisli. Every standard construction is induced by a pair of adjoint functors. *Proceedings of the American Mathematical Society*, 16 (3):544–546, 1965. doi: 10.1090/S0002-9939-1965-0177024-4.

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

Ryan Martin. Valid and efficient imprecise-probabilistic inference with partial priors, ii. general framework. *arXiv preprint arXiv:2211.14567*, 2022.

Ryan Martin. An efficient monte carlo method for valid prior-free possibilistic statistical inference, 2025. URL `https://arxiv.org/abs/2501.10585`.

Paolo Perrone. *Starting Category Theory*. World Scientific, 2024.

References V

Alex Christopher Stutts, Divake Kumar, Theja Tulabandhula, and Amit Trivedi. Invited: Conformal inference meets evidential learning: Distribution-free uncertainty quantification with epistemic and aleatoric separability. In *Proceedings of the 61st ACM/IEEE Design Automation Conference*, DAC '24, New York, NY, USA, 2024. Association for Computing Machinery. ISBN 9798400706011. doi: 10.1145/3649329.3663512. URL https://doi.org/10.1145/3649329.3663512.

Vladimir Vovk. Transductive conformal predictors. In Harris Papadopoulos, Andreas S. Andreou, Lazaros Iliadis, and Ilias Maglogiannis, editors, *Artificial Intelligence Applications and Innovations*, pages 348–360, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg. ISBN 978-3-642-41142-7.

Vladimir Vovk, Alexander Gammerman, and Glenn Shafer. *Algorithmic learning in a random world*, volume 29. Springer, 2005.