Decision-making with possibilistic inferential models

Shih-Ni Prim, Jonathan P. Williams, and Ryan Martin

Information Processing and Management of Uncertainty in Knowledge-Based Systems
Rome, Italy, June 18, 2026

NC STATE UNIVERSITY

Table of contents

Section 1

Decision making

Setup: Inference and Uncertainty Quantification

- Let $X = (X_1, \cdots, X_n)$ denote the observable data.
- A statistical model $\{\mathrm{P}_\theta : \theta \in \mathbb{T}\}$ is introduced to quantify aleatory uncertainty.
- When $X = x$ is observed, $\theta \mapsto p_\theta(x)$ is the likelihood function.
- The "true value" of the unknown model parameter will be denoted by Θ.
- The goal, then, is to quantify uncertainty about Θ, given the observation $X = x$.

Decision-making

- Let $(\theta, a) \mapsto \ell_a(\theta)$ denote a loss function that measures the loss incurred by choosing the action $a \in \mathbb{A}$ when the parameter is θ.
- If the true parameter is known, $a^* = \arg\min_{a \in \mathbb{A}} \ell_a(\Theta)$.
- When the true parameter is unknown, we need a model to quantify uncertainty about Θ and to find the expected loss.
- Decision-making is essentially a ranking strategy—picking the action that **minimizes the expected loss**.
- Two important goals of decision-making are finding the optimal action and reliable uncertainty quantification for risk assessment.

Bayesian (probabilistic) decision-making

- Posterior comes from likelihood and prior:

$$\mathrm{Q}_x(d\theta) \propto p_\theta(x)\mathrm{Q}(d\theta).$$

- Expected loss is integrated from the posterior distribution. For an action a and loss function $\ell_a(\theta)$,

$$Q_x\ell_a = \int \ell_a(\theta)\mathrm{Q}_x(d\theta).$$

- The optimal action is usually the one minimizing the loss:

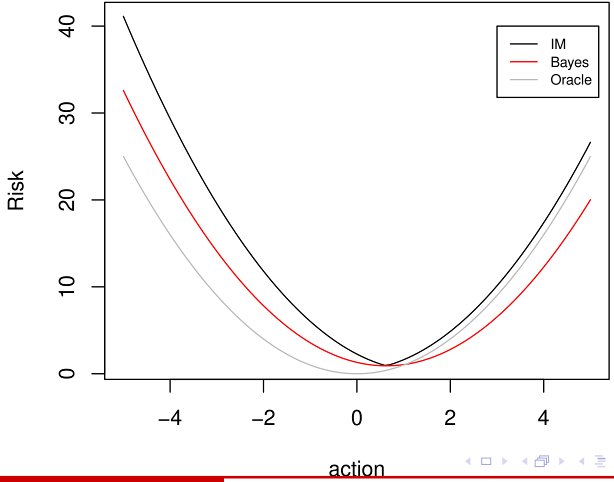
$$a^\star = \arg\min_a Q_x\ell_a.$$

Example: Normal distribution

- Data: $X^n \overset{\text{iid}}{\sim} N(\Theta, 1)$, assume squared loss
- Bayesian model:
  - Prior: $\Theta \sim N(0, \tau^2)$
  - Posterior: $Q_x \sim N\left(\theta_{\text{MAP}}, \frac{\tau^2}{1+n\tau^2}\right)$, where $\theta_{\text{MAP}} = \frac{n\tau^2}{1+n\tau^2}\overline{X}_n$.
  - Expected Bayes loss: $(\theta_{\text{MAP}} - a)^2 + \frac{\tau^2}{1+n\tau^2}$.
- Expected IM loss: $(\overline{X}_n - a)^2 + 2\left|(\overline{X}_n - a)\right| m(p) + v(p)$, where $m(p) = \int |z| p(z) dz$, $v(p) = var(Z)$, $Z \sim N(0, \frac{1}{n})$.
- Oracle loss: $(\Theta - a)^2$

Comparing risk assessment

- One random draw of $n = 10$ from Normal$(0, 1)$ and then calculate losses from some actions $a \in [-5, 5]$.



Problem of over-confidence

- Under-estimating risk is dangerous and can lead to systematic loss.
    - Betting and long-term loss in the sense of de Finetti's framework.
- This is related to the false confidence theorem:
    - Bayes solutions, including those based on default-priors, and fiducial solutions are unreliable in that they can assign high posterior probability to hypotheses that are false. (Balch et al. 2019).
    - The core problem is **additivity**.

How do we fix the problem?

- Additivity is the root cause of false confidence. We need **non-additivity**, which means going outside the probabilistic framework.
- We need some sense of **reliability** guarantee, i.e. **aleatory** uncertainty quantification.
- We will use **possibilistic inferential models**,[1] constructed in the possibilistic framework and not afflicted with false confidence.

[1] Martin & Liu (2013), Martin (2026)

Section 2

Possibilistic Inferential Models

Possibilistic Inferential Models (IMs)[3]

- Relative likelihood:

$$R(x,\theta) = \frac{p_\theta(x)}{\sup_\vartheta p_\vartheta(x)}, \text{ with } \sup_\theta R(x,\theta) = 1.$$

- Possibilistic contour function:

$$\pi_x(\theta) = P_\theta\{R(X,\theta) \le R(x,\theta)\}.$$

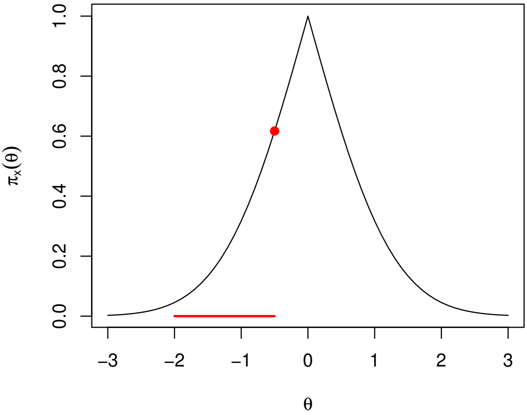
- Then the corresponding possibility measure,[2] or upper probability, is defined as

$$\Pi_x(H) = \sup_{\theta \in H} \pi_x(\theta), \quad H \subseteq \mathbb{T}.$$

[2] Dubois & Prade (1988)

[3] Martin & Liu (2013), Martin (2026)

Possibility contour function



- Possibilistic IMs can quantify epistemic uncertainty.
- For $H : \theta \in [-2, -0.5], \Pi_x(H) = \sup_{-2 \leq \theta \leq -0.5} \pi_x(\theta) \approx 0.617$.

Key properties

- The IM construction has the **validity** property:

$$\sup_{\Theta \in H} P_{\Theta}\{\Pi_X(H) \leq \alpha\} \leq \alpha, \quad \alpha \in [0,1]. \tag{1}$$

  - *It's a small probability event that the IM assigns small upper probability to a true hypothesis.*

  - The upper α set of the possibility contour

$$\mathcal{C}_\alpha(x) = \{\theta \in \mathbb{T} : \pi_x(\theta) > \alpha\}, \alpha \in [0,1]$$

    is a $100(1-\alpha)\%$ frequentist **confidence set**.

"Imprecision" in IMs

- A lower probability (necessity measure) also exists.

$$\text{Ц}_x(H) = 1 - \Pi_x(H^c).$$

  - IMs are data-dependent and imprecise-probabilistic:

$$x \mapsto (\text{Ц}_x, \Pi_x)$$

- A credal set

$$\mathscr{C}(\Pi_x) = \{\mathsf{Q}_x \in \text{probs}(\mathbb{T}) : \mathsf{Q}_x(\cdot) \leq \Pi_x(\cdot)\} \tag{2}$$

characterizes the correspondence between IMs and precise probabilities.

Section 3

Use IMs in decision making

Use IMs in decision making[5]

- Expected IM loss is calculated by the Choquet integral:[4]

$$\Pi_x \ell_a = \int_0^1 \left\{ \sup_{\theta : \pi_x(\theta) > s} \ell_a(\theta) \right\} ds$$
$$= \sup \{ \mathsf{Q}_x \ell_a : \mathsf{Q}_x \in \mathscr{C}(\Pi_x) \}$$

- This is the **upper expectation** and essentially the upper bound of the expected loss from the credal set (2).
- The optimal action is the minimizer of the Choquet integral.

$$\hat{a}(x) = \arg\min_a \Pi_x \ell_a.$$

- This follows the **minimax** principle!

[4] Troffaes & de Cooman (2014)

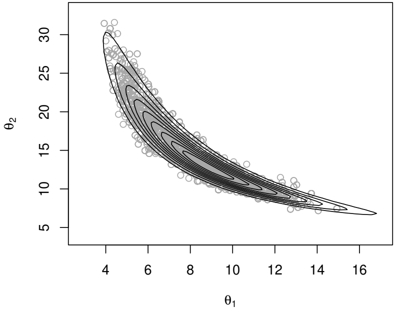
[5] Martin (2025), Martin et al. (2026)

Application 1: Gamma model

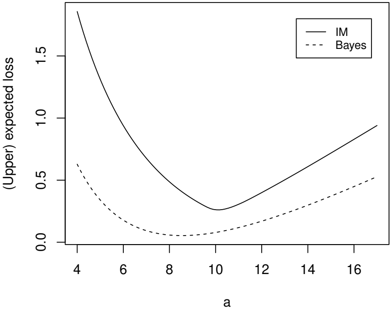
- The gamma distribution is commonly used as a model for time-to-event data, e.g., in survival or reliability analysis.
- $X = (X_1, \ldots, X_n) \sim \text{Gamma}(\Theta_1, \Theta_2), \Theta_1 > 0, \Theta_2 > 0.$
- Consider the data in Example 3 of Fraser et al. (1997) on the survival time (in weeks) for $n = 20$ rats exposed to a certain level of radiation.
- For the decision-making problem, we focus here on estimating the shape using the loss

$$\ell_a(\theta) = (\log a - \log \theta_1)^2, \quad a > 0.$$

Application 1: Gamma model



(a) Contour function



(b) Risk assessments

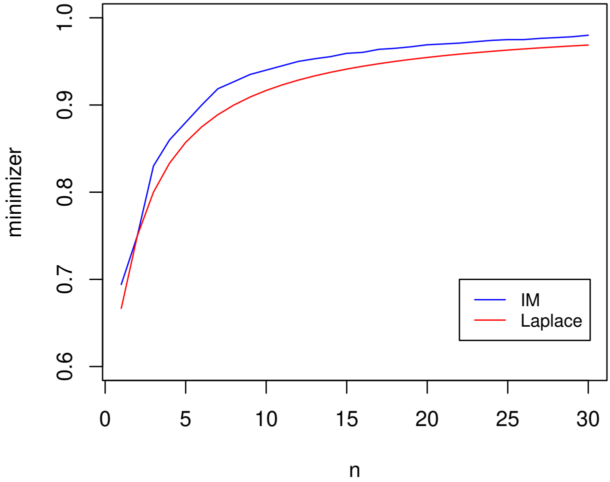
Application 2: Laplace's rule of succession

- Sunrise problem; $X \sim \text{Bin}(n, \theta)$, $X = n$. MLE $= 1$.
- Laplace's rule of succession: $\hat{a}_{LAPLACE} = \frac{n+1}{n+2}$.
- This agrees with the Bayes solution with uniform prior and the squared loss $\ell_a(\theta) = (a - \theta)^2$.
- With IM constructions:
  - $$R(x, \theta) = \left(\frac{n\theta}{x}\right)^x \left(\frac{n(1-\theta)}{x}\right)^{n-x}.$$
  - $$\pi_n(\theta) = \sum_{k=0}^{n} \mathbf{1}\left\{R(k, \theta) \leq R(n, \theta)\right\} \cdot f_k(\theta),$$

  where $f_k(\theta)$ is the probability mass function of a $X = k$ with the probability of success θ.

$$\hat{a}_{IM} = \arg\min_a \Pi_x \ell_a = \arg\min_a \int_0^1 \sup_{\theta:\pi_n(\theta)>s} (a - \theta)^2 \, ds.$$

Application 2: Laplace's rule of succession



Section 4

Theoretic properties

Finite-sample validity

- This is an extension from the validity property in (1).
- Let $\ell_a : \mathbb{T} \to [0, \infty)$ be a non-negative loss function, and define

$$\mathcal{R}(x, \Theta) = \inf_{a \in \mathbb{A}} \frac{\Pi_x \ell_a - \ell_a(\hat{\theta}_x)}{L_a(x, \Theta) - \ell_a(\hat{\theta}_x)} \geq 0, \tag{3}$$

where $L_a(x, \Theta)$ is

$$L_a(x, \Theta) = \sup\{\ell_a(\theta) : \pi_x(\theta) > \pi_x(\Theta)\}, \tag{4}$$

and $\hat{\theta}_x$ is the maximum likelihood estimator of Θ.

Finite-sample validity

Theorem (Validity for Risk Assessment)

$$\sup_{\theta} \mathsf{P}_{\theta}\{\mathcal{R}(X,\theta) \le \alpha\} \le \alpha, \quad \textit{for all } \alpha \in [0,1]. \tag{5}$$

This also implies the simpler-but-weaker result

$$\sup_{\theta} \mathsf{P}_{\theta}\Big\{\inf_{a \in \mathbb{A}} \frac{\Pi_X \ell_a}{L_a(X,\theta)} \le \alpha\Big\} \le \alpha, \quad \text{for all } \alpha \in [0,1].$$

Large sample efficiency

Theorem

Assume the regularity conditions as in Section 3.2 of Martin & Williams (2025) and summarized in Appendix A.2 of Martin et al. (2026). For each $a \in \mathbb{A}$*, let* ℓ_a *be a non-negative and continuous function with compact support. Then* $\Pi_{X^n}\ell_a \to \ell_a(\Theta)$ *in* P_Θ*-probability as* $n \to \infty$*, for each action* $a \in \mathbb{A}$*. Moreover, if the collection* $\{\ell_a : a \in \mathbb{A}\}$ *is uniformly bounded and equicontinuous, then the convergence is uniform in actions, i.e.,* $\sup_{a\in\mathbb{A}} \left|\Pi_{X^n}\ell_a - \ell_a(\Theta)\right| \to 0$ *in* P_Θ*-probability as* $n \to \infty$*.*

Summary

- We developed a framework for decision-making under uncertainty based on **possibilistic IMs**.
    - We define the IM's risk assessment, $a \mapsto \Pi_x \ell_a$, to be an upper expectation or Choquet integral.
    - Different actions a can be compared using a natural variation of the minimize-expected-loss principle.
- **Finite-sample validity property**: We show that $\Pi_X \ell_a$ tends to not be smaller than the oracle assessment, uniformly over actions a, thus giving the IM a sense of reliability in the decision-making context.
- A **large-sample result** establishes that the IM's decision-making safety isn't a consequence of making grossly conservative assessments.
- Lots of numerical results are presented to back up the conceptual and theoretical developments mentioned above.

International Journal of Approximate Reasoning 196 (2026) 109720

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

journal homepage: www.elsevier.com/locate/ijar

Decision-making with possibilistic inferential models *

Ryan Martin *, Shih-Ni Prim, Jonathan P. Williams

Department of Statistics, North Carolina State University, 2311 Stinson Drive, Raleigh, NC, 27695, USA

Thank you!

Website: `https://snprim.github.io/`

Email: `sprim@ncsu.edu`

Email: `prims@wfu.edu`

Section 5

References

Balch, M. S., Martin, R. & Ferson, S. (2019), ‘Satellite conjunction analysis and the false confidence theorem’, *Proc. Royal Soc. A* **475**(2227), 2018.0565.

Dubois, D. & Prade, H. (1988), *Possibility Theory*, Plenum Press, New York.

Fraser, D. A. S., Reid, N. & Wong, A. (1997), ‘Simple and accurate inference for the mean of a gamma model’, *Canad. J. Statist.* **25**(1), 91–99.

Martin, R. (2025), ‘Decision-theoretic properties of possibilistic inferential models’, *International Symposium on Imprecise Probabilities: Theories and Applications* .

Martin, R. (2026), ‘Possibilistic inferential models: a review’, *J. Amer. Statist. Assoc.* **121**(553), 807–826.

Martin, R. & Liu, C. (2013), ‘Inferential models: a framework for prior-free posterior probabilistic inference’, *J. Amer. Statist. Assoc.* **108**(501), 301–313.

Martin, R., Prim, S.-N. & Williams, J. P. (2026), ‘Decision-making with possibilistic inferential models’, *International Journal of Approximate Reasoning* p. 109720.
URL: *https://www.sciencedirect.com/science/article/pii/S0888613X26000964*

Martin, R. & Williams, J. P. (2025), ‘Asymptotic efficiency of inferential models and a possibilistic Bernstein–von Mises theorem’, *Internat. J. Approx. Reason.* **180**, Paper No. 109389.

Troffaes, M. C. M. & de Cooman, G. (2014), *Lower Previsions*, Wiley Series in Probability and Statistics, John Wiley & Sons, Ltd., Chichester.

Section 6

Appendix

Appendix: Finite-sample validity

Proof.

Define the function

$$h_{x,a}(s) = \sup_{\theta:\pi_x(\theta)>s} \ell_a(\theta), \quad s \in [0,1], \tag{6}$$

so that the Choquet integral $\Pi_x \ell_a$ is just a Riemann integral of $h_{x,a}$. It's clear that $s \mapsto h_{x,a}(s)$ is decreasing, which implies that, for any θ,

$$\begin{aligned}
\Pi_x \ell_a &= \int_0^1 h_{x,a}(s)\, ds \\
&= \Big(\int_0^{\pi_x(\theta)} + \int_{\pi_x(\theta)}^1 \Big) h_{x,a}(s)\, ds \\
&\geq \pi_x(\theta)\, h_{x,a}(\pi_x(\theta)) + \{1 - \pi_x(\theta)\}\, h_{x,a}(1),
\end{aligned}$$

Appendix: Finite-sample validity

where $h_{x,a}(1) = \ell_a(\hat{\theta}_x)$. Since $L_a(x,\theta) = h_{x,a}(\pi_x(\theta))$, it follows that

$$\begin{aligned}\mathcal{R}(X,\Theta) \le \alpha \iff & \Pi_X \ell_a - \ell_a(\hat{\theta}_X) \le \alpha \{L_a(X,\Theta) - \ell_a(\hat{\theta}_X)\} \\ \implies & \pi_X(\Theta)\{L_a(X,\Theta) - \ell_a(\hat{\theta}_X)\} \le \alpha \{L_a(X,\Theta) - \ell_a(\hat{\theta}_X)\} \\ \iff & \pi_X(\Theta) \le \alpha.\end{aligned}$$

Validity implies the latter event has P_Θ-probability $\le \alpha$, which proves (3). The simpler, second inequality (5) holds because we can ignore the non-negative "$\{1 - \pi_x(\theta)\}h_{x,a}(1)$" in the lower bound for $\Pi_x \ell_a$ stated above. □

Appendix: Expected Bayes loss

- Data: $X^n \overset{\text{iid}}{\sim} N(\Theta, 1)$
- Prior: $\Theta \sim N(0, \tau^2)$
- Posterior: $Q_x \sim N\left(\theta_{\text{MAP}}, \frac{\tau^2}{1+n\tau^2}\right)$, where $\theta_{\text{MAP}} = \frac{n\tau^2}{1+n\tau^2}\overline{X}_n$.
- Expected Bayes loss:

$$
\begin{aligned}
Q_x \ell_a &= \int (\theta - a)^2 q_x(\theta) d\theta \\
&= \int \left[(\theta - \theta_{\text{MAP}})^2 + (\theta_{\text{MAP}} - a)^2\right] d\theta \\
&= (\theta_{\text{MAP}} - a)^2 + \frac{\tau^2}{1 + n\tau^2}
\end{aligned}
$$

Appendix: Expected IM loss

$$\pi_x(\theta) = 2\left\{1 - P(\sqrt{n}|\overline{X}_n - \theta|)\right\}.$$

$$\begin{aligned}
\overline{\Pi}_x \ell_a &= \int_0^1 \sup_{\theta: 2\left\{1-P(\sqrt{n}|\overline{X}_n-\theta|)\right\}>s} (a-\theta)^2 ds \\
&= \int_0^1 \max\left\{\overline{X}_n - a \pm P^{-1}(1-\frac{s}{2})\right\}^2 ds \\
&= \int_0^1 \left\{|\overline{X}_n - a| + P^{-1}(1-\frac{s}{2})\right\}^2 ds \\
&= \int_{-\infty}^{\infty} (|\overline{X}_n - a|^2 + |z|)^2 p(z) dz \\
&= (\overline{X}_n - a)^2 + 2\left|(\overline{X}_n - a)\right| m(p) + v(p),
\end{aligned}$$

where $m(p) = \int |z| p(z) dz, v(p) = var(Z),\ Z \sim N(0, \frac{1}{n}), \overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i.$

Consider $a = \overline{X}_n$, $\Theta = 0$

- Expected Bayes loss:

$$
\begin{aligned}
\mathrm{E}[Q_x \ell_a(\theta)] &= \frac{\tau^2}{1+n\tau^2} + \mathrm{E}[(\theta_{\text{MAP}} - \overline{X}_n)^2] \\
&= \frac{\tau^2}{1+n\tau^2} + \mathrm{E}\left[\left(\frac{n\tau^2}{1+n\tau^2} - 1\right)^2 \overline{X}_n^2\right] \\
&= \underbrace{\left(\frac{n\tau^2}{1+n\tau^2} + \frac{1}{(1+n\tau^2)^2}\right)}_{=:\, h(n\tau^2)} \cdot \frac{1}{n} < \frac{1}{n}.
\end{aligned}
$$

- $h(n\tau^2) < 1$ for $n\tau^2 \in (0, \infty)$.
- Expected IM loss:

$$
E\left[(\overline{X}_n - a)^2 + 2\left|(\overline{X}_n - a)\right| m(p) + v(p)\right] = v(p) = \frac{1}{n}.
$$

- Oracle loss: $(0 - \overline{X}_n)^2 = \overline{X}_n^2$; $\quad E(\overline{X}_n)^2 = \frac{1}{n}$.