

Imprecise Probabilistic Machine Learning

Winter Semester 2026/27

Saarland University / University of Hamburg / Public Enrolment

Lecture 3: Gambles, Sets of Desirable Gambles, and Lower Prevision

Krikamol Muandet

30 October 2026

Recap

1. Uncertainty in machine learning
2. Precise probability and its pitfalls
3. Beyond precise probability

Outline

1. Gambles
2. Set of desirable gambles
3. Coherent lower prevision and lower probability



From Outcomes to Random Variables



- Consider rolling a fair six-sided die: $\Omega = \{1, 2, 3, 4, 5, 6\}$.
- Define a **random variable** X: $\Omega \to \mathbb{R}$, $X(\omega) = \omega$.

outcome ω	1	2	3	4	5	6
$X(\omega)$	1	2	3	4	5	6

- The random variable then induces a **probability distribution**:

$$P_X(B) = \mathbb{P}(X \in B) = \mathbb{P}\left(X^{-1}(B)\right), \qquad B \in \mathfrak{B}(\mathbb{R})$$

- For a fair die,

$$P_X(\{k\}) = \mathbb{P}(X = k) = \frac{1}{6}, \qquad k = 1, \ldots, 6$$

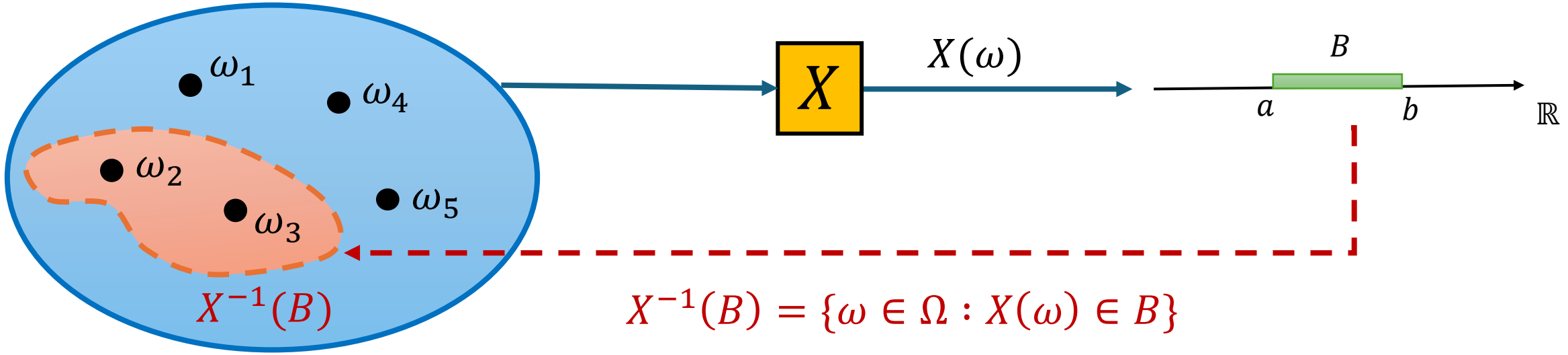
- A random variable is a function that induces a probability distribution P_X from the underlying probability measure $\mathbb{P}$.

Random Variables

Probability Space
$(\Omega, \mathcal{E}, \mathbb{P})$

Random Variable
$\mathrm{X}\colon\ \Omega \to \mathbb{R}$

Induced Distribution
$P_X(B) = \mathbb{P}(X \in B) = \mathbb{P}(X^{-1}(B))$



$$X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\}$$

Indicator Gamble

- **Bet**: You receive €1 if the die shows an even number, and €0 otherwise.
- Let $A = \{2, 4, 6\}$. The corresponding **indicator gamble** is

$$1_A(\omega) = \begin{cases} 1, & \omega \in A \\ 0, & \omega \notin A \end{cases}$$

Outcome ω	1	2	3	4	5	6
$1_A(\omega)$	0	1	0	1	0	1

- The expected payoff is

$$\mathbb{E}_{\mathbb{P}}[1_A] = \mathbb{P}(A) = \frac{1}{2}$$

- An indicator gamble represents the payoff of a **binary event**.

General Gamble

- **Bet**: You receive €5 if the die shows 6, lose €2 if it shows 1, and receive €0 otherwise.

$$f(\omega) = \begin{cases} — & \omega = 1 \\ — & \omega = 6 \\ — & \text{otherwise} \end{cases}$$

Outcome ω	1	2	3	4	5	6
$f(\omega)$	-2	0	0	0	0	5

- The expected payoff under the fair-die model is $\mathbb{E}_P[f] = \frac{-2+5}{6} = 0.5$.
- In IP, a gamble is defined as a **bounded real-valued function** on Ω.

Behaviour Foundations of IP

Concept	Mathematical Object	Interpretation
Random variable X	$X: \Omega \to \mathbb{R}$	Quantity of interest
Indicator gamble	$1_A: \Omega \to \{0,1\}$	Payoff associated with an event
Gamble f	$\boldsymbol{f: \Omega \to \mathbb{R}}$	**State-dependent reward or loss**
Probability measure P	$P: \mathcal{E} \to [0,1]$	Beliefs about events
Distribution P_X	$P_X(B) = P(X \in B)$	Probabilities of value of X

In IP, ***gambles are treated as primitive object*** representing uncertain rewards or losses, without requiring a probability measure.

From Events to Gambles

	Random Variable	Indicator Gamble	General Gamble
Function	$X: \Omega \to \mathbb{R}$	$1_A: \Omega \to \{0,1\}$	$f: \Omega \to \mathbb{R}$
Interpretation	Numerical outcome	Whether an event occurs	Payoff or reward
Example	Face value	Die shows an even number	Reward for each outcome
Values	1,2,3,4,5,6	0,1,0,1,0,1	-3,2,-1,4,0,8

Events $\mathbb{P}(A)$ → Indicator gambles $\mathbb{E}_{\mathbb{P}}[1_A]$ → General gambles $\mathbb{E}_{\mathbb{P}}[f]$

Expectations of gambles provide ***a richer language for reasoning under uncertainty***

Gambles in Machine Learning

Machine learning concept	Interpretation as a gamble	
Loss function	$f_h(z) = \ell(h, z)$	$h = \text{model},\ z = \text{data}$
Classification error	$f_h(x, y) = 1_{\{h(x) \neq y\}}$	$x = \text{input},\ y = \text{output}$
Reward function	$f_\pi(\omega) = R(\pi, \omega)$	$\pi = \text{policy}$
Utility function	$f_a(\omega) = u(a, \omega)$	$a = \text{action}$
Prediction error	$f_h(x, y) = (h(x) - y)^2$	
Treatment benefit	$f(\omega) = Y_1(1, \omega) - Y_0(0, \omega)$	$Y_t = \text{potential outcomes}$

Exercise: Think of any other concepts in machine learning that can be interpreted as a gamble.

Precise Probabilities in Machine Learning

1. **Statistical Learning** – Minimizing Expected Risk
   Assume an underlying data-generating distribution $\boldsymbol{P}$.

$$h^* = \operatorname{argmin}_{h \in \mathcal{H}} \mathbb{E}_{Z \sim \boldsymbol{P}}[\ell(h, Z)], \qquad \hat{h} = \operatorname{argmin}_{h \in \mathcal{H}} \frac{1}{n} \sum_{i=1}^{n} \ell(h, z_i)$$

2. **Bayesian Learning** – Updating Beliefs
   Start with a prior probability distribution over parameters $\boldsymbol{P}(\theta)$.

$$P(\theta \mid D) = \frac{P(D \mid \theta)\boldsymbol{P}(\theta)}{P(D)}, \qquad P(y|x, D) = \int P(y|x, \theta) P(\theta|D)\, d\theta$$

What if probabilities are *unknown*?

Suppose the die may be biased, but probabilities are unknown.

Outcome ω	⚀	⚁	⚂	⚃	⚄	⚅
Probability	?	?	?	?	?	?
$f(\omega)$	-2	-1	0	2	3	5
$g(\omega)$	-3	-2	-1	1	2	4

1) Expected utility

$$\mathbb{E}_P[f] \text{ and } \mathbb{E}_P[g]$$

Require probability distribution

2) Payoff bounds

$$-2 \leq f(\omega) \leq 5$$
$$-3 \leq g(\omega) \leq 4$$

Bounded payoffs

3) Dominance

$$f(\omega) \geq g(\omega)$$

Better in every possible state

Takeaway: We can reason about gambles without knowing P.

Acceptability

- Suppose someone offers you this gamble. Would you accept it?

$$f(\omega) = \begin{cases} 5, & \omega = 6 \\ -1, & \text{otherwise} \end{cases}$$

- How about these gambles?

  $f = (1, 0, 1, 0, 1, 0), g = (-1, -1, -1, -1, -1, -1), h = (3, -8, 0, 1, -2, 10)$

- If you accept $f = (1, 2, 3, -1, -2, -3)$, should you also accept this gamble?

$$g = (2, 4, 6, -2, -4, -6)$$

- Instead of requiring you to specify a probability distribution, we can ask which gambles $f \in \mathbb{D} \subseteq \mathcal{L}(\Omega)$ you consider acceptable.

Coherent Set of Acceptable Gambles

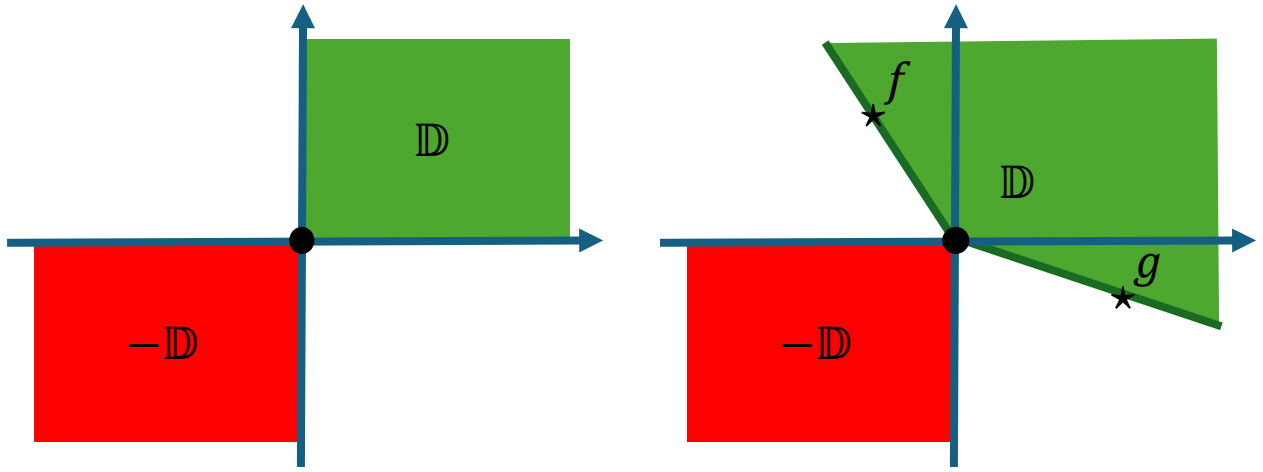
A set of **acceptable gambles** $\mathbb{D} \subseteq \mathcal{L}(\Omega)$ is *coherent* if and only if:

(A1) If $f < 0$, then $f \notin \mathbb{D}$ (avoid partial loss)

(A2) If $f \geq 0$, then $f \in \mathbb{D}$ (accept partial gain)

(A3) If $f \in \mathbb{D}$, then $\lambda f \in \mathbb{D}$ $(\forall \lambda \in \mathbb{R}, \lambda > 0)$ (positive scale invariance)

(A4) If $f \in \mathbb{D}$ and $g \in \mathbb{D}$, then $f + g \in \mathbb{D}$ (combination)

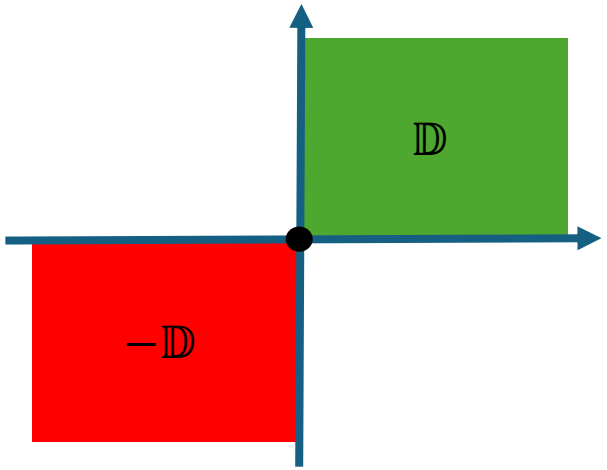


Exercise: Show that these four axioms entail a *monotonicity* property:

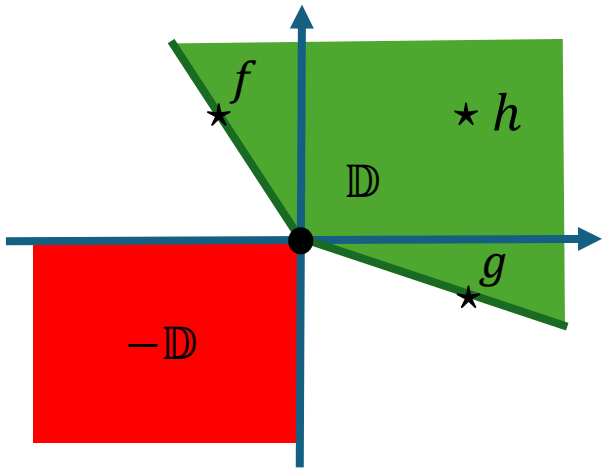
(A5) If $f \in \mathbb{D}$ and $g \geq f$, then $g \in \mathbb{D}$

That is, any bounded gamble g that dominates an acceptable gamble f is an acceptable gamble.

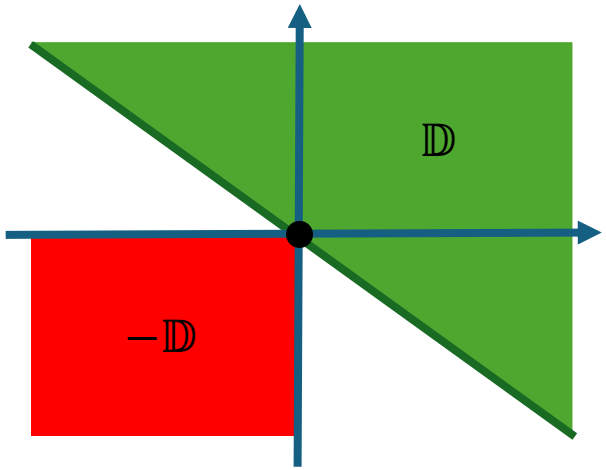
Coherent Set of Acceptable Gambles



Minimal commitment



$\mathbb{A} = \{f, g, h\}$



Maximal commitment

Lower Prevision

Let $\mathbb{D} \subseteq \mathcal{L}(\Omega)$ be a set of acceptable gambles. Then, for any gamble $f \in \mathcal{L}(\Omega)$ and real number μ, the **lower prevision** operator $\underline{P}(\cdot)$ is

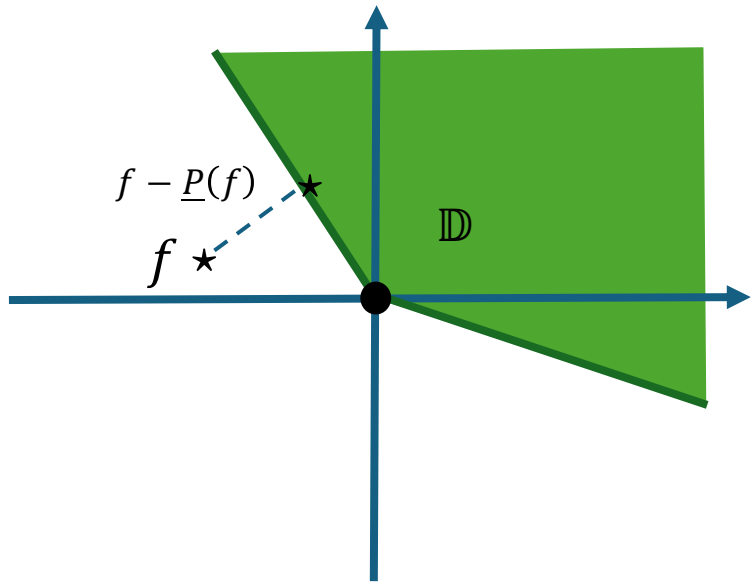
$$\underline{P}(f) := \sup\{\mu \in \mathbb{R} : f - \mu \in \mathbb{D}\}$$

Interpretations:

- The *supremum price* μ you are willing to pay for f.
- For all $\varepsilon > 0$, you are willing to accept the *uncertain reward*

$$f - \underline{P}(f) + \varepsilon$$

where $\underline{P}(f)$ and ε are constant gambles.



Upper Prevision

Let $\mathbb{D} \subseteq \mathcal{L}(\Omega)$ be a set of acceptable gambles. Then, for any gamble $f \in \mathcal{L}(\Omega)$ and real number λ, the **upper prevision** operator $\overline{P}(\cdot)$ is

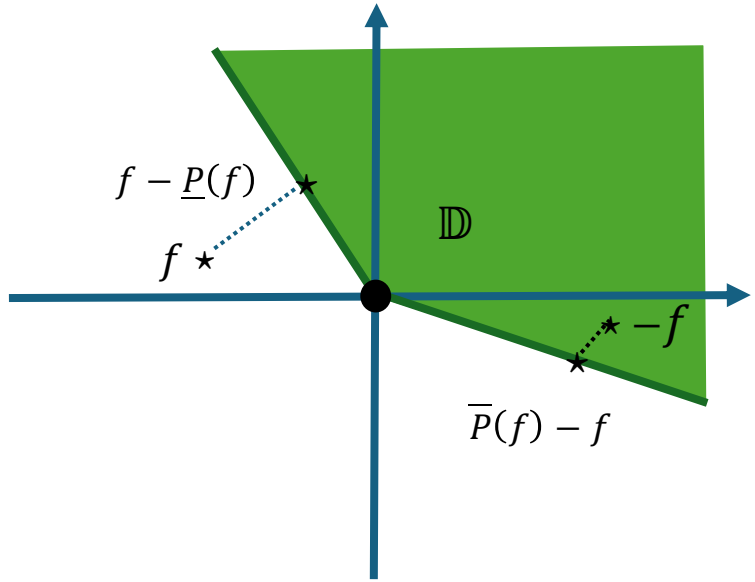
$$\overline{P}(f) := \inf\{\lambda \in \mathbb{R} : \lambda - f \in \mathbb{D}\}$$

Interpretations:

- The *infimum selling price* λ for f that is acceptable to you.
- For all $\varepsilon > 0$, you are willing to accept the *uncertain reward*

$$\overline{P}(f) - f + \varepsilon$$

where $\overline{P}(f)$ and ε are constant gambles.



Lower and Upper Previsions

$$\boxed{\overline{P}(f) = -\underline{P}(-f)}$$

By the definition of upper prevision,

$$\begin{aligned}
\overline{P}(f) &:= \inf\{\, \lambda \in \mathbb{R} : \lambda - f \in \mathbb{D} \,\} && \text{(by definition)} \\
&= \inf\{\, \lambda \in \mathbb{R} : -f - (-\lambda) \in \mathbb{D} \} \\
&= \inf\{\, -\mu \in \mathbb{R} : -f - \mu \in \mathbb{D} \,\} && \text{(setting } \mu = -\lambda\text{)} \\
&= -\sup\{\, \mu \in \mathbb{R} : -f - \mu \in \mathbb{D} \,\} && \text{(}\inf\{-a\} = -\sup\{a\}\text{)} \\
&= -\underline{P}(-f) && \text{(by definition)}
\end{aligned}$$

Unlike precise probabilities, **self-conjugacy** does not generally hold.

$$P(A) = 1 - P(A^c)$$

Coherent Lower Prevision

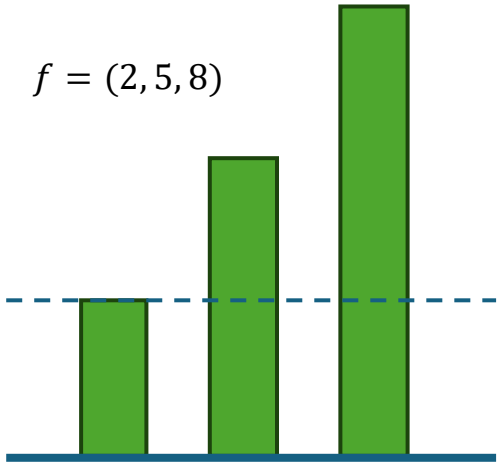
A **lower prevision model** $(\Omega, \mathbb{D}, \underline{P})$ where $\underline{P}$ is a real functional called a lower prevision. For gambles $f, g \in \mathcal{L}(\Omega)$ and any strictly positive number λ, $\underline{P}$ is a *coherent lower prevision* if and only if the following conditions are satisfied:

(LP1). $\underline{P}(f) \geq \inf_{\omega \in \Omega} f(\omega)$ (bounds)
(LP2). $\underline{P}(f + g) \geq \underline{P}(f) + \underline{P}(g)$ (super-additivity)
(LP3). $\underline{P}(\lambda f) = \lambda \underline{P}(f)$ (non-negative homogeneity)

Property	Key Intuition
LP1: Bounds	Respect guaranteed rewards
LP2: Super-additivity	Combining gambles preserves acceptable prices
LP3: Non-negative homogeneity	Value scales with the stakes

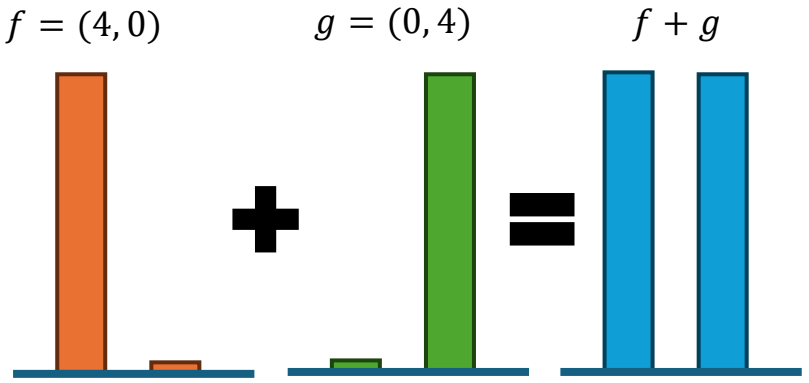
Coherent Lower Prevision

LP1: Bounds
$\underline{P}(f) \geq \inf_{\omega \in \Omega} f(\omega)$



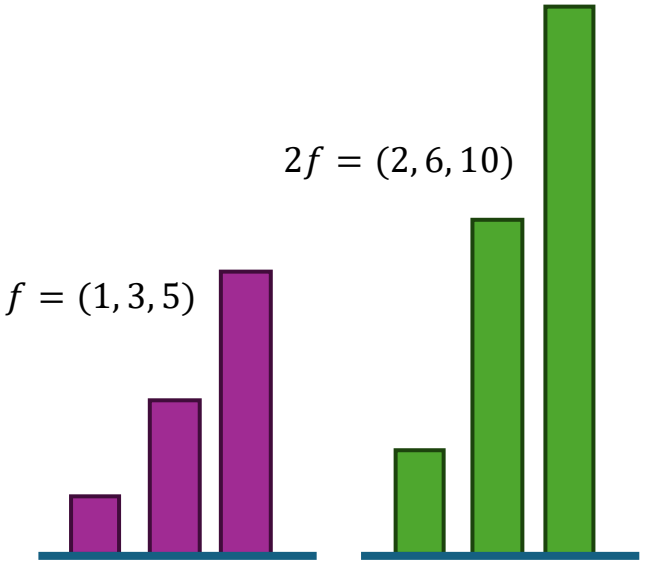
If you are guaranteed to receive at least €2, you should be willing to pay any price below €2.

LP2: Super-additivity
$\underline{P}(f + g) \geq \underline{P}(f) + \underline{P}(g)$



Two individually uncertain gambles can complement each other, producing a more valuable combined gamble.

LP3: Non-negative homogeneity
$\underline{P}(\lambda f) = \lambda \underline{P}(f)$



Doubling the stakes doubles the acceptable buying price.

Coherent Lower Prevision

A lower prevision $\underline{P}$ should ***avoid sure loss***, meaning that

$$\sup_{\omega\in\Omega} \sum_{i=1}^{n} [f_i(\omega) - \underline{P}(f_i)\,] \geq 0 \quad \text{for all } n \geq 0 \text{ and all } f_1, \dots, f_n$$

If this is not satisfied, then there are $n > 0, f_1, \dots, f_n$ and $\delta > 0$ such that

$$\sup_{\omega\in\Omega} \sum_{i=1}^{n} [f_i(\omega) - (\underline{P}(f_i) - \delta)] \leq -\delta$$

Conclusion: the sum of desirable transactions $f_i(\omega) - (\underline{P}(f_i) - \delta)$ results in a loss of at least δ, no matter the outcome of the experiment.

Coherent Lower Prevision

A stronger rationality condition, ***coherence***, requires that supremum acceptable buying price for a gamble f cannot be raised by considering a positive linear combination of a finite number of other acceptable gambles:

$$\sup_{\omega\in\Omega} \sum_{i=1}^{n} \left[f_i(\omega) - \underline{P}(f_i)\right] - m\left[f_0(\omega) - \underline{P}(f_0)\right] \geq 0$$

for all natural $n, m \geq 0$ and all $f_0, f_1, \ldots, f_n$.

Exercise: What if the lower prevision is incoherent?

Linear Prevision

Exercises:
1. If a gamble f costs €2, what should twice the gamble cost? What about the opposite gamble $-f$?
2. Show that the *homogeneity* condition:
   (P3) $P(\lambda f) = \lambda P(f),\ \lambda \in \mathbb{R}$
   follows from P1 and P2.

If P is a **linear prevision** on a linear space $\mathbb{D}$ of bounded gambles, then P is *coherent* if and only if the following conditions are satisfied:

(P1). $P(f) \geq \inf_{\omega \in \Omega} f(\omega)$ (bounds)
(P2). $P(f+g) = P(f) + P(g)$ (additivity)

Interpretation:

- $P(f)$ assigns a **fair price** to every gamble f.
- A coherent linear prevision is an *expectation* under a precise (finitely additive) probability measure.

$$P(f) = \mathbb{E}_{\mathbb{P}}[f]$$

Portfolio	Linear Prevision P		
$f = (0,2)$	€0.80	€0.80	€0.80
$g = (2,0)$	€1.20	€1.20	€1.20
$f + g = (2,2)$	**€2.00**	**€1.50**	**€2.50**
Profit	**€0.00**	**€0.50**	**€0.50**

No-arbitrage: Coherent linear previsions assign prices that cannot be exploited through risk-free trades.

Coherent Lower Previsions: Characterisation

- **Domination**: A linear prevision P *dominates* a lower prevision $\underline{P}$ if

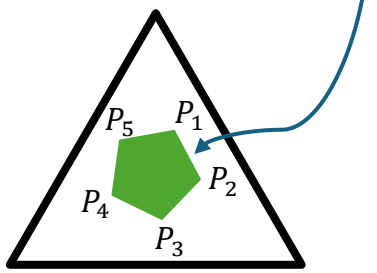
$$P(f) \geq \underline{P}(f), \qquad \forall f \in \mathbb{D}$$

A dominating linear prevision assigns at least as high a value to every gamble.

- **Lower-envelope representation**: A lower prevision is coherent iff it is the lower envelope of a closed, convex set of dominating linear previsions:

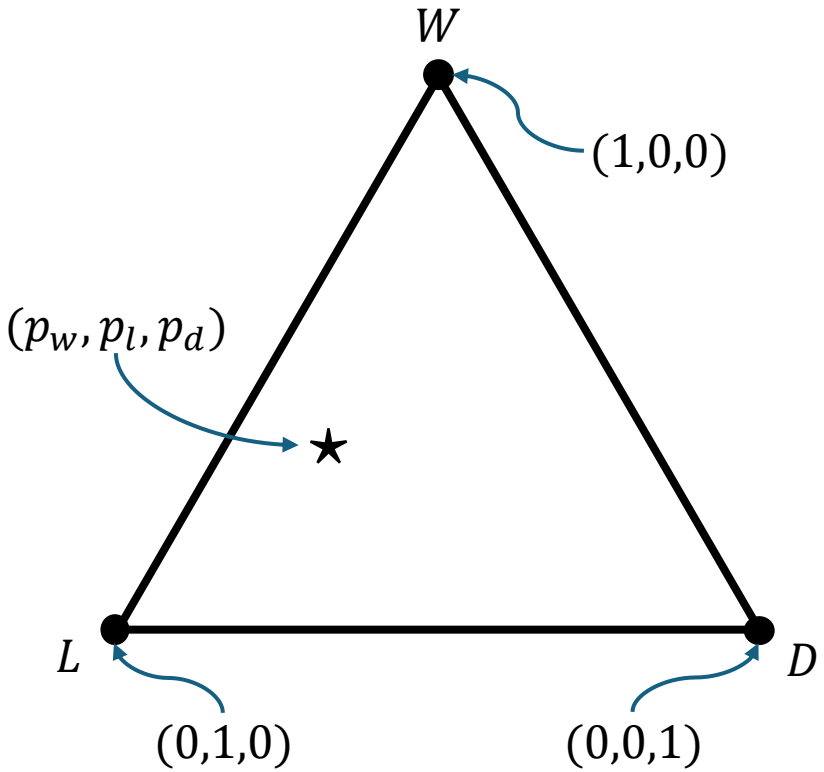
$$\underline{P}(f) = \inf_{P \in \mathcal{M}[\underline{P}]} P(f), \qquad \mathcal{M}[\underline{P}] = \{P \in \mathcal{P} : P(f) \geq \underline{P}(f), \forall f \in \mathbb{D}\}$$

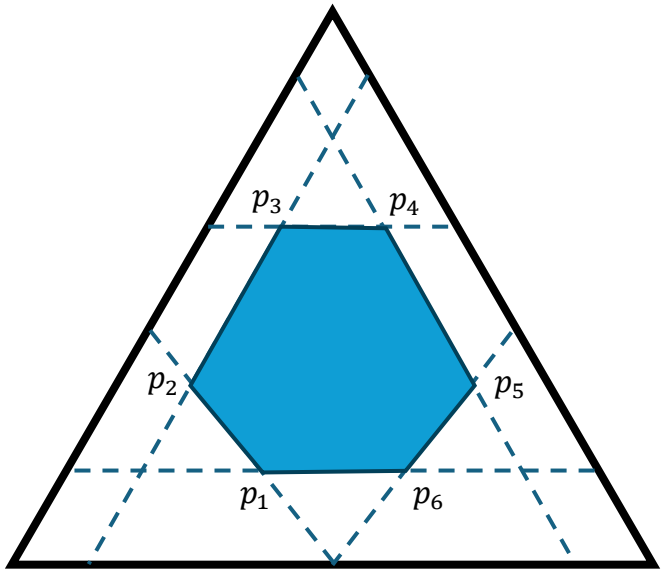
$\Omega = \{\omega_1, \omega_2, \omega_3\}$
$p_1 + p_2 + p_3 = 1, p_i \geq 0$

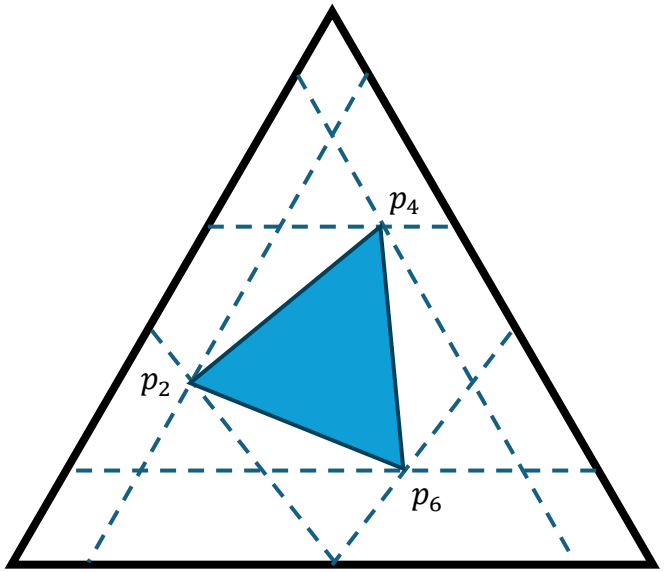


Credal Set

Consider the possibility space $\Omega = \{\text{win, lose, draw}\}$.



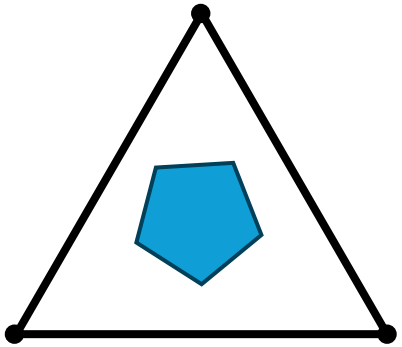




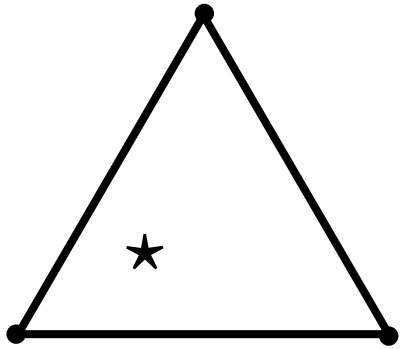
Lower and Upper Previsions

- Let $\mathcal{C}$ be a non-empty subset of a set of all probabilities, i.e., $\mathcal{C} \subseteq \mathcal{P}$. The **lower and upper previsions** can be defined as

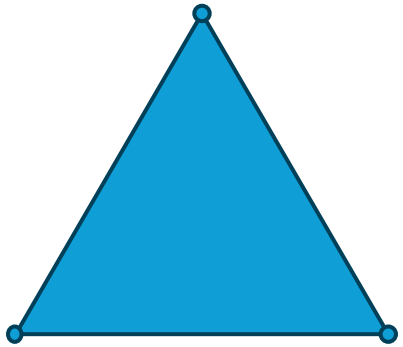
$$\underline{P}(f) \coloneqq \inf_{\mathbb{P}\in\mathcal{C}} \mathbb{E}_{\mathbb{P}}[f], \qquad \overline{P}(f) \coloneqq \sup_{\mathbb{P}\in\mathcal{C}} \mathbb{E}_{\mathbb{P}}[f]$$



$\mathcal{P}^A = \{\mathbb{P} \in \mathcal{P} : \mathbb{P}(A) = 1\}$



$\mathcal{C} = \{\mathbb{P}\}$



$\mathcal{C} = \mathcal{P}$

Probability Intervals

A **probability interval** $(\underline{p}, \overline{p})$ is a pair of lower and upper probability mass functions satisfying:

1. $0 \leq \underline{p}_\omega \leq \overline{p}_\omega \leq 1, \forall \omega \in \Omega$ (bounded)
2. $\sum_{\omega \in \Omega} \underline{p}_\omega \leq 1 \leq \sum_{\omega \in \Omega} \overline{p}_\omega$ (proper)
3. $\underline{p}_\omega \geq 1 - \sum_{z \neq \omega} \overline{p}_z$ and
   $\overline{p}_\omega \leq 1 - \sum_{z \neq \omega} \underline{p}_z$ for all $\omega \in \Omega$ (reachable)

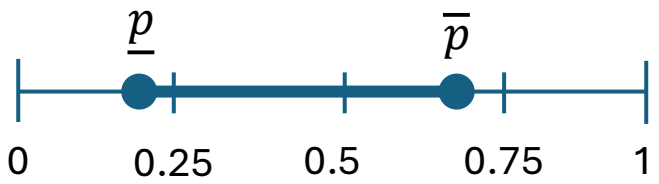
A set of compatible pmfs is called the **credal set**:

$$\mathcal{C} = \{p : \underline{p}_x \leq p_x \leq \overline{p}_x\}$$

Probability Intervals

Boundedness

$0 \leq \underline{p}_\omega \leq \overline{p}_\omega \leq 1, \forall \omega \in \Omega$



Intuition: No probability can be negative or exceed one, and the lower bound cannot exceed the upper bound.

Properness

$\sum_{\omega \in \Omega} \underline{p}_\omega \leq 1 \leq \sum_{\omega \in \Omega} \overline{p}_\omega$

Proper

$[0.2, 0.5]$
$[0.1, 0.4]$
$[0.2, 0.6]$

$0.5 \leq 1 \leq 1.5$

Not proper

$[0.1, 0.2]$
$[0.1, 0.3]$
$[0.1, 0.2]$

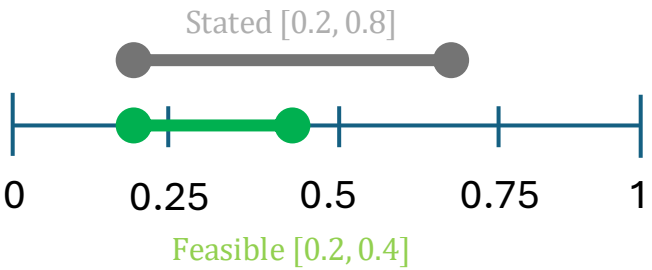
$0.3 \leq 1 \not\leq 0.7$

Intuition: Properness ensures that the credal set is nonempty. There must be at least one probability distribution that is compatible with the intervals.

Reachability

$\underline{p}_\omega \geq 1 - \sum_{z \neq \omega} \overline{p}_z$ and
$\overline{p}_\omega \leq 1 - \sum_{z \neq \omega} \underline{p}_z$ for all $\omega \in \Omega$

$p_1 \in [0.2, 0.8], p_2 \in [0.3, 0.4], p_3 \in [0.3, 0.4]$
$p_2 + p_3 \geq 0.6 \Rightarrow p_1 \leq 1 - 0.6 = 0.4$



Intuition: The bounds must be attainable. It also ensures that the interval bounds are *tight* with respect to the compatible credal set.

Lower and Upper Probabilities

The lower and upper probabilities can be defined as

$$\underline{P}(S) := \max\left\{\sum_{\omega\in S}\underline{p}_\omega\,, 1-\sum_{\omega\in S^c}\overline{p}_\omega\right\}, \qquad \overline{P}(S) := \min\left\{\sum_{\omega\in S}\overline{p}_\omega\,, 1-\sum_{\omega\in S^c}\underline{p}_\omega\right\}$$

Exercise: Consider the following lower and upper pmf:

	X	Y	XY	YY
$\overline{p}$	3/8	3/8	5/8	5/8
$\underline{p}$	1/8	1/8	3/8	0/8

Then, calculate:

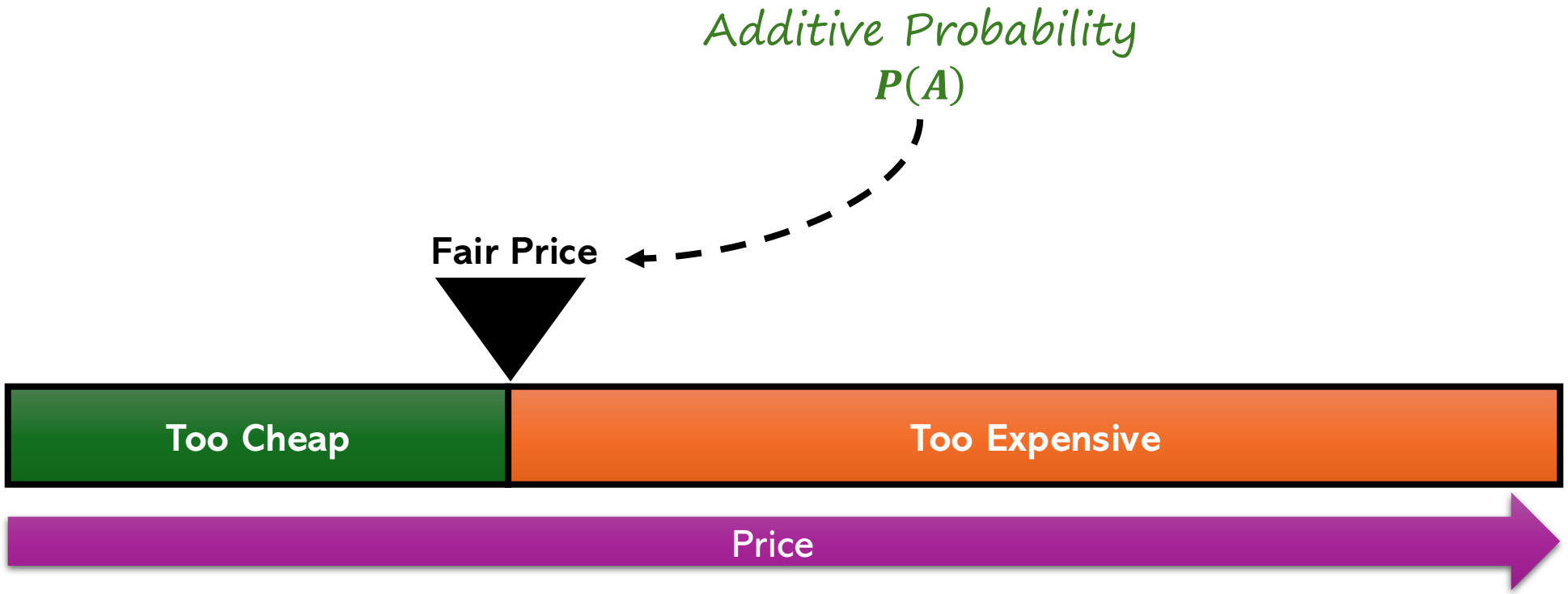
- $\overline{P}(\{Y, XY\}) = ?$
- $\underline{P}(\{X, Y\}) = ?$

Probability Judgement

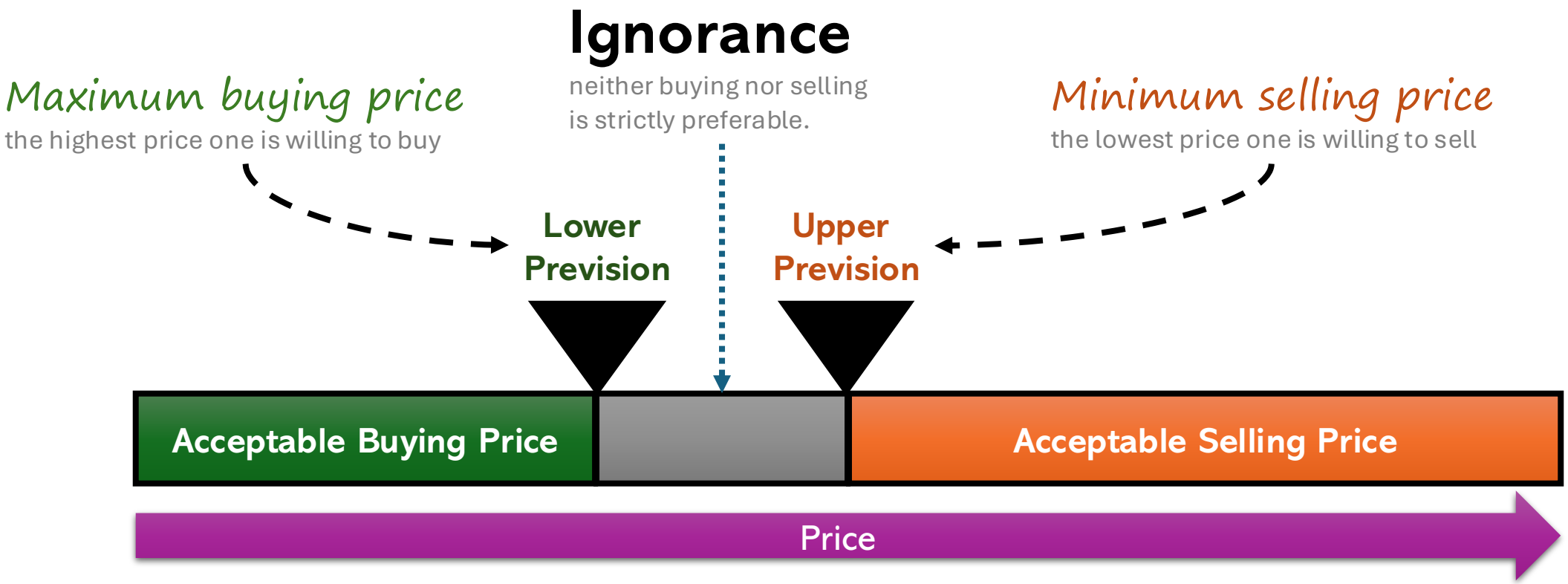
Probability Judgement	Acceptable Gamble	Lower Prevision	
E is probable	$E - \frac{1}{2} \in \mathbb{D}$	$\underline{P}(E) \geq \frac{1}{2}$	
E is at least as probable as F	$E - F \in \mathbb{D}$	$\underline{P}(E - F) \geq 0$	*
E is twice as probable as F	$E - 2F \in \mathbb{D}$	$\underline{P}(E - 2F) \geq 0$	*
Conditional on G, E is at least as probable as F	$G(E - F) \in \mathbb{D}$	$\underline{P}(G(E - 2F)) \geq 0$	*
The odds against E are at least 3 to 1	$E^c - 3E \in \mathbb{D}$	$\overline{P}(E) \leq \frac{1}{4}$	
E without F is improbable	$\frac{1}{2} - E + F \in \mathbb{D}$	$\overline{P}(E - F) \leq \frac{1}{2}$	*
Your degree of belief that E is $\frac{1}{3}$	$E - \frac{1}{3} \in \mathbb{D}$ and $\frac{1}{3} - E \in \mathbb{D}$	$\underline{P}(E) = \frac{1}{3}$ and $\overline{P}(E) = \frac{1}{3}$	

The comparative probability judgments marked by $*$ cannot be modeled by lower probabilities.

Big Picture


Additive Probability
$P(A)$
Fair Price
Too Cheap
Too Expensive
Price

Big Picture


Ignorance
neither buying nor selling is strictly preferable.
Maximum buying price
the highest price one is willing to buy
Minimum selling price
the lowest price one is willing to sell
Lower Prevision
Upper Prevision
Acceptable Buying Price
Acceptable Selling Price
Price

Big Picture

From events to arbitrary gambles

From precise to imprecise assessment

	Events/Indicator Gamble	**General Gambles**
Precise	Probability $P(A)$	Linear prevision $P(f)$
Imprecise	Lower probability $\underline{P}(A)$	Lower prevision $\underline{P}(f)$

Is *imprecision* relevant to modern machine learning?

Distribution shift · Compositional generalization · Continual learning · Multi-agent systems · Causal inference · Open-ended environments · AI for scientific discovery · AI safety · …

Imprecision is an inherent feature of learning under uncertainty.