Robust uncertainty models

Erik Quaeghebeur

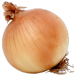
Thursday 8 April 2021

Robust uncertainty models: Main points

- Possibility to express a lack of or limited relevant knowledge
- Interval inferences instead of point inferences
- Set of maximal options instead of single optimal option
- Computational complexity increases with increasing expressivity

Elections in the United Crates of Vegetablia

Office Main Dish 

Candidates
- The Onion 
- The Cabbage 

Pollster (we) Red Tomato 

Elections in the United Crates of Vegetablia

Red Tomato's reputation is at stake!

 made a bad prediction for the previous election,

claiming The Pumpkin  was highly favored to win.

Elections in the United Crates of Vegetablia

Red Tomato's reputation is at stake!

 made a bad prediction for the previous election,

claiming The Pumpkin  was highly favored to win.

This time, we want to avoid such embarrassment:

- Use models that can express uncertainty and lack of knowledge
- Use models that can generate cautious predictions

Elections in the United Crates of Vegetablia

Red Tomato's reputation is at stake!

 made a bad prediction for the previous election,

claiming The Pumpkin  was highly favored to win.

This time, we want to avoid such embarrassment:

- Use models that can express uncertainty and lack of knowledge
- Use models that can generate cautious predictions

Basic setup:

- Random variable X
- Set of possible outcomes: clear win for a candidate or a recount

$\{$  ,  ,  $\}$

Uncertainty Modeling Theories

$$\begin{tikzcd} & \cdots \arrow[d] & \\ & \cdots \arrow[d] & \\ & \cdots \arrow[dl] \arrow[dr] & \\ \textbf{Probability} & & \cdots \end{tikzcd}$$

Probability Theory

Each possible outcome is assigned a probability value.

1. Positive: e.g., $P(\{$$\}) = p$≥ 0
2. Additive: e.g., $P(\{$$,$$\}) = P(\{$$\}) + P(\{$$\})$
3. Normed: $P(\{$$,$$,$$\}) = 1$

Probability Theory

Each possible outcome is assigned a probability value.

1. Positive: e.g., $P(\{$$\}) = p$≥ 0
2. Additive: e.g., $P(\{$$,$$\}) = P(\{$$\}) + P(\{$$\})$
3. Normed: $P(\{$$,$$,$$\}) = 1$

Inferences

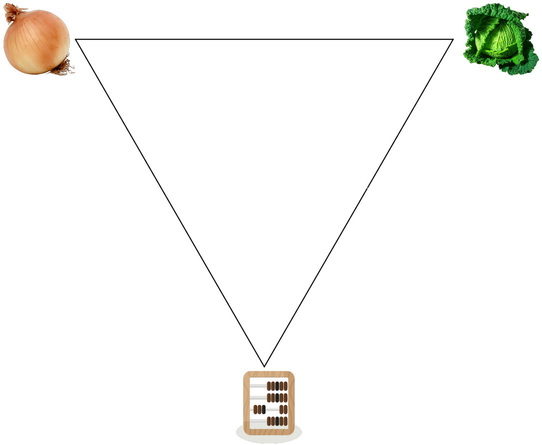
- probability values
- expectations [summation/integration]

$E(f) = p$$f($$) + p$$f($$) + p$$f($$)$

- outcomes with maximal probability
- options minimizing/maximizing expectation

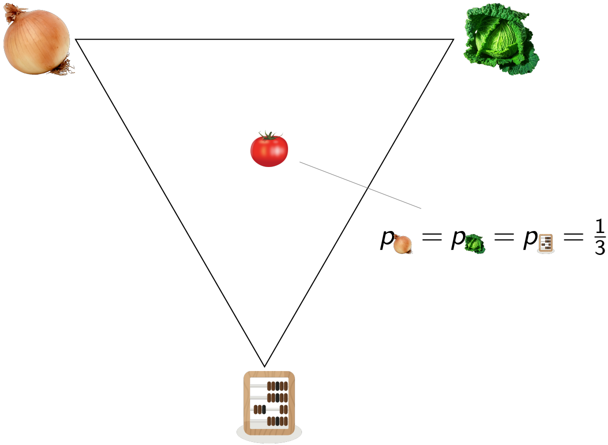
Probability Theory: Illustration

Showing probability mass functions on the probability simplex



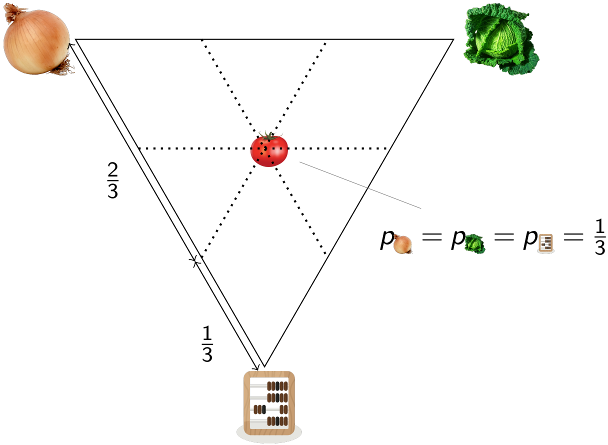
Probability Theory: Illustration

Showing probability mass functions on the probability simplex



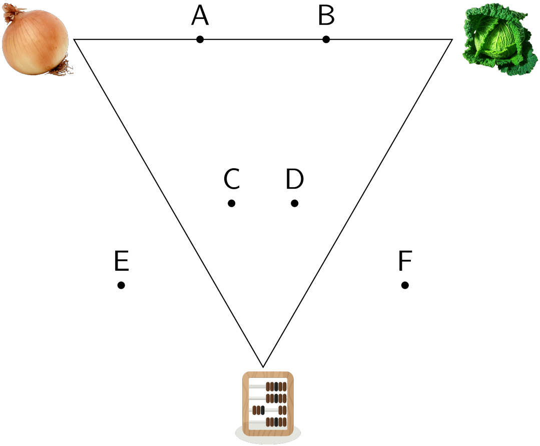
Probability Theory: Illustration

Showing probability mass functions on the probability simplex



Probability Theory: Quiz

Where should we put if $p = \frac{1}{6}$ and $p = \frac{2}{6}$?



Uncertainty Modeling Theories

$$
\begin{tikzcd}
 & \cdots \arrow[d] & \\
 & \cdots \arrow[d] & \\
 & \textbf{Belief Functions} \arrow[dl] \arrow[dr] & \\
\text{Probability} & & \cdots
\end{tikzcd}
$$

Belief Function Theory

(aka Dempster–Shafer Theory, Theory of Completely Monotone Capacities)

Idea & implementation

You do not have to divide all probability mass over outcomes.

Specify a probability mass function m over the 2 $\setminus \varnothing$.

Belief Function Theory

(aka Dempster–Shafer Theory, Theory of Completely Monotone Capacities)

Idea & implementation

You do not have to divide all probability mass over outcomes.

Specify a probability mass function m over the $2^{\{\,,\,,\,\}} \setminus \varnothing$.

Inferences

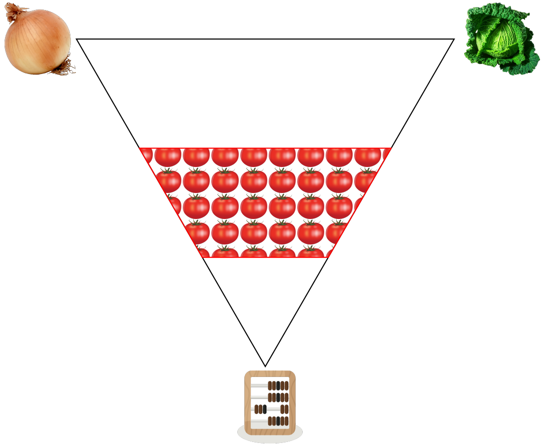
- lower probability values (belief) $\underline{P}(A) = \sum_{B \subseteq A} m_B$
- lower expectations [Choquet integration]

$$\begin{aligned} \underline{E}(f) &= \textstyle\sum_{A \subseteq \{\,,\,,\,\}} m_A \min_{x \in A} f(x) \\ &= \min f + \int_{\min f}^{\max f} \underline{P}(\{x : f(x) \geq v\}) dv \end{aligned}$$

- outcomes with maximal lower/upper probability
- options minimizing/maximizing lower/upper expectation, maximal options in expectation interval order,. . .

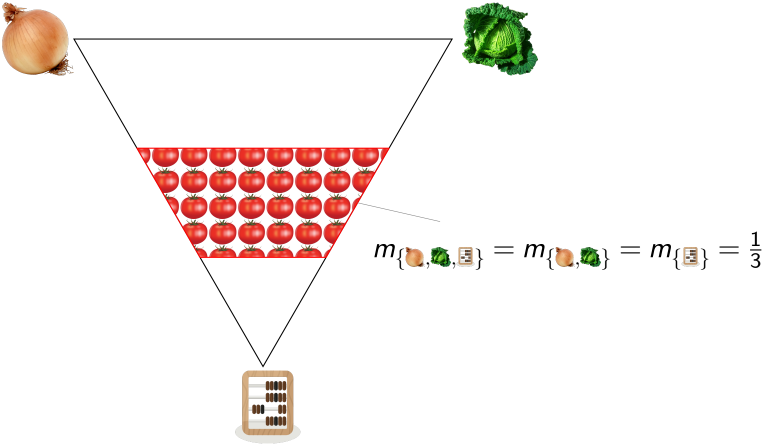
Belief Function Theory: Illustration

Convex combinations of vacuous and degenerate belief models



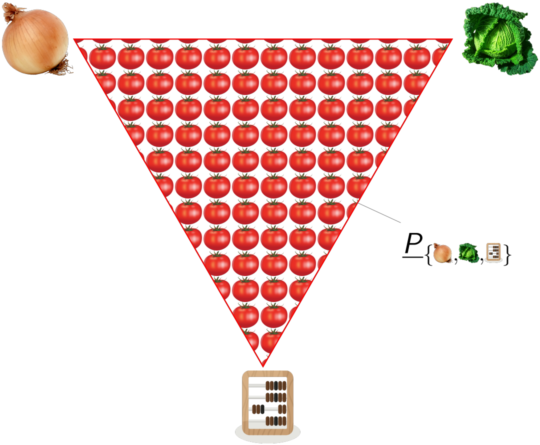
Belief Function Theory: Illustration

Convex combinations of vacuous and degenerate belief models



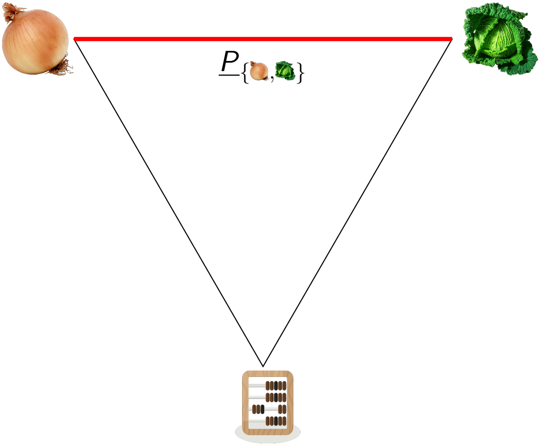
Belief Function Theory: Illustration

Convex combinations of vacuous and degenerate belief models



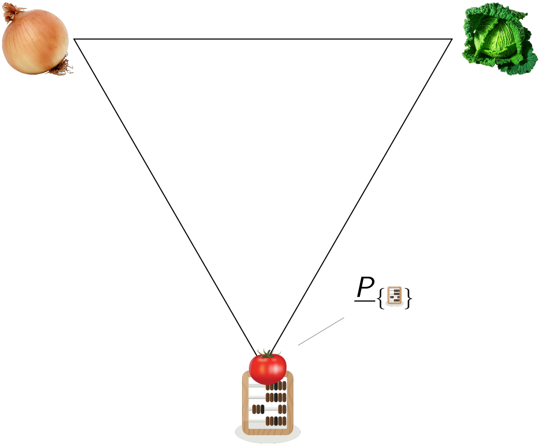
Belief Function Theory: Illustration

Convex combinations of vacuous and degenerate belief models



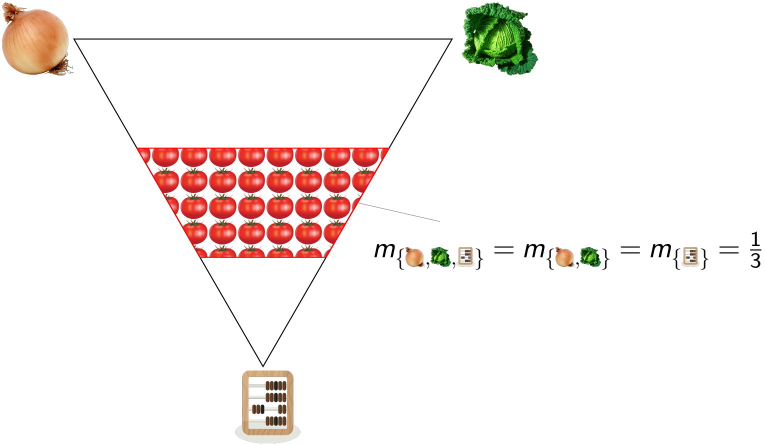
Belief Function Theory: Illustration

Convex combinations of vacuous and degenerate belief models



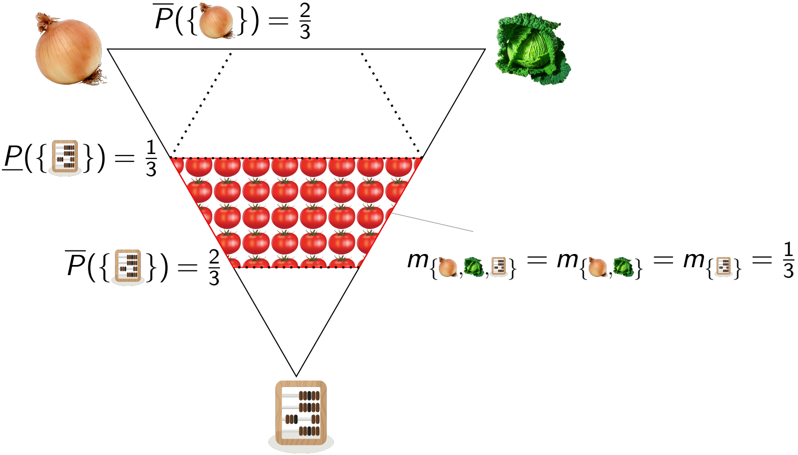
Belief Function Theory: Illustration

Convex combinations of vacuous and degenerate belief models



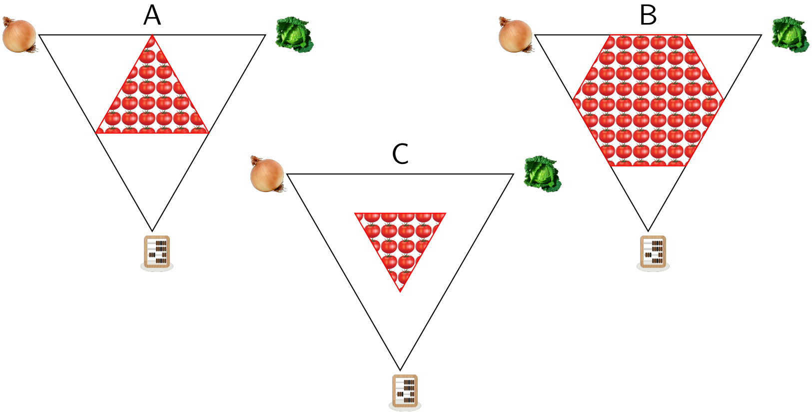
Belief Function Theory: Illustration

Convex combinations of vacuous and degenerate belief models



Belief Function Theory: Quiz

What does the belief function with basic mass assignment
$m_{\{\ ,\ \}} = m_{\{\ ,\ \}} = m_{\{\ ,\ \}} = \frac{1}{3}$ look like in the simplex?



Uncertainty Modeling Theories

$$\begin{tikzcd} & \cdots \arrow[d] & \\ & \textbf{Interval Probability} \arrow[d] & \\ & \text{Belief Functions} \arrow[dl] \arrow[dr] & \\ \text{Probability} & & \cdots \end{tikzcd}$$

Interval Probability Theory

Idea: direct probability assessment

Directly specify lower or upper probabilities for some events.
(Not necessarily bounds on some 'true' probability distribution.)

Constraints

Credal set (bounded probability mass functions) may not be empty.

Interval Probability Theory

Idea: direct probability assessment

Directly specify lower or upper probabilities for some events.
(Not necessarily bounds on some 'true' probability distribution.)

Constraints

Credal set (bounded probability mass functions) may not be empty.

Important property

- Super/Sub-Additivity:

$\underline{P}(\{$  $\}) + \underline{P}(\{$  $\}) \leq \underline{P}(\{$  $,$  $\})$

$\leq \overline{P}(\{$  $,$  $\}) \leq \overline{P}(\{$  $\}) + \overline{P}(\{$  $\})$

Inferences

Lower and upper probabilities and expectations are more complex to derive (discussed later).

Interval Probability Theory

Idea: direct probability assessment

Directly specify lower or upper probabilities for some events.
(Not necessarily bounds on some 'true' probability distribution.)

Constraints

Credal set (bounded probability mass functions) may not be empty.

Important property

- Super/Sub-Additivity:

$\underline{P}(\{$  $\}) + \underline{P}(\{$  $\}) \leq \underline{P}(\{$  $,$  $\})$

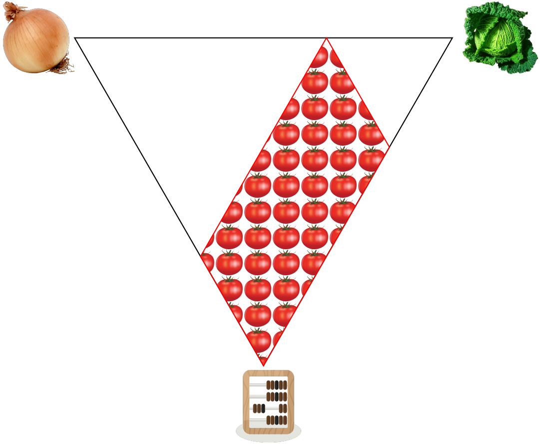
$\leq \overline{P}(\{$  $,$  $\}) \leq \overline{P}(\{$  $\}) + \overline{P}(\{$  $\})$

Intermezzo: Conjugacy relations

$$\overline{P}(A) = 1 - \underline{P}(A^c) \qquad \overline{E}(f) = -\underline{E}(-f)$$

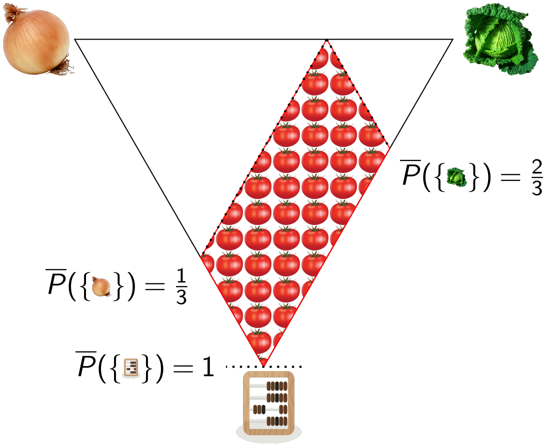
Interval Probability Theory: Illustration

Constraints and conjugacy



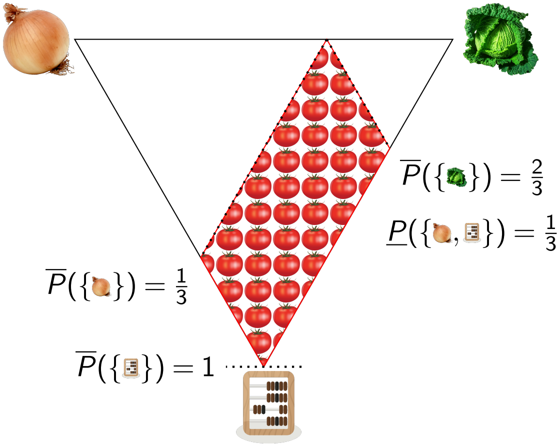
Interval Probability Theory: Illustration

Constraints and conjugacy



Interval Probability Theory: Illustration

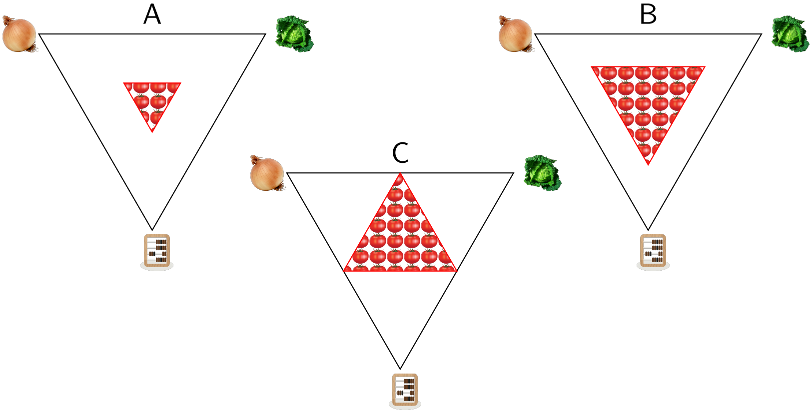
Constraints and conjugacy



Interval Probability Theory: Quiz

What does the credal set for the assessment

$\underline{P}(\{$, $\}) = \underline{P}(\{$, $\}) = \underline{P}(\{$, $\}) = \frac{1}{2}$ look like?



Uncertainty Modeling Theories

$$\begin{tikzcd}
 & \textbf{Interval Expectation} \arrow[d] & \\
 & \text{Interval Probability} \arrow[d] & \\
 & \text{Belief Functions} \arrow[dl] \arrow[dr] & \\
\text{Probability} & & \cdots
\end{tikzcd}$$

Interval Expectation Theory

(aka Imprecise Probability Theory, Theory of Coherent Lower Previsions)

Idea: direct expectation assessment

Directly specify lower expectations for some functions $g \in \mathscr{G}$.

Constraints

- *Avoiding sure loss*: Credal set $\mathscr{C}$ may not be empty
- *Coherence*: Specified expectations must be consistent, i.e., for all $f \in \mathscr{G}$ and $\lambda_g \geq 0$

$$\sum_{g \in \mathscr{G}} \lambda_g \underline{E}(g) - \underline{E}(f) \leq \max\left(\sum_{g \in \mathscr{G}} \lambda_g g - f\right)$$

Interval Expectation Theory

(aka Imprecise Probability Theory, Theory of Coherent Lower Previsions)

Idea: direct expectation assessment

Directly specify lower expectations for some functions $g \in \mathscr{G}$.

Constraints

- *Avoiding sure loss*: Credal set $\mathscr{C}$ may not be empty
- *Coherence*: Specified expectations must be consistent, i.e., for all $f \in \mathscr{G}$ and $\lambda_g \geq 0$

$$\sum_{g \in \mathscr{G}} \lambda_g \underline{E}(g) - \underline{E}(f) \leq \max\left(\sum_{g \in \mathscr{G}} \lambda_g g - f\right)$$

Inferences

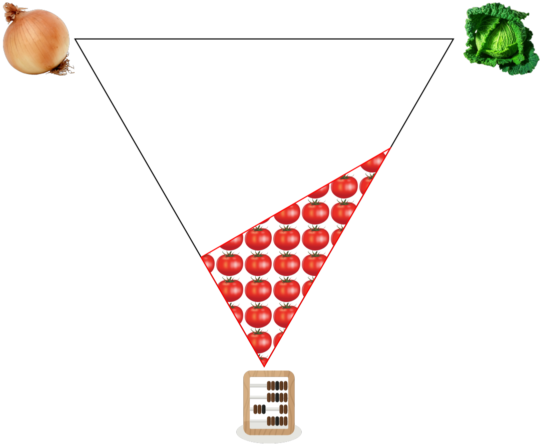
- Natural extension:

$$\underline{E}(f) = \max\left\{\alpha \in \mathbb{R} : f - \alpha \geq \sum_{g \in \mathscr{G}} \lambda_g \left(g - \underline{E}(g)\right), \lambda_g \geq 0\right\}$$

- Lower envelope: $\underline{E}(f) = \min_{p \in \mathscr{C}} E_p(f)$

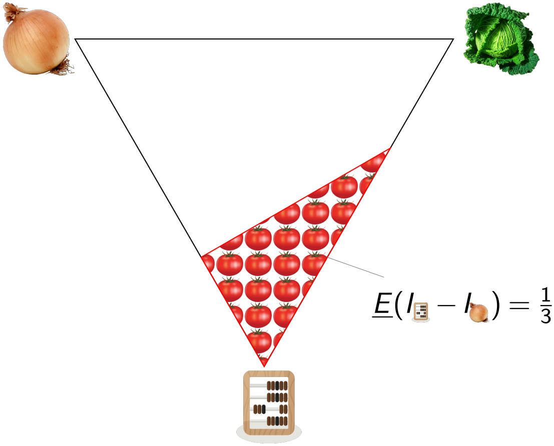
Interval Expectation Theory: Illustration

Constraints, vertices, and envelopes



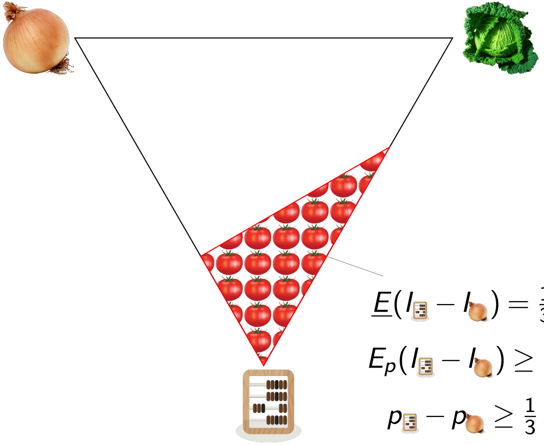
Interval Expectation Theory: Illustration

Constraints, vertices, and envelopes



Interval Expectation Theory: Illustration

Constraints, vertices, and envelopes



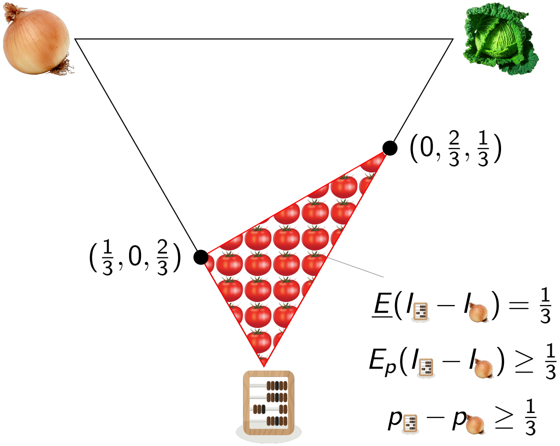
$$\underline{E}(I - I) = \frac{1}{3}$$

$$E_p(I - I) \geq \frac{1}{3}$$

$$p - p \geq \frac{1}{3}$$

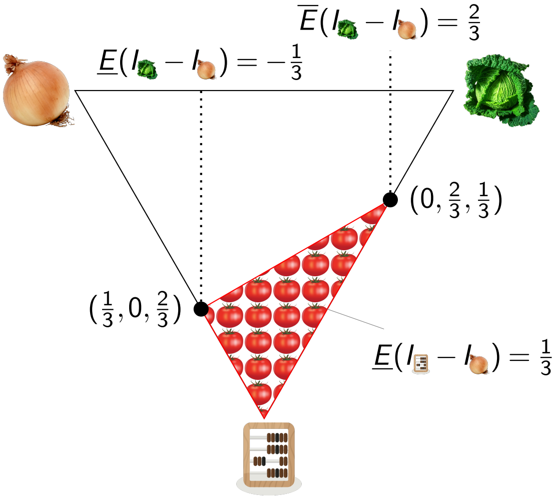
Interval Expectation Theory: Illustration

Constraints, vertices, and envelopes



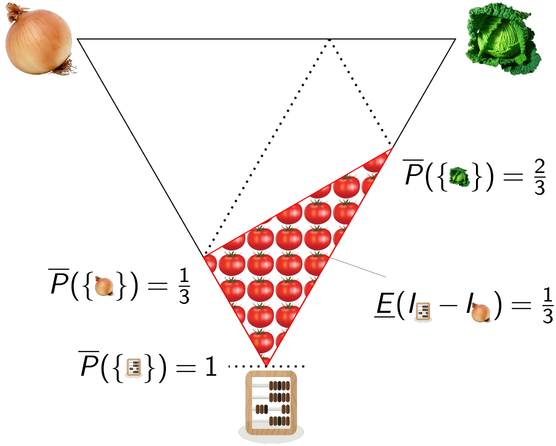
Interval Expectation Theory: Illustration

Constraints, vertices, and envelopes



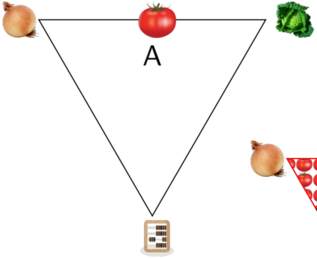
Interval Expectation Theory: Illustration

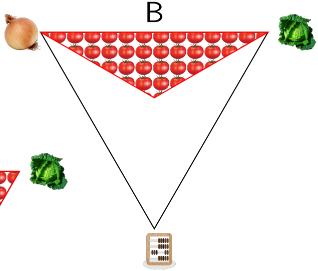
Constraints, vertices, and envelopes

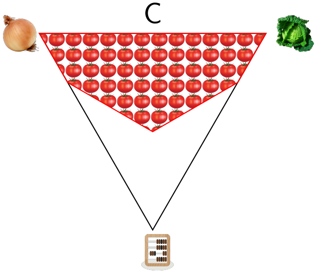


Interval Expectation Theory: Quiz

What does the credal set for the assessment
$\underline{E}(I_{🧅} - I_{🧮}) = \underline{E}(I_{🥬} - I_{🧮}) = \frac{1}{2}$ look like?







Uncertainty Modeling Theories

$$
\begin{tikzcd}
 & \text{Interval Expectation} \arrow[d] & \\
 & \text{Interval Probability} \arrow[d] & \\
 & \text{Belief Functions} \arrow[dl] \arrow[dr] & \\
\text{Probability} & & \textbf{Possibility}
\end{tikzcd}
$$

Possibility Theory

(related to Fuzzy Set Theory)

Subclass of belief functions with m defined on nested sets.

The corresponding upper probability (called *possibility*) has a convenient property:

$$\overline{P}(A \cup B) = \max\{\overline{P}(A), \overline{P}(B)\} \quad \text{or} \quad \overline{P}(A) = \max_{x \in A} \overline{P}(\{x\}) = \max_{x \in A} \pi_x.$$

Possibility Theory
(related to Fuzzy Set Theory)

Subclass of belief functions with m defined on nested sets.

The corresponding upper probability (called *possibility*)
has a convenient property:

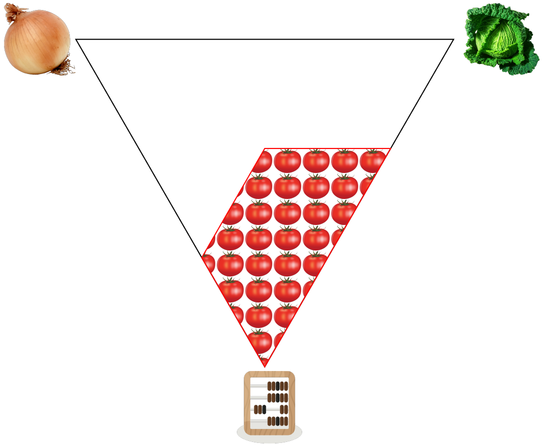
$$\overline{P}(A \cup B) = \max\{\overline{P}(A), \overline{P}(B)\} \quad \text{or} \quad \overline{P}(A) = \max_{x \in A} \overline{P}(\{x\}) = \max_{x \in A} \pi_x .$$

Inferences

Same as for belief function theory
(ignoring questions of interpretation)

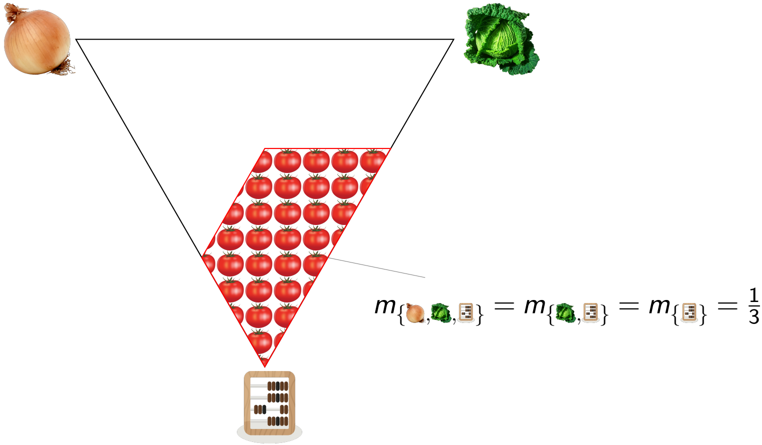
Possibility Theory: Illustration

Nested sets and possibility distributions



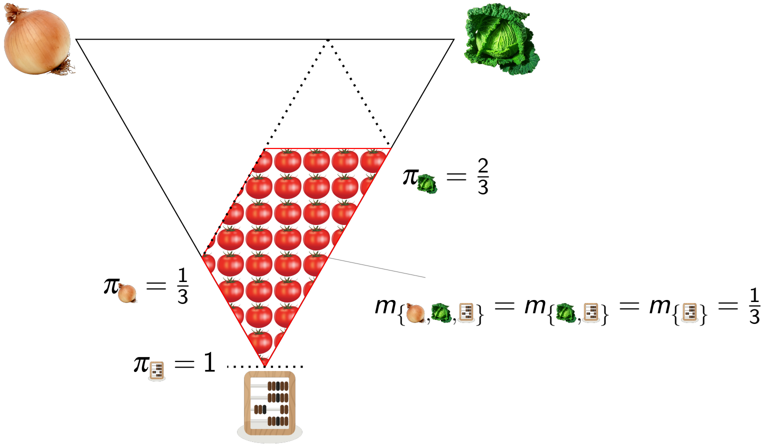
Possibility Theory: Illustration

Nested sets and possibility distributions



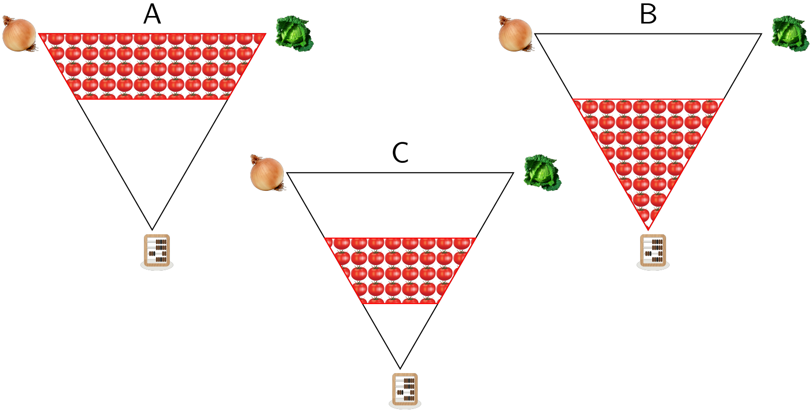
Possibility Theory: Illustration

Nested sets and possibility distributions



Possibility Theory: Quiz

What does the possibility measure with possibility distribution
$\pi = 1$, $\pi = 1$, and $\pi = \frac{1}{3}$ look like in the simplex?



Uncertainty Modeling Theories

$$\begin{tikzcd}
 & \text{Interval Expectation} \arrow[d] & \\
 & \text{Interval Probability} \arrow[d] & \\
 & \text{Belief Functions} \arrow[dl] \arrow[dr] & \\
\text{Probability} & & \text{Possibility}
\end{tikzcd}$$

More Uncertainty Modeling Theories