A gentle introduction to imprecise probability models
and their behavioural interpretation

Gert de Cooman
gert.decooman@ugent.be

SYSTeMS research group, Ghent University
Overview

- General considerations about probability
- Epistemic probability
- Decision making
- Conditioning
Part I

General considerations about probability
Two kinds of probabilities

Aleatory probabilities
- physical property, disposition
- related to frequentist models
- other names: objective, statistical or physical probability, chance

Epistemic probabilities
- model knowledge, information
- represent strength of beliefs
- other names: personal or subjective probability
Part II

Epistemic probability
**First observation**

For many applications, we need theories to represent and reason with certain and uncertain knowledge

- certain → logic
- uncertain → probability theory

One candidate: **Bayesian theory of probability**

I shall:

- argue that it is not general enough
- present the basic ideas behind a more general theory

**imprecise probability theory (IP)**
A theory of epistemic probability

Three pillars:

- how to measure epistemic probability?
- by what rules does epistemic probability abide?
- how can we use epistemic probability in reasoning, decision making, statistics . . . ?

Notice that:

1 and 2 = knowledge representation
3 = reasoning, inference
How to measure personal probability?

- Introspection
  - difficulty: how to convey and compare strengths of beliefs?
  - lack of a common standard
- belief = inclination to act
  - beliefs lead to behaviour, that can be used to measure their strength
- special type of behaviour: accepting gambles
- a gamble is a transaction/action/decision that yields different outcomes (utilities) in different states of the world.
Gambles

- \( \Omega \) is the set of possible outcomes \( \omega \)
- A **gamble** \( X \) is a bounded real-valued function on \( \Omega \)

\[
X : \Omega \rightarrow \mathbb{R} : \omega \mapsto X(\omega)
\]

Example: How did I come to Lugano? By **plane** \( p \), by **car** \( c \) or by **train** \( t \)?

- \( \Omega = \{ p, c, t \} \)
- \( X(p) = -3, X(c) = 2, X(t) = 5 \)

Whether your accept this gamble or not will depend on your knowledge about how I came to Lugano

Denote your set of desirable gambles by

\[
\mathcal{D} \subseteq \mathcal{L}(\Omega)
\]
Modelling your uncertainty

- Accepting a gamble
  = taking a decision/action in the face of uncertainty

- Your set of desirable gambles contains the gambles that you accept

- It is a **model for your uncertainty** about which value $\omega$ of $\Omega$ actually obtains (or will obtain)

- More common models
  - (lower and upper) previsions
  - (lower and upper) probabilities
  - preference orderings
  - probability orderings
  - sets of probabilities
Desirability and rationality criteria

- Rewards are expressed in units of a **linear** utility scale.

- Axioms: a set of desirable gambles $\mathcal{D}$ is **coherent** iff
  
  D1. $0 \not\in \mathcal{D}$
  
  D2. If $X > 0$ then $X \in \mathcal{D}$
  
  D3. If $X, Y \in \mathcal{D}$ then $X + Y \in \mathcal{D}$
  
  D4. If $X \in \mathcal{D}$ and $\lambda > 0$ then $\lambda X \in \mathcal{D}$

- Consequence: If $X \in \mathcal{D}$ and $Y \geq X$ then $Y \in \mathcal{D}$

- Consequence: If $X_1, \ldots, X_n \in \mathcal{D}$ and $\lambda_1, \ldots, \lambda_n > 0$ then $\sum_{k=1}^{n} \lambda_k X_k \in \mathcal{D}$

- A coherent set of desirable gambles is a convex cone of gambles that contains all positive gambles but not the zero gamble.
Definition of lower/upper prevision

- Consider a gamble $X$
- Buying $X$ for a price $\mu$ yields a new gamble $X - \mu$
- the lower prevision $\underline{P}(X)$ of $X$
  - $\sup \{\mu : X - \mu \in \mathcal{D}\}$
  - supremum acceptable price for buying $X$
  - supremum $p$ such that $X - \mu$ is desirable for all $\mu < p$
- Selling $X$ for a price $\mu$ yields a new gamble $\mu - X$
- the upper prevision $\overline{P}(X)$ of $X$
  - $\inf \{\mu : \mu - X \in \mathcal{D}\}$
  - infimum acceptable price for selling $X$
  - infimum $p$ such that $\mu - X$ is desirable for all $\mu > p$
Selling a gamble $X$ for price $\mu$

$= \text{buying } -X \text{ for price } -\mu$: 

$$\mu - X = (-X) - (-\mu)$$

Consequently:

$$\overline{P}(X) = \inf \{\mu: \mu - X \in \mathcal{D}\}$$

$$= \inf \{-\lambda: -X - \lambda \in \mathcal{D}\}$$

$$= -\sup \{\lambda: -X - \lambda \in \mathcal{D}\}$$

$$= -\underline{P}(-X)$$
Lower and upper prevision – 2

\[ P(X) = \sup \{ \mu : X - \mu \in D \} \]

if you specify a lower prevision \( P(X) \), you are committed to accepting

\[ X - P(X) + \epsilon = X - [P(X) - \epsilon] \]

for all \( \epsilon > 0 \) (but not necessarily for \( \epsilon = 0 \)).

\[ \overline{P}(X) = \inf \{ \mu : \mu - X \in D \} \]

if you specify an upper prevision \( \overline{P}(X) \), you are committed to accepting

\[ \overline{P}(X) - X + \epsilon = [\overline{P}(X) + \epsilon] - X \]

for all \( \epsilon > 0 \) (but not necessarily for \( \epsilon = 0 \)).
When lower and upper prevision coincide:

\[ P(X) = \overline{P}(X) = \underline{P}(X) = P(X) \]

is called the (precise) **prevision** of \( X \)

- \( P(X) \) is a prevision, or **fair price** in de Finetti’s sense
- Previsions are the precise, or Bayesian, probability models
- if you specify a prevision \( P(X) \), you are committed to accepting

\[ [P(X) + \epsilon] - X \text{ and } X - [P(X) - \epsilon] \]

for all \( \epsilon > 0 \) (but not necessarily for \( \epsilon = 0 \)).
Allowing for indecision

a) \[ p \quad P(X) \quad q \]

buy \( X \) for price \( p \) \hspace{1cm} sell \( X \) for price \( q \)

b) \[ p \quad P(X) \quad \bar{P}(X) \quad q \]

buy \( X \) for price \( p \) \hspace{1cm} sell \( X \) for price \( q \)

- Specifying a precise prevision \( P(X) \) means that you choose, for essentially any real price \( p \), between buying \( X \) for price \( p \) or selling \( X \) for that price.

- Imprecise models allow for indecision!
Events and lower probabilities

- An event is a subset of $\Omega$
- Example: the event $\{c, t\}$ that I did not come by plane to Lugano.
- It can be identified with a special gamble $I_A$ on $\Omega$

$$I_A(\omega) = \begin{cases} 
1 & \text{if } \omega \in A, \text{ i.e., } A \text{ occurs} \\
0 & \text{if } \omega \notin A, \text{ i.e., } A \text{ doesn’t occur}
\end{cases}$$

- The lower probability $\underline{P}(A)$ of $A$
  = lower prevision $\underline{P}(I_A)$ of indicator $I_A$
  = supremum rate for betting on $A$
  = measure of evidence in favour of $A$
  = measure of (strength of) belief in $A$
Upper probabilities

The upper probability $\overline{P}(A)$ of $A$

$= \text{the upper prevision } \overline{P}(I_A) = \overline{P}(1 - I_{\text{co}A}) = 1 - \underline{P}(I_{\text{co}A})$

of $I_A$

$= 1 - \underline{P}(\text{co}A)$

$= \text{measures lack of evidence against } A$

$= \text{measures the plausibility of } A$

This gives a behavioural interpretation to lower and upper probability

evidence for $A \uparrow \Rightarrow \underline{P}(A) \uparrow$

evidence against $A \uparrow \Rightarrow \overline{P}(A) \downarrow$
Rules of epistemic probability

- Lower and upper previsions represent commitments to act/behave in certain ways.
- Rules that govern lower and upper previsions reflect rationality of behaviour.
- Your behaviour is considered to be irrational when:
  - it is harmful to yourself: specifying betting rates such that you lose utility, whatever the outcome; avoiding sure loss (cf. logical consistency)
  - it is inconsistent: you are not fully aware of the implications of your betting rates; coherence (cf. logical closure)
Avoiding sure loss

- Example: two bets
  
  on $A$: $I_A - \underline{P}(A)$
  on $\text{co}A$: $I_{\text{co}A} - \overline{P}(\text{co}A)$
  
  together: $1 - [\underline{P}(A) + \overline{P}(\text{co}A)] \geq 0$

- Avoiding a sure loss implies

  $\underline{P}(A) + \overline{P}(\text{co}A) \leq 1$,  or  $\underline{P}(A) \leq \overline{P}(A)$
Avoiding sure loss: general condition

A set of gambles $\mathcal{K}$ and a lower prevision $P$ defined for each gamble in $\mathcal{K}$.

**Definition 1.** $P$ avoids sure loss if for all $n \geq 0$, $X_1, \ldots, X_n$ in $\mathcal{K}$ and for all non-negative $\lambda_1, \ldots, \lambda_n$:

$$\sup_{\omega \in \Omega} \left[ \sum_{k=1}^{n} \lambda_k[X_k(\omega) - P(X_k)] \right] \geq 0.$$ 

If it doesn’t hold, there are $\epsilon > 0$, $n \geq 0$, $X_1, \ldots, X_n$ and positive $\lambda_1, \ldots, \lambda_n$ such that for all $\omega$:

$$\sum_{k=1}^{n} \lambda_k[X_k(\omega) - P(X_k) + \epsilon] \leq -\epsilon!$$
Coherence

Example: two bets involving $A$ and $B$ with $A \cap B = \emptyset$

on $A$: $I_A - P(A)$
on $B$: $I_B - P(B)$

Together: $I_{A \cup B} - [P(A) + P(B)]$

Coherence implies that

$P(A) + P(B) \leq P(A \cup B)$
Coherence: general condition

A set of gambles $\mathcal{K}$ and a lower prevision $P$ defined for each gamble in $\mathcal{K}$.

**Definition 2.** $P$ is coherent if for all $n \geq 0$, $X_o, X_1, \ldots, X_n$ in $\mathcal{K}$ and for all non-negative $\lambda_o, \lambda_1, \ldots, \lambda_n$:

$$\sup_{\omega \in \Omega} \left[ \sum_{k=1}^{n} \lambda_k [X_k(\omega) - P(X_k)] - \lambda_o [X_o - P(X_o)] \right] \geq 0.$$ 

If it doesn't hold, there are $\epsilon > 0$, $n \geq 0$, $X_o, X_1, \ldots, X_n$ and positive $\lambda_1, \ldots, \lambda_n$ such that for all $\omega$:

$$X_o(\omega) - (P(X_o) + \epsilon) \geq \sum_{k=1}^{n} \lambda_k [X_k(\omega) - P(X_k)] + \epsilon!$$
Coherence of precise previsions

A (precise) prevision is coherent when it is coherent both as a lower and as an upper prevision.

A precise prevision $P$ on $\mathcal{L}(\Omega)$ is coherent iff

$$P(\lambda X + \mu Y) = \lambda P(X) + \mu P(Y)$$

if $X \geq 0$ then $P(X) \geq 0$

$P(\Omega) = 1$

coincides with de Finetti’s notion of a coherent prevision.

restriction to events is a (finitely additive) probability measure

Let $\mathcal{P}$ denote the set of all linear previsions on $\mathcal{L}(\Omega)$.
Sets of previsions

- Lower prevision $P$ on a set of gambles $\mathcal{K}$
- Let $\mathcal{M}(P)$ be the set of precise previsions that dominate $P$ on its domain $\mathcal{K}$:

$$
\mathcal{M}(P) = \{ P \in \mathcal{P} : (\forall X \in \mathcal{K})(P(X) \geq P(X)) \}. 
$$

- Then avoiding sure loss is equivalent to:

$$
\mathcal{M}(P) \neq \emptyset. 
$$

- and coherence is equivalent to:

$$
\underline{P}(X) = \min \{ P(X) : P \in \mathcal{M}(P) \}, \quad \forall X \in \mathcal{K}. 
$$

- A lower envelope of a set of precise previsions is always coherent
probability measures, previsions à la de Finetti
2-monotone capacities, Choquet capacities
contamination models
possibility and necessity measures
belief and plausibility functions
random set models
Coherent lower/upper previsions – 2

- reachable probability intervals
- lower and upper mass/density functions
- lower and upper cumulative distributions ($p$-boxes)
- (lower and upper envelopes of) credal sets
- distributions (Gaussian, Poisson, Dirichlet, multinomial, …) with interval-valued parameters
- robust Bayesian models
- …
Natural extension

Third step toward a scientific theory
= how to make the theory useful
= use the assessments to draw conclusions about other things [(conditional) events, gambles, . . . ]

Problem: extend a coherent lower prevision defined on a collection of gambles to a lower prevision on all gambles (conditional events, gambles, . . . )

Requirements:
• coherence
• as low as possible (conservative, least-committal)

= NATURAL EXTENSION
Lower probabilities $\underline{P}(A)$ and $\underline{P}(B)$ for two events $A$ and $B$ that are logically independent:

$$A \cap B \neq \emptyset \quad A \cap \text{co}B \neq \emptyset \quad \text{co}A \cap B \neq \emptyset \quad \text{co}A \cap \text{co}B \neq \emptyset$$

For all $\lambda \geq 0$ and $\mu \geq 0$, you accept to buy any gamble $X$ for price $\alpha$ if for all $\omega$

$$X(\omega) - \alpha \geq \lambda[I_A(\omega) - \underline{P}(A)] + \mu[I_B(\omega) - \underline{P}(B)]$$

The natural extension $\underline{E}(X)$ of the assessments $\underline{P}(A)$ and $\underline{P}(B)$ to any gamble $X$ is the highest $\alpha$ such that this inequality holds, over all possible choices of $\lambda$ and $\mu$. 
Natural extension: an example – 2

Calculate $E(A \cup B)$: maximise $\alpha$ subject to the constraints: $\lambda \geq 0$, $\mu \geq 0$, and for all $\omega$:

$$I_{A\cup B}(\omega) - \alpha \geq \lambda [I_A(\omega) - \underline{P}(A)] + \mu [I_B(\omega) - \underline{P}(B)]$$

or equivalently:

$$I_{A\cup B}(\omega) \geq \lambda I_A(\omega) + \mu I_B(\omega) + [\alpha - \lambda \underline{P}(A) - \mu \underline{P}(B)]$$

and if we put $\gamma = \alpha - \lambda \underline{P}(A) - \mu \underline{P}(B)$ this is equivalent to maximising

$$\gamma + \lambda \underline{P}(A) + \mu \underline{P}(B)$$

subject to the inequalities

$$1 \geq \lambda + \mu + \gamma, \quad 1 \geq \lambda + \gamma, \quad 1 \geq \mu + \gamma, \quad 0 \geq \gamma$$

$$\lambda \geq 0, \quad \mu \geq 0$$
This is a linear programming problem, and its solution is easily seen to be:

\[ E(A \cup B) = \max \{ P(A), P(B) \} \]

Similarly, for \( X = I_{A \cap B} \) we get another linear programming problem that yields

\[ E(A \cap B) = \max \{ 0, P(A) + P(B) - 1 \} \]

These are the Fréchet bounds! Natural extension always gives the most conservative values that are still compatible with coherence and other additional assumptions made . . .
Another example: set information

- Information: $\omega$ assumes a value in a subset $A$ of $\Omega$
- This information is represented by the vacuous lower prevision relative to $A$:

$$P_A(X) = \inf_{\omega \in A} X(\omega); \quad X \in \mathcal{L}(\Omega)$$

- $P \in \mathcal{M}(P_A)$ iff $P(A) = 1$
- $P_A$ is the natural extension of the precise probability assessment ‘$P(A) = 1$’; also of the belief function with probability mass one on $A$
- Take any $P$ such that $P(A) = 1$, then $P(X)$ is only determined up to an interval $[P_A(X), \overline{P}_A(X)]$ according to de Finetti’s fundamental theorem of probability
Natural extension: sets of previsions

- Lower prevision $P$ on a set of gambles $\mathcal{K}$

- If it avoids sure loss then $\mathcal{M}(P) \neq \emptyset$ and its natural extension is given by the lower envelope of $\mathcal{M}(P)$:

$$E(X) = \min \{ P(X) : P \in \mathcal{M}(P) \}, \quad \forall X \in \mathcal{L}(\Omega)$$

- $P$ is coherent iff it coincides on its domain $\mathcal{K}$ with its natural extension.
Consider a set $\mathcal{D}$ of gambles you have judged desirable. What are the implications of these assessments for the desirability of other gambles?

The natural extension $\mathcal{E}$ of $\mathcal{D}$ is the smallest coherent set of desirable gambles that includes $\mathcal{D}$. It is the smallest extension of $\mathcal{D}$ to a convex cone of gambles that contains all positive gambles but not the zero gamble.
Natural extension: special cases

Natural extension is a very powerful reasoning method. In special cases it reduces to:

- logical deduction
- belief functions via random sets
- fundamental theorem of probability/prevision
- Lebesgue integration of a probability measure
- Choquet integration of 2-monotone lower probabilities
- Bayes’ rule for probability measures
- Bayesian updating of lower/upper probabilities
- robust Bayesian analysis
- first-order model from higher-order model
Three pillars

1. behavioural definition of lower/upper previsions that can be made operational

2. rationality criteria of
   - avoiding sure loss
   - coherence

3. natural extension to make the theory useful
Gambles and events – 1

- How to represent: event \( A \) is at least \( n \) times as probable as event \( B \)
- Set of precise previsions \( \mathcal{M} \):

\[
P \in \mathcal{M} \Leftrightarrow P(A) \geq nP(B) \Leftrightarrow P(I_A - nI_B) \geq 0
\]
- lower previsions: \( P(I_A - nI_B) \geq 0 \)
- sets of desirable gambles: \( I_A - nI_B + \epsilon \in \mathcal{D}, \forall \epsilon > 0 \).
- \( I_A - nI_B \) is a gamble, generally not an indicator!
- Cannot be expressed by lower probabilities:

\[
\begin{cases}
P(A) \geq P(B), & \bar{P}(A) \geq \bar{P}(B) \\
P(A) \geq \bar{P}(B)
\end{cases}
\]

\begin{align*}
\text{too weak} & \quad \text{too strong}
\end{align*}
Did I come to Lugano by plane, by car or by train?

Assessments:

‘not by plane’ is at least as probable as ‘by plane’
‘by plane’ is at least a probable as ‘by train’
‘by train’ is at least a probable as ‘by car’

Convex set $\mathcal{M}$ of probability mass functions $m$ on $\{p, t, c\}$ such that

$$m(p) \leq \frac{1}{2}, \quad m(p) \geq m(t), \quad m(t) \geq m(c)$$

$\mathcal{M}$ is a convex set with extreme points

$$\left(\frac{1}{2}, \frac{1}{2}, 0\right), \quad \left(\frac{1}{2}, \frac{1}{4}, \frac{1}{4}\right), \quad \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$$
the natural extension $\underline{E}$ is the lower envelope of this set

$$\underline{E}(X) = \min_{m \in \mathcal{M}} [m(p)X(p) + m(t)X(t) + m(c)X(c)]$$

The lower probabilities are completely specified by

$$\underline{E}(\{p\}) = \frac{1}{3} \quad \bar{E}(\{p\}) = \frac{1}{2}$$
$$\underline{E}(\{t\}) = \frac{1}{2} \quad \bar{E}(\{t\}) = \frac{1}{4}$$
$$\underline{E}(\{c\}) = 0 \quad \bar{E}(\{c\}) = \frac{1}{3}$$
the corresponding set of mass functions $\mathcal{M}^*$ is a convex set with extreme points

$$(\frac{1}{2}, \frac{1}{2}, 0), \quad (\frac{1}{2}, \frac{1}{4}, \frac{1}{4}), \quad (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$$

$$(\frac{5}{12}, \frac{1}{4}, \frac{1}{3}), \quad (\frac{1}{3}, \frac{1}{2}, \frac{1}{6})$$

$\mathcal{M}$ is more informative than $\mathcal{M}^*$: $\mathcal{M} \subset \mathcal{M}^*$

with $\mathcal{M}$ we can infer that $E(I_{\{p\}} - I_{\{t\}}) = 0$: ‘by plane’ is at least as probable as ‘by train’

with $\mathcal{M}^*$ this inference cannot be made: we lose information by restricting ourselves to lower probabilities
Gambles and events – 5

\[ A \iff I_A \]

\[ P(A) \iff P(I_A) \]

In precise probability theory:
\[ \rightarrow \text{events are as expressive as gambles} \]

In imprecise probability theory:
\[ \rightarrow \text{events are less expressive than gambles} \]
And by the way

There is a natural embedding of classical propositional logic into imprecise probability theory.

- set of propositions $\rightarrow$ lower probability
- logically consistent $\rightarrow$ ASL
- deductively closed $\rightarrow$ coherent
- deductive closure $\rightarrow$ natural extension
- maximal deductively closed $\rightarrow$ probability

No such embedding exists into precise probability theory.
Part III

Decision making
Consider an action $a$ whose outcome (reward) depends on the actual value of $\omega$ (state of the world)
With such an action we can associate a reward function

$$X_a : \Omega \rightarrow \mathbb{R} : \omega \mapsto X_a(\omega)$$

When do you strictly prefer action $a$ over action $b$:

$$a > b \iff P(X_a - X_b) > 0$$

You almost-prefer $a$ over $b$ if

$$a \geq b \iff P(X_a - X_b) \geq 0$$

We identify an action $a$ with its reward function $X_a$
You are **indifferent** between $a$ and $b$ if

\[ a \approx b \iff a \geq b \text{ and } b \geq a \]

\[ \iff P(X_a - X_b) = \overline{P}(X_a - X_b) = 0 \]

Actions $a$ and $b$ are **incomparable** if

\[ a \parallel b \iff a \succ b \text{ and } b \succ a \text{ and } a \not\approx b \]

In that case there is not enough information in the model to choose between $a$ and $b$: you are **undecided**!

Imprecise probability models **allow for indecision**!

In fact, modelling and allowing for indecision is one of the motivations for introducing imprecise probabilities.
Decision making: maximal actions

Consider a set of actions $\mathbb{A}$ and reward functions $\mathcal{K} = \{X_a : a \in \mathbb{A}\}$

Due to the fact that certain actions may be incomparable, the actions cannot be linearly ordered, only partially!
Ordering of actions
Consider a set of actions $\mathbb{A}$ and reward functions $\mathcal{K} = \{X_a : a \in \mathbb{A}\}$.

Due to the fact that certain actions may be incomparable, the actions cannot be linearly ordered, only partially!

The maximal actions $a$ are actions that are undominated:

$$(\forall b \in \mathbb{A})(b \nsgtr a)$$

or equivalently

$$(\forall b \in \mathbb{A})(\overline{P}(X_a - X_b) \geq 0)$$

Two maximal actions are either indifferent or incomparable!
Decision making: the precise case

- \( a > b \iff P(X_a - X_b) > 0 \iff P(X_a) > P(X_b) \)
- \( a \geq b \iff P(X_a - X_b) \geq 0 \iff P(X_a) \geq P(X_b) \)
- \( a \approx b \iff P(X_a) = P(X_b) \)
- never \( a \parallel b \)!

There is no indecision in precise probability models

- Whatever the available information, they always allow you a best choice between two available actions!

- Actions can always be ordered linearly, maximal actions are unique (up to indifference): they have the highest expected utility.
Decision making: sets of previsions

- $a > b \iff (\forall P \in \mathcal{M}(P))(P(X_a) > P(X_b))$
- $a \geq b \iff (\forall P \in \mathcal{M}(P))(P(X_a) \geq P(X_b))$
- $a \approx b \iff (\forall P \in \mathcal{M}(P))(P(X_a) = P(X_b))$
- $a \parallel b \iff (\exists P \in \mathcal{M}(P))(P(X_a) < P(X_b))$
  and $(\exists Q \in \mathcal{M}(P))(Q(X_a) > Q(X_b))$
- If $\mathcal{K}$ is convex then $a$ is maximal if and only if there is some $P \in \mathcal{M}(P)$ such that
  
  $$(\forall b \in A)(P(X_a) \geq P(X_b))$$
Part IV

Conditioning
Generalised Bayes Rule

Let $\underline{P}$ be defined on a large enough domain, and $B \subseteq \Omega$.

If $\underline{P}(B) > 0$ then coherence implies that $\underline{P}(X|B)$ is the unique solution of the following equation in $\mu$:

$$\underline{P}(I_B[X - \mu]) = 0 \quad \text{(Generalised Bayes Rule)}$$

If $\underline{P} = \underline{P}$ is precise, this reduces to

$$P(X|B) = \mu = \frac{P(XI_B)}{P(B)} \quad \text{(Bayes’ Rule)}$$

Observe that also

$$\underline{P}(X|B) = \inf \{ P(X|B) : P \in \mathcal{M}(\underline{P}) \}$$
Regular extension

If $P(B) = 0$ but $\overline{P}(B) > 0$ then one often considers the so-called regular extension $\underline{R}(X|B)$: it is the greatest $\mu$ such that

$$P(I_B[X - \mu]) \geq 0$$

Observe that also

$$\underline{R}(X|B) = \inf \{ P(X|B) : P \in \mathcal{M}(\overline{P}) \text{ and } P(B) > 0 \}$$

Regular extension is the most conservative coherent extension that satisfies an additional regularity condition.
Questions