Decision Making for Inconsistent Expert Judgments Using Signed Probabilities

J. Acacio de Barros

Liberal Studies Program
San Francisco State University

Ubiquitous Quanta, Università del Salento
Lecce, Italy, 2013
Why probabilities?

- Most ways to think *rationally* lead to probability measures a la Kolmogorov:
  - Pascal (motivated by Antoine Gombaud, Chevalier de Méré).
  - Cox, Jaynes, Ramsey, de Finetti.
  - Venn, von Mises.
- Originally, probabilities were meant to be normative, and not descriptive.
QM observable operators do not fit into a standard boolean algebra (quantum lattice).

Such lattice leads to nonmonotonic upper probability measures or to signed probabilities.¹

Upper probabilities are consequence of strong contextual (inconsistent) correlations.

QM observable operators do not fit into a standard boolean algebra (quantum lattice).

Such lattice leads to nonmonotonic upper probability measures or to signed probabilities.\(^1\)

Upper probabilities are consequence of strong contextual (inconsistent) correlations.

How to think “rationally” about inconsistencies?

- Quantum descriptions?
- Nonstandard (negative) probabilities?

Outline

1. Inconsistent Beliefs

2. Modeling Inconsistent Beliefs
   - Bayesian Model
   - Quantum Model
   - Signed Probability Model

3. Final remarks
Outline

1. Inconsistent Beliefs

2. Modeling Inconsistent Beliefs
   - Bayesian Model
   - Quantum Model
   - Signed Probability Model

3. Final remarks
Inconsistent Beliefs

Inconsistencies

- In logic, any two or more sentences are inconsistent if it is possible to derive from them a contradiction, i.e., if there exists an $A$ such that $(A \land \neg A)$ is a theorem.\(^2\)
- If a set of sentences is inconsistent, then it is trivial.
- To see this, let’s start with $A \land \neg A$ as true. Then $A$ is also true. But since $A$ is true, then so is $A \lor B$ for any $B$. But since $\neg A$ is true, it follows from conjunction elimination that $B$ is necessarily true.
- Paraconsistent logics may be used to deal with inconsistent sentences without exploding.\(^3\)

Take $X$, $Y$, and $Z$ as $\pm 1$-valued random variables.

The above example is equivalent to the deterministic case where

$$E(XY) = E(XZ) = E(YZ) = -1.$$ 

Clearly the correlations are too strong to allow for a joint probability distribution.
Let $X$, $Y$, and $Z$ be $\pm 1$ random variables with zero expectation representing future trends on stocks of companies $X$, $Y$, and $Z$ going up or down.

Three experts, Alice, Bob, and Carlos, have beliefs about the relative behavior of pairs of stocks.

There is no joint\footnote{Suppes, P. and Zanotti, M. (1981) *Synthese* 48(2), 191–199} for $E_A (XY) = -1$, $E_B (XZ) = -1/2$, $E_C (YZ) = 0$, as

$$-1 \leq E (XY) + E (XZ) + E (YZ) \leq 1 + 2 \min \{ E (XY), E (XZ), E (YZ) \}.$$
How to deal with inconsistencies?

- Question: what is the triple moment $E(XYZ)$?
- There are several approaches in the literature.
  - Paraconsistent logics.
  - Consensus reaching.
  - Bayesian.
- Here we will examine two possible alternatives:
  - Quantum.
  - Signed probabilities.
Outline

1. Inconsistent Beliefs

2. Modeling Inconsistent Beliefs
   - Bayesian Model
   - Quantum Model
   - Signed Probability Model

3. Final remarks
Outline

1. Inconsistent Beliefs

2. Modeling Inconsistent Beliefs
   • Bayesian Model
   • Quantum Model
   • Signed Probability Model

3. Final remarks
We start with Alice, Bob, and Carlos as experts, and Deanna Troy as a decision maker.

In the Bayesian approach, Deanna starts with a prior probability distribution.

If we assume she knows nothing about $X$, $Y$, and $Z$, it is reasonable that she sets

$$p^D_{xyz} = p^D_{xyz} = \cdots = p^D_{xyz} = \frac{1}{16}.$$
In order to apply Bayes’s theorem, Deanna needs to have a model of the experts (likelihood function).

Imagine that an oracle tells Deanna that tomorrow the actual correlation $E(\mathbf{XY}) = -1$.

If Deanna thinks her expert is good, knowing that $E(\mathbf{XY}) = -1$ means that she should think that $p_{xy}$ and $p_{xy}$ should be highly improbable for Alice, whereas $p_{xy}$ and $p_{xy}$ highly probable.

For instance, Deanna might propose that the likelihood function is given by

\[ p_{xy} = p_{xy} = 1 - \frac{1}{4} (1 - \epsilon_A)^2, \]

\[ p_{xy} = p_{xy} = -\frac{1}{4} (1 - \epsilon_A)^2, \]

where $E_A(\mathbf{XY}) = \epsilon_A$.

Similarly for Bob and Carlos.
Applying Bayes’s Theorem

Deanna can use Bayes’s theorem to revise her prior belief’s about \( X, \ Y, \) and \( Z. \)

For example,

\[
p_{D|A} = k \left[ 1 - \frac{1}{4} (1 - \epsilon_A)^2 \right] \frac{1}{8},
\]

where

\[
k^{-1} = \left[ 1 - \frac{1}{4} (1 - \epsilon_A)^2 \right] \frac{1}{8} + \left[ \frac{1}{4} (1 - \epsilon_A)^2 \right] \frac{1}{8} + \left[ \frac{1}{4} (1 - \epsilon_A)^2 \right] \frac{1}{8} + \left[ 1 - \frac{1}{4} (1 - \epsilon_A)^2 \right] \frac{1}{8} + \left[ \frac{1}{4} (1 - \epsilon_A)^2 \right] \frac{1}{8} + \left[ \frac{1}{4} (1 - \epsilon_A)^2 \right] \frac{1}{8} + \left[ 1 - \frac{1}{4} (1 - \epsilon_A)^2 \right] \frac{1}{8} + \left[ 1 - \frac{1}{4} (1 - \epsilon_A)^2 \right] \frac{1}{8} + \left[ \frac{1}{4} (1 - \epsilon_A)^2 \right] \frac{1}{8} + \left[ \frac{1}{4} (1 - \epsilon_A)^2 \right] \frac{1}{8} + \left[ 1 - \frac{1}{4} (1 - \epsilon_A)^2 \right] \frac{1}{8},
\]

\[
= \frac{1}{2}.
\]
Incorporating Bob and Carlos’s opinion

- Deanna can now revise her posterior $p_{xyz}^{D|A}$ using once again Bayes’s theorem.
- She gets
  \[
  p_{xyz}^{D|AB} = \frac{1}{32} \left[ (\epsilon_A^2 - 2\epsilon_A - 3) \epsilon_B^2 + (-2\epsilon_A^2 + 4\epsilon_A + 6) \epsilon_B - 3\epsilon_A^2 + 6\epsilon_A + 9 \right].
  \]
- A third application of the theorem gives us $p_{xyz}^{D|ABC}$.
- Similar computations can be carried out for the other atoms.
Example

- If $\epsilon_A = 0$, $\epsilon_B = -\frac{1}{2}$, $\epsilon_C = -1$, we have

$$p_{xyz}^D|ABC = p_{\bar{x}yz}^D|ABC = p_{\bar{x}y\bar{z}}^D|ABC = p_{x\bar{y}z}^D|ABC = 0,$$

$$p_{\bar{x}yz}^D|ABC = p_{x\bar{y}z}^D|ABC = \frac{7}{68},$$

and

$$p_{\bar{x}y\bar{z}}^D|ABC = p_{x\bar{y}z}^D|ABC = \frac{27}{68}.$$

- From the joint, we obtain, e.g.,

$$E(\textit{XYZ}) = 0.$$
The Bayesian approach is the standard probabilistic approach for decision making.

It is extremely sensitive on the prior distribution.

Depends on the model of experts (likelihood function).

Allows to compute a proper joint probability distribution.
Outline

1. Inconsistent Beliefs

2. Modeling Inconsistent Beliefs
   - Bayesian Model
   - Quantum Model
   - Signed Probability Model

3. Final remarks
Theorem

Let $\hat{X}$, $\hat{Y}$, and $\hat{Z}$ be three observables in a Hilbert space $\mathcal{H}$ with eigenvalues $\pm1$, and let them pairwise commute, and let the $\pm1$-valued random variable $X$, $Y$, and $Z$ represent the outcomes of possible experiments performed on a quantum system $|\psi\rangle \in \mathcal{H}$. Then, there exists a joint probability distribution consistent with all the possible outcomes of $X$, $Y$, and $Z$. 
Quantum model

### Theorem

Let $\hat{X}$, $\hat{Y}$, and $\hat{Z}$ be three observables in a Hilbert space $\mathcal{H}$ with eigenvalues $\pm 1$, and let them pairwise commute, and let the $\pm 1$-valued random variable $X$, $Y$, and $Z$ represent the outcomes of possible experiments performed on a quantum system $|\psi\rangle \in \mathcal{H}$. Then, there exists a joint probability distribution consistent with all the possible outcomes of $X$, $Y$, and $Z$.

- Bell: “The only thing proved by impossibility proofs is the author’s lack of imagination.”
Inserting different contexts: measurement

- If we want to model the above correlations, we need to explicitly include the context.
- E.g.
  \[ E_A(XY) = \langle \psi_{xy} | \hat{X} \hat{Y} | \psi_{xy} \rangle, \]
  where \( |\psi\rangle_{xy} \neq |\psi\rangle_{yz} \neq |\psi\rangle_{xz} \).
- For instance, consider the three orthonormal states \( |A\rangle, |B\rangle, \) and \( |C\rangle \), and let
  \[ |\psi\rangle = c_{xy} |\psi_{xy}\rangle \otimes |A\rangle + c_{xz} |\psi_{xz}\rangle \otimes |B\rangle + c_{yz} |\psi_{yz}\rangle \otimes |C\rangle. \]
- We can compute a joint, and therefore \( E(XYZ) \), from \( |\psi\rangle \).
- There are infinite number of \( |\psi\rangle \) satisfying the correlations, and
  \[ -1 \leq E(XYZ) \leq 1. \]
Summary: quantum

- Provides a way to compute the triple moment from a context-dependent vector.
- Imposes no constraint on the relative weights or triple moment.
- Doesn’t tell us what is our best bet.
Outline

1. Inconsistent Beliefs

2. Modeling Inconsistent Beliefs
   - Bayesian Model
   - Quantum Model
   - Signed Probability Model

3. Final remarks
Kolmogorov model

- Kolmogorov axiomatized probability in a set-theoretic way, with the following simple axioms.

\[
\begin{align*}
A1. \quad 1 & \geq P(A) \geq 0 \\
A2. \quad P(\Omega) & = 1 \\
A3. \quad P(A \cup B) & = P(A) + P(B)
\end{align*}
\]
How do we deal with inconsistencies?

- de Finetti: relax Kolmogorov’s axiom A2:

\[ P^* (A \cup B) \geq P^* (A) + P^* (B) \]

or

\[ P^* (A \cup B) \leq P^* (A) + P^* (B). \]

- Subjective meaning: bounds of best measures for inconsistent beliefs (imprecise probabilities).
Consequence:

\[ M^* = \sum_i P_i^* > 1, \]

\[ M_* = \sum_i P_{{*}i} < 1. \]

- \( M^* \) and \( M_* \) should be as close to one as possible.
- Inequalities and nonmonotonicity make it hard to compute upper and lowers for practical problems.
Workaround?

- Define $M^T = \sum_i |p(A_i)|$.
- Instead of violating A2, relax A1:

  A’1. $p_i$ are such that $M^T$ is minimum.

  A’2. $p(A_i \cup A_j) = p(A_i) + p(A_j), \ i \neq j,$

  A’3. $\sum_i p(A_i) = 1.$

- $A_i$ (probability of atom $i$)\(^5\) can now be negative.
- $p$ defines an optimal upper probability distribution by simply setting all negative probability atoms to zero.
- Atoms with negative probability are thought subjectively as impossible events.

\(^5\)The definition of atoms might be difficult once we relax A1, but for finite probability spaces this is not a problem.
Why negative probabilities?

- We can compute them easily (compared to uppers/lowers).
- May be helpful to think about certain contextual problems (e.g. non-signaling conditions, counterfactual reasoning).
- They have a meaning in terms of subjective probability.
Marginals from Alice, Bob, and Carlos

\[ p_{xyz} + p_{x\bar{y}z} + p_{x\bar{y}z} + p_{xy\bar{z}} + p_{\bar{x}yz} + p_{\bar{x}y\bar{z}} + p_{\bar{x}y\bar{z}} = 1, \]  
(1)

\[ p_{xyz} + p_{x\bar{y}z} + p_{x\bar{y}z} + p_{xy\bar{z}} - p_{\bar{x}yz} - p_{\bar{x}y\bar{z}} - p_{\bar{x}y\bar{z}} = 0, \]  
(2)

\[ p_{xyz} + p_{x\bar{y}z} - p_{x\bar{y}z} + p_{xy\bar{z}} + p_{\bar{x}yz} - p_{\bar{x}y\bar{z}} - p_{\bar{x}y\bar{z}} = 0, \]  
(3)

\[ p_{xyz} + p_{x\bar{y}z} + p_{x\bar{y}z} - p_{xy\bar{z}} - p_{\bar{x}yz} + p_{\bar{x}y\bar{z}} - p_{\bar{x}y\bar{z}} = 0, \]  
(4)

\[ p_{xyz} - p_{x\bar{y}z} - p_{x\bar{y}z} + p_{xy\bar{z}} - p_{\bar{x}yz} + p_{\bar{x}y\bar{z}} + p_{\bar{x}y\bar{z}} = 0, \]  
(5)

\[ p_{xyz} - p_{x\bar{y}z} + p_{x\bar{y}z} - p_{xy\bar{z}} + p_{\bar{x}yz} - p_{\bar{x}y\bar{z}} + p_{\bar{x}y\bar{z}} = -\frac{1}{2}, \]  
(6)

\[ p_{xyz} + p_{x\bar{y}z} - p_{x\bar{y}z} - p_{xy\bar{z}} - p_{\bar{x}yz} - p_{\bar{x}y\bar{z}} + p_{\bar{x}y\bar{z}} = -1, \]  
(7)
Signed Probabilities

\[ p_{xyz} = -p_{\bar{x}yz} = -\frac{1}{8} - \delta, \]
\[ p_{x\bar{y}z} = p_{\bar{x}y\bar{z}} = \frac{3}{16}, \]
\[ p_{xy\bar{z}} = p_{\bar{x}y\bar{z}} = \frac{5}{16}, \]
\[ p_{x\bar{y}z} = -p_{x\bar{y}z} = -\delta, \]

\[ E(XYZ) = -\frac{1}{4} - 4\delta. \]

From A’1, we have as constraint

\[ -\frac{1}{8} \leq \delta \leq 0, \text{ which implies } -\frac{1}{4} \leq E(XYZ) \leq \frac{1}{2}. \]
Summary: signed probabilities

- Signed probabilities have a natural interpretation in terms of (subjective) upper probabilities.
- Minimization of $M^-$ requires the improper distributions to approach as best as possible the rational proper jpd.
- This has a normative constraint on the choices of triple moment.
Outline

1. Inconsistent Beliefs

2. Modeling Inconsistent Beliefs
   - Bayesian Model
   - Quantum Model
   - Signed Probability Model

3. Final remarks
Summary

- Standard Bayesian approach is sensitive to choices of prior and likelihood function (well-known problem).
  - “It ain’t what you don’t know that gets you into trouble. It’s what you know for sure that just ain’t so.” -Mark Twain
  - E.g. say that Deanna starts with $E(\text{XYZ}) = \epsilon$ as her prior. The posterior will give $E(\text{XYZ}) = \epsilon$ regardless of Alice, Bob, and Carlos’s opinions.

- The quantum-like approach, using vectors on a Hilbert space, seems to be too permissive, and to not have normative power. (Is it the only quantum model for it?)
  - Can we find some additional principle in QM to help with this?

- Negative probabilities, with the minimization of the negative mass, offers a lower and upper bound for values of triple moment.