Decision Making for Inconsistent Expert Judgments Using Signed Probabilities

J. Acacio de Barros

Liberal Studies Program
San Francisco State University

Purdue University, February 24, 2014

What's up with 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 Finneti.
  - Venn, von Mises.
- Originally, probabilities were meant to be normative, and not descriptive.

Quantum Social Sciences?

- Human decision-making does not seem to satisfy the rules of classical probability theory[1]
- To model such cases, many researchers have used the mathematical formalism of QM: "quantum probabilities"[2]
- Feynman proposed the use of negative probabilities in QM[3]

[1] Kahneman, D. (2003) *American Psychologist* **58(9)**, 697–720

[2] Busemeyer, J. R. and Bruza, P. D. (2012) Quantum models of cognition and decision, Cambridge University Press, Cambridge, UK and references therein.

[3] Feynman, R. P. (1987) Negative probability In B. J. Hiley and F. David Peat, (ed.), Quantum implications: essays in honour of David Bohm, pp. 235–248 Routledge London and New York

Outline

1. Inconsistent Beliefs
2. Modeling Inconsistent Beliefs
   - Bayesian Model
   - Quantum Model
   - Signed Probability Model
3. Final remarks

Outline

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 $(A \wedge \neg A)$ is a theorem.[4]
- If a set of sentences is inconsistent, then it is trivial.
  - Start with $A \wedge \neg A$ as true. Then A is true. But since A is true, then, for any B, so is $A \vee B$. But since $\neg A$ is true, it follows from conjunction elimination that B is necessarily true.

[4] Suppes, P. (1999) Introduction to Logic, Dover Publications, Mineola, New York.

With probabilities

- Take **X**, **Y**, and **Z** as ± 1-valued random variables.
- The above example is equivalent to the deterministic case where

$$E(\mathbf{XY}) = E(\mathbf{XZ}) = E(\mathbf{YZ}) = -1.$$

- Clearly the correlations are too strong to allow for a joint probability distribution.

A subtler case

- Let **X**, **Y**, and **Z** be ± 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.
- No direct disagreement between experts: all about $E(\mathbf{X}) = E(\mathbf{Y}) = E(\mathbf{Z}) = 0$
- But there is no joint[5] for $E_A(\mathbf{XY}) = 0$, $E_B(\mathbf{XZ}) = -1/2$, $E_C(\mathbf{YZ}) = -1$, as

$$-1 \leq E(\mathbf{XY}) + E(\mathbf{XZ}) + E(\mathbf{YZ}) \leq 1 + 2\min\{E(\mathbf{XY}), E(\mathbf{XZ}), E(\mathbf{YZ})\}.$$

[5] Suppes, P. and Zanotti, M. (1981) *Synthese* **48(2)**, 191–199

How to deal with inconsistencies?

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

Outline

Outline

Bayesian Model: Priors

- 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

$$p^D_{xyz} = p^D_{\bar{x}yz} = \cdots = p^D_{\overline{xyz}} = \frac{1}{8}.$$

Model of experts

- 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 $p_{xy\cdot}$ and $p_{\overline{xy}\cdot}$ should be highly improbable for Alice, whereas $p_{\overline{x}y\cdot}$ and $p_{x\overline{y}\cdot}$ highly probable.
- For instance, Deanna might propose that the likelihood function is given by

$$p_{xy\cdot} = p_{\overline{xy}\cdot} = 1 - \frac{1}{4}(1-\epsilon_A)^2,$$

$$p_{\overline{x}y\cdot} = p_{\overline{x}y\cdot} = \frac{1}{4}(1-\epsilon_A)^2,$$

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,

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

where

$$\begin{aligned} k^{-1} &= \left[1-\frac{1}{4}\left(1-\epsilon_A\right)^2\right]\frac{1}{8}+\left[\frac{1}{4}\left(1-\epsilon_A\right)^2\right]\frac{1}{8}+\left[\frac{1}{4}\left(1-\epsilon_A\right)^2\right]\frac{1}{8} \\ &+\left[1-\frac{1}{4}\left(1-\epsilon_A\right)^2\right]\frac{1}{8}+\left[\frac{1}{4}\left(1-\epsilon_A\right)^2\right]\frac{1}{8}+\left[\frac{1}{4}\left(1-\epsilon_A\right)^2\right]\frac{1}{8} \\ &+\left[1-\frac{1}{4}\left(1-\epsilon_A\right)^2\right]\frac{1}{8}+\left[1-\frac{1}{4}\left(1-\epsilon_A\right)^2\right]\frac{1}{8} \\ &= \frac{1}{2}. \end{aligned}$$

Incorporating Bob and Carlos's opinion

- Deanna can now revise her posterior $p_{xyz}^{D|A}$ using once again Bayes's theorem.
- She gets

$$p_{xyz}^{D|AB} = \frac{1}{32}\left[\left(\epsilon_A^2 - 2\epsilon_A - 3\right)\epsilon_B^2 + \left(-2\epsilon_A^2 + 4\epsilon_A + 6\right)\epsilon_B - 3\epsilon_A^2 + 6\epsilon_A + 9\right].$$

- 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

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

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

and

$$p_{xy\overline{z}}^{D|ABC} = p_{\overline{xy}z}^{D|ABC} = \frac{27}{68}.$$

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

$$E\left(\mathbf{XYZ}\right) = 0.$$

Summary: Bayesian

- The Bayesian approach is the standard probabilistic approach for decision making.
- It is extremely dependent 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

Quantum model

- In quantum models, random variables are replaced with observables in a Hilbert space $\mathcal{H}$.
  - **X**, **Y**, and **Z** are modeled by the linear Hermitian operators $\hat{X}$, $\hat{Y}$, and $\hat{Z}$ on $\mathcal{H}$.
  - A state vector $|\psi\rangle \in \mathcal{H}$ codes the state of the system.
  - Expectations are given by
  $$\langle\psi|\hat{A}|\psi\rangle,$$
  where $\hat{A}$ is an observable (Hermitian operator).
  - E.g. $E(\mathbf{X}) = \langle\psi|\hat{X}|\psi\rangle$, $E(\mathbf{XY}) = \langle\psi|\hat{X}\hat{Y}|\psi\rangle$, etc.
  - Note that $\hat{X}\hat{Y}$ is Hermitian if $\left[\hat{X},\hat{Y}\right] = 0$.

Ops!

Theorem

Let $\hat{X}$, $\hat{Y}$, and $\hat{Z}$ be three observables in a Hilbert space $\mathcal{H}$ with eigenvalues ± 1 and that pairwise commute, and let the ± 1-valued random variables $\mathbf{X}$, $\mathbf{Y}$, and $\mathbf{Z}$ 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 $\mathbf{X}$, $\mathbf{Y}$, and $\mathbf{Z}$.

Ops!

Theorem

Let $\hat{X}$, $\hat{Y}$, and $\hat{Z}$ be three observables in a Hilbert space $\mathcal{H}$ with eigenvalues ± 1 and that pairwise commute, and let the ± 1-valued random variables $\mathbf{X}$, $\mathbf{Y}$, *and* $\mathbf{Z}$ *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* $\mathbf{X}$, $\mathbf{Y}$, *and* $\mathbf{Z}$.

- "The only thing proved by impossibility proofs is the author's lack of imagination." J. S. Bell

How to have different contexts? Include explicitly!

- If we want to model the Alice, Bob, and Carlos's correlations, we need to explicitly include the context.
- E.g.
$$E_A(\mathbf{XY}) = \langle \psi_{xy} | \hat{X} \hat{Y} | \psi_{xy} \rangle,$$
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 $E(\mathbf{XYZ})$, from $|\psi\rangle$.
- There are infinite number of $|\psi\rangle$ satisfying the correlations, and $-1 \leq E(\mathbf{XYZ}) \leq 1$.

Summary: quantum

- Makes context explicit.
- Imposes no constraint on the relative weights or triple moment.
- Doesn't tell us what is our best bet.

Outline

Kolmogorov model

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

$$\text{K1. } 1 \geq P(A) \geq 0$$

$$\text{K2. } P(\Omega) = 1$$

$$\text{K3. } P(A \cup B) = P(A) + P(B)$$

Upper and lower probabilities

- 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).

Upper and lower probabilities

- Consequence:

$$M^* = \sum_i P^*(\{\omega_i\}) > 1,$$

$$M_* = \sum_i P_*(\{\omega_i\}) < 1.$$

- 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 $M^T = \sum_i |p(\{\omega_i\})|$, $\omega_i \in \Omega$.
- Instead of violating K3, relax K1:

N1. p_i are such that M^T is minimum.

$$\text{N2.} \quad \sum_i p(\{\omega_i\}) = 1,$$

$$\text{N3.} \quad p(\{\omega_i\} \cup \{\omega_j\}) = p(\{\omega_i\}) + p(\{\omega_j\}), \; i \neq j.$$

- $p(\{\omega_i\})$ (probability of atom i) can now be *negative*.

Why negative probabilities?

- May be helpful to think about certain contextual problems (e.g. non-signaling conditions, counterfactual reasoning in physics).
- May have a meaning in terms of subjective probability.
  - p can define an upper probability distribution by simply setting $P^*(\omega_i) = p(\omega_i) + |p_{\text{min}}|$.
- If nothing else, it is a good computational device.
  - We can compute them easily (compared to uppers/lowers).

Example: Marginals from Alice, Bob, and Carlos

$$p_{xyz} + p_{\overline{x}yz} + p_{x\overline{y}z} + p_{xy\overline{z}} + p_{x\overline{yz}} + p_{\overline{x}y\overline{z}} + p_{\overline{xy}z} + p_{\overline{xyz}} = 1, \tag{1}$$

$$p_{xyz} + p_{\overline{x}yz} + p_{x\overline{y}z} + p_{xy\overline{z}} - p_{x\overline{yz}} - p_{\overline{x}y\overline{z}} - p_{\overline{xy}z} - p_{\overline{xyz}} = 0, \tag{2}$$

$$p_{xyz} + p_{\overline{x}yz} - p_{x\overline{y}z} + p_{xy\overline{z}} - p_{x\overline{yz}} + p_{\overline{x}y\overline{z}} - p_{\overline{xy}z} - p_{\overline{xyz}} = 0, \tag{3}$$

$$p_{xyz} + p_{\overline{x}yz} + p_{x\overline{y}z} - p_{xy\overline{z}} - p_{x\overline{yz}} - p_{\overline{x}y\overline{z}} + p_{\overline{xy}z} - p_{\overline{xyz}} = 0, \tag{4}$$

$$p_{xyz} - p_{\overline{x}yz} - p_{x\overline{y}z} + p_{xy\overline{z}} - p_{x\overline{yz}} - p_{\overline{x}y\overline{z}} + p_{\overline{xy}z} + p_{\overline{xyz}} = 0, \tag{5}$$

$$p_{xyz} - p_{\overline{x}yz} + p_{x\overline{y}z} - p_{xy\overline{z}} - p_{x\overline{yz}} + p_{\overline{x}y\overline{z}} - p_{\overline{xy}z} + p_{\overline{xyz}} = -\frac{1}{2}, \tag{6}$$

$$p_{xyz} + p_{\overline{x}yz} - p_{x\overline{y}z} - p_{xy\overline{z}} + p_{x\overline{yz}} - p_{\overline{x}y\overline{z}} - p_{\overline{xy}z} + p_{\overline{xyz}} = -1. \tag{7}$$

Signed Probabilities

The general solution for the system of equations is

$$\begin{aligned} p_{xyz} &= -p_{\overline{x}yz} &&= -\frac{1}{8} - \delta, \\ p_{x\overline{y}z} &= p_{\overline{x}y\overline{z}} = && \frac{3}{16}, \\ p_{xy\overline{z}} &= p_{\overline{xy}z} = && \frac{5}{16}, \\ p_{x\overline{yz}} &= -p_{\overline{xyz}} = && -\delta, \end{aligned}$$

which gives

$$E(\mathbf{XYZ}) = -\frac{1}{4} - 4\delta.$$

Minimizing total probability mass

- But not all values of δ satisfy N1, i.e., minimize $M^{-} = \sum |p(\omega_i)|$.
- If we impose this, we we have

$$-\frac{1}{8} \leq \delta \leq 0$$

and

$$-\frac{1}{4} \leq E(\mathbf{XYZ}) \leq \frac{1}{2}.$$

Summary: signed probabilities

- Signed probabilities have a possible 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

Bayesian approach

- Standard Bayesian approach is sensitive to choices of prior and likelihood function (well-known issue).

Bayesian approach

- Standard Bayesian approach is sensitive to choices of prior and likelihood function (well-known issue).
- "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

Bayesian approach

- Standard Bayesian approach is sensitive to choices of prior and likelihood function (well-known issue).
- "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
- Say that Deanna starts with $E(\mathbf{XYZ}) = \epsilon$ as her prior.
    - The posterior will be $E(\mathbf{XYZ}) = \epsilon$ regardless of Alice, Bob, and Carlos's opinions.
    - Triple moment is unchanged by lower moment revisions.

Quantum approach

- The quantum-like approach, using vectors on a Hilbert space, seems to be too permissive.
  - No normative power.
- But at least it is explicit!
- Perhaps additional principles could be used.

Negative probability approach

- Negative probabilities (with the minimization of the negative mass) offer a lower and upper bound for values of triple moment (normative).
- They are not as constrained as QM mathematical structures.
- Offer a unifying framework for “rationality” and “irrationality.”

Negative probability approach

- Negative probabilities (with the minimization of the negative mass) offer a lower and upper bound for values of triple moment (normative).
- They are not as constrained as QM mathematical structures.
- Offer a unifying framework for "rationality" and "irrationality."

Thank you!