# Subjective multi-prior probability: a representation of a partial likelihood relation

Shiri Alon and Ehud Lehrer

TAU, 22-11-2011

Outline of the talk:

- Motivation
- Background
- Examples
- Axioms
- Result
- Related literature
- Outline of the proof

# Motivation - Likelihood relation

*Why is it interesting to explore likelihood comparisons among events?*

# Motivation - Likelihood relation

*Why is it interesting to explore likelihood comparisons among events?*

Examples: IAEA reports, US' intelligence reports, assessments of market trends supplied by committees of experts or consulting firms, likelihood estimations of natural events.

# Motivation - Likelihood relation

*Why is it interesting to explore likelihood comparisons among events?*

Examples: IAEA reports, US' intelligence reports, assessments of market trends supplied by committees of experts or consulting firms, likelihood estimations of natural events.

Characteristics:

- Output = likelihood judgements among events.
- Assessments rely on 'objective' data.
- Statements of the form 'event $A$ is more likely than event $B$' are key in producing the output.

# Motivation - Partial relation

- In the examples: ambiguity (arising from lack of information or knowledge), multiple opinions.

# Motivation - Partial relation

- In the examples: ambiguity (arising from lack of information or knowledge), multiple opinions.
- Implication: indecisiveness among certain pairs of events; There is no way to compare likelihood of events, and judge whether event $A$ is at least as likely as event $B$ or vice versa.

# Motivation - Partial relation

- In the examples: ambiguity (arising from lack of information or knowledge), multiple opinions.
- Implication: indecisiveness among certain pairs of events; There is no way to compare likelihood of events, and judge whether event $A$ is at least as likely as event $B$ or vice versa.

*Some comparisons among events may thus be left undetermined.*

# Motivation - Axiomatic treatment

*Why axioms, and why formulated on an 'at least as likely as' binary relation over events?*

# Motivation - Axiomatic treatment

*Why axioms, and why formulated on an 'at least as likely as' binary relation over events?*

- The notion of 'at least as likely as' judgement calls among events seem relatively easy to grasp.
- Convinces to use a functional form, which may be easier to formulate.
- Subscribes means to complete some of the comparisons among pairs of events.
- Lends normative justification to methods already using the functional form (later).

# Motivation - Our goal

**Our goal: suggest an axiomatic treatment of an 'at least as likely as' relation over events, that does not have to be complete.**

What would be an adequate probabilistic representation of such a relation?.......

# Background - Likelihood relation

Ramsey (1931), De Finetti (1931,1937) and Savage (1954):

- Claimed that probability is a subjective notion, emerging from likelihood comparisons among events, and expressing an individual's degree of confidence in the occurrence of events.
- Worked in a purely subjective environment, with no exogenous probabilities assumed.
- Formulated conditions on a likelihood relation over events that imply that the relation may be represented by a probability measure.

# Background - Partial relation

Critique of the completeness assumption:

- Von-Neumann and Morgenstern (1944) and Aumann (1962):
  - Preference over lotteries
  - Questioned completeness on descriptive as well as normative grounds
- Models where alternatives are mappings from events to consequences:
  - Criticized completeness based on ambiguity.
  - Most expressed ambiguity through an inability to formulate one prior probability, and described belief using a set of prior probabilities.

# Conclusion

The same ambiguity considerations seem to apply for an ‘at least as likely as’ relation over events.

# Conclusion

The same ambiguity considerations seem to apply for an ‘at least as likely as’ relation over events.

**If we are to represent a partial relation over events, a single prior probability will not do.**

# Conclusion (cont.)

In order to represent a partial 'at least as likely as' relation over events, we employ a *set* of prior probabilities. Our result:

**A binary relation over events, $\succsim$, satisfies axioms ... , iff, there exists a set of prior probabilities, $\mathcal{P}$, such that for any two events $A$ and $B$,**

$$A \succsim B \quad \Leftrightarrow \quad \mu(A) \geq \mu(B) \ \ \textbf{for all } \mu \in \mathcal{P}.$$

# Example

Let $S = [0,1)$, $\Sigma$ the algebra generated by all intervals $[a,b)$ contained in $[0,1)$, and $\succsim$ the likelihood relation over $\Sigma$ defined by:

$$A \succsim B \quad \Leftrightarrow \quad \lambda(A \setminus B) \geq 3\lambda(B \setminus A).$$

A set of probabilities $\mathcal{P}$ representing $\succsim$:

For $E \in \Sigma$ such that $\lambda(E) = \frac{1}{4}$, let $\pi_E$ be the probability measure defined by the density:

$$f_E(s) = \begin{cases} 2 & s \in E \\ \frac{2}{3} & s \notin E \end{cases}.$$

Let $\mathcal{P}$ be the convex and closed set generated by all measures $\pi_E$.

# Notations

$S$ a nonempty set, with typical elements $s, t, \ldots$.

$\Sigma$ an algebra of *events* over $S$, with typical elements $A, B, \ldots$.

$\mathbf{1}_A$ the indicator function of event $A$.

$\succsim$ A binary relation over $\Sigma$. A statement $A \succsim B$ is interpreted as '$A$ is at least as likely as $B$'.

A probability measure $P$ *agrees* with $\succsim$ if
$A \succsim B \quad \Leftrightarrow \quad P(A) \geq P(B)$.

A probability measure $P$ *almost agrees* with $\succsim$ if
$A \succsim B \quad \Rightarrow \quad P(A) \geq P(B)$.

# Axiomatization - Background

De Finetti introduced four basic postulates:

**Complete Order:** The relation $\succsim$ is complete and transitive.

**Cancellation:** For any three events, $A, B$ and $C$, such that $A \cap C = B \cap C = \emptyset$, $A \succsim B \quad \Leftrightarrow \quad A \cup C \succsim B \cup C$.

**Positivity:** For every event $A$, $A \succsim \emptyset$.

**Non Triviality:** $S \succ \emptyset$.

# Axiomatization - Background

De Finetti introduced four basic postulates:

| | |
|---|---|
| **Complete Order:** | The relation $\succsim$ is complete and transitive. |
| **Cancellation:** | For any three events, $A, B$ and $C$, such that $A \cap C = B \cap C = \emptyset$, $A \succsim B \quad \Leftrightarrow \quad A \cup C \succsim B \cup C$. |
| **Positivity:** | For every event $A$, $A \succsim \emptyset$. |
| **Non Triviality:** | $S \succ \emptyset$. |

Cancellation: any event has *the same marginal contribution of likelihood*, unrelated to other, disjoint events.

# Axiomatization - Background

De Finetti introduced four basic postulates:

**Complete Order:** The relation $\succsim$ is complete and transitive.

**Cancellation:** For any three events, $A, B$ and $C$, such that $A \cap C = B \cap C = \emptyset$, $A \succsim B \quad \Leftrightarrow \quad A \cup C \succsim B \cup C$.

**Positivity:** For every event $A$, $A \succsim \emptyset$.

**Non Triviality:** $S \succ \emptyset$.

Cancellation: any event has *the same marginal contribution of likelihood*, unrelated to other, disjoint events.

The four assumptions are necessary for the relation to have an agreeing probability.

# Axiomatization - Background

De Finetti introduced four basic postulates:

**Complete Order:** The relation $\succsim$ is complete and transitive.

**Cancellation:** For any three events, $A, B$ and $C$, such that $A \cap C = B \cap C = \emptyset$, $A \succsim B \quad \Leftrightarrow \quad A \cup C \succsim B \cup C$.

**Positivity:** For every event $A$, $A \succsim \emptyset$.

**Non Triviality:** $S \succ \emptyset$.

Cancellation: any event has *the same marginal contribution of likelihood*, unrelated to other, disjoint events.

The four assumptions are necessary for the relation to have an agreeing probability. De Finetti asked: are they also *sufficient* for the relation to have an agreeing probability, or even an almost agreeing probability?

# Axiomatization - Background (cont.)

The answer – NO – was given by Kraft, Pratt and Seidenberg (1959). Kraft et al., Scott (1964), Kranz et al. (1971) and Narens (1974), suggested a strengthening of Cancellation:

## Finite Cancellation

*For two sequences of events, $(A_i)_{i=1}^n$ and $(B_i)_{i=1}^n$,*

$$\begin{aligned} &\textit{If} \quad \sum_{i=1}^n \mathbf{1}_{A_i}(s) = \sum_{i=1}^n \mathbf{1}_{B_i}(s) \quad \textit{for all } s \in S, \\ &\textit{and} \quad A_i \succsim B_i \quad \textit{for } i = 1, \ldots, n-1 \ , \\ &\textit{then} \quad B_n \succsim A_n \quad . \end{aligned}$$

# Axiomatization - Background (cont.)

$\sum_{i=1}^{n} \mathbf{1}_{A_i}(s) = \sum_{i=1}^{n} \mathbf{1}_{B_i}(s)$ for all $s \in S$:

each state appears the same number of times in each sequence.

# Axiomatization - Background (cont.)

$\sum_{i=1}^{n} \mathbf{1}_{A_i}(s) = \sum_{i=1}^{n} \mathbf{1}_{B_i}(s)$ for all $s \in S$:

each state appears the same number of times in each sequence.

$+$ *the same marginal contribution of likelihood*

# Axiomatization - Background (cont.)

$\sum_{i=1}^{n} \mathbf{1}_{A_i}(s) = \sum_{i=1}^{n} \mathbf{1}_{B_i}(s)$ for all $s \in S$:

each state appears the same number of times in each sequence.

$+$ *the same marginal contribution of likelihood*

$\Rightarrow$ it cannot be that the $A$-sequence weighs more than the $B$-sequence.

# Axiomatization - Background (cont.)

$\sum_{i=1}^{n} \mathbf{1}_{A_i}(s) = \sum_{i=1}^{n} \mathbf{1}_{B_i}(s)$ for all $s \in S$:

each state appears the same number of times in each sequence.

$+$ *the same marginal contribution of likelihood*

$\Rightarrow$ it cannot be that the $A$-sequence weighs more than the $B$-sequence.

likelihood weights - double-entry booking:

| **credit column** | **debit column** |
|---|---|
| $A_1$ | $B_1$ |
| $\dots$ | $\dots$ |
| $A_{n-1}$ | $B_{n-1}$ |

# Axiomatization - Background (cont.)

$\sum_{i=1}^{n} \mathbf{1}_{A_i}(s) = \sum_{i=1}^{n} \mathbf{1}_{B_i}(s)$ for all $s \in S$:

each state appears the same number of times in each sequence.

$+$ *the same marginal contribution of likelihood*

$\Rightarrow$ it cannot be that the $A$-sequence weighs more than the $B$-sequence.

likelihood weights - double-entry booking:

| **credit column** | **debit column** | |
|---|---|---|
| $A_1$ | $B_1$ | |
| $\dots$ | $\dots$ | |
| $A_{n-1}$ | $B_{n-1}$ | $\Rightarrow$ |
| $B_n$ | $A_n$ | |

# Axiomatization - Main axiom

Our main axiom, Generalized Finite Cancellation (GFC), is a simple extension.

**P4. Generalized Finite Cancellation:**

For two sequences of events, $(A_i)_{i=1}^n$ and $(B_i)_{i=1}^n$, and an integer $k \in \mathbb{N}$,

$$\text{If} \quad \sum_{i=1}^{n-1} \mathbf{1}_{A_i}(s) + k\mathbf{1}_{A_n}(s) = \sum_{i=1}^{n-1} \mathbf{1}_{B_i}(s) + k\mathbf{1}_{B_n}(s) \text{ for all } s \in S,$$

and $A_i \succsim B_i$ for $i = 1, \ldots, n-1$,

then $B_n \succsim A_n$.

# Axiomatization - Additional basic axioms

**P1. Reflexivity:**

For all $A \in \Sigma$, $A \succsim A$.

**P2. Positivity:**

For all $A \in \Sigma$, $A \succsim \emptyset$.

**P3. Non Triviality:**

$\neg(\emptyset \succsim S)$.

# Finite case - Result

**Theorem**

*Suppose that $S$ is finite, and let $\succsim$ be a binary relation over events in $S$. Then statements (i) and (ii) below are equivalent:*

(i) *$\succsim$ satisfies Reflexivity, Positivity, Non Triviality and Generalized Finite Cancellation.*

(ii) *There exists a nonempty set $\mathcal{P}$ of additive probability measures over events in $S$, such that for every $A, B \subseteq S$,*

$$A \succsim B \quad \Leftrightarrow \quad \mu(A) \geq \mu(B) \;\; \textit{for every } \mu \in \mathcal{P}.$$

# Finite case - Result (cont.)

Remarks:

- The set of prior probabilities need not be unique.
- Even if completeness is added, uniqueness is not implied.
- A possible set of priors to consider - the maximal w.r.t inclusion set $=$ the union of all representing sets. Contains all the almost agreeing probabilities.

# Finite case - Example

Let $S = \{H, T\}$ and $\Sigma = 2^S$. Suppose that $H \succsim T$ .

Any probability measure of the sort $(H : p \,;\; T : 1 - p)$ , for $0.5 \leq p \leq 1$ , represents the relation.

# Finite case - Example

Let $S = \{H, T\}$ and $\Sigma = 2^S$. Suppose that $H \succsim T$ .

Any probability measure of the sort $(H : p \; ; \; T : 1 - p)$ , for $0.5 \leq p \leq 1$ , represents the relation.

$\Rightarrow$ No uniqueness, even in the complete case.

# Infinite case - Richness assumption

## Definition

For two events $A, B \in \Sigma$, the notation $A \succ\succ B$ states that there exists a finite partition $\{G_1, \ldots, G_r\}$ of $S$, such that $A \setminus G_i \succsim B \cup G_j$ for all $i, j$.

## P5. Non Atomicity:

If $\neg(A \succsim B)$ then there exists a finite partition of $A^c$, $\{A'_1, \ldots, A'_m\}$, such that for all $i$, $A'_i \succ\succ \emptyset$ and $\neg(A \cup A'_i \succsim B)$.

**Definition**

A set $\mathcal{P}$ of probability measures is said to be *uniformly absolutely continuous*, if:

(a) For any event $B$, $\mu(B) > 0 \Leftrightarrow \mu'(B) > 0$ for every pair of probabilities $\mu, \mu' \in \mathcal{P}$.

(b) For every $\varepsilon > 0$, there exists a finite partition $\{G_1, \ldots, G_r\}$ of $S$, such that for all $j$, $\mu(G_j) < \varepsilon$ for all $\mu \in \mathcal{P}$.

In particular, every probability measure $\mu$ in $\mathcal{P}$ is *locally dense*: for every event $B$, the set $\{\mu(A) \mid A \subset B\}$ is dense in $[0, \mu(B)]$.

# Infinite case - Result

**Theorem**

*Let $\succsim$ be a binary relation over $\Sigma$. Then statements (i) and (ii) below are equivalent:*

(i) *$\succsim$ satisfies axioms* **P1**-**P5**.

(ii) *There exists a nonempty, uniformly absolutely continuous set $\mathcal{P}$ of additive probability measures over $\Sigma$, such that for every $A, B \in \Sigma$,*

$$A \succsim B \quad \Leftrightarrow \quad \mu(A) \geq \mu(B) \text{ for every } \mu \in \mathcal{P}.$$

# Background - Related literature

**Bewley (2002)**

- Alternatives are mappings from an abstract set of *states* to *lotteries* over an abstract set of *consequences* (as in Anscombe and Aumann, 1963).
- Axiomatized a partial relation that admits a multi-prior expected utility representation: for every pair of alternatives $f$ and $g$,

$$f \succsim g \quad \Leftrightarrow \quad \mathbb{E}_\mu(u(f)) \geq \mathbb{E}_\mu(u(g)) \text{ for every } \mu \in \mathcal{P},$$

  for a unique convex and closed set of priors $\mathcal{P}$ and a vN-M utility function $u$.
- Axioms = AA axioms minus completeness ; Result = multi-prior EU instead of (a single prior) EU.

# Background - Related literature (cont.)

**Bewley (2002)**

Bewley's model with only two consequences (better/worse) yields an 'at least as likely as' relation over events.

# Background - Related literature (cont.)

**Bewley (2002)**

Bewley's model with only two consequences (better/worse) yields an 'at least as likely as' relation over events.

Downside: his setup is not purely subjective, as it employs mixtures with exogenous probabilities.

# Background - Related literature (cont.)

**Ghirardato et al (2003)**

- Alternatives are mappings from an abstract set of *states* to an abstract set of *consequences* = purely subjective environment (no exogenous probabilities).
- An *á* – *la* Bewley result.

# Background - Related literature (cont.)

**Ghirardato et al (2003)**

- Alternatives are mappings from an abstract set of *states* to an abstract set of *consequences* = purely subjective environment (no exogenous probabilities).
- An *á – la* Bewley result.

Downside: Impossible to apply the model for two consequences alone, hence cannot be applied to identify solely an 'at least as likely' relation over events.

# Background - Related literature (cont.)

**Nehring (2009)**

- The primitive is a relation over events, not multi-valued alternatives.
- No exogenous probabilities assumed.
- A multi-prior probability result: there exists a set of prior probabilities, $\mathcal{P}$, such that for any two events $A$ and $B$,

$$A \succsim B \quad \Leftrightarrow \quad \mu(A) \geq \mu(B) \text{ for all } \mu \in \mathcal{P}.$$

for a unique convex and closed set $\mathcal{P}$.

# Background - Related literature (cont.)

**Nehring (2009)**

Downside: explicit assumption that any event can be divided into two equally-likely events. That is,

for any event $A$ there exists an event $B \subset A$, such that $\mu(B) = \mu(A)/2$ for all $\mu \in \mathcal{P}$.

Implications:

- In particular, all the priors agree on a rich algebra of events: the one generated by dividing the entire space into n $2^n$ equally likely events, for any integer $n$.

  (Note: not satisfied in the example presented earlier.)
- Result cannot be applied to a finite state space.

# Outline of the proof

- Define the closed convex cone generated by $\{ \mathbf{1}_A - \mathbf{1}_B \mid A \succsim B \}$ .
- Show that if $\mathbf{1}_A - \mathbf{1}_B = \sum_{i=1}^{n} \alpha_i(\mathbf{1}_{A_i} - \mathbf{1}_{B_i})$ for $\alpha_i > 0$, $A_i \succsim B_i$ , then $A \succsim B$ – This is precisely GFC (first for rational coefficients, then for real coefficients).
- In the finite case, the convex cone is generated by a finite number of indicator differences, therefore is closed. A separation theorem yields the result.
- In the infinite case, Non Atomicity yields that there cannot be $\mathbf{1}_A - \mathbf{1}_B$ on the boundary of the cone with $\neg(A \succsim B)$ . A separation theorem again yields the result.