What chance-credence norms should not be

Richard Pettigrew

Department of Philosophy
University of Bristol

29th May 2012
The question

How should our credences in propositions concerning objective chances relate to our credences in other propositions?

- Enumerate the possible chance-credence norms.
- Show that one *prima facie* plausible one in fact behaves very badly in the circumstances in which it is designed to be used.
Terminology: credences

- Let $F$ be the algebra of propositions about which our agent has an opinion. (Assume $F$ is finite.)
- Let $b_t : F \rightarrow [0, 1]$ be her *credence function* at $t$.
- Let $E_t$ be her total evidence at $t$. 
Terminology: chances

The *ur-chance function* at world $w$ is the probability function $ch_w$ such that, if $H_{tw}$ is the history of $w$ up to time $t$, then the chances in $w$ at $t$ are given by $ch_w(\cdot|H_{tw})$.

Given a probability function $ch$, let

$$C_{ch} \equiv \text{The ur-chances are given by } ch.$$  

Thus, $C_{ch}$ is true at $w$ iff $ch = ch_w$.

(Assume our agent has an opinion about only finitely many possible ur-chance functions.)
The putative chance-credence norms

\[(\text{PP}) \quad b_t(A|C_{ch}) = ch(A|E_t). \quad \text{(Lewis 1980)}\]

\[(\text{NP}) \quad b_t(A|C_{ch}) = ch(A|E_t \land C_{ch}). \quad \text{(Hall 1994)}\]

\[(\text{IP}) \quad b_t(A) = \sum_{ch} b_t(C_{ch}) ch(A|E_t). \quad \text{(Ismael 2008)}\]
Suppose we know that the world contains only four coin tosses.

Sixteen possible worlds:

- HHHH, HHHT, HHTH, \ldots, TTTH, TTTT.

Five possible ur-chance functions for the reductionist:

\[
\begin{align*}
ch_0(\text{Heads}) &= 0 & ch_1(\text{Heads}) &= \frac{1}{4} & ch_2(\text{Heads}) &= \frac{1}{2} \\
ch_3(\text{Heads}) &= \frac{3}{4} & ch_4(\text{Heads}) &= 1
\end{align*}
\]

\[
\begin{align*}
C_{ch_0} &\equiv TTTT \\
C_{ch_1} &\equiv TTTT \lor TTHT \lor THTT \lor HTTT \\
C_{ch_2} &\equiv HHTT \lor HTHT \lor TTHH \lor THHT \lor HTTH \\
C_{ch_3} &\equiv HHHT \lor HHTH \lor HTHH \lor THHH \\
C_{ch_4} &\equiv HHHH
\end{align*}
\]
Definition
An ur-chance function \( ch \) is **self-undermining** in the presence of evidence \( E \) if \( ch(C_{ch}|E) < 1 \).

In our example, the self-undermining ur-chance functions are: \( ch_1, ch_2, ch_3 \).

For example:

\[
ch_3(C_{ch_1}) = ch_3(TTTH) + \ldots + ch_3(HTTT) \\
= 4 \times \left( \frac{3}{4} \right) \times \left( \frac{1}{4} \right)^3 \\
= \frac{3}{64} > 0.
\]

So \( ch_3(C_{ch_3}) < 1 \).
Theorem

If there is at least one chance function that is self-undermining in the presence of $E_t$, then (PP) cannot be satisfied at $t$.

Proof. If $ch$ is self-undermining in the presence of $E_t$, then

$$ch(C_{ch}|E_t) < 1 = b_t(C_{ch}|C_{ch})$$

Theorem

Whatever the ur-chance functions are like, (NP) can be satisfied at any time.

Theorem

Whatever the ur-chance functions are like, (IP) can be satisfied at any time.
Three problems for (IP)

The reductionist must choose between (NP) and (IP).

- **The Problem of Updating**
  There is no satisfactory updating rule that is consistent with (IP).

- **The Problem of Determinism**
  - In the absence of evidence, (IP) demands certainty in determinism.
  - In the presence of little evidence, (IP) demands certainty about future chance events.

- **The Problem of Deference**
  If (IP) formalizes deference, then ur-chance functions don’t defer to themselves.

Thus, the reductionist ought to choose (NP).
The Problem of Updating

Bayesian Conditionalization (BC)
It ought to be the case that:

\[ b_{t'}(A) = b_t(A | E_{t'}) \]

providing \( b_t(E_{t'}) > 0 \).
The Problem of Updating

Theorem

*If*

- $b_t$ satisfies (NP);
- $b_{t'}$ is obtained from $b_t$ in accordance with (BC)

*then*

- $b_{t'}$ satisfies (NP).
The Problem of Updating

Theorem
There are $b_t$ and $b_{t'}$ such that

- $b_t$ satisfies (IP);
- $b_{t'}$ is obtained from $b_t$ in accordance with (BC)

and yet

- $b_{t'}$ does not satisfy (IP).
What’s so good about (BC)?

Definition

\( b \) is *immodest* if, for all \( c \neq b \),

\[
\sum_{w \in W} b(w)EU(c, w) < \sum_{w \in W} b(w)EU(b, w)
\]

Theorem (Greaves and Wallace)

*If* \( b_t(\cdot | E_{t'}) \) *is immodest, then, for all* \( c \neq b_t(\cdot | E_{t'}) \),

\[
\sum_{w \in E_{t'}} b_t(w)EU(c, w) < \sum_{w \in E_{t'}} b_t(w)EU(b_t(\cdot | E_{t'}), w)
\]
The Problem of Updating

The Brier score

\[ B(b, w) := 1 - \sum_{A \in F} (b(A) - v_w(A))^2 \]

Theorem
Relative to \( B \),

- \( b_t \) is immodest over \( E_t \) \( \iff \) \( b_t \) is a probability function and \( b_t(E_t) = 1 \).

- (BC) maximizes expected epistemic utility.
The Problem of Updating

The Chance Brier score

$$C^E_I(b, w) := 1 - \sum_{A \in \mathcal{F}} (b(A) - ch_w(A|E))^2$$

Theorem
Relative to $C^E_I$,

- $b_t$ is immodest over $E_t$ iff $b_t$ satisfies (IP).
- The following updating rule minimizes expected epistemic utility:

$$b_{t'}(A) = \sum_{ch} b_t(C_{ch}|E_{t'})ch(A|E_{t'})$$

Call it Ismael Conditionalization or (IC).
The Problem of Updating

The victory is shortlived...

Theorem

There are credence functions $b_t$ and $b_{t'}$ such that

- $b_t$ satisfies (IP),
- $b_{t'}$ is obtained from $b_t$ in accordance with (IC) and yet
- $b_{t'}$ does not satisfy (IP).
The Problem of Determinism

Theorem
Suppose \( ch \neq ch' \) and
(i) \( ch \) is not self-undermining in the presence of \( E_t \)
(ii) \( ch'(C_{ch}|E_t) > 0 \)
Then, if \( b_t \) satisfies (IP), then \( b_t(C_{ch'}) = 0 \).
The Problem of Determinism

Suppose we know that the world contains only four coin tosses. Sixteen possible worlds:

HHHH, HHHT, HHTH, ..., TTTH, TTTT.

Five possible ur-chance functions for the reductionist:

\[ ch_n(\text{Heads}) = \frac{n}{4} \quad n = 0, 1, 2, 3, 4 \]

- Self-undermining in the presence of \( E_t = \top \): \( ch_1, ch_2, ch_3 \).
- \( ch_i(C_{ch_0}), ch_i(C_{ch_4}) > 0 \), for \( i = 1, 2, 3 \).
- Therefore, \( b_t(C_{ch_i}) = 0 \), for \( i = 1, 2, 3 \).
- Therefore, \( b_t(\text{Determinism}) = b_t(C_{ch_0} \lor C_{ch_4}) = 1 \).
The Problem of Determinism

Suppose we know that the world contains only four coin tosses. Sixteen possible worlds:

\[
\text{HHHH, HHHT, HHTH, \ldots, TTTT, TTTT.}
\]

Five possible ur-chance functions for the reductionist:

\[
ch_n(\text{Heads}) = \frac{n}{4} \quad n = 0, 1, 2, 3, 4
\]

- Self-undermining in the presence of \(E_t = H\): \(ch_1, ch_2, ch_3\).
- \(ch_i(C_{ch_4} | H) > 0\), for \(i = 1, 2, 3\).
- Therefore, \(b_t(C_{ch_i}) = 0\), for \(i = 1, 2, 3\).
- Therefore, \(b_t(C_{ch_4}) = b_t(\text{HHHH}) = 1\).
The Problem of Determinism

There is no analogous problem for (NP):

**Theorem**

Suppose $\lambda_{ch} \geq 0$ for all $ch$ and $\sum_{ch} \lambda_{ch} = 1$. Then define $b_t$ as follows:

$$b_t(A) = \sum_{ch} \lambda_{ch} ch(A|C_{ch} \land E_t)$$

Then $b_t$ satisfies (NP).
The Problem of Deference

Do the ur-chance functions satisfy (IP)? Not all of them.

\[
ch_0(A) = \sum_{i=0}^{4} ch_0(C_{ch_i}) ch_i(A)
\]

\[
ch_4(A) = \sum_{i=0}^{4} ch_4(C_{ch_i}) ch_i(A)
\]

\[
ch_1(HHHH) = \frac{1}{256} \neq \frac{2128}{65,536} = \sum_{i=0}^{4} ch_1(C_{ch_i}) ch_i(HHHH)
\]

\[
ch_2(HHHH) = \frac{1}{16} \neq \frac{15}{256} = \sum_{i=0}^{4} ch_2(C_{ch_i}) ch_i(HHHH)
\]

\[
ch_3(HHHH) = \frac{81}{256} \neq \frac{24,528}{65,536} = \sum_{i=0}^{4} ch_3(C_{ch_i}) ch_i(HHHH)
\]
A chance-credence norm is supposed to express the intuition that agents ought to defer to the chances when they set their credences.

If deference to the chances involves satisfying (IP) and if the chances violate (IP), then the chances do not defer to themselves.

Meta-Normative Principle
An agent ought not to defer to an epistemic expert that does not defer to itself.
The Problem of Deference

No analogous problem for (NP) (under certain assumptions):

**Theorem**

*Suppose the possible ur-chance functions are $ch_0, \ldots, ch_n$. Suppose that for all worlds $w, w'$ such that $ch_w = ch_{w'}$, we have $ch(w) = ch(w')$, for all $ch$. Then each possible ur-chance function satisfies (NP).*
Objection  \( ch(C_{ch'}) \) is not defined.

Reply  Yes, it is. Consider the example from above:

- \( C_{ch_1} \equiv \text{TTTH} \lor \text{TTHT} \lor \text{THTT} \lor \text{HTTT} \).
- Each \( ch_i \) is defined at \( \text{TTTH} \), \( \text{TTHT} \), \( \text{THTT} \), and \( \text{HTTT} \).

And in general:

- Chance hypotheses (of the form \( C_{ch} \)) are disjunctions of world histories.
- Chances must be defined on world histories in order to define the notion of ‘fit’ required by the Best-System Analysis of chance.
Conclusion

Which chance-credence norm should we adopt?

- (IP): implausible consequences in the presence of self-undermining chances.
- (NP): no analogous problems.
References


Draft of paper available at:
http://eis.bris.ac.uk/~rp3959/papers/