WHEN CAN A PRIOR BE RECOVERED FROM A POSTERIOR?

Carl Wagner

1. Strict Conditioning.

Let $A$ be an algebra of subsets of a set $\Omega$ of possible states of the world. Suppose that you are given a finitely additive probability measure (henceforth, “probability”) $q$ on $A$, and are told that $q$ has come from some probability $p$ on $A$ by conditioning on the event $E$. Can you determine $p$?

Well, yes, if $E=\Omega$ (in which case, it must be true that $p=q$), but not if $E$ is a proper subset of $\Omega$. For given the fully specified posterior $q$, along with $E$, there exist infinitely many priors that yield $q$ by conditioning on $E$. Here’s why: Choose any number $\nu \in (0,1]$ and any $\omega \in E^c$. Define the probabilities

$$m_\omega (A) = 1 \text{ if } \omega \in A \text{ and } m_\omega (A) = 0 \text{ if } \omega \not\in A,$$

and

$$p_\nu (A) = \nu q(A \cap E) + (1-\nu)m_\omega (A \cap E^c).$$

It is straightforward to check that $m_\omega$ and $p_\nu$ are indeed probabilities on $A$. Furthermore,

$$p_\nu (A \mid E) = \frac{p_\nu (A \cap E)}{p_\nu (E)} = \frac{\nu q(A \cap E)}{\nu} = q(A \cap E) = q(A),$$

since $q(E^c) = 0$ and, hence, $q(A \cap E^c) = 0$.

2. Jeffrey Conditioning.

With $\Omega$ and $A$ as above, suppose that $E = \{E_1, \ldots, E_n\}$ is a measurable partition of $\Omega$ (i.e., a set of nonempty, pairwise disjoint events in $A$, with union equal to $\Omega$), where $n \geq 2$. Suppose that you are given a probability $q$ on $A$, and you are told that $q$ has come from some probability $p$ on $A$ by Jeffrey conditioning (henceforth, “JC”) on $E$, i.e., that for all $A \in A$,

$$q(A) = \sum_{i=1}^n e_i p(A \mid E_i)$$

for some probability $p$ such that $p(E_i) > 0$, $i = 1, \ldots, n$, with each $e_i (= q(E_i)) > 0$, and $e_1 + \cdots + e_n = 1$. It is easy to check that a formula of type (2.1) holds with the posited conditions if and only if

$$q(A \mid E_i) = p(A \mid E_i) \text{ for all } A \in A, \text{ and each } i = 1, \ldots, n.$$
Condition (2.2) is variously termed the rigidity, sufficiency, and invariance. Can \( p \) be recovered from \( q \), along with knowledge of the values \( e_1, \ldots, e_n \) and the fact that \( q \) has come from \( p \) by JC on \( \mathcal{E} \)? Again, no. To take a simple illustration, suppose that \( \Omega = \{1,2,3,4\} \), \( \mathcal{A} = 2^\Omega \) (the set of all subsets of \( \Omega \), \( E_1 = \{1,2\} \) and \( E_2 = \{3,4\} \). Let \( q(\{1\}) = 1/9 \), \( q(\{2\}) = 2/9 \), \( q(\{3\}) = q(\{4\}) = 1/3 \), extending \( q \) to the remaining subsets of \( \Omega \) in the obvious way. We may construct infinitely many probabilities \( p_v \) on \( \mathcal{A} \), such that \( q \) comes from \( p_v \) by JC on \( \{E_1, E_2\} \) with \( q(E_1) = e_1 = 1/3 \) and \( q(E_2) = e_2 = 2/3 \). For each \( \nu \in (0,1) \), let \( p_v(\{1\}) = \nu / 3 \), \( p_v(\{2\}) = 2\nu / 3 \), \( p_v(\{3\}) = p_v(\{4\}) = (1-\nu) / 2 \). It is easily checked that each \( p_v \) has the desired property.\(^1\)

**Exercise.** Let \( q \) be a probability on an algebra \( \mathcal{A} \) of subsets of the set \( \Omega \), and let \( \mathcal{E} = \{E_1, \ldots, E_n\} \) be a measurable partition of \( \Omega \), with \( q(E_i) = e_i \) for \( i = 1, \ldots, n \). Let \( f_1, \ldots, f_n \) be any sequence of positive real numbers such that \( f_1 + \cdots + f_n = 1 \). For all \( A \in \mathcal{A} \), let

\[
(2.3) \quad p(f_j)(A) = \sum_{i=1}^n f_i q(A \mid E_i).
\]

Then \( q \) comes from \( p(f_j) \) by JC on \( \mathcal{E} \), with \( q(E_i) = e_i \), \( i = 1, \ldots, n \).

3. An Alternative Parameterization of Jeffrey Conditioning.

Let \( \Omega, \mathcal{A}, \) and \( \mathcal{E}= \{E_1, \ldots, E_n\} \) be as above, and let \( p \) be a probability on \( \mathcal{A} \) such that \( p(E_i) > 0 \) for \( i = 1, \ldots, n \). Let \( u_1, \ldots, u_n \) be any sequence of positive real numbers, and consider revising the prior \( p \) to a posterior \( q \) by the formula

\[
(3.1) \quad q(A) = \frac{\sum_{i=1}^n u_i p(A \cap E_i)}{\sum_{i=1}^n u_i p(E_i)}, \text{ for all } A \in \mathcal{A}.
\]

It is straightforward to check that the set function \( q \) is indeed a probability on \( \mathcal{A} \). Moreover, initial appearances notwithstanding, formula (3.1) furnishes no new and exotic method of probability revision. For, for all \( A \in \mathcal{A} \), and \( j = 1, \ldots, n \),

\[
(3.2) \quad q(A \mid E_j) = \frac{q(A \cap E_j)}{q(E_j)} = \frac{u_j p(A \cap E_j)}{u_j p(E_j)} = p(A \mid E_j).
\]
So \( q \) simply comes from \( p \) by JC on \( E \). But what do the parameters \( u_i \) represent? Recall that if \( q \) is a revision of the probability \( p \) and \( A \) and \( B \) are events, then the Bayes factor \( \beta_p^q(A:B) \) is simply the ratio of the new odds on \( A \) against \( B \) to the old such odds, i.e.,

\[
(3.3) \quad \beta_p^q(A:B) = \frac{q(A)}{q(B)} \cdot \frac{p(A)}{p(B)}.
\]

When \( q \) comes from \( p \) by conditioning on \( E \), then \( \beta_p^q(A:B) \) is simply the likelihood ratio \( p(E|A)/p(E|B) \).

**Exercise.** From formula (3.1) it follows that, for all \( i, j \in \{1, \ldots, n\} \),

\[
(3.4) \quad \frac{u_i}{u_j} = \beta_p^q(E_i:E_j).
\]

Interestingly, given a posterior \( q \), the partition \( E \), the parameters \( u_1, \ldots, u_n \), and the fact that \( q \) has come from some probability by JC on \( E \), this information determines a unique prior \( p \) satisfying formula (3.1), namely the probability \( p \) defined for all \( A \in \mathcal{A} \) by

\[
(3.5) \quad p(A) = \frac{\sum_{i=1}^n u_i^{-1}q(A \cap E_i)}{\sum_{i=1}^n u_i^{-1}q(E_i)}.
\]

It is straightforward to check that (3.5) implies (3.1). But there is more work to be done to show that \( p \), as defined by (3.5), is the only prior that yields \( q \) by means of formula (3.1). For this we must show that (3.1) implies (3.5).

From (3.1) and its consequence (3.4),

\[
(3.6) \quad \frac{u_j}{u_i} = \beta_p^q(E_j:E_i) = \frac{q(E_j)p(E_i)}{q(E_i)p(E_j)}, \quad \text{and so}
\]

\[
(3.7) \quad p(E_j) = \frac{u_iq(E_j)p(E_i)}{u_jq(E_i)}, \quad \text{whence}
\]

\[
(3.8) \quad \frac{p(E_j)}{p(E_i)} = \frac{u_iq(E_j)}{u_jq(E_i)}.
\]

Summing each side of (3.8) from \( j = 1 \) to \( j = n \) yields
\[
\frac{1}{p(E_i)} = \sum_{j=1}^{n} \frac{u_j}{q(E_i)} q(E_j) = \sum_{i=1}^{n} u_i q(E_i), \quad \text{whence}
\]
\[
p(E_i) = \left( \sum_{i=1}^{n} u_i q(E_i)^{-1} \right)^{-1},
\]
and substituting the right-hand side of (3.10) for \( p(E_i) \) in (3.7) yields
\[
p(E_j) = \frac{u_j^{-1} q(E_j)}{\sum_{i=1}^{n} u_i^{-1} q(E_i)},
\]
which establishes (3.5) when \( A = E_j \).

But by (3.2), \( p(A \mid E_j) = q(A \mid E_j) \) for all \( A \in \mathbf{A} \) and \( j = 1, ..., n \). So
\[
p(A) = \sum_{j=1}^{n} p(E_j) p(A \mid E_j) = \sum_{j=1}^{n} p(E_j) q(A \mid E_j) = \sum_{j=1}^{n} \frac{u_j^{-1} q(E_j) q(A \mid E_j)}{\sum_{i=1}^{n} u_i^{-1} q(E_i)} = \]
\[
\frac{\sum_{j=1}^{n} u_j^{-1} q(A \cap E_j)}{\sum_{i=1}^{n} u_i^{-1} q(E_i)}.
\]

**Remark.** Special cases of formula (3.1) occur in Field (1978), where
\[
u_i = G_i := (\prod_{j=1}^{n} \beta_x^p(E_j : E_j))^{
u_u},
\]
and Jeffrey and Hendrickson (1988/89) and Wagner (2002), where
\[
u_i = B_i := \beta^q_p(E_i : E_i).
\]

**References**


Notes

1. For every finite $\Omega = \{\omega_1, \ldots, \omega_n\}$, and any probabilities $p$ and $q$ on $2^\Omega$ for which $p(\{\omega_i\}) > 0$ and $q(\{\omega_i\}) > 0$ for $i = 1, \ldots, n$, it is (trivially) the case that $q$ comes from $p$ by JC on $E = \{E_1, \ldots, E_n\}$, where $E_i = \{\omega_i\}$ and $e_i = q(\{\omega_i\})$. That is, each positive probability $q$ on $2^\Omega$ comes from every positive probability $p$ on $2^\Omega$ by JC on $E$. In such cases $q$ obliterates all traces of the prior $p$ from which it came by JC, including any nontrivial information about the conditional probabilities $p(A \mid E_i) = q(A \mid E_i)$, which take only the values zero and one here.