WHEN CAN A PRIOR BE RECOVERED FROM A POSTERIOR?

Carl Wagner

1. Strict Conditioning.

Let **A** be an algebra of subsets of a set Ω 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 Ω. 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_ω (called the *point mass at* ω) and p_ν for each $A \in \mathbf{A}$ by

(1.1) $m_\omega(A)=1$ if $\omega \in A$ and $m_\omega(A)=0$ if $\omega \notin A$, and

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

It is straightforward to check that m_ω and p_ν are indeed probabilities on **A**. Furthermore,

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

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

2. Jeffrey Conditioning.

With Ω and **A** as above, suppose that **E** = $\{E_1,...,E_n\}$ is a *measurable partition of* Ω (i.e., a set of nonempty, pairwise disjoint events in **A**, with union equal to Ω), where $n \ge 2.$ 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 $A \in \mathbf{A}$,

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

for some probability p such that $p(E_i)>0,\ i=1,...,n$, with each $e_i\left(=q(E_i)\right)>0$, 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

(2.2) $q(A|E_i)=p(A|E_i)$ for all $A \in \mathbf{A}$, and each $i=1,...,n.$

Condition (2.2) is variously termed the *rigidity*, *sufficiency*, and *invariance*. Can p be recovered from q, along with knowledge of the values $e_1,...,e_n$ and the fact that q has come from p by JC on **E** ? Again, no. To take a simple illustration, suppose that $\Omega = \{1,2,3,4\}$, **A** = 2^Ω (the set of all subsets of Ω, $E_1 = \{1,2\}$ 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 Ω in the obvious way. We may construct infinitely many probabilities p_ν on **A**, such that q comes from p_ν 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 $p_\nu(\{1\}) = \nu/3$, $p_\nu(\{2\}) = 2\nu/3$, $p_\nu(\{3\}) = p_\nu(\{4\}) = (1-\nu)/2$. It is easily checked that each p_ν has the desired property.[1]

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

(2.3) $$p_{(f_i)}(A) = \sum_{i=1}^{n} f_i q(A \mid E_i).$$

Then q comes from $p_{(f_i)}$ by JC on **E**, with $q(E_i) = e_i$, $i = 1,...,n$.

3. An Alternative Parameterization of Jeffrey Conditioning.

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

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

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

(3.2) $$q(A \mid E_j) = \frac{q(A \cap E_j)}{q(E_j)} = \frac{u_j p(A \cap E_j)}{u_j p(E_j)} = p(A \mid E_j).$$

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) $$\beta_p^q(A:B) = \frac{q(A)/q(B)}{p(A)/p(B)} .$$

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 $i, j \in \{1,...,n\}$,

(3.4) $$\frac{u_i}{u_j} = \beta_p^q(E_i : E_j)$$

Interestingly, given a posterior q, the partition **E,** the parameters $u_1,...,u_n$, 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 $A \in \mathbf{A}$ by

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

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) $$\frac{u_j}{u_1} = \beta_p^q(E_j : E_1) = \frac{q(E_j)p(E_1)}{q(E_1)p(E_j)}$$, and so

(3.7) $$p(E_j) = \frac{u_1 q(E_j) p(E_1)}{u_j q(E_1)}$$, whence

(3.8) $$\frac{p(E_j)}{p(E_1)} = \frac{u_1 q(E_j)}{u_j q(E_1)}.$$

Summing each side of (3.8) from $j=1$ to $j=n$ yields

(3.9) $$\frac{1}{p(E_1)} = \sum_{j=1}^{n} \frac{u_1}{q(E_1)} \frac{q(E_j)}{u_j} = \sum_{i=1}^{n} \frac{u_1 q(E_i)}{u_i q(E_1)}$$, whence

(3.10) $$p(E_1) = (\sum_{i=1}^{n} \frac{u_1 q(E_i)}{u_i q(E_1)})^{-1} ,$$

and substituting the right-hand side of (3.10) for $p(E_1)$ in (3.7) yields

(3.11) $$p(E_j) = \frac{u_j^{-1} q(E_j)}{\sum_{i=1}^{n} u_i^{-1} q(E_i)},$$

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

(3.12) $$p(A) = \sum_{j=1}^{n} p(E_j) p(A \mid E_j) = \sum_{j=1}^{n} p(E_j) q(A \mid E_j) = \sum_{j=1}^{n} \frac{u_j^{-1} q(E_j) q(A \mid E_j)}{\sum_{i=1}^{n} u_i^{-1} q(E_i)} =$$

$$\frac{\sum_{i=1}^{n} u_i^{-1} q(A \cap E_i)}{\sum_{i=1}^{n} u_i^{-1} q(E_i)}.$$

Remark. Special cases of formula (3.1) occur in Field (1978), where

(3.13) $$u_i = G_i := (\prod_{j=1}^{n} \beta_p^q (E_i : E_j))^{1/n} ,$$

and Jeffrey and Hendrickson (1988/89) and Wagner (2002), where

(3.14) $$u_i = B_i := \beta_p^q (E_i : E_1).$$

References

1. Hartry Field (1978), A note on Jeffrey conditionalization, *Philosophy of Science* **45**: 361-367.

2. Richard Jeffrey and Michael Hendrickson (1988/89), Probabilizing pathology, *Proceedings of the Aristotelian Society* **89** (Part 3), 211-225.

3. Carl Wagner (2002), Probability kinematics and commutativity, *Philosophy of Science* **69**: 266-278.

Notes

1. For every finite $\Omega = \{\omega_1, \ldots, \omega_n\},$ and any probabilities p and q on 2^{Ω} 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^{Ω} comes from *every* positive probability p on 2^{Ω} 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.