Postscript to my paper, Probability Kinematics and Commutativity
(*Philosophy of Science* **69** (June 2002), pp. 266-278)

Carl Wagner (1 December 2020)

These notes elaborate on material appearing in the above paper, in which proofs of the more elementary results were omitted. I have written them for students and others new to formal epistemology, should they find a detailed exposition of such proofs to be helpful.

1. Notation and terminology. In what follows, Ω denotes a set of possible states of the world, and $\mathbf{A}$ is a sigma algebra of subsets (called *events*) of Ω. $\mathbf{E} = \{E_i\}$ and $\mathbf{F} = \{F_j\}$ are countable partitions of Ω, with $E_i, F_j \in \mathbf{A}$ for all i, j. To avoid having constantly to postulate the positivity of various probabilities, we adopt the convention that all probability measures p on $\mathbf{A}$ are assumed to be *strictly coherent*, in the sense that $p(A) > 0$ whenever $A \neq \varnothing$.

If p and q are probability measures on $\mathbf{A}$ and $A \in \mathbf{A}$, the *relevance quotient*, $R_p^q(A)$, is defined by the formula $R_p^q(A) := q(A)/p(A)$. Typically, q is thought of as resulting from the revision of p as a result of encountering new evidence. In such cases, p is called the *prior probability*, and q the *posterior probability*. If q comes from p by conditioning on the event E, then

$$(1.1) \qquad R_p^q(A) = p(A \mid E)/p(A) = p(AE)/p(A)p(E) = p(E \mid A)/p(E).$$

Above, and in what follows, AE is an abbreviation for $A \cap E$.

If p and q are as above, and A and B are events, the *Bayes factor*, $B_p^q(A:B)$, is defined by the formula $B_p^q(A:B) := \dfrac{q(A)/q(B)}{p(A)/p(B)}$, which is simply the ratio of the posterior odds on A against B to the prior such odds. Relevance quotients and Bayes factors are clearly connected by the formula

$$(1.2) \qquad B_p^q(A:B) = \frac{R_p^q(A)}{R_p^q(B)} .$$

When q comes from p by conditioning on the event E, then $B_p^q(A:B)$ reduces to the familiar *likelihood ratio* $p(E \mid A)/p(E \mid B)$.

A probability measure q is said to come from p by *probability kinematics* (or *Jeffrey conditioning*) *on* $\mathbf{E}$ if,

$$(1.3) \qquad q(A) = \sum_{i=1} e_i p(A \mid E_i), \text{ with each } e_i > 0, \text{ and } \sum_i e_i = 1.$$

It is easy to see that (1.3) is equivalent to the *rigidity condition*

(1.4) $$q(A \mid E_i) = p(A \mid E_i), \text{ along with } q(E_i) = e_i.$$

2. Successive probability-kinematical updating: the general model

Figure 2.1

$$\begin{tikzcd}
p \arrow[rr, "\mathbf{E}, e_i"] \arrow[d, "\mathbf{F}, f'_j"'] & & q \arrow[d, "\mathbf{F}, f_j"] \\
q' \arrow[r, "\mathbf{E}, e'_i"'] & r' & r
\end{tikzcd}$$

Suppose that p is revised to q by formula (1.3), and then q is revised to r, where

(2.1) $$r(A) = \sum_j f_j q(A \mid F_j), \text{ with each } f_j > 0, \text{ and } \sum_j f_j = 1.$$

Expressing the conditional probabilities in (2.1) by means of (1.3) yields the formula

(2.2) $$r(A) = \sum_{i,j} \frac{e_i f_j}{p(E_i) q(F_j)} p(AE_i F_j).$$

Now consider reversing the order of updating, first revising p to q' by the formula

(2.3) $$q'(A) = \sum_{j=1}^{m} f'_j p(A \mid F_j) \text{ , with each } f'_j > 0, \text{ and } \sum_j f'_j = 1.$$

and then revising q' to r' by the formula

(2.4) $$r'(A) = \sum_{i=1}^{n} e'_i \, q'(A \mid E_i), \text{ with each } e'_i > 0, \text{ and } \sum_i e'_i = 1.$$

Note that, as our notation suggests, e'_i may differ from e_i, and f'_j from f_j.

Combining (2.3) with (2.4) yields

(2.5) $$r'(A) = \sum_{i,j} \frac{e'_i f'_j}{q'(E_i) p(F_j)} p(AE_i F_j).$$

Comparing (2.2) and (2.5), it is obvious, and unsurprising, that r' may differ from r. Under what conditions will the identity $r' = r$ hold ? In Wagner (2002) I proved that the *Bayes factor identities*

(2.6) $$B_p^{q'}(F_{j_1} : F_{j_2}) = B_q^{r}(F_{j_1} : F_{j_2}) \text{ and } B_{q'}^{r'}(E_{i_1} : E_{i_2}) = B_p^{q}(E_{i_1} : E_{i_2})$$

are sufficient and, under mild regularity conditions, necessary for $r' = r$. My paper also contains a number of observations about the special case of successive updating in which $e_i' = e_i$ and $f_j' = f_j$, as explored earlier in a beautiful paper by Diaconis and Zabell (1982). The following section contains proofs of those observations, which, though elementary, may be of interest to students and others new to the subject.

3. Successive probability-kinematical updating: the restricted model

Figure 3.1

$$\begin{tikzcd}
p \arrow[rr, "\mathbf{E}, e_i"] \arrow[d, "\mathbf{F}, f_j"'] & & q \arrow[d, "\mathbf{F}, f_j"] \\
q' \arrow[r, "\mathbf{E},\ e_i"'] & r' & r
\end{tikzcd}$$

In the restricted model, we have from (2.2) and (2.5) that

(3.1) $$r(A) = \sum_{i,j} \frac{e_i f_j}{p(E_i)q(F_j)} p(AE_iF_j) \text{ and } r'(A) = \sum_{i,j} \frac{e_i f_j}{q'(E_i)p(F_j)} p(AE_iF_j).$$

So it is obvious that what Diaconis and Zabell term *Jeffrey independence*, namely, the conditions,

(3.2) $$(a) \quad q'(E_i) = p(E_i) \qquad \text{and} \qquad (b) \quad q(F_j) = p(F_j),$$

are sufficient to ensure that $r' = r$. Indeed, Diaconis and Zabell proved that these conditions are necessary as well. In Wagner (2001) I noted, without proof, several equivalents of Jeffrey independence.

Theorem 3.1. In the restricted model, the following are equivalent: (*i*) Jeffrey independence, as expressed above in (3.2); (*ii*) the relevance quotient identities

(3.3) $\quad (a)\ R^{r'}_{q'}(E_i) = R^q_p(E_i) \quad$ and $\quad (b)\ R^{q'}_p(F_j) = R^r_q(F_j)$;

and (*iii*) the Bayes factor identities,

(3.4) $\quad$ (a) $\quad B^{r'}_{q'}(E_i : E_1) = B^q_p(E_i : E_1) \quad$ and $\quad (b)\ B^{q'}_p(F_j : F_1) = B^r_q(F_j : F_1)$.[1]

Proof. We show only that $(3.2a) \Rightarrow (3.3a) \Rightarrow (3.4a) \Rightarrow (3.2a)$. The proofs that $(3.2b) \Rightarrow (3.3b) \Rightarrow (3.4b) \Rightarrow (3.2b)$ follow from analogous reasoning.

(i) $(3.2a) \Rightarrow (3.3a)$: from $q'(E_i) = p(E_i)$ and the restricted model identity $r'(E_i) = e'_i = e_i = q(E_i)$, it follows immediately that $R^{r'}_{q'}(E_i) = \dfrac{r'(E_i)}{q'(E_i)} = \dfrac{q(E_i)}{p(E_i)} = R^q_p(E_i)$.

(ii) $(3.3a) \Rightarrow (3.4a)$: immediate, based on formula (1.2).

(iii) $(3.4a) \Rightarrow (3.2a)$: By the restricted model identity $r'(E_i) = q(E_i)$, (3.4a) reduces to the equation $q'(E_1)p(E_i) = p(E_1)q'(E_i)$, which, summed over all i, yields $q'(E_1) = p(E_1)$, and, hence, $q'(E_i) = p(E_i)$. $\square$

Diaconis and Zabell (1982) also note the following:

Theorem 3.2. If the partitions **E** and **F** are p-independent, i.e., $p(E_iF_j) = p(E_i)p(F_j)$, for all i, j, then the Jeffrey independence conditions hold.

Proof. We first prove (3.2a). By (2.3) and p-independence, we have $q'(E_i) = \sum_j f'_j p(E_i \mid F_j) = p(E_i)\sum_j f'_j = p(E_i)$. To prove (3.2b) we apply (1.3) to get

$q(F_j) = \sum_i e_i p(F_j \mid E_i) = p(F_j)\sum_i e_i = p(F_j)$. $\square$

Remark 3.1. Note that our proof that p-independence implies Jeffrey independence actually holds in the general model. However, this result is of no interest outside the confines of the restricted model. For, as will be seen below, far from ensuring that $r' = r$ in the general model, under mild regularity conditions, Jeffrey independence actually implies that $r' \neq r$ if there exists an i such that $e'_i \neq e_i$ or there exists a j such that $f'_j \neq f_j$.

4. Jeffrey independence is otiose outside the restricted model

Theorem 4.1. Suppose that in the general model described in Section 2 above, the following regularity conditions hold:

(4.1) For all i_1 and all i_2, there exists a j such that $p(E_{i_1}F_j)p(E_{i_2}F_j) > 0$, and

(4.2). For all j_1 and all j_2, there exists an i such that $p(E_iF_{j_1})p(E_iF_{j_2}) > 0$.

(In particular, since we are assuming for simplicity here that all probability measures are strictly coherent, (4.1) and (4.2) will hold if the partitions **E** and **F** are *qualitatively independent*, in the sense that, for all *i* and all *j*, $E_iF_j \neq \varnothing$). If there exists an *i* such that $r'(E_i) \neq q(E_i)$ or there exists a j such that $q'(F_j) \neq r(F_j)$, and Jeffrey independence holds, then $r' \neq r$.

Proof. We prove the contrapositive, namely, that if $r' = r$, and Jeffrey independence holds, then (1) $r'(E_i) = q(E_i)$ for all i and (2) $q'(F_j) = r(F_j)$ for all *j*. We'll just prove (1), since the proof of (2) is nearly identical. Since $r' = r$, and the regularity conditions (4.1) and (4.2) hold, the Bayes factor identities $B_{q'}^{r'}(E_i : E_1) = B_p^q(E_i : E_1)$ hold by Wagner (2002, Theorem 4.1), i.e.,

$$\frac{q(E_i)p(E_1)}{q(E_1)p(E_i)} = \frac{r'(E_i)q'(E_1)}{r'(E_1)q'(E_i)}. \tag{4.3}$$

By Jeffrey independence, $q'(E_i) = p(E_i)$, and so (4.3) simplifies to

$$r'(E_1)q(E_i) = q(E_1)r'(E_i). \tag{4.4}$$

Summing each side of (4.4) over all *i* yields $r'(E_1) = q(E_1)$, which, with (4.4), implies that $r'(E_i) = q(E_i)$.

Notes

1. It is easy to show that (3.4a) is equivalent to $B_{q'}^{r'}(E_{i_1} : E_{i_2}) = B_p^q(E_{i_1} : E_{i_2})$, for all i_1, i_2, and, similarly, that (3.4b) is equivalent to $B_p^{q'}(F_{j_1} : F_{j_2}) = B_q^r(F_{j_1} : F_{j_2})$, for all j_1, j_2. Use, inter alia, the fact that $B_p^q(E_{i_1} : E_{i_2}) = B_p^q(E_{i_1} : E_1) / B_p^q(E_{i_2} : E_1)$.

References

Persi Diaconis and Sandy Zabell (1982), Updating subjective probability, *Journal of the American Statistical Association* **77**. 822-830.

Richard Jeffrey (1965), *The Logic of Decision*, McGraw-Hill (2nd ed, 1983, University of Chicago Press).

Richard Jeffrey (1988), Conditioning, kinematics, and exchangeability, in Brian Skyrms and William Harper (eds.), *Causation, Chance, and Credence*, vol. 1, Kluwer, 221-255 (reprinted in Richard Jeffrey (1992), *Probability and the Art of Judgment*, Cambridge Studies in Probability, Induction, and Decision Theory, Cambridge University Press.

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