Independence (Probabilistic) and Independence (Logical)

David Miller
Department of Philosophy
University of Warwick
COVENTRY CV4 7AL UK
http://www.warwick.ac.uk/go/dwmiller

Abstract

This paper is concerned with the problem of defining the relation of *probabilistic independence* in a way that reveals how, if at all, it is a generalization of one of the various relations of *logical independence* that are familiar to logicians. It is well known that in this regard not all is well within the classical theory of probability, in which probability is intrinsically a singulary function, or within the modern variant of that theory that was given currency in the axiom system of Kolmogorov (1933). More incisiveness might have been hoped for in those subtler axiomatizations in which a relative (binary) probability function is adopted as primitive, especially the axiomatic theory analysed in appendix *iv, and presented systematically in appendix *v, of *The Logic of Scientific Discovery* (Popper 1959). But it turns out that the classical definition is even less satisfactory here than it is in its classical setting. Near the end of his life Popper became aware that something was amiss, and in the final appendix to the 10th edition of *Logik der Forschung* (Popper 1994) he proposed an alternative definition of probabilistic independence. This appendix, on which Dorn (2002), §2.4, reports, has, predictably enough, been neglected, anyway in English-language publications (by me as much as by anyone). A recent paper (Fitelson & Hájek 2017) devoted to the same topic does not notice it. It is the primary purpose of the present paper to show to what extent Popper's definition is an advance on the definitions commended by Fitelson & Hájek. But I shall also address once more the question of what logical significance the relation of probabilistic independence en joys, and shall broach the idea that there is a genuinely attractive alternative.

Probabilistic independence In the classical theory, where absolute probability $\mathbf{p}(a)$ is primitive, the *probabilistic independence* of elements a, b in the domain of the $\mathbf{p}$ is defined by

DU $$\cup(a, b) \;\leftrightarrow\; \mathbf{p}(ab) = \mathbf{p}(a)\mathbf{p}(b).$$

This is a symmetric relation. An immediate and untoward consequence of this definition is that every element b is independent of every element a for which $\mathbf{p}(a) = 0$; in particular every b is probabilistically independent of the contradictory (or zero) element s, even though b, being deducible from s, is logically dependent on s. It is rather less immediate that every element a for which $\mathbf{p}(a) = 1$ is independent of every element b. But if $\mathbf{p}(a) = 1$, then $\mathbf{p}(a \vee b) \leq \mathbf{p}(a)$, whence by the addition and monotony laws $\mathbf{p}(b) \leq \mathbf{p}(ab) \leq \mathbf{p}(b)$, and therefore $\mathbf{p}(ab) = \mathbf{p}(a)\mathbf{p}(b)$. This implies that the tautological (or unit) element t is probabilistically independent of every element b, even though t, being deducible from b, is logically dependent on b. These anomalies in the classical theory are perfectly well known, and some effort has been made to surmount them. There is, in particular, a stronger asymmetric sense of the *probabilistic independence* of a from b that uses the relative (often called *conditional*) probability $\mathbf{p}(a, b)$, which, provided that $\mathbf{p}(b) \neq 0$, is defined by $\mathbf{p}(a, b) = \mathbf{p}(ab)/\mathbf{p}(b)$. This stronger definiens is

DV $$\vee(a, b) \;\leftrightarrow\; \mathbf{p}(a, b) = \mathbf{p}(a).$$

Since the first term on the right here is not defined if $\mathbf{p}(b) = 0$, the conclusion that every element a is probabilistically (but not logically) independent of s is thwarted. It remains true, nonetheless, that the element s is independent of any element b that is not identical with s itself, and that the element t and any element b for which $\mathbf{p}(b) \neq 0$ are mutually independent.

In appendix *XX to *Logik der Forschung* Popper acknowledged these shortcomings of the classical definitions: s and t ought not to be probabilistically independent of other elements. But he thought that it must be possible for some other contingent statements with zero or unit probability — his examples of the latter were 'There exists a white raven' and 'There exists a golden mountain' — to be counted as probabilistically independent of each other. His solution was to introduce two new relations of *weak independence* $\mathtt{W}(a, b)$ and *independence* $\mathtt{I}(a, b)$,

DW $\qquad \mathtt{W}(a, b) \leftrightarrow \mathbf{p}(a, b) = \mathbf{p}(a, b')$

DI $\qquad \mathtt{I}(a, b) \leftrightarrow \mathtt{W}(a, b) \,\&\, \mathtt{W}(a', b) \,\&\, \mathtt{W}(b, a) \,\&\, \mathtt{W}(b', a)$

(Popper called these definitions D_0 and D_1 respectively). The main theorem (*Haupttheorem*) of appendix *XX demonstrates that neither t nor s bears the (symmetric) relation $\mathtt{I}$ to any element. Fitelson & Hájek also dismiss DU because it requires that 'anything with extreme probability has the peculiar property of being probabilistically independent of *itself*'; this is a state of affairs that, they judge, may perhaps be acceptable for contingent events with unit probability (§6), but it is intolerable for contingent events with zero probability (§8). They propose, as successors to $\mathtt{U}(a, b)$, two other definientia of the probabilistic independence of a from b: the weaker one is just $\mathtt{V}(a, b)$ released from the restriction that $\mathbf{p}(b) \neq 0$, while the stronger one is just $\mathtt{W}(a, b)$ itself. Popper had shown that, among other things, $\mathtt{W}(a, b)$ implies $\mathtt{V}(a, b)$, and hence that $\mathbf{p}(a, b') = \mathbf{p}(a)$; and, near the end of the appendix, that $\mathtt{I}(a, b)$ implies $\mathtt{U}(a, b)$, so that independence, newly defined by DI, 'implies classical independence'. It is easily checked that none of $\mathtt{U}(a, b)$, $\mathtt{V}(a, b)$, $\mathtt{W}(a, b)$, $\mathtt{I}(a, b)$ is equivalent to any of the others, and hence that they provide stronger and stronger definitions of independence. It is evident too that $\mathtt{I}$, like $\mathtt{U}$, is a symmetric relation, whereas $\mathtt{V}$ and $\mathtt{W}$ are asymmetric. Fitelson & Hájek appear to regard this as a discovery, rejoicing that 'on a Popperian account of independence', as they audaciously label their approach, 'we must specify a *direction* of independence' (§8).

Popper did not give an explicit construction to show that the relation $\mathtt{I}(a, b)$ can obtain between two contingent elements a, b with probability 1, such as the existential statements lately mentioned, but it is easily done. There are some other nice results that he did not assert (let alone prove). One, which he would surely have been pleased about is that, according to DI, no element a is independent of itself, or of its negation a'. He would perhaps have been less pleased to learn that probabilistic independence $\mathtt{I}(a, b)$ may obtain even when the element b logically implies the element a; that is, when they are, in the usual sense, *logically dependent*.

Logical Independence If we are to decide between the relations $\mathtt{V}(a, b)$ and $\mathtt{W}(a, b)$ commended by Fitelson & Hájek, which may be too weak, and the relation $\mathtt{I}(a, b)$ commended by Popper, which may be too strong, we shall have be clearer about the job that probabilistic independence is being asked to do. At one point Fitelson & Hájek note that 'We may well want inductive logic, understood as probability theory, to be continuous with deductive logic' (§6). Popper too, has in several places (for example 1957, point 3) been inclined to 'identify logical independence with probabilistic independence', and the vague idea that logically independent elements are probabilistically independent, or approximately so, lurks unacknowledged behind many judgements of probabilistic independence. But there are several different kinds of logical independence too, and some of them turn out to be more appropriate than others.

Simple (logical) *independence* is simply non-deducibility. It is what is asserted when it is said that the axiom of parallels is independent of the other Euclidean axioms, and that the axiom of choice AC is independent of ZF set theory. More generally, a set K of statements is *simply independent* if and only if no a in K is deducible from the other elements. There are two well known extensions of this idea. The set K is *completely independent* (Moore 1910) if and only if for every subset A of K, all the elements of A can be true while all the other elements are false. It is immediate that a completely independent set K is both consistent and simply independent. It is to be noted that Popper did not attempt to show that the four conjuncts in the definition DI of the relation $\mathtt{I}(a, b)$ are consistent, which of course they are, or simply independent, which they are, or completely independent, which they are not: it is impossible that exactly one of the four conjuncts $\mathsf{W}(a, b)$, $\mathsf{W}(a', b)$, $\mathsf{W}(b, a)$, and $\mathsf{W}(b', a)$ is true. Although this fact deftly intimates that the relation $\mathsf{W}(a, b)$ is needlessly weak, the virtues of complete independence are really far from obvious. It is known that, in classical logic, there are infinite sets that are not equivalent to any completely independent set (Kent 1975). What is decisive, however, is that, unless all probabilities equal 0 or 1, there exist completely independent sets whose elements in pairs are not probabilistically independent, even according to DU (Popper & Miller 1987, note 2). A brave attempt to solve this problem has been made by Mura (2006).

A set K is *maximally independent* (Sheffer 1926) if and only if it is simply independent and no (non-tautological) consequence of any element a in K is derivable from the other elements. The elements of a maximally independent K have no content in common, and are what was traditionally called *subcontraries*. Tarski (1930, Theorem 17) showed that, in classical logic, every set K is equivalent to a maximally independent set. Maximally independent elements are not in general probabilistically independent (Popper & Miller *ibidem*, Theorem 1). Nonetheless, it is attention to maximal logical independence that looks like the way forward.

References

Dorn, G. J. W. (2002). 'Induktion und Wahrscheinlichkeit — ein Gedankenaustausch mit Karl Popper'. In E. Morscher, editor, pp.13-79. *Was wir Karl R. Popper und sie Philosophie verdanken*. *Zu seinem 100. Geburtstag*. Sankt Augustin: Academia Verlag.

Fitelson, B. & Hájek, A. (2017). 'Declarations of Independence'. *Synthese* **194**, 10, pp.3979—3995.

Kent, C. F. (1975). 'Independence versus Logical Independence in the Countable Case'. In H. E. Rose and J. C. Shepherdson, editors, pp.399—408. *Logic Colloquium '73*. Amsterdam: North-Holland Publishing Company

Kolmogorov, A. N. (1933). *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Berlin: Springer. English translation 1950. *Foundations of the Theory of Probability*. New York: Chelsea Publishing Company.

Moore, E. H. (1910). 'Introduction to a Form of General Analysis'. In E. H. Moore, E. J. Wilczynski, & M. Mason, editors, pp.1—150. *The New Haven Mathematical Colloquium* (New Haven CT, 5—8 September 1906). New Haven CT: Yale University Press.

Mura, A. M. (2006). (2006): 'Deductive Probability, Physical Probability, and Partial Entailment'. In M. Alai & G. Tarozzi, editors, pp. 181—202. *Karl Popper Philosopher of Science*. Soveria Mannelli: Rubbettino.

Popper, K. R. (1935). *Logik der Forschung*. Vienna: Julius Springer Verlag. 10th edition (1994). Tübingen: J. C. B. Mohr (Paul Siebeck). 11th edition (2005), edited by H. Keuth. Tübingen: Mohr Siebeck.

—————— (1957). 'A Second Note on Degree of Confirmation'. *The British Journal for the Philosophy of Science* **7,** 28, pp.350—353. CORRECTIONS: *ibidem* **8,** 32, pp.294f. Reprinted in appendix *ix of Popper (1959).

—————— (1959). *The Logic of Scientific Discovery*. Expanded English translation of Popper (1935). London: Hutchinson Educational.

Popper, K. R. & Miller, D. W. (1987). 'Why Probabilistic Support Is Not Inductive'. *Philosophical Transactions of the Royal Society of London*, Series A **321**,1562, 30.iv.1987, pp.569—91.

Sheffer, H. M. (1926). 'The General Theory of Notational Relativity'. In E. S. Brightman, editor, pp.348—351. *Proceedings of the Sixth International Congress of Philosophy*. Cambridge MA: Harvard University Press. Original mimeographed version: Department of Philosophy, Harvard University, 1921.

Tarski, A. (1930). 'Über einige fundamentale Begriffe der Metamathematik'. *Comptes Rendus des séances de la Societé des Sciences et des Lettres de Varsovie* **23**, cl. iii, pp. 22—29. English translation: Chapter III of A. Tarski, (1956). *Logic, Semantics, Metamathematics. Papers from 1923 to 1938*. Oxford: Clarendon Press.