doi: 10.5007/1808-1711.2011v15n2p271

Probabilistic Substitutivity at a Reduced Price

David Miller
University of Warwick

Abstract. One of the many intriguing features of the axiomatic systems of probability investigated in Popper (1959), appendices ∗iv, ∗v, is the different status of the two arguments of the probability functor with regard to the laws of replacement and commutation. The laws for the first argument, (rep1) and (comm1), follow from much simpler axioms, whilst (rep2) and (comm2) are independent of them, and have to be incorporated only when most of the important deductions have been accomplished. It is plain that, in the presence of (comm1), the principle (sub), which says that terms that are intersubstitutable in the first argument are intersubstitutable also in the second argument, implies (comm2), and in Popper's systems the converse implication obtains. It is naturally asked what is needed in an axiomatic theory of probability in order to enforce this equivalence. Leblanc (1981) offered a rather weak set of axioms, containing (comm1) and (comm2), that suffice for the derivation of (sub). In this paper Leblanc's result is improved in a number of different ways. Three weaker systems, one of which is incomparable with the other two, are shown to admit the same implication.

Keywords: Probability axiomatics; probabilistic indistinguishability; commutative law, substitution principle; Popper; Leblanc.

Introduction

The axiomatic system $\mathbb{B}$ of probability of Popper (1959), appendices ∗iv and ∗v, presumes a language containing the usual supply of variables and elementary logical and arithmetical apparatus (connectives, quantifiers, $=, 0, 1, +, \cdot$) along with three special functors. Two of them, represented by concatenation and the overbar $^{-}$ (shown eventually to stand for meet and complementation operations respectively), produce new terms of the same category as the variables. If X and Z are such terms, then the expression $\mathfrak{p}(X, Z)$ is a numerical term. All other terms are defined terms. The concern of the present paper is with some fragments of the theory $\mathbb{B}$ that omit entirely the operator $^{-}$. In line with standard practice in the theory of probability, we shall throughout write $\mathfrak{p}(X \mid Z)$ rather than $\mathfrak{p}(X, Z)$.

Sentences that, given some suitably modest theory of the real numbers, are interderivable will be identified and called *statements*. For the purposes of most of this paper, an (axiomatic) *system* (of probability) is a set of statements of the language based on concatenation and $\mathfrak{p}$. A *theory* (of probability) is a set of statements closed under logical derivability (relative to the chosen theory of real numbers). The relation of axiomatic system to (axiomatic) theory is accordingly many-one, since in

Published by NEL — Epistemology and Logic Research Group, Federal University of Santa Catarina (UFSC), Brazil.

general one theory may be axiomatized in many ways. The expressions *subsystem* and *subtheory* will be used analogously.

It is not necessary to state here the axioms of the system $\mathbb{B}$, since the full system $\mathbb{B}$ is not our business. But it may be noted that the crucial product or multiplication law

$$\text{(prod)} \qquad \mathfrak{p}(xz\,|\,y) = \mathfrak{p}(x\,|\,zy)\mathfrak{p}(z\,|\,y)$$

of $\mathbb{B}$ belongs to all the subtheories of $\mathbb{B}$ that we discuss. Nor need we list yet any of the main theorems of $\mathbb{B}$, except the laws of idempotence, commutation, and association, in the first argument of $\mathfrak{p}$:

$$\text{(id1)} \qquad \mathfrak{p}(xx\,|\,z) = \mathfrak{p}(x\,|\,z),$$
$$\text{(comm1)} \qquad \mathfrak{p}(xz\,|\,y) = \mathfrak{p}(zx\,|\,y),$$
$$\text{(assoc1)} \qquad \mathfrak{p}(x(yz)\,|\,w) = \mathfrak{p}((xy)z\,|\,w).$$

These results indicate that when the relation $\sim$ of (probabilistic) *indistinguishability* (in the first argument of $\mathfrak{p}$) is defined by

$$\text{(df}\sim\text{)} \qquad x \sim z =_{\text{Df}} \forall y[\mathfrak{p}(x\,|\,y) = \mathfrak{p}(z\,|\,y)],$$

the terms xx and x, xz and zx, and $x(yz)$ and $(xy)z$, are pairwise indistinguishable. If we identify elements denoted by indistinguishable terms, the domain of interpretation of $\mathbb{B}$ is thus shown to have at least the structure of a lower semilattice. For further details of this development, along with related ones, the reader is directed to Popper & Miller (1994).

Before the identification of indistinguishable terms can be authorized, it must be established that the relation of indistinguishability, which is manifestly an equivalence relation, is also a congruence. That is to say, if $x \sim z$ and Z differs from X at most by containing occurrences of z where X contains occurrences of x, then both

$$\text{(rep1)} \qquad x \sim z \implies \mathfrak{p}(X\,|\,w) = \mathfrak{p}(Z\,|\,w),$$
$$\text{(rep2)} \qquad x \sim z \implies \mathfrak{p}(w\,|\,X) = \mathfrak{p}(w\,|\,Z),$$

hold. The law (rep1) is easily proved by induction along the length of Z, since both

$$x \sim z \Rightarrow \mathfrak{p}(xy \mid w) = \mathfrak{p}(zy \mid w)$$

(whose proof requires only a couple of simple applications of (product)) and

$$x \sim z \Rightarrow \mathfrak{p}(yx \mid w) = \mathfrak{p}(yz \mid w)$$

(whose proof is then obtained by two applications of (comm1)) are straightforwardly proved, and similar results (which are not relevant to the present investigation) hold also for the the operator $^{-}$. But it is not possible to prove within $\mathbb{B}$ the corresponding replacement law (rep2) for the second argument, that is, that any term occurring in the second argument of $\mathfrak{p}$ can be replaced by an indistinguishable term. In particular it is not possible to prove in $\mathbb{B}$ the substitution principle

$$\text{(sub)} \qquad \forall y[\mathfrak{p}(x \mid y) = \mathfrak{p}(z \mid y)] \Longrightarrow \forall y[\mathfrak{p}(y \mid x) = \mathfrak{p}(y \mid z)]$$

or the law of commutation in the second argument

$$\text{(comm2)} \qquad \mathfrak{p}(y \mid xz) = \mathfrak{p}(y \mid zx).$$

This is shown by a counterexample (Example 0) given by Popper *op.cit.*, pp. 339f. (Popper 2002, p. 345), whose relevant features will be sketched in §1 below. It is plain that, given (rep1), the substitution principle (sub) ensures (rep2). It is plain too that the conditional (sub) ⇒ (comm2) is a theorem of any theory, such as $\mathbb{B}$, that contains (comm1) as a theorem. Popper showed that the converse conditional (comm2) ⇒ (sub) is a theorem of $\mathbb{B}$ too, though he seems never to have published a full proof. The equivalence of (comm2) and (sub) was implicitly asserted in Popper *op.cit.*, p. 335 (Popper 2002, p. 340), and explicitly in Popper (1994), p. 278. In this connection, it should be mentioned that appendices ∗iv and ∗v of the later (German) editions of *Logik der Forschung*, up to Popper (1994), are in some important respects different from the corresponding appendices in Popper (1959).

Neither (sub) nor (comm2) involves the operator $^{-}$ of $\mathbb{B}$, and it seems fairly obvious that the equivalence of these two statements should be demonstrable in theories much weaker than $\mathbb{B}$. In (1981) Leblanc offered a straightforward proof of (sub) from (comm2) within a transparent axiomatic system $\mathbb{L}$, to be discussed in the next section. He stressed the weakness of his four axioms, one of which is (comm1), and conjectured (p. 320, note 3) that no proper subset of $\mathbb{L}\cup$\{comm2\} suffices for the proof of (sub). We shall confirm, but not verify, this conjecture. The main purpose of the present paper is then to improve Leblanc's result, by giving three alternative axiomatic systems, $\mathbb{L}-, \mathbb{M}=$, and $\mathbb{N}$, all weaker than $\mathbb{L}$, in which the derivation of (sub) from (comm2) may be carried out.

It is worth recording that each of the systems to be presented is a subtheory of the system $\mathbb{B}$, and therefore holds in any non-trivial Boolean algebra if $\mathfrak{p}(x,z)$ is assigned the value 1 when $z \preceq x$ and the value 0 otherwise. That is to say, none of the systems, by implying that $\mathfrak{p}$ is single-valued, makes both the antecedent and the consequent of (sub) a vacuous truth.

1. Leblanc's Proof

The set $\mathbb{L}$ of axioms, displayed in Table 0, constitutes a proper subtheory of the system $\mathbb{B}$. It is the system proposed by Leblanc (1981) for the derivation of (sub) from (comm2).

(bounds)	$0 \leq \mathfrak{p}(x \mid z) \leq 1$
(unity)	$\mathfrak{p}(y \mid y) = 1$
(comm1)	$\mathfrak{p}(xz \mid y) = \mathfrak{p}(zx \mid y)$
(prod)	$\mathfrak{p}(xz \mid y) = \mathfrak{p}(x \mid zy)\mathfrak{p}(z \mid y)$

Table 0: *The system* $\mathbb{L}$

A slightly simplified version of Leblanc's derivation within $\mathbb{L}$ of (sub) from (comm2) proceeds like this. We first use (bounds), (unity), (product), and again (bounds), to show that

$$\mathfrak{p}(x \mid yx) \leq 1 = \mathfrak{p}(yx \mid yx) = \mathfrak{p}(y \mid x(yx))\mathfrak{p}(x \mid yx) \leq \mathfrak{p}(x \mid yx). \tag{1.0}$$

Hence $\mathfrak{p}(x \mid yx) = 1$. The identity (1.0) just proved allows us to infer from the assumption that $x \sim z$ that

$$\mathfrak{p}(y \mid x) = \mathfrak{p}(x \mid yx)\mathfrak{p}(y \mid x) = \mathfrak{p}(z \mid yx)\mathfrak{p}(y \mid x) = \mathfrak{p}(zy \mid x) = \mathfrak{p}(yz \mid x) \tag{1.1}$$

by (product) and (comm1). Another use of (product), of the assumption $x \sim z$, and of (unity), then yields

$$\mathfrak{p}(y \mid x) = \mathfrak{p}(y \mid zx)\mathfrak{p}(z \mid x) = \mathfrak{p}(y \mid zx)\mathfrak{p}(x \mid x) = \mathfrak{p}(y \mid zx). \tag{1.2}$$

By exchanging x and z we obtain $\mathfrak{p}(y \mid z) = \mathfrak{p}(y \mid xz)$; so by (comm2)

$$\mathfrak{p}(y \mid x) = \mathfrak{p}(y \mid z), \tag{1.3}$$

which tells us that if the elements x and z are probabilistically indistinguishable, then they are intersubstitutable in the second argument of $\mathfrak{p}$. It follows that (comm2) ⇒ (sub) is a theorem of $\mathbb{L}$.

It is worth noting that this proof proves much more than (comm2) ⇒ (sub), which is a conditional of the form $\forall \mathrm{u}\Phi\mathrm{u} \Rightarrow \forall \mathrm{u}\Psi\mathrm{u}$, where u is a string of variables. It proves also the stronger conditional $\forall \mathrm{u}(\Phi\mathrm{u} \Rightarrow \Psi\mathrm{u})$. In other words, commutation in the second argument of $\mathfrak{p}$ pointwise implies in the system $\mathbb{L}$ the intersubstitutability in the second argument of $\mathfrak{p}$ of elements that are indistinguishable according to (df$\sim$). For each x, z such that $x \sim z$, the identity of $\mathfrak{p}(y \mid xz)$ and $\mathfrak{p}(y \mid zx)$ is equivalent, for each y, to the identity of $\mathfrak{p}(y \mid x)$ and $\mathfrak{p}(y \mid z)$, whether or not (comm2) and (sub) hold universally.

Before proceeding further, we shall examine briefly Leblanc's conjecture that no proper subset of $\mathbb{L} \cup \{\text{comm2}\}$ suffices for the proof of (sub). We begin with (comm2). A simple counterexample, already mentioned, shows that (comm2) is not a theorem of $\mathbb{L}$, and that if it fails then (sub) may fail too. The domain of interpretation is any three-element set $\{\mathrm{a}, \mathrm{b}, \mathrm{c}\}$. Concatenation in Example 0 is interpreted by the operation $\bullet$, and $\mathfrak{p}$ by the numerical function μ, specified in the displayed matrices.

$$\begin{array}{c|ccc} \bullet & \mathrm{a} & \mathrm{b} & \mathrm{c} \\ \hline \mathrm{a} & \mathrm{a} & \mathrm{b} & \mathrm{c} \\ \mathrm{b} & \mathrm{b} & \mathrm{b} & \mathrm{b} \\ \mathrm{c} & \mathrm{a} & \mathrm{b} & \mathrm{c} \end{array} \qquad \begin{array}{c|ccc} \mu & \mathrm{a} & \mathrm{b} & \mathrm{c} \\ \hline \mathrm{a} & 1 & 1 & 1 \\ \mathrm{b} & 0 & 1 & 1 \\ \mathrm{c} & 1 & 1 & 1 \end{array}$$

Example 0

As usual, the rows provide the first arguments, and the columns provide the second arguments, for the operations $\bullet$ and μ; for example $\mathrm{c} \bullet \mathrm{a} = \mathrm{a} \neq \mathrm{c} = \mathrm{a} \bullet \mathrm{c}$ and $\mu(\mathrm{b}, \mathrm{a}) = 0 \neq 1 = \mu(\mathrm{b}, \mathrm{c})$. Of the axioms of $\mathbb{L}$, (bounds) and (unity) are immediately checked. Since $\bullet$ is a commutative operation except for the pair $\{\mathrm{a}, \mathrm{c}\}$, the axiom (comm1) can fail only if $\mu(\mathrm{c} \bullet \mathrm{a}, y) \neq \mu(\mathrm{a} \bullet \mathrm{c}, y)$; but actually $\mu(\mathrm{a}, y) = \mu(\mathrm{c}, y)$ for all y. Similar, though more involved, considerations show that (product) holds. In short, the axioms of $\mathbb{L}$ are all satisfied. But (comm2) fails, since $\mu(\mathrm{b}, \mathrm{c} \bullet \mathrm{a}) \neq \mu(\mathrm{b}, \mathrm{a} \bullet \mathrm{c})$; and since a and c are probabilistically indistinguishable, and yet $\mu(\mathrm{b}, \mathrm{a}) \neq \mu(\mathrm{b}, \mathrm{c})$, the substitution principle (sub) fails too.

To show that (unity) and (product) are individually necessary to the derivation of (sub) from (comm2), we resort to the functions μ and ν defined on the four-element Boolean algebra $\{0, \mathrm{b}, \mathrm{b}', 1\}$ in Examples 1 and 2.

$$\begin{array}{c|cccc} \mu & 0 & \mathrm{b} & \mathrm{b}' & 1 \\ \hline 0 & 0 & 0 & 0 & 0 \\ \mathrm{b} & 0 & 0 & 0 & 0 \\ \mathrm{b}' & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{array} \qquad \begin{array}{c|cccc} \nu & 0 & \mathrm{b} & \mathrm{b}' & 1 \\ \hline 0 & 1 & 1 & \beta & 0 \\ \mathrm{b} & 1 & 1 & 1 & 0 \\ \mathrm{b}' & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{array}$$

Example 1 Example 2

In Example 2 the value of β may be anywhere in the interval $[0,1)$. Since $\{0,\mathrm{b},\mathrm{b}',1\}$ is a Boolean algebra (in which the law of commutation holds for meet), both (comm1) and (comm2) hold for both μ and ν. It is plain that (bounds) holds for each function, and that (unity) holds for ν but not for μ. It is an unusually simple task to show that μ satisfies (product). But ν does not satisfy (product), since

$$\nu(0\bullet \mathrm{b},\mathrm{b}') = \nu(0,\mathrm{b}') = \beta \neq 1$$

whilst

$$\nu(0,\mathrm{b}\bullet \mathrm{b}')\nu(\mathrm{b},\mathrm{b}') = \nu(0,0)\nu(\mathrm{b},\mathrm{b}') = 1.$$

In each matrix there are two identical rows (0 and b for μ; b and b′ for ν) whose corresponding columns are not identical. That is to say, in each case (sub) is not satisfied.

$$\begin{array}{c|ccc} \bullet & \mathrm{a} & \mathrm{b} & \mathrm{c} \\ \hline \mathrm{a} & \mathrm{a} & \mathrm{a} & \mathrm{c} \\ \mathrm{b} & \mathrm{a} & \mathrm{b} & \mathrm{c} \\ \mathrm{c} & \mathrm{a} & \mathrm{c} & \mathrm{c} \end{array} \qquad \begin{array}{c|ccc} \mu & \mathrm{a} & \mathrm{b} & \mathrm{c} \\ \hline \mathrm{a} & 1 & 0 & 1 \\ \mathrm{b} & 1 & 1 & 1 \\ \mathrm{c} & 1 & 1 & 1 \end{array}$$

Example 3

To show that (comm1) cannot be dispensed with in the derivation of (sub) from (comm2), we need a model that is not a Boolean algebra. Example 3 (from Popper 1959, p. 340; Popper 2002, p. 346; Popper 1994, p. 289) does the trick. The proofs of (bounds) and (unity) are immediate. The proof of (product) is tedious. The operation $\bullet$ in Example 3 is not commutative, since $\mathrm{a}\bullet\mathrm{c} \neq \mathrm{c}\bullet\mathrm{a}$, and it follows that (comm1) fails:

$$\mu(\mathrm{a}\bullet \mathrm{c},\mathrm{b}) = \mu(\mathrm{c},\mathrm{b}) = 1 \neq 0 = \mu(\mathrm{a},\mathrm{b}) = \mu(\mathrm{c}\bullet \mathrm{a},\mathrm{b}).$$

There is, however, no infringement of (comm2). That could happen only if $\mu(y,\mathrm{a}\bullet\mathrm{c}) \neq \mu(y,\mathrm{c}\bullet\mathrm{a})$, which is impossible by the definition of μ. Yet (sub) fails, since rows b and c are the same in the matrix for μ, yet columns b and c are different. We may conclude that (bounds), (unity), and (product) hold, that (comm1) fails, and that the conditional (comm2) $\Rightarrow$ (sub) also fails.

In other words, none of the systems obtained from $\mathbb{L}$ by discarding one of (unity), (product), or (comm1), is strong enough to permit the derivation of (comm2) $\Rightarrow$ (sub). This does not imply, of course, that it is not possible to replace one of these axioms by some weaker statement. All it implies is that it is not possible to replace one of these axioms by a logical truth, or by no statement at all. This point is nicely illustrated by the axiom (bounds).

I do not know whether (bounds) is needed within $\mathbb{L}$ for deriving the conditional (comm2) ⇒ (sub), in the sense that the derivation fails in $\mathbb{L} \setminus \{\text{bounds}\}$. (It can be shown that there is no two-element or three-element model in which (unity), (comm1), (comm2), and (product) hold but (sub) fails.) We may note, however, that in the proof outlined at the beginning of this section (bounds) was used only at the start, to establish (as a consequence of the compound formula (1.0)) that $\mathfrak{p}(x \mid yx) = 1$. Yet all that is needed in the proof of (1.1) is the weaker identity $\mathfrak{p}(y \mid x) = \mathfrak{p}(x \mid yx)\mathfrak{p}(y \mid x)$, which may be derived if we use, instead of (bounds), the intuitively much weaker statement

$$\text{(id2)} \qquad \mathfrak{p}(x \mid zz) = \mathfrak{p}(x \mid z).$$

For we need only apply (unity), (id2), (product), (comm1), and (product) again to obtain

$$\begin{aligned}\text{(1.4)} \qquad \mathfrak{p}(y \mid x) &= \mathfrak{p}(y \mid x)\mathfrak{p}(x \mid x) = \\ &= \mathfrak{p}(y \mid xx)\mathfrak{p}(x \mid x) = \mathfrak{p}(yx \mid x) = \mathfrak{p}(xy \mid x) = \mathfrak{p}(x \mid yx)\mathfrak{p}(y \mid x).\end{aligned}$$

The theory $\mathbb{L}- = \mathbb{L} \setminus \{\text{bounds}\} \cup \{\text{id2}\}$, whose axioms are stated explicitly in Table 1, therefore suffices for the derivation of (sub) from (comm2).

(unity)	$\mathfrak{p}(y \mid y) = 1$
(id2)	$\mathfrak{p}(x \mid zz) = \mathfrak{p}(x \mid z).$
(comm1)	$\mathfrak{p}(xz \mid y) = \mathfrak{p}(zx \mid y)$
(product)	$\mathfrak{p}(xz \mid y) = \mathfrak{p}(x \mid zy)\mathfrak{p}(z \mid y)$

Table 1: *The system* $\mathbb{L}-$

We conclude this section by showing that $\mathbb{L}-$ is a theory properly weaker than $\mathbb{L}$. A little work is needed to show that (id2) is a theorem of $\mathbb{L}$, but fortunately most of it has been done in the proofs of Theorems 0 and 1 of Popper & Miller *op.cit.* We first use (unity) and (product) to show that

$$\text{(1.5)} \qquad \mathfrak{p}(x \mid zz) = \mathfrak{p}(x \mid zz)\mathfrak{p}(z \mid z) = \mathfrak{p}(xz \mid z);$$

and then (comm1), (product), and (bounds), to show that

$$\text{(1.6)} \qquad \mathfrak{p}(xz \mid z) = \mathfrak{p}(zx \mid z) = \mathfrak{p}(z \mid xz)\mathfrak{p}(x \mid z) \leq \mathfrak{p}(x \mid z);$$

and finally (unity), (product), (comm1), (product), and (bounds), to show that

$$(1.7)\qquad \begin{aligned}\mathfrak{p}(x \mid z) &= \mathfrak{p}(xz \mid xz)\mathfrak{p}(x \mid z) = \mathfrak{p}((xz)x \mid z)\\ &= \mathfrak{p}(x(xz) \mid z) = \mathfrak{p}(x \mid (xz)z)\mathfrak{p}(xz \mid z) \leq \mathfrak{p}(xz \mid z).\end{aligned}$$

Together (1.5), (1.6), and (1.7) imply (id2).

Example 4 is a two-element Boolean algebra $\{0,1\}$ that suffices to show that (bounds) is not a theorem of $\mathbb{L}-$.

μ	0	1
0	1	β
1	1	1

Example 4

Here β is any number outside the unit interval $[0,1]$. The truth of (product) is easily checked. A four-element model, also a Boolean algebra, in which the two halves of (bounds) are simultaneously infringed is to be found in Popper 1959, p. 343 (Popper 2002, p. 349; Popper 1994, p. 291).

In § 2 we turn to an alternative way of weakening $\mathbb{L}$ without disturbing the derivability of (sub) from (comm2). In § 3 it will be shown that a system even weaker than $\mathbb{L}-$ is sufficient.

2. A System Equivalent to $\mathbb{L}$, and a Weaker One

The following system $\mathbb{M}$ was isolated from $\mathbb{B}$ by Popper & Miller *op.cit.*, § 1, in an attempt to write down a minimal set of recognisable axioms sufficient for the derivation of the probabilistic lattice laws (id1), (comm1), and (assoc1).[1]

(nontrivial)	$\exists x \exists z \mathfrak{p}(x \mid z) \neq 0$
(downbound)	$0 \leq \mathfrak{p}(x \mid z)$
(upbound)	$\mathfrak{p}(x \mid z) \leq \mathfrak{p}(y \mid y)$
(monotony)	$\mathfrak{p}(xz \mid y) \leq \mathfrak{p}(x \mid y)$
(product)	$\mathfrak{p}(xz \mid y) = \mathfrak{p}(x \mid zy)\mathfrak{p}(z \mid y)$

Table 2: *The system* $\mathbb{M}$

The principal theorems of interest in the system $\mathbb{M}$ are (unity), (bounds), (id1), (id2), (comm1), and (assoc1), which are proved as (1.1), (1.2), (1.6), (1.6), (1.11), and (1.12), in Popper & Miller *op.cit.* It follows without further ado that $\mathbb{L}$ is a subtheory of $\mathbb{M}$. Each axiom of $\mathbb{M}$, moreover, is a theorem of $\mathbb{L}$. No work is needed to establish this fact for (nontrivial), (downbound), (upbound), and (product); and for (monotony), it suffices to note that

$$\text{(2.0)} \qquad \mathfrak{p}(xz \mid y) = \mathfrak{p}(zx \mid y) = \mathfrak{p}(z \mid xy)\mathfrak{p}(x \mid y) \leq \mathfrak{p}(x \mid y)$$

by successive application of (comm1), (product), and (bounds). This means that $\mathbb{L}$ and $\mathbb{M}$, though different axiomatic systems, axiomatize the same theory.

There is a fruitful way of weakening $\mathbb{M}$, however, that is inapplicable to $\mathbb{L}$. Let us write $\mathbb{M}- = \mathbb{M} \setminus \{\text{nontrivial}\}$. Since the axioms of $\mathbb{M}-$ allow all probabilities to be 0, (unity) is not derivable in $\mathbb{M}-$, and hence $\mathbb{L}$ (and $\mathbb{M}$) are properly stronger than $\mathbb{M}-$. Yet the theories $\mathbb{L}$ and $\mathbb{M}-$ are of equal strength as far as the derivation of any formula of the form $\Phi \Rightarrow$ (sub) is concerned. For suppose that $\mathbb{M}- \nvdash \Phi \Rightarrow$ (sub). This means that in some model of $\mathbb{M}-$ the formula $\Phi \Rightarrow$ (sub) fails; hence (sub) fails, and so its consequent $\forall y[\mathfrak{p}(y \mid x) = \mathfrak{p}(y \mid z)]$ fails. But the falsity in a model of such a (universalized) equation ensures that (nontrivial) is true. It is therefore a model of $\mathbb{M}$ also, which implies that $\mathbb{M} \nvdash \Phi \Rightarrow$ (sub). It follows that if $\Phi \Rightarrow$ (sub) is derivable in either of $\mathbb{L}$ and $\mathbb{M}-$, then it is derivable in both. In general the presence of (nontrivial) in a theory is irrelevant to the derivability of a formula of the form $\Phi \Rightarrow$ (sub).

We may go further. By (upbound) and (monotony),

$$\text{(2.1)} \qquad \mathfrak{p}(y \mid yy) \leq \mathfrak{p}(yy \mid yy) \leq \mathfrak{p}(y \mid yy)$$

$$\text{(2.2)} \qquad \mathfrak{p}(y \mid y(yy)) \leq \mathfrak{p}(y(yy) \mid y(yy)) \leq \mathfrak{p}(y \mid y(yy)).$$

It is evident that, according to (upbound), the value of $\mathfrak{p}(y \mid y)$ is independent of y, whence $\mathfrak{p}(yy \mid yy) = \mathfrak{p}(y \mid y) = \mathfrak{p}(y(yy) \mid y(yy))$, from which we may infer that $\mathfrak{p}(y \mid yy) = \mathfrak{p}(y \mid y) = \mathfrak{p}(y \mid y(yy))$. By (product),

$$\text{(2.3)} \qquad \mathfrak{p}(y \mid y) = \mathfrak{p}(yy \mid yy) = \mathfrak{p}(y \mid y(yy))\mathfrak{p}(y \mid yy) = \mathfrak{p}(y \mid y)^2$$

so that the value of $\mathfrak{p}(y \mid y)$ equals either 0 or 1. Suppose that $\mathfrak{p}(y \mid y) = 0$ for all y. If (product) and (monotony) both hold then $0 = \mathfrak{p}(x \mid yy)\mathfrak{p}(y \mid y) = \mathfrak{p}(xy \mid y) \leq \mathfrak{p}(x \mid y)$, which means that (downbound) holds. But if (upbound) and (nontrivial) hold (downbound) must fail. In short, $\mathfrak{p}(y \mid y) = 1$ follows from (nontrivial), (upbound), (monotony), and (product) alone. The axiom (downbound) is not required. (This is Theorem 0 of Popper & Miller 1994.)

In the rest of this section it will be shown that it is possible entirely to eliminate (downbound), as well as (nontrivial), from the proof in $\mathbb{M}$ of (comm2) $\Rightarrow$ (sub). We shall therefore limit our attention to the system $\mathbb{M}=$, whose axioms are listed in Table 3.

(upbound)	$\mathfrak{p}(x \mid z) \leq \mathfrak{p}(y \mid y)$
(monotony)	$\mathfrak{p}(xz \mid y) \leq \mathfrak{p}(x \mid y)$
(product)	$\mathfrak{p}(xz \mid y) = \mathfrak{p}(x \mid zy)\mathfrak{p}(z \mid y)$,

Table 3: *The system* $\mathbb{M}=$

This system $\mathbb{M}=$ is definitely weaker than $\mathbb{L}$. Plainly (unity) is not a theorem of $\mathbb{M}=$ (which allows all probabilities to be 0). If in the matrix for the function μ in Example 4 the value of β is negative, then we have a simple proof that the first conjunct of (bounds), that is, (downbound), is not a theorem of $\mathbb{M}=$. But Theorem 0 of Popper & Miller *op.cit.*, which shows that $\mathfrak{p}(y \mid y)$ equals either 0 or 1, shows in consequence that the second conjunct of (bounds) does hold in $\mathbb{M}=$. But it is not a theorem of $\mathbb{L}-$, as may be shown by setting $\beta > 1$ in Example 4. In brief, $\mathbb{M}=$ is not a subtheory of $\mathbb{L}-$.

Nor is $\mathbb{L}-$ a subtheory of $\mathbb{M}=$, since in the latter system, in contrast to all the other systems mentioned in this paper, (comm1) is not a theorem. It holds, of course, when (unity) fails, since in such a model (nontrivial) also fails. But Example 5 provides a primitive model that shows that in general (comm1) is not true. The validity of (upbound) in Example 5 is immediate, and that of (monotony) too, since $x \bullet z = x$ for all x, z . The validity of the remaining axiom (product) may be checked without fuss. But (comm1) fails, since $\mu(\mathrm{a} \bullet \mathrm{c}, \mathrm{c}) = -1 \neq 1 = \mu(\mathrm{c} \bullet \mathrm{a}, \mathrm{c})$.

•	a	c
a	a	a
c	c	c

μ	a	c
a	1	−1
c	−1	1

Example 5

Interestingly enough, however, (comm1) is a consequence within $\mathbb{M}=$ of (comm2). If (nontrivial) fails then there is nothing to prove. If not, then, as already noted, (unity) holds. By (comm2), $\mathfrak{p}(z \mid x(zy)) = \mathfrak{p}(z \mid (zy)x)$ and by (unity), (upbound), and two applications of (monotony),

$$\text{(2.4)} \qquad \mathfrak{p}(z \mid (zy)x) \leq 1 = \mathfrak{p}((zy)x \mid (zy)x) \leq \mathfrak{p}(zy \mid (zy)x) \leq \mathfrak{p}(z \mid (zy)x),$$

and hence $\mathfrak{p}(z \mid x(zy)) = 1$. Then by (product),

$$\text{(2.5)} \qquad \mathfrak{p}(xz \mid y) = \mathfrak{p}(x \mid zy)\mathfrak{p}(z \mid y) = \mathfrak{p}(z \mid x(zy))\mathfrak{p}(x \mid zy)\mathfrak{p}(z \mid y),$$

and so by two more applications of (product), and one of (monotony),

$$\mathfrak{p}(xz \mid y) = \mathfrak{p}(zx \mid zy)\mathfrak{p}(z \mid y) = \mathfrak{p}((zx)z \mid y) \leq \mathfrak{p}(zx \mid y). \tag{2.6}$$

By exchanging x and z, we may conclude that $\mathfrak{p}(zx \mid y) \leq \mathfrak{p}(xz \mid y)$, and the derivation of (comm1) from (comm2) is complete.

It may also be shown that (id2) is a theorem of $\mathbb{M}=$. Again (nontrivial), and therefore (unity), may be assumed. By (unity), (product), and two applications of (monotony),

$$\mathfrak{p}(x \mid z) = \mathfrak{p}(xz \mid xz)\mathfrak{p}(x \mid z) = \mathfrak{p}((xz)x \mid z) \leq \mathfrak{p}(xz \mid z) \leq \mathfrak{p}(x \mid z), \tag{2.7}$$

so that $\mathfrak{p}(x \mid z) = \mathfrak{p}(xz \mid z)$, and hence by (product) and (unity) again,

$$\mathfrak{p}(x \mid z) = \mathfrak{p}(x \mid zz)\mathfrak{p}(z \mid z) = \mathfrak{p}(x \mid zz). \tag{2.8}$$

That (comm2) ⇒ (comm1) and (id2) are derivable in $\mathbb{M}=$ allows us to give a revealing proof that (comm2) ⇒ (sub) is derivable too.

We assume (comm2), which entitles us also to (comm1). Once again we may assume (nontrivial), and therefore (unity), since (sub) is immediate when (nontrivial) fails. We assume further the antecedent of (sub), which is to say, $\mathfrak{p}(x \mid y) = \mathfrak{p}(z \mid y)$ for all y. What must be shown is the consequent of (sub), which is to say, $\mathfrak{p}(y \mid x) = \mathfrak{p}(y \mid z)$ for every y. By (id2), (product), and (comm1), then (product), the assumption, and (product), then (comm1), (product), and the assumption,

$$\begin{aligned} \mathfrak{p}(y \mid x)\mathfrak{p}(x \mid x) &= \mathfrak{p}(y \mid xx)\mathfrak{p}(x \mid x) = \mathfrak{p}(yx \mid x) = \mathfrak{p}(xy \mid x) \\ &= \mathfrak{p}(x \mid yx)\mathfrak{p}(y \mid x) = \mathfrak{p}(z \mid yx)\mathfrak{p}(y \mid x) = \mathfrak{p}(zy \mid x) \\ &= \mathfrak{p}(yz \mid x) = \mathfrak{p}(y \mid zx)\mathfrak{p}(z \mid x) = \mathfrak{p}(y \mid zx)\mathfrak{p}(x \mid x). \end{aligned} \tag{2.9}$$

Since (unity) holds, we may divide by $\mathfrak{p}(x \mid x)$, yielding $\mathfrak{p}(y \mid x) = \mathfrak{p}(y \mid zx)$. We now exchange x and z, and use (comm2) to conclude that $\mathfrak{p}(y \mid x) = \mathfrak{p}(y \mid z)$. In this way (sub) may be derived from (comm2) within $\mathbb{M}=$.

As for the converse, Example 5 demonstrates also that (comm2) cannot be derived from (sub). For since there are no indistinguishable elements in the model, (sub) holds trivially, and yet $\mu(\mathfrak{c}, \mathfrak{a} \bullet \mathfrak{c}) = -1 \neq 1 = \mu(\mathfrak{c}, \mathfrak{c} \bullet \mathfrak{a})$. It may well be thought, therefore, that the system $\mathbb{M}=$ is of only limited interest.

3. A System Weaker Still

To authorize division by $\mathfrak{p}(x \mid x)$ in the derivation just completed we cited (unity), according to which $\mathfrak{p}(y \mid y) = 1$ for all y; but since (2.9) is a string of identities, it is plainly enough if $\mathfrak{p}(y \mid y)$ is never equal to zero. The above proof of (comm2) ⇒ (sub), which uses otherwise only (id2), (comm1), and (product), therefore demonstrates also that the conclusion is derivable in the axiomatic system $\mathbb{N}$, whose axioms are given in Table 4.

(nonzero)	$\mathfrak{p}(y \mid y) \neq 0$
(id2)	$\mathfrak{p}(x \mid zz) = \mathfrak{p}(x \mid z)$
(comm1)	$\mathfrak{p}(xz \mid y) = \mathfrak{p}(zx \mid y)$
(product)	$\mathfrak{p}(xz \mid y) = \mathfrak{p}(x \mid zy)\mathfrak{p}(z \mid y)$

Table 4: *The system* $\mathbb{N}$

The intuitive weakness of these axioms for probability is remarkable. No doubt the axiom (product) is substantial, but (nonzero), though indispensable, has rather little to say, for $\mathbb{N}$ does not assume, and does not imply, that $\mathfrak{p}(y \mid y)$ is the same for every y.

This is harder to show than might have been expected; indeed, there is no finite model of $\mathbb{N}$ in which $\mathfrak{p}(y \mid y)$ varies, or even one that, through the failure of (unity), is not a model of $\mathbb{L}-$. For in such a model there would have to be at least one element y that $0 \neq \mathfrak{p}(y \mid y) \neq 1$. That is, an element b for which $\mu(\mathrm{b},\mathrm{b}) \neq \mu(\mathrm{b},\mathrm{b})^2$. Yet we have by (product) and (id2)

$$\mathfrak{p}(yy \mid y) = \mathfrak{p}(y \mid yy)\mathfrak{p}(y \mid y) = \mathfrak{p}(y \mid y)^2, \tag{3.0}$$

so that the element $\mathrm{b} \bullet \mathrm{b}$ would be probabilistically distinct from b. Indeed, by (id2) again, $\mathfrak{p}(yy \mid yy) = \mathfrak{p}(y \mid y)^2$, which means that $\mu(\mathrm{b},\mathrm{b})$ and $\mu(\mathrm{b} \bullet \mathrm{b}, \mathrm{b} \bullet \mathrm{b})$ would be different, and different also from 0 and 1. But then $y = (\mathrm{b} \bullet \mathrm{b}) \bullet (\mathrm{b} \bullet \mathrm{b})$ would provide another value for $\mathfrak{p}(y \mid y)$, and so on.

This gives us a clue to the construction of Example 6, which shows that $\mathbb{N}$ is genuinely weaker than $\mathbb{L}-$ is. The model contains infinitely many elements $\mathrm{b}_1, \mathrm{b}_2, \ldots$, with the degenerate conjunction rule $\mathrm{b}_i \bullet \mathrm{b}_k = \mathrm{b}_k \bullet \mathrm{b}_i = \mathrm{b}_{i+k}$, and an almost degenerate allocation of values to μ that makes $\mu(x,z)$ independent of z. If β is equal to neither 0 nor 1 then (unity) fails, but the truth of all the axioms of $\mathbb{N}$ is immediate.

We may easily establish in addition that both $\mathbb{L}-$ and $\mathbb{N}$ are logically incomparable with $\mathbb{M}{=}$. Since $\mathbb{M}{=}$ permits all probabilities to equal 0, it cannot imply (unity) or (nonzero). On the other hand, if in Example 4 the value of β is greater than 1, then all the axioms of $\mathbb{L}-$ and $\mathbb{N}$ hold, whilst (upbound) and (monotony) both fail.

$\bullet$	b_1	b_2	b_3	$\dots$
b_1	b_2	b_3	b_4	$\dots$
b_2	b_3	b_4	b_5	$\dots$
b_3	b_4	b_5	b_6	$\dots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

μ	b_1	b_2	b_3	$\dots$
b_1	β	β	β	$\dots$
b_2	β^2	β^2	β^2	$\dots$
b_3	β^3	β^3	β^3	$\dots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

Example 6

Since $\mathbb{N}$ has (comm1) as one of its axioms, it is not a subtheory of $\mathbb{M}=$. Since $\mathbb{M}=$ is not a subtheory of $\mathbb{L}-$, it is not a subtheory of $\mathbb{N}$ either.

It remains to show that each of the four axioms of $\mathbb{N}$ is necessary to the proof of (sub) from (comm2). The indispensability of each of (nonzero), (comm1), and (product) is established by those models (Examples 1, 2, 3) used in § 1 to prove the indispensability of (unity), (comm1), and (product) in the system $\mathbb{L}$. Example 0 shows likewise that (sub) is not derivable in $\mathbb{N}$ without the intervention of (comm2). As for (id2), the commutative but not idempotent model of Example 7 suffices. Here β is any number distinct from 0 and from 1.

$\bullet$	a	b	c
a	c	b	a
b	b	b	b
c	a	b	c

μ	a	b	c
a	β	1	$1/\beta$
b	β	1	$1/\beta$
c	1	1	1

Example 7

Note that (id2) fails because $\mu(a, a \bullet a) = 1/\beta \neq \beta = \mu(a, a)$. The element b may be omitted from Example 7 if all that is wanted is a proof of the independence of (id2) from the other axioms. Its only purpose here is to head a row that is probabilistically identical to that headed by a, and a column that is different from that headed by a, thus violating (sub). To check the validity of (product), it is easiest to show first that it holds when the domain is restricted to $\{a, c\}$; and then to use the unimaginative properties assigned to the dummy element b (that is, $b \bullet y = y \bullet b = b$; $\mu(y, b) = 1$; $\mu(b, y) = \mu(a, y)$), to show that the insertion of b into the domain cannot make any difference.

To sum up the main results. The four systems $\mathbb{L}$ (which is equivalent to $\mathbb{M}$), $\mathbb{L}-$, $\mathbb{M}=$, and $\mathbb{N}$, whose axioms are given Table 5, are all sufficient for the derivation of (sub) from (comm2). Each of $\mathbb{L}-$ and $\mathbb{M}=$ is a proper non-trivial subtheory of $\mathbb{L}$; and $\mathbb{N}$ is a proper non-trivial subtheory of $\mathbb{L}-$. Neither $\mathbb{L}-$ nor $\mathbb{N}$ is logically comparable with $\mathbb{M}=$.

$\mathbb{L}$	$\mathbb{L}-$	$\mathbb{M}=$	$\mathbb{N}$
(bounds)			
		(upbound)	
(unity)	(unity)		
			(nonzero)
		(monotony)	
	(id2)		(id2)
(comm1)	(comm1)		(comm1)
(product)	(product)	(product)	(product)

Table 5: The four systems compared

If, as in Leblanc's original treatment, we adjoin (comm2) to each theory (and therefore inquire about the derivability of (sub)) then, as the same models show, all these relations of proper inclusion are preserved. Example 6 shows, for instance, that $\mathbb{N} \cup \{(\text{comm2})\}$ does not imply $\mathbb{L}- \cup \{(\text{comm2})\}$.[2]

References

da Costa, N. C. A. 1993. *Lógica Indutiva e Probabilidade*. 2nd edition. São Paulo: Editora Hucitec & Editora da Universidade de São Paulo.

Leblanc, H. 1981. What Price Substitutivity? A Note on Probability Theory. *Philosophy of Science* **48**: 317–22.

Popper, K. R. 1934. *Logik der Forschung*. Vienna: Julius Springer Verlag.

———. 1959. *The Logic of Scientific Discovery*. London: Hutchinson. English translation of Popper 1934.

———. 1994. *Logik der Forschung*. Tübingen: Mohr Siebeck. 10th edition of Popper 1934.

———. 2002. *The Logic of Scientific Discovery*. Classics edition of Popper 1959. London & New York: Routledge.

Popper, K. R. & Miller, D. W. 1994. Contributions to the Formal Theory of Probability. In P. W. Humphreys (ed.) *Patrick Suppes: Scientific Philosopher*, Volume I. Dordrecht: Kluwer Academic Publishers, pp. 3–23. (The copyright in this paper belongs to the authors and their heirs and assigns, and not, as stated on the first page, to Kluwer Academic Publishers.)

DAVID MILLER
Department of Philosophy
University of Warwick
COVENTRY CV4 7AL UK
dwmiller57@yahoo.com

Resumo. Um dos muitos aspectos intrigantes dos sistemas axiomáticos da probabilidade que foram investigados em Popper (1959), apêndices ∗iv, ∗v, são as situações diferentes dos dois argumentos do functor de probabilidade quanto às leis de substituição e de comutação. As leis pelo primeiro argumento, (rep1) e (comm1), seguem a partir de axiomas muitos simples, enquanto (rep2) e (comm2) são independentes deles, e devem incorporar-se só quando a maior parte das deduções importantes tenham sido executadas. É evidente que, na presença de (comm1), o princípio (sub), que diz que termos que podem substituir-se no primeiro argumento podem substituir-se também no segundo argumento, implica (comm2), e nos sistemas de Popper a implicação conversa está em vigor. É natural perguntar do que precisa uma teoria axiomática da probabilidade para aplicar esta equivalência. Leblanc (1981) ofereceu um conjunto bastante fraco de axiomas, contendo (comm1) e (comm2), que são suficiente para a derivação de (sub). O artigo presente aperfeiçoa o resultado de Leblanc em vários modos diferentes. Demonstra-se que três sistemas mais fracos, dos quales um é incomparável com os outros dois, permitem a mesma implicação.

Palavras-chave: Probabilidade; indistinguibilidade probabilística; lei de comutação; princípio de substituição; Popper; Leblanc.

Notes

[1] For the purpose of deriving (id1), (comm1), and (assoc1), the system $\mathbb{M}$ is unnecessarily strong. For reasons given in the text, the system $\mathbb{M}- = \mathbb{M}\backslash\{\text{nontrivial}\}$ suffices.

$\mathbb{M}$ is a subtheory, but not an axiomatic subsystem, of $\mathbb{B}$. The axioms (monotony) and (product) are axioms also of $\mathbb{B}$, whilst (nontrivial) is equivalent (within $\mathbb{B}$, but not within $\mathbb{M}$) to $\mathbb{B}$'s axiom $\exists y \exists w \mathfrak{p}(x \mid z) \neq \mathfrak{p}(y \mid w)$. The axioms (upbound) and (downbound) are theorems of $\mathbb{B}$ (Popper *op.cit.*, appendix $*$v, lines (1) and (18)).

Although $\mathbb{M}$ omits the complementation operator $^{-}$ of $\mathbb{B}$, it fails in several respects to capture that part of $\mathbb{B}$ that does not involve $^{-}$. In the first place, $\mathbb{B}$ contains the axiom (sub) and implies (comm2) and other equivalents. Secondly, $\mathbb{M}$ deliberately omits the existential theorem

(zero) $$\exists x \exists z \mathfrak{p}(x \mid z) = 0$$

(which is a consequence of line (27) in the development of $\mathbb{B}$ in Popper *op.cit.*, appendix $*$v), and is equivalent, within $\mathbb{M}$, to $\mathbb{B}$'s axiom $\exists y \exists w \mathfrak{p}(x \mid z) \neq \mathfrak{p}(y \mid w)$ cited above. But there are further theorems of $\mathbb{B}$ that do not involve the operator $^{-}$ and are not theorems of the system $\mathbb{M}{+}0 = \mathbb{M} \cup \{(\text{sub})\} \cup \{(\text{zero})\}$, one example being the formula

(weakadd) $$\mathfrak{p}(x \mid y) + \mathfrak{p}(z \mid y) \leq \mathfrak{p}(xz \mid y) + 1$$

Other examples are the formulas (1.15), (1.16), and (2.9) of Popper & Miller *op.cit.*, which are interderivable with each other (though not with (weakadd)) in $\mathbb{M}{+}0$. It is an open question whether the adjunction of any or all of these formulas to the system $\mathbb{M}{+}0$ captures all the complementation-free formulas of $\mathbb{B}$; that is, whether $\mathbb{B}$ is a conservative extension of any of these systems. For my part, I doubt it.

[2] This paper was commissioned in the (northern) summer of 1999 for the special issues of *Synthese* planned to celebrate Newton da Costa's 70th birthday on 17th September 1999.

Unfortunately it was not possible to reach an agreement with the publishers, Kluwer Academic Publishers, concerning ownership of rights in the paper, and it was withdrawn. It has since been made more readable, and a few small independence results have been added, but it is really the same paper as the paper of twelve years ago. I offer it to da Costa anew as a belated 80th birthday present. I should like it now to serve also as a small memorial to Hugues Leblanc, who died on 10th September 1999, a few days after the original version was completed.

The axiomatics of probability theory has not been one of Newton's main pursuits, but §6.3 of his little book (1993) does concern itself with one version of Popper's axioms. I hope that the present work may enhance his appreciation of the logical subtleties of these elegant systems, which are still little known.

An abbreviated version of what appears here was presented as a contributed paper to the ASL European Summer Meeting LOGIC COLLOQUIUM '97 at the University of Leeds in July 1997. Most of the principal results of §§1-2 were obtained in 1988 and 1989, or even earlier, but were then organized very differently.