Modus Tollens Probabilized

CARL G. WAGNER

University of Tennessee, U. S. A.

Abstract

We establish a probabilized version of modus tollens, deriving from $p(E|H) = a$ and $p(\bar{E}) = b$ the best possible bounds on $p(\bar{H})$. In particular, we show that $p(\bar{H}) \to 1$ as $a, b \to 1$, and also as $a, b \to 0$.

Keywords

conditional probability, modus ponens, modus tollens, probability logic

1   Introduction

Given an exact or approximate specification of the probabilities of propositions $A_1, \ldots, A_n$, what are the narrowest bounds that can be placed on the probability of some other proposition B? This problem was first investigated by George Boole [B54], whose work was revived by Theodore Hailperin [H65, H84, H96] in a sequence of publications that offered a fully rigorous, comprehensive treatment of the problem from the standpoint of linear programming. We follow Hailperin in calling this problem *Boole's General Probability Problem*.

While the propositions in question may be arbitrary, special interest naturally attaches to cases in which the inference from the conjunction of the "premisses" A_i to the "conclusion" B is truth-functionally valid, symbolized in what follows by $A_1, \ldots, A_n \vDash B$. For solutions to Boole's General Probability Problem in such cases may be expected to reveal the degree of confidence in the premisses A_i sufficient to ensure some desired degree of confidence in B, as measured by their subjective probabilities. Elementary examples include the valid argument

$$A_1, \ldots, A_n \vDash A_1 \vee \cdots \vee A_n,$$

which has the probabilistic counterpart

$$p(A_i) = a_i \ \Rightarrow \ \max\{a_i\} \leqslant p(A_1 \vee \cdots \vee A_n) \leqslant \min\{1, a_1 + \cdots + a_n\}, \quad (1)$$

and the valid argument

$$A_1, \ldots, A_n \vDash A_1 \wedge \cdots \wedge A_n,$$

1

which has the probabilistic counterpart

$$p(A_i) = a_i \; \Rightarrow \; a_1 + \cdots + a_n - (n-1) \leqslant p(A_1 \wedge \cdots \wedge A_n) \leqslant \min\{a_i\}. \quad (2)$$

The right-hand inequality in (1) and the left-hand inequality in (2) are, respectively, the familiar inequalities of Boole and Bonferroni. That the bounds in (1) and (2) are the best possible was apparently first shown by Maurice Fréchet [F35].

A number of valid arguments of propositional logic involve the truth-functional connective $\longrightarrow$, variously termed *the conditional, the indicative conditional, and material implication*, the truth table of $A \rightarrow B$ being identical with that of $\bar{A} \vee B$. In the next two sections we consider probabilistic counterparts of *modus ponens* ($E \rightarrow H$, $E \vDash H$) and *modus tollens* ($H \rightarrow E$, $\bar{E} \vDash \bar{H}$). In section 2, a premiss such as $E \rightarrow H$ is simply replaced by its probability $p(E \rightarrow H)$. In section 3, it is replaced by the arguably more salient conditional probability $p(H|E)$. Our results on the conditional probability counterpart of modus tollens are new, and exhibit a feature that is unusual in these sorts of problems, namely, that a high probability of $\bar{H}$ is entailed not only when $p(E|H)$ and $p(\bar{E})$ are high, but also when they are low.

2   Probabilities of Conditionals

The valid argument $E \rightarrow H$, $E \vDash H$, known as *modus ponens* or the *rule of detachment*, has the probabilistic counterpart

$$p(E \rightarrow H) = a \text{ and } p(E) = b \; \Rightarrow \; a + b - 1 \leqslant p(H) \leqslant a, \quad (3)$$

for all a, b such that $a + b \geqslant 1$, the bounds in (3) being the best possible. These results follow immediately from the observations that

$$p(E \rightarrow H) = p(\bar{E} \vee H) \leqslant p(\bar{E}) + p(H)$$

and

$$p(H) = p(E \wedge H) + p(\bar{E} \wedge H) \leqslant p(E \wedge H) + p(\bar{E}) = P(E \rightarrow H),$$

along with easily constructed examples of probabilities p satisfying the antecedent conditions of (3) and, respectively, (i) $p(\bar{E} \wedge H) = 0$ and (ii) $p(\bar{E} \wedge \bar{H}) = 0$. Hailperin [H84] has also given an interesting proof using linear programming. Note that the condition $a + b \geqslant 1$ is necessary and sufficient to ensure that the set of probabilities satisfying the antecedent conditions of (3) is nonempty.

It follows from (3) that

$$p(E \rightarrow H) \geqslant 1 - \varepsilon \text{ and } p(E) \geqslant 1 - \varepsilon \; \Rightarrow \; p(H) \geqslant 1 - 2\varepsilon \quad (4)$$

for all $\varepsilon \leqslant 1/2$, as noted by Patrick Suppes [S66].

From the fact that $p(H \to E) = p(\bar{E} \to \bar{H})$ it follows trivially from the above that the valid argument $H \to E, \bar{E} \vDash \bar{H}$, known as *modus tollens*, has the probabilistic counterpart

$$p(H \to E) = a \text{ and } p(\bar{E}) = b \ \Rightarrow \ a+b-1 \leqslant p(\bar{H}) \leqslant a, \qquad (5)$$

for all a, b such that $a + b \geqslant 1$. In particular,

$$p(H \to E) \geqslant 1 - \varepsilon \text{ and } p(\bar{E}) \geqslant 1 - \varepsilon \ \Rightarrow \ p(\bar{H}) \geqslant 1 - 2\varepsilon \qquad (6)$$

for all $\varepsilon \leqslant 1/2$. Understandably, not much is usually made of (5) and (6), since they are trivial variants of (3) and (4). We include them here only for purposes of comparison with results in the next section.

3 Conditional Probabilities

3.1 Adam's Thesis

Truthful speakers do not assert propositions that they do not take to be highly probable. As David Lewis [L76] puts it, "assertibility goes by subjective probability." But Ernest Adams [A65, A75] has pointed out that conditionals seem to be an exception to this rule, with the assertibility of $E \to H$ going not by the probability $p(E \to H)$ of the conditional but, rather, by the conditional probability $p(H|E)$.

Adam's evidence consists of a number of cogent examples that accord with his thesis. Lewis [L76] offers an additional argument: When $p(E \to H)$ is high because $p(\bar{E})$ is high and, moreover, $p(\bar{H}|E)$ is high, then it is pointless, and in fact misleading, to assert $E \to H$, rather than $\bar{E}$. Taking the product $p(\bar{E})p(\bar{H}|E)$ as a simple measure of the lessening of assertibility of $E \to H$, we then arrive at $p(E \to H) - p(\bar{E})p(\bar{H}|E)$, easily shown to be equal to $p(H|E)$, as a suitable measure of the assertibility of $E \to H$.

Frank Jackson [J79] offers a further argument in support of Adam's thesis: Suppose that we have assessed the prior probability $p(E \to H)$. Further inquiry convinces us of the truth of E, and in such a way that the revision of p to the posterior $q(\cdot) = p(\cdot|E)$ is judged to be reasonable. Then

$$q(E \to H) = p(E \to H|E) = p(H|E) \text{ and } q(E) = 1$$

and so $q(H) \geqslant p(H|E)$ by (3). A large value of $p(E \to H)$ is of no utility in such dynamic cases of modus ponens. What matters is the value of $P(H|E)$.

Even apart from the above discussion, it is clear that conditional probabilities are salient in a way that probabilities of conditionals are not. For the probability that H is true, given that E is true, is conceptually simpler and easier to assess than the probability that E is false or that E and H are both true. Moreover, when

the probabilities in question are observed or estimated relative frequencies, probabilities of conditionals are of virtually no interest.

For the above reasons it is of interest to investigate probabilistic versions of valid arguments such as modus ponens and modus tollens in which conditional probabilities play a central role, a subject to which we turn in what follows.

3.2 Modus Ponens for Conditional Probabilities

Replacing $p(E \to H)$ by $p(H|E)$ yields the following variant of (3): If $0 \leqslant a \leqslant 1$ and $0 < b \leqslant 1$, then

$$p(H|E) = a \text{ and } p(E) = b \ \Rightarrow \ ab \leqslant p(H) \leqslant ab + 1 - b, \tag{7}$$

the bounds in (7) being the best possible. This result follows immediately from the fact that

$$p(E)p(H|E) = p(E \wedge H) \leqslant p(H) = p(E \wedge H) + p(\bar{E} \wedge H) \leqslant p(E \wedge H) + p(\bar{E}),$$

along with easily constructed examples of probabilities p satisfying the antecedent conditions of (7) and, respectively, (i) $p(\bar{E} \wedge H) = 0$ and (ii) $p(\bar{E} \wedge \bar{H}) = 0$ Hailperin [H96] has also given an interesting proof using linear programming.

It follows from (7) that if $0 \leqslant \varepsilon < 1$, then

$$p(H|E) \geqslant 1 - \varepsilon \text{ and } p(E) \geqslant 1 - \varepsilon \ \Rightarrow \ p(H) \geqslant (1 - \varepsilon)^2, \tag{8}$$

as noted by Suppes [S66].

3.3 Modus Tollens for Conditional Probabilities

Unlike the case of modus tollens for probabilities of conditionals, we cannot make short work of the present subject by invoking the results of section 3.2. For it is not generally true that $p(E|H) = p(\bar{H}|\bar{E})$. Indeed, from $p(E|H) = a$, with $0 \leqslant a < 1$, one can conclude only that $0 \leqslant p(\bar{H}|\bar{E}) \leqslant 1$, as shown by Hailperin [H96]. The following theorem, which seems not to have appeared in the literature, delineates what might reasonably be termed *modus tollens for conditional probabilities*.

Theorem 1 *Let* $p(E|H) = a$ *and* $p(\bar{E}) = b$.
  If (i) $0 < a, b < 1$, *then*

$$\max\{(1 - a - b)/(1 - a), (a + b - 1)/a\} \leqslant p(\bar{H}) < 1. \tag{9}$$

  If (ii) $a = 0$ *and* $0 < b \leqslant 1$, *then*

$$1 - b \leqslant p(\bar{H}) < 1, \text{ and} \tag{10}$$

  If (iii) $a = 1$ *and* $0 \leqslant b < 1$, *then*

$$b \leqslant p(\bar{H}) < 1, \tag{11}$$

and these bounds are the best possible.

Proof. Let $\mathcal{A}$ be the algebra of propositions generated by E and H. Under each of the conditions (i), (ii), (iii) above (and only under these conditions) the set $\mathcal{P}$ of probabilities p on $\mathcal{A}$ satisfying $p(E|H) = a$ and $p(\bar{E}) = b$ is nonempty and consists of precisely those p such that $p(E \wedge H) = at$, $p(E \wedge \bar{H}) = 1 - b - at$, $p(\bar{E} \wedge H) = (1-a)t$, and $p(\bar{E} \wedge \bar{H}) = b - (1-a)t$, where the parameter t is subject to restrictions ensuring that these quantities are all nonnegative.

Under condition (i), we have

$$0 < t \leqslant \min\{b/(1-a), (1-b)/a\}. \tag{12}$$

Since $p(\bar{H}) = 1 - t$, the inequalities in (9) follow from those in (12).

It is clear that the upper bound in (9) is the best possible. To show that this is also true for the lower bound requires consideration of two cases. (1). If $a + b \leqslant 1$ then $b/(1-a) \leqslant (1-b)/a$, and so $t = b/(1-a)$ is an allowable value of the parameter t. For the probability p defined by this value of t, we have $p(\bar{H}) = (1-a-b)/(1-a)$. Since $a + b \leqslant 1$, we have $(1-a-b)/(1-a) \geqslant (a+b-1)/a$, and so $p(\bar{H})$ attains the lower bound in (9). (2). If $a + b > 1$, then $(1-b)/a < b/(1-a)$, and so $t = (1-b)/a$ is an allowable value of t. For the probability p defined by this value of t, we have $p(\bar{H}) = (a+b-1)/a$. Since $a + b > 1$, we have $(a+b-1)/a > (1-a-b)/(1-a)$, and so $p(\bar{H})$ again attains the lower bound in (9).

Under condition (ii), we have

$$0 < t \leqslant b, \tag{13}$$

which, with $p(\bar{H}) = 1 - t$, implies (10), the lower bound in (10) being attained when $t = b$.

Under condition (iii), we have

$$0 < t \leqslant 1 - b, \tag{14}$$

which, with $p(\bar{H}) = 1 - t$, implies (11), the lower bound in (11) being attained when $t = 1 - b$. $\qquad\square$

The above analysis reveals an interesting feature of modus tollens for conditional probabilities, namely, that $p(\bar{H})$ may be made as close to 1 as we wish not only by making $p(E|H)$ and $p(\bar{E})$ sufficiently large, but also by making them sufficiently small. This is of course intuitively reasonable. To take a very concrete example, suppose that we have a finite set of spheres, each either hollow (H) or solid ($\bar{H}$), and each painted either ecru (E) or white ($\bar{E}$). If most of the hollow spheres are ecru, but most of the spheres are white, then most of the spheres must be solid. Similarly, if few of the hollow spheres are ecru and few of the spheres are white, then most of the spheres must be solid.

We conclude with an elaboration of the above in the spirit of Suppes' inequalities (4), (6), and (8).

Corollary 1 *Let* $0 < \varepsilon \leqslant 1/2$. *If*

$$p(E|H) \geqslant 1 - \varepsilon \text{ and } p(\bar{E}) \geqslant 1 - \varepsilon, \text{ or} \tag{15}$$

$$p(E|H) \leqslant \varepsilon \text{ and } p(\bar{E}) \leqslant \varepsilon, \text{ then} \tag{16}$$

$$p(\bar{H}) \geqslant (1 - 2\varepsilon)/(1 - \varepsilon). \tag{17}$$

Proof.   That (15) implies (17) follows from setting $a = b = 1 - \varepsilon$ in (9) and verifying by partial differentiation that $(a + b - 1)/a$ is increasing in both a and b. That (16) implies (17) follows from setting $a = b = \varepsilon$ in (9) and verifying by partial differentiation that $(1 - a - b)/(1 - a)$ is decreasing in a and b.   □

Note that in each of the probabilistic inferences (4), (6), and (8) there is some degradation in probability in passing from premisses to conclusion. The same is true for the inference from (15) to (17), but this case of modus tollens for conditional probabilities always involves less degradation in probabilities than modus ponens or modus tollens for probabilities of conditionals and, if $\varepsilon < (3 - \sqrt{5})/2 \doteq 0.38$, it involves less degradation than modus ponens for conditional probabilities.

References

[A65] E. Adams, The logic of conditionals, *Inquiry* 8: 166–197, 1965.

[A75] E. Adams, *The Logic of Conditionals*, Reidel, Dordrecht, 1975.

[B54] G. Boole, *An Investigation of the Laws of Thought, on Which are Founded the Mathematical Theories of Logic and Probabilities*, Walton and Maberly, London, 1854.

[F35] M. Fréchet, Généralizations du théoreme des probabilités totales, *Fundamenta Mathematicae* 25: 379–387, 1935.

[H65] T. Hailperin, Best possible inequalities for the probability of a logical function of events, *American Mathematical Monthly* 72: 343–359, 1965.

[H84] T. Hailperin, Probability logic, *Notre Dame Journal of Formal Logic* 25: 198–212, 1984.

[H96] T. Hailperin, *Sentential Probability Logic*, Lehigh University Press, Bethlehem, 1996.

[J79] F. Jackson, On assertion and indicative conditionals, *The Philosophical Review* 88: 565–589, 1979.

[L76] D. Lewis, Probabilities of conditionals and conditional probabilities, *The Philosophical Review* 85: 297–315, 1976.

[L86] D. Lewis, Probabilities of conditionals and conditional probabilities II, *The Philosophical Review* 95: 581–589, 1986.

[S66] P. Suppes, Probabilistic inference and the concept of total evidence, in *Aspects of Inductive Logic*, J. Hintikka and P. Suppes, eds, North-Holland, Amsterdam, 49–55, 1966.

Carl G. Wagner is with the Department of Mathematics, University of Tennessee, Knoxville, TN, U.S.A. 37996-1300. Email: `wagner@math.utk.edu`