Article

On Basic Probability Logic Inequalities †

Marija Boričić Joksimović

Faculty of Organizational Sciences, University of Belgrade, Jove Ilića 154, 11000 Belgrade, Serbia;
marija.boricic@fon.bg.ac.rs

† The conclusions given in this paper were partially presented at the European Summer Meetings of the Association for Symbolic Logic, Logic Colloquium 2012, held in Manchester on 12–18 July 2012.

Abstract: We give some simple examples of applying some of the well-known elementary probability theory inequalities and properties in the field of logical argumentation. A probabilistic version of the hypothetical syllogism inference rule is as follows: if propositions A, B, C, $A \rightarrow B$, and $B \rightarrow C$ have probabilities a, b, c, r, and s, respectively, then for probability p of $A \rightarrow C$, we have $f(a,b,c,r,s) \leq p \leq g(a,b,c,r,s)$, for some functions f and g of given parameters. In this paper, after a short overview of known rules related to conjunction and disjunction, we proposed some probabilized forms of the hypothetical syllogism inference rule, with the best possible bounds for the probability of conclusion, covering simultaneously the probabilistic versions of both modus ponens and modus tollens rules, as already considered by Suppes, Hailperin, and Wagner.

1. Introduction

The main part of probabilization of logical inference rules is defining the corresponding best possible bounds for probabilities of propositions. Some of them, connected with conjunction and disjunction, can be obtained immediately from the well-known Boole's and Bonferroni's inequalities, but those related to implication need a deeper specific argumentation. Wagner's paper [1] on probabilistic versions of modus ponens and modus tollens inference rules influenced this work immediately. This paper presents the first step in the process of the probabilization of a complete inference rule system for classical propositional calculus. As stated by Hailperin [2] "one of the examples Boole works out (see [3], Example 7, p. 284) is that of finding the probability of the conclusion of a hypothetical syllogism given the probabilities of its premises. Unfortunately, he confuses the probability of a conditional (probability of: if A, then B) with conditional probability (probability of B, given A) and solves not the stated problem (which nevertheless his method could handle) but one involving conditional probabilities." Using linear programming methods Hailperin, developed an approach to the best possible bounds' computation of the probability of a logical function of events (see [4]); according to him (see also [1]), it was first shown by Fréchet [5] that the well-known Boole's and Bonferroni's inequalities are the best possible.

A modern introduction to the conditional probability's deductive properties was given by Frisch and Haddawy [6] and Wagner [1], connecting their work with the corresponding philosophical background of earlier authors and attempts. An extensive and very valuable general overview of probability logics was presented by Z. Ognjanović, M. Rašković, and Z. Marković [7].

In this paper, we present some probabilistic forms of the hypothetical syllogism rule covering both Hailperin's approach, dealing with the case when probability $P(A)$ of a proposition A belongs to the interval $[a,b] \subseteq [0,1]$ (see [8]), and Suppes' approach, considering probabilities $P(A)$ that are greater than or equal to $1 - \varepsilon$, for a given small

real $\varepsilon > 0$ (see [1,9,10]). The given statement generalizes and contains both results of Hailperin's modus ponens probabilized and Wagner's modus tollens probabilized forms (see [1,10]). Furthermore, we obtained the probabilistic versions of Suppes-style modus ponens and modus tollens rules (see [1,10]) as an immediate consequence of our result. In [11], presenting just a conference abstract, only a formulation of our statement without any argumentation was given, but here, we give a complete proof in detail. Our result will be of importance in the context of statistical, probabilistic, and approximate reasoning.

2. Inference Rules of Pure Propositional Classical Logic

The most famous inference rules in propositional logics are modus ponens (B follows from A and $A \rightarrow B$) and modus tollens ($\neg A$ follows from $\neg B$ and $A \rightarrow B$). The rules treating the behavior of conjunction and disjunction are simplification rules: A follows from $A \wedge B$, and B follows from $A \wedge B$; adjunction rule: $A \wedge B$ follows from A and B; addition rules: $A \vee B$ follows from A, and $A \vee B$ follows from B; and disjunctive syllogism rule: B follows from $A \vee B$ and $\neg A$. The rule containing both, modus ponens and modus tollens, is the rule of hypothetical syllogism: $A \rightarrow C$ follows from $A \rightarrow B$ and $B \rightarrow C$. These are the basic and most popular natural deduction rules defining formally the essence of logical connectives (see [12]).

In attempts to obtain the statistical or probabilistic versions of those pure logical rules, each rule of the form 'C follows from A and B, will be transformed as follows: if A is realized in p of m cases and B is realized in q of n cases, then C is realized in $f(p,q;m,n)$ of $g(p,q;m,n)$ cases or, in more probabilistic manner: if the probabilities of propositions A and B are a and b, respectively, then the probability of C is $c \in [f(a,b),g(a,b)]$, for some functions f and g of given parameters. Instead of the last sentence, we used its shorter form: if $P(A) = a$ and $P(B) = b$, then $f(a,b) \leq P(C) \leq g(a,b)$, where $P(A)$ denotes the probability of a proposition A. Obviously, in both cases, we have to deal with inequalities. We also used Suppes' style expressing proposition probability: if $P(A) \geq 1 - \varepsilon$ and $P(B) \geq 1 - \varepsilon$, then $P(C) \geq 1 - n\varepsilon$, for some $\varepsilon > 0$ and some natural number n, presenting a particularly elegant form of probabilistic inference rules (see [1,9,10]), where, for a fixed value of ε, the probabilities of sentences can be given as a pure function of natural numbers.

3. Inference Rules of Probabilistic Propositional Classical Logic

First, let us revisit the elementary properties of random events:

Lemma 1. *Let A and B be random events. If $P(A) = a$ and $P(B) = b$, then:*
(a) $P(\overline{A}) = 1 - a$;
(b) (C. E. Bonferroni) $\max\{0, a + b - 1\} \leq P(A \cap B) \leq \min\{a, b\}$;
(c) (G. Boole) $\max\{a, b\} \leq P(A \cup B) \leq \min\{1, a + b\}$.

Frechét (see [5]) was the first to show that bounds given in above inequalities are the best possible.

According to Carnap's and Popper's sentence probability axioms (see [13,14]), the function P, defined over propositions, satisfies the following conditions:
(i) if A is a tautology, then $P(A) = 1$;
(ii) if $P(A \wedge B) = 0$, then $P(A \vee B) = P(A) + P(B)$;
(iii) if A is equivalent to B, then $P(A) = P(B)$.
Variations of these conditions were considered by Leblanc and van Fraassen (see [15–17]). In the same context, the probabilistic counterpart of logical rules, bearing in mind the Carnap–Popper approach, the function P, defined over propositions, satisfies the following conditions:

Lemma 2. *Let A and B be any propositions. If $P(A) = a$ and $P(B) = b$, then:*
(a) $P(\neg A) = 1 - a$;
(b) $\max\{0, a + b - 1\} \leq P(A \wedge B) \leq \min\{a, b\}$ (adjunction rule);
(c) $\max\{a, b\} \leq P(A \vee B) \leq \min\{1, a + b\}$ (addition rule);

(d) $\max\{1 - a, b\} \leq P(A \to B) \leq \min\{1, 1 - a + b\}$;
(e) if $P(A \to B) = 1$, then $a \leq b$ *(monotonicity)*.
The bounds in (b), (c), (d), and (e) are the best possible.

Proof. Immediately, by Boole's and Bonferroni's inequalities, for instance: (e) from $1 = P(A \to B) = P(\neg A \vee B) = P(\neg A) + P(B) - P(\neg A \wedge B)$, we have $P(B) = P(A) + P(\neg A \wedge B) \geq P(A)$. In the case when $P(\neg A \wedge B) = 0$, we have $a = b$. $\square$

Lemma 3. *Simplification rule probabilized:*
(a) If $P(A \wedge B) = p$, then $p \leq P(A) \leq \min\{1, P(A) + P(B)\}$;
(b) If $P(A \wedge B) \geq 1 - \varepsilon$, then $P(A) \geq 1 - \varepsilon$, for any $0 \leq \varepsilon \leq 1$.
The bounds in (a) and (b) are the best possible.

Note that the best possible lower bound is reached for $P(A) = P(A \wedge B)$, and the upper bound is reached when $A \leftrightarrow B$. Part (b) follows from (a).

Lemma 4. *Disjunctive syllogism probabilized:*
(a) If $P(A \vee B) = q$ and $P(\neg A) = t$, then $q + t - 1 \leq P(B) \leq q$;
(b) If $P(A \vee B) \geq 1 - \varepsilon$ and $P(\neg A) \geq 1 - \varepsilon$, then $P(B) \geq 1 - 2\varepsilon$, for any $0 \leq \varepsilon \leq \frac{1}{2}$.
The bounds in (a) and (b) are the best possible.

Proof. (a) From Lemma 2, Part (c), we have $P(B) \leq P(A \vee B) = q$. On the other hand, $q = P(A \vee B) \leq P(A) + P(B)$ and $P(A) = 1 - t$ imply $P(B) \geq q - P(A) = q + t - 1$. The best possible upper bound is reached for $P(B) = P(A \vee B)$, and the lower bound is reached when $A \leftrightarrow B$. Part (b) follows from (a). $\square$

4. The Probabilistic Counterparts of the Hypothetical Syllogism Rule

The hypothetical syllogism inference rule, in its pure logical form, covers both modus ponens and modus tollens rules. We generalized known results regarding statistical and probabilistic forms of modus ponens and modus tollens and obtained the best possible bounds for the conclusion when the bounds for hypotheses are given.

Let A be a propositional formula with the corresponding probability $P(A)$ of its truthfulness. The central point of our paper is the following statement, as formulated in [11], dealing with the *hypothetical syllogism* rule.

Theorem 1. *Let* $P(A) = a$, $P(B) = b$, $P(C) = c$, $P(A \to B) = r$ *and* $P(B \to C) = s$. *Then:*
(a) (T. Hailperin [2]) (modus ponens probabilized):

$$a + r - 1 \leq P(B) \leq r;$$

(b) (C. G. Wagner [1]) (modus tollens probabilized):

$$r - b \leq P(\neg A) \leq r;$$

(c) (M. Boričić [11]) (hypothetical syllogism rule probabilized):

$$r + s - 1 \leq P(A \to C) \leq r + s + b - a;$$

(d) (M. Boričić [11]) (hypothetical syllogism rule probabilized):

$$\max\{1 - a, b\} + \max\{1 - b, c\} - 1 \leq P(A \to C) \leq b + c + 2(1 - a);$$

(e) (T. Hailperin [2]) (hypothetical syllogism rule probabilized):

$$\max(0, r + s - 1) \leq P(A \to C) \leq \min(r + s, 1).$$

The bounds in (a), (b), (c), (d) and (e) are the best possible.

Let us explain the differences of various "best possible bounds". In each considered case, we have functions depending on different sets of variables. For instance, in Case (c), the bounds depend on a, b, r, and s, while in Case (d), they depend on a, b, c, etc. Therefore, it is natural to obtain different best possible bounds for the same proposition.

Proof. (a) Suppose $P(A \to B) \le P(\neg A) + P(B)$ and $P(B) \le P(A \to B)$. Then, we can infer $P(A \to B) - P(\neg A) \le P(B) \le P(A \to B)$, i.e., $a + r - 1 \le P(B) \le r$. The bounds are the best possible, because for $P(\neg A \wedge \neg B) = 0$, we have:

$$r = P(A \to B) = P(\neg A) + P(B) - P(\neg A \wedge B) = P(B) - P(\neg A \wedge \neg B) = P(B).$$

Similarly, for $P(\neg A \wedge B) = 0$, we have $P(B) = a + r - 1$;

(b) From $P(A \to B) = P(\neg B \to \neg A)$, we directly infer $r - b \le P(\neg A) \le r$.

(c) Bearing in mind that $(A \to B) \wedge (B \to C)) \to (A \to C)$ is a tautology, we have $P(((A \to B) \wedge (B \to C)) \to (A \to C)) = 1$. Using Lemma 2(e), we inferred $P((A \to B) \wedge (B \to C)) \le P(A \to C)$. Furthermore, by Lemma 2(b), we have:

$$P((A \to B) \wedge (B \to C)) = P(A \to B) + P(B \to C) - P((A \to B) \vee (B \to C)) = r + s - 1;$$

The upper bound is obtained from the following:

$$\begin{aligned}
P(A \to C) &= 1 - P(\neg(A \to C)) \\
&= 1 - P(A \wedge \neg C) \\
&= 1 - P(A \wedge \neg C \wedge B) - P(A \wedge \neg C \wedge \neg B) \\
&\le 1 - P(A) - P(\neg C \wedge B) + 1 - P(\neg B) - P(A \wedge \neg C) + 1 \\
&= 1 - P(A) + P(B \to C) - P(\neg B) + P(A \to C) \\
&= r + s + b - a.
\end{aligned}$$

We show that the lower bound is reached in the case $P(A \wedge \neg B \wedge C) = P(\neg A \wedge B \wedge \neg C) = 0$ and the upper bond in the case $P(A \vee (B \wedge \neg C)) = P(\neg B \vee (A \wedge \neg C)) = 1$. Bearing in mind the two following equations:

$$P(A \to C) = 1 + P(A \wedge \neg B \wedge C) + P(\neg A \wedge B \wedge \neg C) - P(A \wedge \neg B) - P(B \wedge \neg C),$$

$$P((A \to B) \wedge (B \to C)) = P(A \to B) + P(B \to C) - 1 = 1 - P(A \wedge \neg B) - P(B \wedge \neg C)$$

and the fact that the lower bound is reached if $P(A \to C) = P((A \to B) \wedge (B \to C))$, we concluded that the lower bound is reached if $P(A \wedge \neg B \wedge C) + P(\neg A \wedge B \wedge \neg C) = 0$, i.e., $P(A \wedge \neg B \wedge C) = P(\neg A \wedge B \wedge \neg C) = 0$. Similarly, we obtained the conditions for the upper bound, and for example, if A is any tautology and for propositions B and C $P(B \wedge C) = 0$ holds, then the upper bound is reached;

(d) From part (c), we conclude $P(A \to C) \ge P(A \to B) + P(B \to C) - 1 \ge \max\{1 - a, b\} + \max\{1 - b, c\} - 1$. On the other hand, for the upper bound, we have $P(A \to C) \le r + s + b - a \le 2 + b + c - 2a$.

The lower bound is reached in the case $P(A \wedge \neg B \wedge C) = P(\neg A \wedge B \wedge \neg C) = 0, P(A \wedge \neg B) = \min\{a, 1 - b\}$ and $P(B \wedge \neg C) = \min\{b, 1 - c\}$. This will be the case when, for instance, B is any contradiction, and A and C propositions such that $P(A \wedge C) = 0$ holds. The upper bond is reached when $P(A \vee (B \wedge \neg C)) = P(\neg B \vee (A \wedge \neg C)) = 1$ and $P(A \wedge \neg B) = P(B \wedge \neg C) = 0$.

□

As immediate consequences of the previous Lemma, we obtained the following Suppes-style probabilistic rules, similarly as in [9]:

Corollary 1. *Let $P(A) \geq 1 - i\varepsilon$, $P(B) \geq 1 - j\varepsilon$, $P(C) \geq 1 - k\varepsilon$, $P(A \to B) \geq 1 - m\varepsilon$, and $P(B \to C) \geq 1 - n\varepsilon$, with $2(x + y)\varepsilon > 1$, for $x, y \in \{i, j, k, m, n\}$. Then:*
(a) (P. Suppes [10]) (modus ponens probabilized):

$$P(B) \geq 1 - (i + m)\varepsilon;$$

(b) (C. G. Wagner [1]) (modus tollens probabilized):

$$P(\neg A) \geq 1 - (j + m)\varepsilon;$$

(c) (M. Boričić [11]) (hypothetical syllogism rule probabilized):

$$P(A \to C) \geq 1 - (j + k)\varepsilon;$$

(d) (M. Boričić [11]) (hypothetical syllogism rule probabilized):

$$P(A \to C) \geq 1 - (m + n)\varepsilon.$$

The bounds in (a), (b), (c), and (d) are the best possible.

It is worth noting that Wagner (see [1]) was the first who emphasized the difference between Hailperin's (see [4,5]) and Suppes' (see [10]) approaches. Suppes' approach is appropriate for dealing with propositions having high probabilities, with the promise of very elegant syntactic forms of probabilistic axioms and rules (see [9]).

Finally, it is important to point out that probabilistic versions of modus ponens and modus tollens were considered in the case when implication $A \to B$ was interpreted as conditional probability $P(B|A)$ (see [1,10]). The argumentation is quite intuitive and natural. On the other hand, if the hypothetical syllogism inference rule is considered in the same way, it loses its usual logical sense. Namely, it is not difficult to find examples when probabilities of hypotheses $P(B|A)$ and $P(C|B)$ are not close to one, but the conclusion $P(C|A)$ has a very high or a very low probability. Let us note that the similar forms of the probabilistic hypothetical syllogism rule were used in [8,9,18], but with essentially different argumentation and context.

5. Concluding Remarks

The probabilization of logical inference rules treating all propositional connectives, except implication, follows the usual principles of probability theory with the identification of intersection with conjunction and union with disjunction. The case with implication is somehow more complex. For instance, the inference rule of hypothetical syllogism (or transitivity), by which we can infer $A \to C$, from $A \to B$ and $B \to C$, loses proposition B in its conclusion appearing in both hypotheses. This means that this rule, as well as, modus ponens and modus tollens, its particular subcases do not possess the so-called subformula property. This is an additional problem to define the probability $P(A \to C)$ essentially depending on proposition B, which does not occur in $A \to C$. This rule is also a good opportunity to highlight subtle differences between the probabilities of conditionals and conditional probabilities (see [1]). We determined the best possible bounds for $P(A \to C)$, covering simultaneously known bounds for both modus ponens and modus tollens rules.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Wagner, C.G. Modus tollens probabilized. *Br. J. Philos. Sci.* **1955**, *54*, 747–753. [CrossRef]
2. Hailperin, T. Probability logic. *Notre Dame J. Form. Log.* **1984**, *25*, 198–212. [CrossRef]
3. Boole, G. *The Laws of Thought*; Prometheus Books: New York, NY, USA, 2003.
4. Hailperin, T. Best possible inequalities for the probability logical function of events. *Am. Math. Mon.* **2017**, *72*, 343–359. [CrossRef]
5. Fréchet, M. Généralisations du théoréme des probabilités totals. *Fundam. Math.* **1935**, *25*, 379–387. [CrossRef]
6. Frisch, A.M.; Haddawy, P. Anytime deduction for probabilistic logic. *Artif. Intell.* **2017**, *69*, 93–122. [CrossRef]
7. Ognjanović, Z.; Rašković, M.; Marković, Z. *Probability Logics*; Springer: Berlin/Heidelberg, Germany, 2016.
8. Boričić, M. Sequent calculus for classical logic probabilized. *Arch. Math. Log.* **2017**, *58*, 119–138. [CrossRef]
9. Boričić, M. Suppes-style sequent calculus for probability logic. *J. Log. Comput.* **2017**, *27*, 1157–1168. [CrossRef]
10. Suppes, P. Probabilistic inference and the concept of total evidence. In *Aspects of Inductive Inference*; Hintikka, J., Suppes, P., Eds.; North-Holland Publ. Company: Amsterdam, The Netherlands, 1966; pp. 49–55.
11. Boričić, M. Hypothetical syllogism rule probabilized. *Bull. Symb. Log.* **2014**, *20*, 401–402
12. Bonevac, D. *Deduction*; Blackwell Publishing: Malden, MA, USA, 2003.
13. Carnap, R. *Logical Foundations of Probability*; University of Chicago Press: Chicago, IL, USA, 1950.
14. Popper, K.R. Two autonomous axiom systems for the calculus of probabilities. *Br. J. Philos. Sci.* **1955**, *6*, 51–57. 176, 351. [CrossRef]
15. Leblanc, H.; van Fraassen, B.C. On Carnap and Popper probability functions. *J. Symb. Log.* **1979**, *44*, 369–373. [CrossRef]
16. Leblanc, H. Popper's 1955 axiomatization of absolute probability. *Pac. Philos. Q.* **1982**, *69*, 133–145. [CrossRef]
17. Leblanc, H. Probability functions and their assumption sets—The singular case. *J. Philos. Log.* **1983**, *12*, 382–402. [CrossRef]
18. Boričić, M. Inference rules for probability logic. *Publ. l'Institut Mathematique* **2016**, *100*, 77–86. [CrossRef]