abstracta Linguagem, Mente & Ação ISSN 1807-9792 Volume 9, Number 1, pp. 51–68, 2016

What is Logical about the Logical Interpretation of Probability?

Parzhad Torfehnezhad

Département de Philosophie
Université de Montréal
Pavillon 2910, boul. Édouard-Montpetit
Montréal, QC, H3C 3J7, Canada
parzhad.torfeh-nezhad@umontreal.ca

Abstract
My goal, in this paper, is to critically assess the categorization of "interpretations of probability" as it appears in the literature. In some sources only Carnap's treatment of probability is understood to be the best example of "logical" probability. This is surprisingly narrow and I will here suggest otherwise. In fact, I believe that certain forms of Baysianism should also be included in the logical camp.

Introduction

In the development of scientific theories, many pre-theoretical concepts, such as weight, volume, density, etc., acquire a precise theoretical meaning together with a systematic numerical assignment. The assignment normally takes place by defining a functor[1] that relates the domain of the physical properties such as weight to that of numbers. It is by the means of this apparatus that we make sense of the sentences such as "the weight of X is n". Clearly, such systems presuppose the basic logical and mathematical syntax of the relevant number theory they employ. In addition to these basic elements, certain primitive terms (e. g., m for mass, a for acceleration, etc.) and axioms (e. g., laws of classical mechanics) are added and constitute the theory. Semantical rules of designation are then laid down allowing us to interpret the primitive term m as mass, a as acceleration, and so on. In short, we may regard the full interpretations of pre-theoretic concepts as being constituted in the following two separate rudiments: Full Interpretation = (axioms & primitive terms) + (descriptive & pure semantical rules of designation)[2]. There are some important points that one should notice here. First, it is well known that the symbolic calculus is to be developed independently of any interpretation of it. Second, both alternative set of axioms and alternative set of designation rules could lead to alternative interpretations.

[1]The term "functor" is used in a Carnapian sense (Carnap, 1942: 17) not as it is commonly used in category theory. To see other examples of the use of axiomatic methods in science, see the section "Physical Calculi and Their Interpretations" in (Carnap, 1939: 56).

[2]Despite using Carnap's terminology for the sake of coherence in this paper, I should clarify that the use of axiomatic method and the relationship between the primitive terms, axioms, and their interpretations as described in this paragraph, are not limited to Carnap's. It is a commonly accepted concept in today's mathematics and in mathematical textbooks (Lee, 2013; Novikov, 2001).

These considerations apply mutatis mutandis to the concept of probability. Indeed, what does it mean to say: “the probability of X is p”? Probability is a vague concept we use in ordinary language that needs to be explicated[3].

My goal, in this paper, is to critically assess the categorization of “interpretations of probability” as it appears in the literature, as, for instance, in the entry of the Stanford Encyclopedia of Philosophy (Hájek, 2012)[4]. In section 3 of “Interpretations of Probability” in the Stanford Encyclopedia of Philosophy, the author categorizes interpretations of probability as follows: 3.1 Classical probability, 3.2 Logical probability, 3.3 Subjective probability, 3.4 Frequency Interpretations, 3.5 Propensity Interpretations, and 3.6 Best-System Interpretations. According to the author, “early proponents of logical probability include (Johnson, 1921), (Keynes, 1921), and (Jeffreys, 1939). However, by far the most systematic study of logical probability was by Carnap”. I should mention that my objection to this categorization is not merely terminological, but, as we will see, it is directed to a seemingly-overlooked and important issue in identifying interpretations in axiomatic systems in general, being the possibility of having different interpretations on the basis of different sets of syntactic rules as opposed to having different interpretations due to different sets of semantic rules. Interestingly enough, Carnap himself may have pointed out this very issue, and if true, his philosophical position on probability should be re-evaluated, regardless of his proposal for inductive logic. It should be underlined that there is also a general point to be made about the relationship between interpretations, syntax and semantics in the axiomatic systems that is not limited to Carnap, and is a well-known fact in today’s mathematics.

As I have just said, only Carnap’s treatment of probability is considered to be the prime example of “logical” probability. This is surprisingly narrow. Can it be that only Carnap well defended a logical interpretation of probability? I will here suggest otherwise. In fact, I believe that certain forms of Baysianism should also be included in the logical camp. I will also suggest another basis for classifying interpretations of probability according to which the subjective interpretation can also be understood as being “logical”. I should clarify here that the point of this paper is not to give a final answer to what probability is; it is neither the point to compare the advantages and disadvantages of different interpretations of probability. The main issue is to establish a framework, consistent with the present axiomatic methods, according to which one could clearly situate and group different interpretations of probability, and thereby have a better understanding of what logical probability is and how we should employ the term.

I will proceed as follows: I will start by considering probability theory as an axiomatized theory. Then, two possible and different sets of axioms in the theory will be given: one with the logical relation of consequence included, and one without. I will argue that at the very basic level, even before reaching the full interpretation of what probability is (that is to say considering semantic rules), there are two basic possibilities for final interpretation of probability. I will then conclude that this basic split should be the primary basis for categorizing interpretations of probability.

[3]The whole process of replacing vague concepts such as one-place predicates like “heavy” in “this thing is heavy” or a two-place predicate “heavier” in “this thing is heavier than that” by a functor that allows us to say, in a more precise way, “the weight of this thing is 7kg” is what Carnap calls an explication. In the given example, the vague concept “heavy” is called an explicandum and the clearer concept “weight (in kilogram)” is called an explicatum. At the axiomatic level, since there is no factual content involved, all that matters is the mathematical properties (such as consistency, completeness, independence, etc.) of the calculus upon which the number assignment is delivered. See (Carnap, 1947a: 7).

[4]This is but one important and typical example. For other examples see (Hájek & Joyce, 2008; Lyon, 2009; Sarkar & Pfeifer, 2006).

I will then proceed to Carnap's general analysis of probability. In this section, I will draw reader's attention to the two layers of Carnap's analysis of probability. The first layer concerns a linguistic approach to the term "probability" in ordinary language[5]. This first layer of the discussion acknowledges the existence of two fundamentally different concepts (two explicanda) of probability, one logical and one factual. I believe that this fundamental distinction speaks to the same point I am trying to make here concerning the two possible arrangements of axioms in probability theory. The second layer of the discussion, however, involves Carnap's own proposal for covering the logical concept. Although I will give a summary of Carnap's proposal for the logical aspect of probability (inductive logic), I will not defend, nor oppose, Carnap's position in this regard here[6].

Probability Theory: A Brief Historical Background

In the mathematical sense, a probability space is a measure space. Measurements, in general, consist of assigning numerical values to different possibilities of measurable objects (in a continuous way). The first axiomatization of the theory was proposed by Andrey Kolmogorov in 1933 (Hájek, 2012), according to which probability P is considered to be a function from the set F to the set of real numbers satisfying the following conditions:

1. (Non-negativity): $P(A) \geq 0$, for all $A \in F$.
2. (Normalization): $P(\Omega) = 1$.
3. (Finite additivity): $P(A \cup B) = P(A) + P(B)$ for all $A, B \in F$ such that $A \cap B = \emptyset$.

Where Ω is a non-empty set (universal set), and F is an algebra[7], closed under complementation and union, which is a subset of the power set of Ω. Now, let us take a look at a very brief background of some of the interpretations of probability. According to Bernoulli-Laplace's classical interpretation[8], in a random process, if N is the number of equally-likely and mutually-exclusive outcomes, and N_A is the number of outcomes in which the desired event A occurs, then the probability of A would be calculated as the ratio of N_A to N.

$$\frac{P(A) = N_A}{N}$$

Frequentists, instead, focus on finding probability for actual events disregarding calculating probabilities prior to trial experiments. If a frequentist wants to know the probability of coming up a head in the process of coin-tossing, for example, she would actually toss the coin for a number of times and record the results in order to establish what is called a "reference class"; the greater the population of the reference class, the more accurate the probability. Then she would compare the limit of the ratio of frequency of the desired event in the reference class population to the total population of the reference class (when the population approaches to infinity) in order to calculate the probability of the desired event occurring. Therefore, if n_A

[5]This layer of discussion is directly related to Carnap's general analysis of linguistic frameworks in "the Foundations of Logic and Mathematics" (Carnap, 1939).

[6]This discussion would be a lengthy one in which one has to discuss more detailed subject matters such as whether or not the value of λ, and hence the confirmation function is regarded as completely arbitrary.

[7]An algebra, here, broadly speaking, is considered as a set along with some operations satisfying certain conditions. An n-ary operation on Ω is a *function* that takes n elements of Ω and gives a single element of Ω.

[8]This interpretation is famously circular for the use of the adjective "equally-likely" in the definition of probability, which itself is a measurement of likelihood. The problem would not be solved, even by appealing to the "principle of indifference": whenever there is no evidence favoring one possibility over another, they have the same probability (Hájek, 2012). It is worth mentioning that Carnap strongly rejects both the classical interpretation and any appeal to this principle: "We regarded the classical conception of probability, represented chiefly by Jacob Bernoulli and Laplace, as definitely refuted by the criticism of the frequentists. The classical conception was essentially based on the principle of insufficient reason or indifference according [...]", Carnap in (Schilpp, 1963: 70).

is the number of occurrences of A in the total number of n_t trial, then the probability of A would be *P(A)* in the following formula where p is a real number:

$$P(A) = \lim_{n_t \to \infty} \frac{n_A}{n_t} = p$$

As practical as this interpretation might be, there are problems (both practical and theoretical) associated with it, *if* we limit ourselves just to this interpretation of probability. For example, in the case of the infinite ways of establishing a reference class, what would justify our choice of a reference class? In the case of coin tossing, is the reference class the class of infinite numbers of tossing the same coin, or it is infinite tossing of different coins of the same type (each only once)? Therefore, what would be the reference class of singular events? There are other problems associated with this interpretation that are discussed at length in the literature (Hájek, 2009).

On the other hand, total reliance on experimentation in order to capture the whole concept of probability may raise other problems, too. For instance, the obviousness of the sentence "when there is a 10% chance for the occurrence of A, then 90% of the time A would not occur" does not rely on any sort of empirical experiment. One can clearly see that there are two different applications of probability here; in one application, probability can be assigned to events synthetically via observation, whereas, in the other application, the assignment of probability follows some well-defined analytic structures[9]. We will have a more detailed discussion on this topic in the subsequent sections.

1 Logicality and Conditional Probability

One of the philosophical issues, with regard to conditional probability, is the issue of possibility versus probability. One may say "possibility" is a stronger concept than probability in the sense that the probability of impossible events might be considered as zero, but zero probability does not imply impossibility. According to Fitelson et al., Kolmogorov himself was certain that "probability 0 does not imply impossible" (Fitelson et al., 2006). For Kolmogorov, as we saw above, conditional probability $P(A|B)$ can be derived from unconditional probability (as represented above in the form of a one-place function *P(_)*).

It is obvious that, according to this formulation the conditional probability is undefined if either $P(B) = 0$ or any unconditional probabilities are undefined.

This situation, in which "zero-probability" is undefined, poses a problem for the theories in which probability is considered as a one-place function because, as Fitelson (Ibid) points out, "in uncountable spaces there can be genuine, non-trivial events whose probabilities are undefined (so-called 'non-measurable sets'), and others whose probabilities are 0".

Let's take a short detour in order to clarify the zero-probability issue and why countable additivity becomes an important issue. Suppose we define the following:

- Sample space $\equiv \Omega$
- Sample point $\equiv$ possible outcome $\equiv \omega \in \Omega$
- Event $\equiv$ (a subset of Ω) $\equiv E$

Therefore, if $\Omega = [0, 1]$ then an event E is a subinterval $[a, b]$ of $[0, 1]$. Now, the property that the measuring function $P(x)$ ought to satisfy with respect to the subintervals is that if $[a, b] \subseteq [0, 1]$, then $P([a, b]) = b - a$. Accordingly, for a single possible outcome ω we have

$$\forall\, \omega \in \Omega, P(\{\omega\}) = P([\omega, \omega]) = \omega - \omega = 0$$

[9]The terms "analytic" and "synthetic", here, are meant to be in the Carnapian sense.

This means that every possible outcome is a zero-probability event. This conclusion seems contradictory because if $\Omega = \cup\{\omega\}$, then $P(\Omega) = P(\cup\{\omega\}) = \sum P(\omega) = 0$, and we know $P(\Omega)$ is supposed to be 1. Here is where the axiom of countable additivity becomes important because the axiom only applies to countable objects and Ω is not countable.

Nevertheless, the zero-probability situation is a mathematically non-trivial question and it is meaningful in an infinite context; and if we are to assume probability as a one-place function, we have to face this situation in one way or another.

An alternative way of axiomatizing the notion of probability is to consider it not as a one-place function but as a two-place function, *P* (_, _), and take this conditional probability as a primitive term (as Carnap does; see below). It has been shown (Stalnaker, 1970) that, in this case, it is possible to derive the unconditional probability of *A* as *P (A,T)*, where *T* is a logical truth (a *L*-true proposition, a tautology). According to Fitelson (Fitelson et al., 2006), various axiomatizations in which conditional probability is taken to be primitive have been defended in the literature[10]; and there are also canonical models with proven completeness and soundness in which probability is taken to be a two-place function (Lepage & Morgan, 2003).

A group of interpretations, also known as Bayesianism (Hájek, 2012), adopts this alternative axiomatization. Subjective interpretation is one of the examples in which conditional probability is primitively taken to be a two-place function. In the subjective interpretation of probability, it is the amount of the agent's knowledge (who is facing an uncertain situation) that plays a crucial role in the probability assignments. In other words, the source of uncertainty is considered to be the epistemic state of the agents. Therefore, the assignment of probability ought to be relativized to the amount of the knowledge the agent has. Then, in fact, what probability represents is the degree of the agent's belief. And, probability theory can be regarded as a guide to rational beliefs in every particular situation. Consequently, subjectivists heavily lean toward the concept of conditional probability (Joyce, 2008). Bayes' theorem[11] is central to subjectivists' theories of confirmation "both because it simplifies the calculation of conditional probabilities and because it clarifies significant features of subjectivist position" (Joyce, 2008).

Nowadays, Bayesian interpretation is regarded to be an extension of propositional logic[12]. Hájek shows (Hájek, 2008) that even at the axiomatic level we may see the reliance of probability theory on deductive logic[13]. At the axiomatic level, we could, as well, attach probabilities to members of a collection *S* of *sentences* of a formal language, closed under (countable) truth-functional combinations, and consider counterparts of the above-mentioned Kolmogorov's axioms as the following:

I. $P(A) \geq 0$ for all $A \in S$.

II. If *T* is a logical truth (in classical logic), then $P(T) = 1$.

---

[10]There are also examples in which both conditional and unconditional probabilities are simultaneously axiomatized (Goosens, 1979) but the system is considered as an extension of Kolmogorov probability theory, and has the usual definitions linking conditional and unconditional probabilities as theorems.

[11]According to the Bayes' theorem conditional probability can be considered as follows:

$$\frac{P(A,B) = P(B,A)P(A)}{P(B)}$$

It can be proven that Bayes' theorem is equivalent to the following formulation (Fitelson et al., 2006) (where $P(A^c)$ is the probability of the events in which *A* would not occur), which would not face the zero-probability problem:

$$\frac{P(A|B) = P(B,A)P(A)}{P(B,A)P(A) + P(B,A^c)P(A^c)}$$

[12]See (Leitgeb, 2014), (Pearl, 1991), (Cowell, 1999), (Jensen, 2001), (Kersting, 2007).

[13]For an extensive and detailed discussion on the alternative set of axioms including the conditional probability function as primitive see §4.2.1 (Leitgeb, 2015).

III. $P(A \vee B) = P(A) + P(B)$ for all $A \in S$ and $B \in S$ such that A and B are logically incompatible.

In this case, since the notions of "logical truth", "logical incompatibility" and "implication" are well defined in deductive logic, one can consider probability theory as a dependent theory on classical logic.

If we consider the evidence *e* and the hypothesis *h* as the sentential arguments of the probability function $P(e, h)$, the relativized Kolmogorov's set of axioms would look like the following:

Kolmogorov's Axioms	Relativized Kolmogorov Axioms
I. $P(E) \geq 0$, for all $E \in F$. II. $P(\Omega) = 1$. III. $P(E_1 \cup E_2) = P(E_1) + P(E_2)$ for all $E_1, E_2 \in F$ if $E_1 \cap E_2 = \emptyset$.	I. $P(e) \geq 0$ for all $e \in Z$. II. If T is a logical truth, then $P(T) = 1$. III. $P(h \vee e) = P(e) + P(h)$ for all $e \in Z$ and $h \in Z$ such that e and h are logically incompatible.

Table 1: Relativized Kolmogorov's Axioms to Classical Logic.

Van Fraassen (Van Fraassen, 1995), for example, offers a probability system in which a two-place function is taken as primitive and the basic axioms for conditional probability were adopted rather than conditional probability is being defined in the usual way from monadic probability. Costa and Parikh (Costa & Parikh, 2005) offer an extended version of van Fraassen's model and use it to model Bayesian update in a non-monotonic way. As quoted by Costa and Parikh (Ibid), van Fraassen studies two place probability functions $P(.|.)$ defined on a σ-field F over some set U with the following requirements:

(I) For any fixed A, the function $P(X|A)$ as a function of X is either a (countably additive) probability measure, or has constant value 1.

(II) $P(B \cap C|A) = P(B|A)\, P(C|B \cap A)$ for all A, B, C in F. If $C \subseteq B \subseteq A$, then (II) can be simplified as

(II') $P(C|A) = P(C|B)\, P(B|A)$.

As we see, mathematically speaking, as far as the conditionality is concerned, subjective interpretation can be well regarded as an extension of propositional logic; because, probability, right from scratch, is taken to be a conditional function. According to the above explanations, one can see that what can be deemed as "logical" in all of the possible interpretations on the basis of taking a two-place function as a primitive function, is the primitive inclusion of the logical consequence relationship in the axioms (i. e., conditional probability). In fact, Leitgeb (Leitgeb, 2014), based on similar argumentation, proposes a new category of logical systems called "probabilistic logic". Maher (Maher, 2006, 2010), also talks about the legitimacy of using "inductive probability" as an umbrella term to cover the logical aspect of probability as opposed to what he calls "physical probability".

So, in terms of axiomatization, we have basically two options with regard to conditional probability. We are either to take the absolute probability (a one-place function) as a primitive term and derive conditional probability from it as a theorem, or to take conditional probability primitively (as a two-place function) and derive absolute probability from it by adding some other definitions. Mathematicians have tried both of these options. It is fairly obvious that the choices are different and not completely equivalent, not only mathematically but also philosophically. In one case, probability is taken to be an intrinsically conditional concept (requiring two arguments for a two-place function) and it is a primitive part of axiomatization, while, in the other case, conditional probability is not a part of axiomatization, and is derivable from an unconditional concept of probability (a one-place function). Now, one may easily see that regardless of the possible alternative semantic set-ups for each possibility, just at the axiomatic level, we can be sure that there exist at least two different (incompatible) possibilities for interpreting probability without knowing what the interpretations actually are going to be. This fact, as we will see in the subsequent section, might be related to what Carnap calls "two explicanda of probability".

It should be fairly clear by now that if we are to categorize interpretations of probability at the end, the primary division should be on the basis of inclusion or exclusion of a two-place function in the axiom set as our measuring function: hence, the division would be between logical (in the sense of primitive inclusion of consequence relationship in the axiom set) versus non-logical [14] (in the sense of primitive exclusion of the conditional relation) probability. Once again, I would like to emphasize that this division is totally independent of any full interpretation of the concept; it is just the result of some mathematical manipulations regardless of the content. Therefore, it should be well regarded as a fundamental split rather than a semantical one. In the subsequent section, we will see some evidence, which strongly suggests that Carnap was already well aware and assertive of this very primary division with regard to the concept of probability, and perhaps it is according to this fundamental split that he finds the subjectivists'-frequentists' debate futile.

In the literature, Carnap's position is categorized as being a "logical interpretation" of probability, meaning that it is neither Bayesian, nor frequentist (Fitelson et al., 2006; Hájek, 2012; Lyon, 2009; Sarkar & Pfeifer, 2006). As we discussed above, being logical (or not) is more a theoretical matter rather than a matter of descriptive semantics, hence, considering this fact, certain forms of Bayesianism can also be equally considered as "logical" for their theoretical positions with regard to conditional probability. Therefore, the common categorization of interpretations of probability in the present literature does not seem to be completely correct[15]. It seems to me that the underlying assumption behind this way of categorizing the various interpretations of probability (contrary to Carnap's analysis) is that probability is a unique and single explicandum that needs to be explicated. And I assume it is probably due to this assumption that the list of interpretations of probability is given generally in the literature, e. g., (Lyon, 2009; Sarkar & Pfeifer, 2006), and in particular in the above-mentioned entry in Stanford Encyclopedia of Philosophy. Thus, we read in the very first lines of the first paragraph.

---

[14]Instead of non-logical, one may use the term "frequential" or "physical".

[15]I should mention here that Hájek himself is against using the term "interpretation" in the sense we normally use with respect to formal systems on two grounds: First, there are other interpretations of probability that are not exactly based on Kolmogorov's Axioms yet are categorized under it (perhaps under the logical setting we talked about). Secondly, Kolmogorov's Axioms can be equally employed for measuring spaces other than probability.

Normally, we speak of interpreting a formal system, that is, attaching familiar meanings to the primitive terms in its axioms and theorems, usually with an eye to turning them into true statements about some subject of interest. However, there is no single formal system that is 'probability', but rather a host of such systems. To be sure, Kolmogorov's axiomatization [...] has achieved the status of orthodoxy [...]. Nevertheless, several of the leading 'interpretations of probability' fail to satisfy all of Kolmogorov's axioms, yet they have not lost their title for that. Moreover, various other quantities that have nothing to do with probability do satisfy Kolmogorov's axioms, and thus are interpretations of it in a strict sense: normalized mass, length, area, volume, and other quantities that fall under the scope of measure theory, the abstract mathematical theory that generalizes such quantities. Nobody seriously considers these to be 'interpretations of probability', however, because they do not play the right role in our conceptual apparatus. (Hájek, 2012)But there is no need to modify the meaning of "interpretation" if we accept that there are fundamentally two different sets of axioms for measuring probability and they are both related to the concept of measurable space.

> 'Interpreting probability' is a commonly used but misleading characterization of a worthy enterprise. The so-called 'interpretations of probability' would be better called 'analyses of various concepts of probability', and 'interpreting probability' is the task of providing such analyses. Or perhaps better still, if our goal is to transform inexact concepts of probability familiar to ordinary folk into exact ones suitable for philosophical and scientific theorizing, then the task may be one of 'explication' in the sense of Carnap.' (Hájek, 2012: 1)

Probability in Carnap's Philosophy

After establishing the existence of two fundamentally different settings for interpreting probability, let us look at Carnap's treatment of probability. One should be aware that there are two levels in Carnap's discussion on probability. At the first level Carnap takes a linguistic approach and considers the term "probability" as an explicandum. This approach is, of course, closely correlated to his general approach in analyzing philosophical problems with respect to the concept of linguistic framework and his theory of meaning, which would require more elaboration (see (my own paper) for details). It is here, at this level, that I believe he wants to authenticate what we have established so far, namely, the existence of two settings for interpreting one concept, probability, or, using his own vocabulary, the existence of two explicanda for one term "probability". Once he establishes that there is a logical setting (logical explicandum) for the term, *then*, we go to the second level of his analysis where he proposes his method for explicating this explicandum via what he calls "inductive logic".

Carnap's philosophy, as we will see, suggests that some concept cannot be explicated as one single explicandum and, probability is one of them. For Carnap both logical and factual explicanda of probability need to be explicated.

Carnap believes that there are two fundamentally different explicanda of probability, each needing to receive explications separately, and hence, each amount to a different kind of confirmation. This, in my view, is the first level in Carnap's philosophy of probability, as I mentioned above. But, if we accept this premise, then we ought to accept that no sole theory of probability can deliver the whole meaning of the term: to capture the whole meaning of probability one needs at least two different theories.

> 'The various theories of probability are attempts at an explication of what is regarded as the prescientific concept of probability. In fact, however, there are two fundamentally different concepts for which the term 'probability' is in general use. [...] (i) Probability**1** is the degree of confirmation of a hypothesis *h* with respect to an evidence statement *e*, (ii) Probability**2** is the relative frequency (in the long run) of one property of events or things with respect to another. A sentence about this concept is factual, empirical. Both concepts are important for science. Many authors who take one of the two concepts as explicandum are not aware of the importance or even of the existence of the other concept. This has led to futile controversy. (Carnap, 1962: 19)

It is well-known that for Carnap "There are two explicanda, both called 'probability': (1) logical or inductive probability (probability$_1$), (2) statistical probability (probability$_2$)" (Carnap, 1973: 269). Carnap is quite clear in that "there is no one meaning of the term 'probability' which is applied with perfect consistency throughout his work by any of the classical authors" (Carnap, 1962: 50). He believes that in some cases, when we have large-enough reference classes, the relative frequency can be regarded as the representative of the ultimate relative frequency, but this concept still ought to be regarded as the explicandum of the factual probability (probability$_2$).

> 'If we take a sufficiently large unknown class K, then the relative frequency of M in K may be regarded as representing the relative frequency "in the long run". But this is the explicandum of probability2, the statistical concept of probability. (Carnap, 1962: 173)

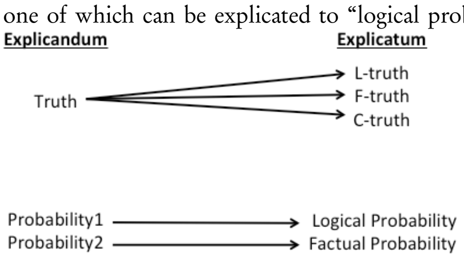
It is worth mentioning that Carnap makes a sharp distinction between "logical" and "factual truth", and also between "pure" and "physical geometry". Making those distinctions are not quite similar to the way he distinguishes between explicanda of probability[16] (as "a concept mostly used by scientists and a concept suited for philosophical investigations"). But, sorting out concepts in a factual-theoretical continuum is characteristic of Carnap's approach towards many philosophical problems. And it is tied to his overall position with regard to linguistic framework, abstraction and his whole theory of meaning. Making these distinctions is not only philosophically important for Carnap but also quite beneficial, because:

> In this way we obtain also a clear distinction between questions about contingent facts and questions about meaning relations. This difference seems to me philosophically important; answering questions of the first kind is not part of the philosopher's task, though he may be interested in analyzing them; but answers to questions of the second kind lie often within the field of philosophy or applied logic. Carnap in (Schilpp, 1963: 63)

There are many places in which Carnap refers to frequentist interpretations as legitimate and acceptable interpretations of factual probability. The main point is that the job of dealing with this interpretation is not primarily the job of philosophers. It is true that, as a philosopher, Carnap tends to work on theories of logical probability, but that does not mean that the concept of factual probability is *persona non grata* in his philosophy. On the contrary, for Carnap, factual probability is a legitimate concept, but responsible for only one part of the whole meaning of probability, which may well receive its own treatment via a frequentist interpretation in light of the empirical data.

> On the other hand, probability2 is an empirical concept; it is the relative frequency in the long run of one property with respect to another. The controversy between the so-called logical conception of probability, as represented e. g. by Keynes, and Jeffreys, and others, and the frequency conception, maintained e. g. by v. Mises and Reichenbach, seems to me futile. These two theories deal with two different probability concepts which are both of great importance for science. Therefore, the theories are not incompatible, but rather supplement each other. (Carnap, 1945: 591)

[16]The difference between explicating "probability" and "truth", for instance, is that "truth" is considered as a *single*-explicandum (and Carnap explicates it into three different explicata F-truth, L-truth, and C-truth). "Probability", in contrast, is considered as a concept that has essentially two explicanda (or a *double*-explicandum, if you wish), one of which can be explicated to "logical probability" and the other to "factual probability" in a parallel manner.



> The fundamental difference is rather this 'probability2' designates an empirical function, viz., relative frequency, while 'probability1' designates a certain logical relation between sentences; these sentences, in turn, may or may not refer to frequencies.(Carnap, 1946: 72)
>
> The statistical concept of probability is well known to ail those who apply in their scientific work the customary methods of mathematical statistics. [...] In the simplest cases, probability in this sense means the relative frequency [...] Thus the statistical concept of probability is not essentially different from other disposition concepts, which characterize the objective state of a thing by describing reactions to experimental conditions [...]. (Carnap, 1955: 1)

Thus, not only does Carnap have no objection to the frequentist theory, but also he even sees it as a necessary complementary part of the general theory of probability. However, frequentist theory alone is incapable in delivering the whole meaning of probability, and the same is true for logical theories of probability. Understandably, Carnap gives an extensive treatment on the logical aspect of probability, which is briefly presented in the next section, but one has to keep in mind that in this treatment Carnap by no means claims that this theory alone would deliver the whole meaning of probability or that this is *the* interpretation of probability, since there is none.

2 Inductive Logic

In this section, we move to the second level of discussion about probability in Carnap's philosophy. For Carnap, inductive logic is not understood in its traditional, Aristotelian sense, where one deals with inferences from particulars to universals, but as an extension of deductive logic that deals with uncertain propositions. Carnap's reasoning for taking a linguistic stance towards probability rests on the assumption that our knowledge about the facts eventually ought to be expressed in the form of propositions. If the propositions are considered to be certain then deductive logic is sufficient, otherwise one should employ an inductive logic to study their relationships. The difference is depicted in the following diagram:

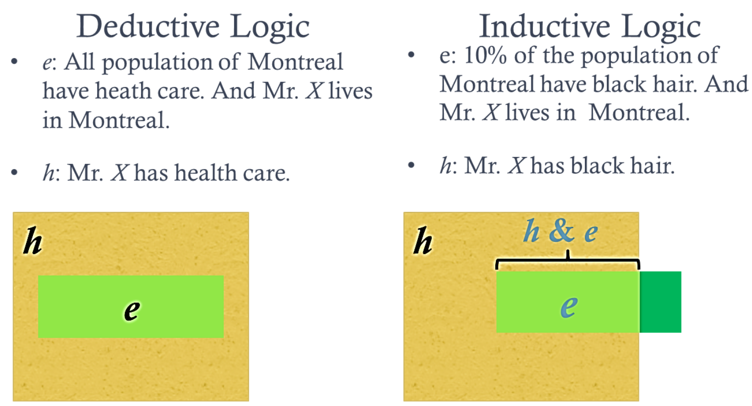


Figure 2.1: Deductive vs. Inductive Logic.

Subsequently, changes in the new evidence that may, in effect, change the previous probability assignment eventually will fall under one of the following events: subtraction or addition of (i) a sentence, (ii) a predicate, or (iii) an object name. We will see how the probability assignment will be subjected to change following each of these events.

For Carnap, "logical probability" is the central concept of inductive logic, in a similar way that "logical implication" is central to deductive logic.

> Inductive logic is here understood as a theory based on a definition of the logical concept of probability, [...]. Its basic concept, the degree of confirmation, is in a certain sense a weak analogue of the concept of logical implication, the basic concept of deductive logic. (Carnap, 1947b: 133)

As we established above, subjective interpretations rest on deductive logic because of their dependency on logical concepts at the axiomatic level. The same is true for logical probability in Carnapian sense; "inductive logic is constructed from deductive logic by the adjunction of a definition of *c* [measuring function]. Hence inductive logic presupposes deductive logic" (Carnap, 1962: 192).

So, Carnap's measuring function is called *c*-function and it is considered to be the explicatum for the vague concept of "logical probability" as its explicandum. *C*-function is considered to be intrinsically conditional and it is a function of two arguments: evidence *e* and hypothesis *h*. This concept, in turn, can play one of the following roles as:

I. The *classificatory* concept (confirming evidence): *c (h, e)* $>$ *b.*

II. The *comparative* concept (higher confirmation): *c (h, e)* $>$ *c (h', e'), or c (h, e)* $>$ *c (h, e'), or c (h, e)* $>$ *c (h', e).*

III. The *quantitative* concept (degree of confirmation): *c (h, e)* = *u*

Carnap then proposes the following five fundamental axioms that can be applied to any pairs of sentences *e* and *h* in a given language L (finite or infinite)(Carnap, 1962: 295).

Carnap's Axioms for the *c*-function	**Relativized Kolmogorov's Axioms**
I. Range of values: $0 \leq c\ (h, e) \leq 1$ II. L-implication: If $\vdash e \supset h$, then $c\ (h, e) = 1$ III. Special addition: If *h¿ h′¿ e* is L-false, then c (*h¿ h′*, e) = $c\ (h, e) + c\ (h', e)$ IV. General multiplication: c (*h¿ h′*, e) = $c\ (h, e) \times c\ (h', h \wedge e$ V. L-equivalence: If $\vdash e \equiv e'$ and $\vdash h \equiv h'$, then $c\ (h, e) = c\ (h', e')$	I. $P\ (e) \geq 0$ for all $e \in Z$. II. If T is a logical truth, then $P(T) = 1$. III. $P\ (h \vee e) = P\ (e) + P\ (h)$ for all $e \in Z$ and $h \in Z$ such that e and h are logically incompatible.

Table 2: Carnap's vs. Relativized Kolmogorov's Axioms.

In order to measure the degree of confirmation Carnap introduces another measuring function called *m*-function that its regular version will be defined for a language L of *N* objects and containing state-description *Zi* (Carnap, 1973), as follows:

m is a regular *m*-function for L $N \overset{\text{def}}{=}$

(a) For every Z_i in L_N, $m\ (Z_i) > 0$

(b) $\sum_i m(Z_i) = 1$

(c) If *j* is *L*-false, $m\ (j) = 0$

(d) If *j* is not *L*-false, $m\ (j) = \sum_i m(Z_i)$ for all Z_i in the *range* of *j*

A state-description describes a (possible) state of affairs or model. Example: if a language system L contains symbols for only three objects *a*, *b*, *c*, and one monadic predicate *F*, then there exist three atomic sentences *i*:*Fa*, *j*:*Fb*, and *k*: *Fc* along with their negations $\sim i$, $\sim j$, $\sim k$. A *state-description Zi* is any conjunction of the mentioned atomic sentences that contains either an atomic sentence or its negation but not both. Example: $Z_1 = i \ \& \sim j \ \& \ k$.

Therefore, the disjunction of all state-descriptions ($\vee Z_i$) gives us a universal class, and the conjunction of them ($\wedge Z_i$) gives us the empty class (see Q-predicates below). The *range* of *i* in

L is the class of those Z in L in which i holds. So, the situation for state-descriptions in our example language of three objects (*a, b, c,*) and one monadic predicate (*F*) can be summarized in the following table:

Z_i	State-description	*m*	Structure-description	Logical Weight	*m**
Z_1	*Fa & Fb & Fc*	1/8	*Everything is F*	1/4	1/4
Z_2	*¬Fa & Fb & Fc*	1/8			1/12
Z_3	*Fa & ¬Fb & Fc*	1/8	*Only two are F*	1/4	1/12
Z_4	*Fa & Fb & ¬Fc*	1/8			1/12
Z_5	*¬Fa & ¬Fb & Fc*	1/8			1/12
Z_6	*¬Fa & Fb & ¬Fc*	1/8	*Only one F*	1/4	1/12
Z_7	*Fa & ¬Fb & ¬Fc*	1/8			1/12
Z_8	*¬Fa & ¬Fb & ¬Fc*	1/8	*Everything is not F*	1/4	1/4

Table 3: State-Descriptions for L_{3^1}.

Using *c*-function, Carnap then calculates the degree of confidence *c* (or *c** considering the logical weight) for the hypothesis *h* given the evidence *e* as follows:

$$c(h,e) = \frac{m(e \bigwedge h)}{m(e)} \qquad \text{or} \qquad c^*(h,e) = \frac{m^*(e \bigwedge h)}{m^*(e)}$$

Let's first examine the *c*-function (regular, i. e., all state-descriptions have equal *m*-value). Take the hypothesis $h{=}Fa$, thus the range is Z_1, Z_3, Z_4, Z_7, and $m(h){=}1/2$. Let's the evidence be $e{=}Fb$, then its range is Z_1, Z_2, Z_4, Z_6, and $m(e)$ =1/2. Then, the range of $h{\wedge}e$ Z_1, Z_4 is $m(h{\wedge}e){=}1/4$. Hence, $c(h,\ e)$ = (¼) /(½) =1/2 which is equal to $m(h)$ and thus it does not confirm it. On the other hand, if we consider all structure-descriptions we will calculate the *c**-function as follows:

$m^*(h){=}$ 1/4+1/12+1/12+1/12=1/2

$m^*(e)$ =1/2

$m^*(h{\wedge}e){=}1/4{+}1/12{=}1/3$

Hence $c^*(h/e)$ $=m^*(h{\wedge}e)/m^*(e)$ = (1/3)/ (1/2) =2/3. Then, by this function, the evidence confirms the hypothesis.

In our example, in which we have only one monadic predicate *F*, the only possible form of proposition is $P = F(x)$. This proposition may either have the property of *Q1* (meaning *P*) or Q_2 (meaning ¬*P*). Knowing these properties, known as *Q-predicates* or *Q-properties*, is essential in determining the logical width (w).

Predicate expression	Logical Nature	Q-Predicate	Width (w)	Relative width (w/κ)
P∧ ¬P	*Empty*		0	0
P	*Factual*	*Q1*	1	½
P ∨ ¬P	*Universal*	*Q1 ∨ Q2*	2	½

Table 4: Q-predicates for L_{3^1}.

Structure-description Str_{i} is the disjunction of isomorphic *Zi* in L that can be identified by Q-numbers, and expressed by a Q-predicate. In the case of having only three monadic predicates, for example, we may have three propositions: $P_1{=}F(x)$, $P_2{=}$ $G(X)$ and $P_3{=}H(x)$. Accordingly we will have the following Q-properties:

Str_i	Q-Predicates	Expression	Logical Nature	Width (w)	w/κ
1		$Pi \wedge \neg Pi$	Empty	0	0
2	Q_1	$Pi \wedge Pj \wedge Pk$	Factual	1	1/8
3	$Q_1 \vee Q_2$	$Pi \wedge Pj$		2	1/4
4	$Q_1 \vee Q_2 \vee Q_3$	$Pi \wedge (Pj \vee Pk)$		3	3/8
5	$Q_1 \vee Q_2 \vee Q_3 \vee Q_4$	Pi		4	1/2
6	$Q_1 \vee Q_2 \vee Q_3 \vee Q_4 \vee Q_5$	$Pi \vee (Pj \wedge Pk)$		5	5/8
7	$Q_1 \vee Q_2 \vee Q_3 \vee Q_4 \vee Q_5 \vee Q_6$	$Pi \vee Pj$		6	3/4
8	$Q_1 \vee Q_2 \vee Q_3 \vee Q_4 \vee Q_5 \vee Q_6 \vee Q_7$	$Pi \vee Pj \vee Pk$		7	7/8
9	$Q_1 \vee Q_2 \vee Q_3 \vee Q_4 \vee Q_5 \vee Q_6 \vee Q_7 \vee Q_8$	$Pi \vee \neg Pi$	Universal	8	1

Table 5: *Q*-properties for L with 3 primitive monadic predicates.

In general, for a language L^n_π where *n* is the number of individual constants, and π is the number of primitive monadic predicates the following holds:

a. The number of atomic sentences is $\beta = \pi n$
b. The number of Q-predicates is $\kappa = 2^\pi$
c. The number of state-descriptions is $2^\beta = 2^{\pi n} = (2^\pi)^n = \kappa^n$

So far we were considering the finite situation. Now, if *n* and π equal to the cardinality of natural numbers $\aleph_0$, then κ forms a continuum, so does the number of state-descriptions.

In the case of continuum, Carnap introduces the function λ(κ) that takes a non-negative real number to characterize the confirmation function *c*λ (*h*, *e*). In the formula for a singular predictive inference[17] to be given below the evidence statement, *e*Q, says that s_1 individuals have property Q_1, s_2 have Q_2,...,s_κ have Q_κ. The hypothesis *hi* says that some individual not mentioned in the evidence has Q_i property ($s_1 + s_2 + \ldots + s_\kappa = s$).

$$c(h_i, e_Q) = \frac{s_i + (\frac{\lambda(\kappa)}{\kappa})}{s + \lambda(\kappa)}$$

Example: In a language that has two objects and one monadic predicate L^2_1 we have $n = 2$; $\pi = 1$; $\kappa = 2^\pi = 2$, and for the *c**-functions $\lambda^*(\kappa) = \kappa = 2$. If $e_Q = Fa$ and $h = Fb$ then $s_1 = 1$, $s_2 = 0$; $s = s_1 + s_2 = 1$. Accordingly, $c^*(h, e) = (1 + 1)/(1 + 2) = 2/3$. For this language, Q-predicates and state-descriptions (Z_i) are:

Predicate Expression	Logical Nature	Q-Predicate
$P \wedge \neg P$	Empty	
P	Factual	Q_1
$P \vee \neg P$	Universal	$Q_1 \vee Q_2$

Z_i
$Fa \wedge Fb$
$Fa \wedge \neg Fb$
$\neg Fa \wedge Fb$
$\neg Fa \wedge \neg Fb$

[17] According to Carnap, main kinds of inductive inferences for a language L^n_π are:

- Direct inference (from the population to a sample),
- Predictive inference (from one sample to another),
- Inference by analogy,
- Inverse inference (from a sample to the population),
- Universal inference (from a sample to a universal law).

There are two methods of considering λ. In the first method, λ is considered to be independent of κ. In that case, $c_\lambda(h_M, e_M) = (s_M + (w/\kappa)\lambda)/(s + \lambda)$. In the second method, λ can be considered as proportional to κ by a constant factor C, which means $\lambda(\kappa) = C\kappa$. In that case $c_\lambda(h_M, e_M) = (s_M + Cw)/(s + C\kappa)$. Regardless of the method, in both of these *c*-functions, Carnap recognizes the ratio of s_M/s as the *empirical factor* and that of w/κ as the *logical factor*.

In the exact same way as we saw above, Carnap derives unconditional probability (for a single statement *j*) from conditional probability given the evidence *e* is a tautological truth, and he calls it "null confirmation" *c0* (Carnap, 1973):

$$c_0(j) \stackrel{\text{def}}{=} c(j, t), \text{ where } t \text{ is a tautological evidence}$$

I need not go into further detail about Carnap's interpretation to establish the point I want to make here. The point is that there are basic similarities between Carnap's interpretation and certain forms of Bayesianism including subjective interpretation, namely: (a) taking probability as intrinsically a two-place relation between two claims, (b) the dependency on implication relationship, (c) deriving unconditional probability by assuming a logical tautology. And because of these similarities they should all be called "logical interpretations" of probability. In fact, we may say that any interpretation relying on any axiomatization of probability that satisfies the three mentioned conditions should be understood as being a logical interpretation of probability. As we saw, all such interpretations, at the axiomatic level (independent of any interpretation), are heavily dependent on the "implication relationship" which is understood to be a purely logical relationship, and this is the reason that allows us to put all these interpretations under the same umbrella. Whether or not subjective probability is a special case of Carnap's logical analysis, or whether or not it can be derived from Carnap's set of axioms, is irrelevant to the point that subjective interpretation, nevertheless, gives us a logical interpretation of probability. Therefore, if we agree that there is a basic split in the concept of probability axiomatically, then we have no choice but to accept Carnap's first point about the essential duality of the concept of probability. And we will find, firstly, that the classifications of interpretations of probability we may see in sources such as the "Stanford Encyclopedia of Philosophy" need a substantial reevaluation, and secondly, that the controversy over whether subjective or frequentist interpretations give us the complete meaning of probability, is futile.

All in all, Carnap is clear in saying that a logical interpretation via conditional probability in the form of the "degree of confirmation", in practice, would provide us the best estimation of a hypothesis, with respect to the given evidence that can be considered as a guide in life (Carnap, 1947c). And, that the assigned probability values to several hypotheses "can be interpreted as the estimate of the relative frequency of truth among them" (Carnap, 1962: 172).

3 The Proposal and Conclusion

To be clear, when we call an interpretation "non-logical" it does not mean that the interpretation in question is illogical or does not follow the rules of logic. As I mentioned in the introduction, any well-defined functor that assigns a number to a property necessarily assumes a logical structure. Therefore, in this sense all interpretations are logical. That being said, let us consider the separation criterion. According to our discussion thus far, one can say every interpretation of probability that satisfies the following three conditions should be understood as a logical interpretation. If an interpretation fails to satisfy these conditions, then it is a non-logical interpretation:

a) Probability is a two-place relation between two claims.

b) The axioms rely on the implication relation.

c) Unconditional probability is derived by assuming a tautology.

Interesting consequences follow from this. For example, Bayesianism turns out to be logical in the same way as Carnap's inductive logic. One of the advantages of this way of classifying interpretations of probability is that it would allow one to primarily and fundamentally group interpretations of probability into two basic "logical" and "non-logical" categories. This would confirm the existence of the two explicanda for probability discussed above. Mathematically speaking, this means there exists at least two non-equivalent (not mutually exclusive) ways for delivering interpretations of probability. Linguistically speaking, this means that there are two different meanings associated with the word "probability". Thus, the concept of probability can be explicated in different ways. Accordingly, one cannot say there is one single interpretation that constitutes *the* interpretation of probability. This may finally put an end to some long lasting philosophical debates such as the subjectivists-frequentists debate over the true interpretation of probability. Of course there remain important questions concerning the nature of the relation between these meanings (e. g., whether they are complimentary or not) and how we are to accommodate these interpretations into our overall theory of probability. Satisfactory answers to these questions require some technical and theoretical adjustments.

The ramification of accepting two explicanda for probability will affect other fields of study as well. One major consequence concerns the confirmation of scientific (probabilistic) theories. One might be tempted to say that such a scientific theory may undergo two different kinds of evaluations. If probability is understood to be logical, then it means we are facing two claims, derived from the same theory, and we evaluate them with respect to each other. On the other hand, in the case of non-logical understanding, we may only consider one claim and evaluate that claim with respect to an established reference class (considering mathematical features such as stability, normalization, and the like). In the second case one considers rules and elements that are not necessarily embedded (or integrated) inside the scientific theory in question. Depending whether our interpretation is understood to be logical or not, the two evaluations are not the same. In the first case the result of the evaluation is responsive to the scientific theory alone. In the second case, the result is also responsive to an auxiliary mathematical theory. The latter is fundamentally different from the scientific theory. For example, in the case where we assign a 0.5 probability to a coin toss, one may evaluate this assignment by considering the description of all the circumstances in the actual process along with the two possible state-descriptions (as seen above in the section 4). On the other hand, one may establish a large enough reference class and observe that it would stabilize at 0.5. In the second case the stability in question is derived from a mathematical theory and disregards the empirical circumstances. That is to say we arrive at the stability on the assumption that this is not limited to the process of coin tossing (identified specifically by the corresponding state-descriptions), and that any set of data that behaves in this manner would be assigned the same ratio. Therefore, although the results of both evaluations are the same in this case, the evaluations themselves do not mean the same thing. There might be cases where only one type of evaluation is possible (e. g., cases in which one cannot possibly establish a reference class or a large enough reference class), and there are cases in which both types of evaluation are possible. If the evaluation of a probability statement is taken to be logical, then what is at stake is the logical consistency of the theory in light of the embedded implication relation (upon which rests the probability). When it is construed to be non-logical, only the validity of the reference class in question (or the sampling method, base lines, etc.) is at stake. It is obvious that discussing confirmation is well beyond the scope of this paper; nevertheless, I just wanted to give the reader a sense of how this discussion might be affected given the validity of both logical and non-logical interpretations of probability.

I tried to establish, in this paper, that there are two different possibilities for an axiomatic setting with regard to probability at the *theoretical* level. The first is one in which the absolute probability is taken as primitive and conditional probability is derived as a theorem (the initial setting of Kolmogorov). The second is one is which it is the conditional probability that is taken as primitive and the absolute probability is derived (though not in the exact same way). Because of the inclusion of the logical consequence relationship in the later setting, one may meaningfully attribute the descriptor "logical" to all the interpretations based on this second setting. Hence the primary division in analysis of probability appears at the theoretical level, which is independent of any further interpretations of the systems. I believe that if we are to classify interpretations of probability, we ought to consider this fundamental split in our classification. And if we do so we will find Carnap's inductive logic along with all forms of Bayesianism on the same side of the division.

I have also tried to show that there are two levels in Carnap's philosophy of probability. At the first level, Carnap speaks of probability from a meaning-theoretic point of view, which bears strong resemblance to our discussion about the fundamental split in analyzing the concept of probability. In the second level, we deal with explicating *one* of the possible explicanda of probability. And, if we are to assign a name for this second part, we may call it "propositional interpretation of probability"[18].

The exact nature of the relationship between the meanings of complete interpretations (based on each axiomatic setting) remains an open philosophical question. Nevertheless, Carnap's answer to this question is that the meanings would be complimentary and there is no reason to adopt one at the expense of losing the other. If we agree with Carnap on this point, then, for the coherence of the theory, we ought to either adopt a principle that says the meanings of all such concepts (concepts with double explicanda) are complimentary (provided that each is based on only one of the axiomatic settings), or provide a criterion for identifying under which circumstances the meanings are complementary.

References

Carnap, R. (1939). *Foundations of Logic and Mathematics* (Vol. 1, no. 3). Chicago: University of Chicago Press.

Carnap, R. (1942). *Introduction to Semantics*. Cambridge, Mass.: Harvard University Press.

Carnap, R. (1945). "On Inductive Logic", *Philosophy of Science*, 12(2), 72–97.

Carnap, R. (1946). "Remarks on Induction and Truth", *Philosophy and Phenomenological Research*, 6(4), 590–602.

Carnap, R. (1947a). *Meaning and Necessity: A Study in Semantics and Modal Logic*. University of Chicago Press, Chicago.

Carnap, R. (1947b). "On the Application of Inductive Logic", *Philosophy and Phenomenological Research*, 8(1), 133–148.

Carnap, R. (1947c). Probability as a Guide in Life. *The Journal of Philosophy*, 44(6), 141–148.

Carnap, R. (1955). "Statistical and Inductive Probability", in: R. Carnap (Ed.), *Statistical and Inductive Probability & Inductive Logic and Science*. The Galois Institute of Mathematics and Art, Brooklyn, NY.

[18]This way, at least, one side of the debate between subjectivists and Carnapians might be reduced to a debate on the epistemic difference between propositions and beliefs.

Carnap, R. (1962). *Logical Foundations of Probability* (Vol. 2d Chicago). University of Chicago Press, Chicago.

Carnap, R. (1973). "Notes on Probability and Induction", *Synthese*, 25(3/4), 269–298.

Costa, H. A. & Parikh, R. (2005). "Conditional Probability and Defeasible Inference", *Journal of Philosophical Logic*, 34(1), 97–119.

Cowell, R. G., Dawid, P., Lauritzen, S.L., Spiegelhalter, D.J. (1999). *Probabilistic Networks and Expert Systems*. Springer, New York.

Fitelson, B., Hájek, A. & Hall, N. (2006). "Probability", in: S. R. Jessica Pfeiffer, S. Sarkar (Ed.), *The Routledge Encyclopedia of Philosophy of Science*: Routledge, New York, 599–609.

Goosens, W. (1979). "Alternative Axiomatizations of Elementary Probability Theory", *Notre Dame Journal of Formal Logic*, 20, 227–239.

Hájek, A. (2008). "Probability - a Philosophical Overview", in: R. B. Gold, R. Simons (Eds.), *Current Issues in the Philosophy of Mathematics from the Perspective of Mathematicians*: Mathematical Association of America, Washington DC, 323–339.

Hájek, A. (2009). "Fifteen Arguments against Hypothetical Frequentism", *Erkenntnis*, 70, 211–235.

Hájek, A. (2012). "Interpretations of Probability", *The Stanford Encyclopedia of Philosophy* (Winter 2012 Edition), Edward N. Zalta (ed.), http://plato.stanford.edu/archives/win2012/entries/probability-interpret/

Hajek, A. & Joyce, J. M. (2008). "Confirmation", in: S. Psillos & M. Curd (Eds.), *The Routledge Companion to the Philosophy of Science*: Routledge, New York.

Jeffreys, H. (1939). *The Theory of Probability* (3 ed.). Oxford University Press, Oxford.

Jensen, F. V. (2001). *Bayesian Networks and Decision Graphs*: Springer, New York.

Johnson, W. E. (1921). *Logic*. Cambridge University Press, New York.

Joyce, J. (2008). "Bayes' Theorem",*The Stanford Encyclopedia of Philosophy* (Fall 2008 Edition), Edward N. Zalta (ed.), http://plato.stanford.edu/archives/fall2008/entries/bayes-theorem/

Kersting, K., De Raedt, L. (2007). "Bayesian Logic Programming: Theory and Tool", in: L. T. Getoor, B. Taskar (Ed.), *Intoduction to Statistical Relational Learning*. MIT Press, Cambridge, MA, 291–318.

Keynes, J. M. (1921). *A Treatise of Probability*. Macmillan and Co., Ltd, London.

Lee, J. M. (2013). *Axiomatic Geometry*. Providence, Rhode Island: American Mathematical Society.

Leitgeb, H. (2014). "Probability in Logic", in: A. Hájek & C. Hitchcock (Eds.), *The Oxford Handbook of Probability and Philosophy*: Oxford University Press, Oxford.

Leitgeb, H. (2015). Preface to The Stability of Belief: An Essay on Rationality and Coherence. First draft published online https://www.academia.edu/11274685/Stability_of_Belief_Preface_and_Table_of_Contents_Draft, last access 28th Sep 2016.

Lepage, F. & Morgan, C. (2003). Probabilistic Canonical Models for Partial Logics. *Notre Dame Journal of Formal Logic*, 44(3), 125–138.

Lyon, A. (2009). "Philosophy of Probablilty", in: F. Allhoff (Ed.), *Philosophies of the Sciences: A Guide*, Wiley-Blackwell, UK, 92–125.

Maher, P. (2006). "The Concept of Inductive Probability", *Erkenntnis*, 65, 185–206.

Maher, P. (2010). "Explication of Inductive Probability", *Journal of Philosophical Logic*, 39(6), 593–616.

Novikov, P. S. (2001). "Axiomatic Method", in: M. Hazewinkel (Ed.), *Encyclopedia of Mathematics*: Kluwer Academic Publishers.

Pearl, J. (1991). "Probabilistic Semantics for Nonmonotonic Reasoning", in: R. C. J. Pollock & R. Cummins (Eds.), *Philosophy and Ai: Essays at the Interface*. MIT press, Cambridge, MA, 157–188.

Sarkar, S. & Pfeifer, J. (2006). *The Philosophy of Science, an Encyclopedia*. Routledge, Taylor & Francis Group, New York.

Schilpp, P. A. (1963). *The Philosophy of Rudolf Carnap* (Vol. 11). Open Court, La Salle, Illinois.

Stalnaker, R. C. (1970). "Probability and Conditionals", *Philosophy of Science*, 37(1), 64–80.

Van Fraassen, B. A. S. (1995). "Fine-Grained Opinion, Probability, and the Logic of Full Belief", *Journal of Philosophical Logic*, 24(4), 349-377.