Carnap's Logical Theory

Conor Mayo-Wilson

Philosophy of Probability
June 10th, 2013

Review: Three objective Interpretations of Probability

- Frequentist
- Propensity
- Logical - Probability is the measure of the degree to which a set of sentences support a conclusion.

Review: Three objective Interpretations of Probability

- Frequentist
- Propensity
- Logical - Probability is the measure of the degree to which a set of sentences support a conclusion.
    - Keynes' theory belongs to this last category.

Today: Carnap's Logical Theory - Another objective interpretation of probability.

Disclaimer: Today, I will present a caricature of the earlier Carnap.

- Carnap's views on probability changed during his life. In particular:
    - Although I discuss Carnap's views on **logical** probability (or what he sometimes calls "strength of evidence"), Carnap argued that logical probability was closely related to interpretations of probabilities involving (i) betting odds and (ii) frequencies. See pp. 162-175 of Carnap [1962]
    - In his later years, Carnap begins to think his definition of logical probability was not as precise, and he more-or-less adopts the subjectivist interpretation of probability.
- For more on Carnap's theory of probability in a historical setting, see Zabell [2011].
- For defenses of Carnap's view from criticisms in [Hájek, 2003], see Maher [2010]

Outline

1. Review

Outline

Outline

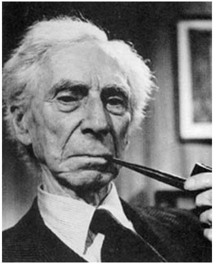
1. Review
2. Properties of Logical Probability
   - Formal Preliminaries
3. Carnap's Interpretation of Probability
   - Admissibility

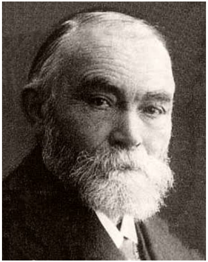
Outline

1. Review
2. Properties of Logical Probability
   - Formal Preliminaries
3. Carnap's Interpretation of Probability
   - Admissibility
4. Axiomatic Confirmation
   - Objectivity and Symmetry

Outline

Early 1900s:
Frege and Russell introduced a collection of formal, logical languages capable of expressing all of mathematics and much of natural language.





Here's the basic idea:

- Let $P(c)$ mean that the object denoted by "c" has property denoted by "P".
- Example: If P means "red" and c refers to my shirt, then $P(c)$ means that "My shirt is red."
- In analogy with natural language, the letter P is called a predicate, which is why the collection of formal languages with introduced by Frege and Russell are called **predicate logic**.

Predicate vs. Propositional Logic

- In propositional logic, different sentences in natural language are represented by completely different variables:
  - "My shirt is red" might be represented by the variable p.
  - "The apple is red" might be represented by q.

Predicate vs. Propositional Logic

- In propositional logic, different sentences in natural language are represented by completely different variables:
  - "My shirt is red" might be represented by the variable p.
  - "The apple is red" might be represented by q.
- In contrast, in predicate logic, different sentences in natural language might be represented in similar ways.
  - "My shirt is red" might be represented by $P(c)$.
  - "The apple is red" might be represented by $P(d)$.

- Like Keynes:
  - Carnap assigns probabilities to sentences in a formal language.

Probability and Logic

- Like Keynes:
  - Carnap assigns probabilities to sentences in a formal language.
  - Carnap is interested in conditional probability $c(h, e)$ of some "hypothesis" h on "evidence" e.
    - Carnap calls $c(h, e)$ a **confirmation** function.
    - Both h and e are sentences in the formal language.

- Like Keynes:
  - Carnap assigns probabilities to sentences in a formal language.
  - Carnap is interested in conditional probability $c(h, e)$ of some "hypothesis" h on "evidence" e.
    - Carnap calls $c(h, e)$ a **confirmation** function.
    - Both h and e are sentences in the formal language.
- Unlike Keynes:
  - Carnap considers sentences in predicate logic.
    - Recall, all of Keynes' axioms concern $a/h, \overline{a}/h, (a+b)/h$, etc. The sentences he considers are expressible in propositional logic.

- Like Keynes:
  - Carnap assigns probabilities to sentences in a formal language.
  - Carnap is interested in conditional probability $c(h, e)$ of some "hypothesis" h on "evidence" e.
    - Carnap calls $c(h, e)$ a **confirmation** function.
    - Both h and e are sentences in the formal language.
- Unlike Keynes:
  - Carnap considers sentences in predicate logic.
    - Recall, all of Keynes' axioms concern $a/h, \bar{a}/h, (a + b)/h$, etc. The sentences he considers are expressible in propositional logic.
  - Predicate logic allows Carnap to formulate particular symmetry principles that cannot be expressed in sentential logic.

Last week, we ended with a central question for Keynes' theory of probability.

Motivation:

- Keynes stipulates that a/h must obey particular mathematical axioms.
- These axioms entail Kolmogorov's axioms if a/h is always a real number.
- a/h is intended to represent the "strength of evidence" for a provided by h, or the degree to which h supports a.

Question: Why should the informal concept of "strength of evidence" obey Keynes' mathematical axioms for a/h?

To put it a different way:

- Keynes simply stipulates that strength of evidence satisfies probability axioms.
- But one could do the same for any interpretation of probability:
  - One could say a/h represents the propensity for a to occur given h has happened. So propensities are probabilities.
  - One could assert that degree a/h represents one's degree of belief in a given one assumes h. So degrees of belief are probabilities.
- Keynes would not accept these as good arguments. So why does a/h represent strength of evidence?

Carnap provides two types of arguments that confirmation $c(h, e)$, or "strength of evidence", satisfies the probability axioms:

- **Indirect:** In statistical settings, confirmation is equal to relative frequencies. Since said frequencies obey the probability axioms, so does confirmation.
  - E.g. e might be the evidence that 99.9% of ravens are black, and h is the hypothesis that most ravens are black.

Confirmation is a Conditional Probability

Carnap provides two types of arguments that confirmation $c(h, e)$, or "strength of evidence", satisfies the probability axioms:

- **Indirect:** In statistical settings, confirmation is equal to relative frequencies. Since said frequencies obey the probability axioms, so does confirmation.
  - E.g. e might be the evidence that 99.9% of ravens are black, and h is the hypothesis that most ravens are black.
- **Direct:** Logical facts about entailment and the syntactic structure of sentences entail that confirmation must satisfy probability axioms.

Let's look at Carnap's direct arguments.

First, why is probability bounded?

- That is, why are there maximal degrees to which a hypothesis can be confirmed or disconfirmed by evidence?
- Why is probability not like length, which is not bounded from above?

Logical Probability is Normalized

Carnap argues that strength of evidence is bounded as follows:

> *Since* $h \vee \neg h$ *is . . . true, no sentence can be more certain on any evidence. Therefore, the strength of* $h \vee \neg h$ *on* e *must have the highest possible value . . .*

Carnap [1962], Section 41A. pp. 165.

Logical Probability is Normalized

- An analogous argument, that no sentence can be less certain than a contradiction, indicates that probabilities are bounded from below.

Logical Probability is Normalized

- An analogous argument, that no sentence can be less certain than a contradiction, indicates that probabilities are bounded from below.
- If one assumes strength of evidence is represented by real numbers, then one can arbitrarily stipulate that the lower and upper bounds on strength of evidence are 0 and 1 respectively.

Are evidential strengths probabilities?

So that leaves two of Kolmogorov's axioms:

1. Probability is numerical.
2. Probability is finitely additive.

Are evidential strengths probabilities?

Carnap provides a rather unconvincing argument for the first and simply stipulates the second.

Is Logical Probability Numerical?

> *A physicist would say, for instance, that he made six experiments to test a certain law and found it confirmed in all six cases. A physician would report that he tried out a new drug in twenty cases of a certain disease and found it successful in twelve cases, not successful in five, while in three cases the result was not clear either way . . .*
> *Thus, let us assume, as most scientists do implicitly, that the concept of a confirming case can be defined . . . Then we can determine the number of confirming cases in the observational report e. If the confirming cases are of different kinds, we can determine the number of confirming cases of each kind . . .*

Carnap [1962], Section 42. pp. 224-225.

Logical Probability Numerical?

Carnap concludes:

> *We have seen that . . . it is rather plausible that [logical factors upon which confirmation depends] can be evaluated numerically.*

Carnap [1962], Section 42. pp. 225-226.

Carnap recognizes this argument is not definitive, and in the next section, he discusses five reasons why quantifying a "confirming case" is difficult.

Additivity of Evidence

What about the additivity assumption?

Like Keynes, Carnap axiomatizes properties of any acceptable "confirmation" function.

One axiom is finite additivity.

Axioms for Confirmation

Carnap's axioms for any acceptable confirmation function are:

1. If h and h' are logically equivalent, then $c(h, e) = c(h', e)$.
2. If e and e' are logically equivalent, then $c(h, e) = c(h, e')$.
3. Multiplication Principle: $c(h \wedge j, e) = c(j, e \wedge h) \cdot c(h, e)$
4. Special Addition Principle: If $e \wedge h \wedge j$ is a contradiction, then

$$c(h \vee j, e) = c(h, e) + c(j, e)$$

The first two are plausibly restrictions for any "logical" confirmation function. What about the latter two?

> *[The last two assumptions] are generally accepted in all theories of probability$_1$ (and incidentally, their analogues occur in all theories of probability$_2$) ... [They] are in accordance with what reasonable people think in terms of probability, and in particular, what they are willing to bet on particular assumptions.*

[Carnap, 1962], Section 53, pp. 285.

A Different Axiomatization of Logical Probability

Maher [2010] provides a different axiomatization of logical probability that entails the additivity axiom:

1. $p(A|B) \geq 0$
2. $p(A|B) = 1$
3. $p(A|B) + p(\neg A|B) = 1$
4. The multiplication axiom
5. The axiom that confirmation is invariant under substitution of logically equivalent sentences.

Is Strength of Evidence a Probability?

- Should we criticize Carnap for merely stipulating that evidential strengths (i.e. confirmation or inductive probability) is numerical and additive?
- Of course, ideally, it would be nice to given a direct argument that "strength of evidence" satisfies the probability axioms (and Carnap would do so if he thought it were possible).

Carnap's goal is to explicate informal concepts of probability:

- One goal of explication is to stay close to an informal concept
- But another goal is to develop a precise, useful concept. So the explication often is more precise than the informal one.
    - Compare: The mathematical definitions of continuity is quite a bit more precise than the informal concept.
    - Compare: Space, considered as a manifold in physics, is quite a bit more precise than the informal concept of space.

Explication

Carnap's goal is to explicate informal concepts of probability:

- One goal of explication is to stay close to an informal concept
- But another goal is to develop a precise, useful concept. So the explication often is more precise than the informal one.
  - Compare: The mathematical definitions of continuity is quite a bit more precise than the informal concept.
  - Compare: Space, considered as a manifold in physics, is quite a bit more precise than the informal concept of space.
- One way of making a concept more precise is to employ mathematics, but in fact, Carnap offers both a qualitative and quantitative explication of confirmation.

Objectivity

- Even if we accept Carnap's axioms of confirmation for the sake of explication, they do not determine a **unique** confirmation function.
- In fact, the axioms don't narrow the space of probability measures at all.

So in what sense is logical probability objective?

Objectivity

Carnap argues for an objective logical probability in two steps:

STEP 1: Reduce confirmation to unconditional probability.

STEP 2: Argue that unconditional probability must obey symmetry principles.

Step 1: Reduction

- Let $\top$ denote any tautology.
- Define $P(h) = c(h, \top)$.

Step 1: Reduction

$$c(h \wedge j, e) = c(j, e \wedge h) \cdot c(h, e)$$

Multiplication principle

Step 1: Reduction

$$
\begin{aligned}
c(h \wedge j, e) &= c(j, e \wedge h) \cdot c(h, e) \\
&\quad \text{Multiplication principle} \\
c(h \wedge j, \top) &= c(j, \top \wedge h) \cdot c(h, \top) \\
&\quad \text{Letting } e = \top
\end{aligned}
$$

Step 1: Reduction

$$
\begin{aligned}
c(h \wedge j, e) &= c(j, e \wedge h) \cdot c(h, e) \\
&\quad \text{Multiplication principle} \\
c(h \wedge j, \top) &= c(j, \top \wedge h) \cdot c(h, \top) \\
&\quad \text{Letting } e = \top \\
c(h \wedge j, \top) &= c(j, h) \cdot c(h, \top) \\
&\quad \text{As } \top \wedge h \leftrightarrow h
\end{aligned}
$$

Step 1: Reduction

$$\begin{aligned}
c(h \wedge j, e) &= c(j, e \wedge h) \cdot c(h, e) \\
&\quad \text{Multiplication principle} \\
c(h \wedge j, \top) &= c(j, \top \wedge h) \cdot c(h, \top) \\
&\quad \text{Letting } e = \top \\
c(h \wedge j, \top) &= c(j, h) \cdot c(h, \top) \\
&\quad \text{As } \top \wedge h \leftrightarrow h \\
P(h \wedge j) &= c(j, h) \cdot P(h) \\
&\quad \text{Definition of } P
\end{aligned}$$

Step 1: Reduction

$$
\begin{aligned}
c(h \wedge j, e) &= c(j, e \wedge h) \cdot c(h, e) \\
&\quad \text{Multiplication principle} \\
c(h \wedge j, \top) &= c(j, \top \wedge h) \cdot c(h, \top) \\
&\quad \text{Letting } e = \top \\
c(h \wedge j, \top) &= c(j, h) \cdot c(h, \top) \\
&\quad \text{As } \top \wedge h \leftrightarrow h \\
P(h \wedge j) &= c(j, h) \cdot P(h) \\
&\quad \text{Definition of } P \\
c(j, h) &= \frac{P(h \wedge j)}{P(h)} \\
&\quad \text{Rearranging}
\end{aligned}
$$

So confirmation is just the ratio definition of conditional probability!

Step 2: Use symmetry principles to define P.

State Descriptions

- Suppose there are only finitely many distinct objects are under investigation $c_1, c_2, \ldots c_n$.
  - That is, we know $(\forall x)(x = c_1 \vee x = c_2 \vee \ldots c_n)$
  - And $c_1 \neq c_2 \wedge c_1 \neq c_3 \wedge \ldots c_{n-1} \neq c_n$.

State Descriptions

- Suppose there are only finitely many distinct objects are under investigation $c_1, c_2, \ldots c_n$.
  - That is, we know $(\forall x)(x = c_1 \vee x = c_2 \vee \ldots c_n)$
  - And $c_1 \neq c_2 \wedge c_1 \neq c_3 \wedge \ldots c_{n-1} \neq c_n$.
- Similarly, there are only finitely many predicates $Q_1, Q_2, \ldots Q_m$ that can hold of these objects.
  - Note: Some predicates may be relations.
  - E.g. $Q_1(c_1, c_2)$ may mean c_1 is taller than c_2.

State Descriptions

- By these assumptions, all sentences in the language are equivalent to disjunctions of state descriptions, which tell us, for any predicate Q_i, which sequences of objects it describes.
  - More on state descriptions in a second.

State Descriptions

- By these assumptions, all sentences in the language are equivalent to disjunctions of state descriptions, which tell us, for any predicate Q_i, which sequences of objects it describes.
  - More on state descriptions in a second.
- So if one defines a probability P on all state descriptions, the additivity axiom determines the probabiity of disjunctions and hence, of all sentences in the language.

State Descriptions

- By these assumptions, all sentences in the language are equivalent to disjunctions of state descriptions, which tell us, for any predicate Q_i, which sequences of objects it describes.
  - More on state descriptions in a second.
- So if one defines a probability P on all state descriptions, the additivity axiom determines the probabiity of disjunctions and hence, of all sentences in the language.
- Since $c(\cdot, \cdot)$ is defined in terms of P, this determines a confirmation function.

State Descriptions

Here's a concrete example.

- Suppose there are two distinct objects are under investigation c_1 and c_2.

Here's a concrete example.

- Suppose there are two distinct objects are under investigation c_1 and c_2.
- Suppose there are two predicates Q_1 and Q_2, which are both unary.

State Descriptions

Here's a concrete example.

- Suppose there are two distinct objects are under investigation c_1 and c_2.
- Suppose there are two predicates Q_1 and Q_2, which are both unary.
- Then there are sixteen state descriptions, which you can think of as descriptions of possible worlds:
  - $Q_1(c_1) \wedge Q_1(c_2) \wedge Q_2(c_1) \wedge Q_2(c_2)$
  - $\neg Q_1(c_1) \wedge Q_1(c_2) \wedge Q_2(c_1) \wedge Q_2(c_2)$
  - $Q_1(c_1) \wedge \neg Q_1(c_2) \wedge Q_2(c_1) \wedge Q_2(c_2)$
  - $\vdots$
  - $\neg Q_1(c_1) \wedge \neg Q_1(c_2) \wedge \neg Q_2(c_1) \wedge \neg Q_2(c_2)$
  - Here, each state description is just a binary string. There are sixteen of length four.

- **Carnap's Intuition**: If we don't know any relevant facts about c_1 and c_2 that distinguish the objects other than their names, then we shouldn't draw any different inferences about them.

- **Carnap's Intuition**: If we don't know any relevant facts about c_1 and c_2 that distinguish the objects other than their names, then we shouldn't draw any different inferences about them.
- **Example:** Suppose I tell you that Jim and John are men between the ages of 25-40. Would you estimate their probabilities of heart disease differently?

> *Suppose X has found by observation that individuals a and b are P; the individuals may be physical objects and P may be an observable property. Let e be the sentence expressing these results: "$P(a) \wedge P(b)$." X considers two hypotheses h and h'; h is the prediction that another object is likewise P ('$P(c)$') and h' says the same for still another object d ('$P(d)$') . . . We should find it entirely implausible if he were to ascribe different values [to $c(h, e)$ and $c(h', e)$]. The reason is that the logical relation between e and h is just the same as that between e and h'.*

Carnap [1962], Section 90. pp. 484.

State Descriptions

Formally: Isomorphic state descriptions should have identical probabilities under P.

- Two Isomorphic State Descriptions: Just switch c_1 and c_2's names.
  - $\neg Q_1(c_1) \wedge Q_1(c_2) \wedge Q_2(c_1) \wedge Q_2(c_2)$
  - $Q_1(c_1) \wedge \neg Q_1(c_2) \wedge Q_2(c_1) \wedge Q_2(c_2)$

State Descriptions

Formally: Isomorphic state descriptions should have identical probabilities under P.

- Two Isomorphic State Descriptions: Just switch c_1 and c_2's names.
  - $\neg Q_1(c_1) \wedge Q_1(c_2) \wedge Q_2(c_1) \wedge Q_2(c_2)$
  - $Q_1(c_1) \wedge \neg Q_1(c_2) \wedge Q_2(c_1) \wedge Q_2(c_2)$
- Two Non-Isomorphic State Descriptions:
  - $\neg Q_1(c_1) \wedge Q_1(c_2) \wedge \neg Q_2(c_1) \wedge Q_2(c_2)$
  - $Q_1(c_1) \wedge \neg Q_1(c_2) \wedge Q_2(c_1) \wedge Q_2(c_2)$

Still not a snowflake . . .

This still isn't enough to determine a unique probability distribution P, but it's a significant constraint.

Carnap's last stipulation requirement for P is much like the principle of indifference.

Structure Descriptions

- Define a structure description to be a set of isomorphic state descriptions.

Structure Descriptions

- Define a structure description to be a set of isomorphic state descriptions.
- Example: A structure description if there are two objects c_1 and c_2, and two unary predicates Q_1, Q_2, then one structure description consists of two sentences:
  - $\neg Q_1(c_1) \wedge Q_1(c_2) \wedge Q_2(c_1) \wedge Q_2(c_2)$
  - $Q_1(c_1) \wedge \neg Q_1(c_2) \wedge Q_2(c_1) \wedge Q_2(c_2)$

Structure Descriptions

- Define a structure description to be a set of isomorphic state descriptions.
- Example: A structure description if there are two objects c_1 and c_2, and two unary predicates Q_1, Q_2, then one structure description consists of two sentences:
  - $\neg Q_1(c_1) \wedge Q_1(c_2) \wedge Q_2(c_1) \wedge Q_2(c_2)$
  - $Q_1(c_1) \wedge \neg Q_1(c_2) \wedge Q_2(c_1) \wedge Q_2(c_2)$
- Carnap claims that all structure descriptions ought to have the same probability, even if different structure descriptions contain different numbers of state descriptions.

- Carnap originally argued that all structure descriptions ought to have the same probability, even if different structure descriptions contain different numbers of state descriptions.

- Carnap originally argued that all structure descriptions ought to have the same probability, even if different structure descriptions contain different numbers of state descriptions.
- He later dropped this claim, as it's not clear why this a feature of **logical** probability.

Discussion: How does Carnap's view fare with respect to the Salmon's criteria?

References I

Carnap, R. (1962). *Logical Foundations of Probability*. Second edition.

Hájek, A. (2003). Interpretations of probability.

Maher, P. (2010). Explication of inductive probability. *Journal of philosophical logic*, 39(6):593616.

Zabell, S. L. (2011). Carnap and the logic of inductive inference. *Inductive Logic*, 10:265.