

Confirmation and Induction

The term “confirmation” is used in epistemology and the philosophy of science whenever observational data and evidence “speak in favor of” or support scientific theories and everyday hypotheses. Historically, confirmation has been closely related to the problem of induction, the question of what to believe regarding the future in the face of knowledge that is restricted to the past and present. One relation between confirmation and inductive logic is that the conclusion H of an [inductively strong argument](#) with premise E is confirmed by E . If inductive strength comes in degrees and the inductive strength of the argument with premise E and conclusion H is equal to r , then the degree of confirmation of H by E is likewise said to be equal to r .

This article begins by briefly reviewing [Hume](#)’s formulation of the problem of the justification of induction. Then we jump to the middle of the twentieth century and [Hempel](#)’s pioneering work on confirmation. After looking at Popper’s falsificationism and the hypothetico-deductive method of hypotheses testing, the notion of probability, as it was defined by Kolmogorov, is introduced. Probability theory is the main mathematical tool for [Carnap](#)’s inductive logic as well as for Bayesian confirmation theory. Carnap’s inductive logic is based on a logical interpretation of probability, which will be discussed at some length. However, his heroic efforts to construct a logical probability measure in purely syntactical terms can be considered to have failed. Goodman’s new riddle of induction will serve to illustrate the shortcomings of such a purely syntactical approach to confirmation. Carnap’s work is nevertheless important because today’s most popular theory of confirmation – Bayesian confirmation theory – is to a great extent the result of replacing Carnap’s logical interpretation of probability with a subjective interpretation as degree of belief *qua* fair betting ratio. The rest of the article will be concerned mainly with Bayesian confirmation theory, although the final section will mention some alternative views on confirmation and induction.

[Table of Contents](#) (Clicking on the links below will take you to those parts of this article)

- [1. Introduction: Confirmation and Induction](#)
- [2. Hempel and the Logic of Confirmation](#)
 - [a. The Ravens Paradox](#)
 - [b. The Logic of Confirmation](#)
- [3. Popper’s Falsificationism and Hypothetico-Deductive Confirmation](#)
 - [a. Popper’s Falsificationism](#)
 - [b. Hypothetico-Deductive Confirmation](#)
- [4. Inductive Logic](#)
 - [a. Kolmogorov’s Axiomatization](#)
 - [b. Logical Probability and Degree of Confirmation](#)
 - [c. Absolute and Incremental Confirmation](#)
 - [d. Carnap’s Analysis of Hempel’s Conditions](#)
- [5. The New Riddle of Induction and the Demise of the Syntactic Approach](#)
- [6. Bayesian Confirmation Theory](#)
 - [a. Subjective Probability and the Dutch Book Argument](#)
 - [b. Confirmation Measures](#)
 - [c. Some Success Stories](#)
- [7. Taking Stock](#)

1. Introduction: Confirmation and Induction

Whenever observational data and evidence speak in favor of, or support, scientific theories or everyday hypotheses, the latter are said to be *confirmed* by the former. The positive result of a pregnancy test speaks in favor of or confirms the hypothesis that the tested woman is pregnant. The dark clouds on the sky support or confirm the hypothesis that it will be raining.

Confirmation takes a qualitative and a quantitative form. Qualitative confirmation is usually construed as a relation, among other things, between three sentences or propositions: evidence E confirms hypothesis H relative to background information B . Quantitative confirmation is, among other things, a relation between evidence E , hypothesis H , background information B , and a number r : E confirms H relative to B to degree r . (Comparative confirmation – H_1 is more confirmed by E_1 relative to B_1 than H_2 by E_2 relative to B_2 – is usually derived from a quantitative notion of confirmation, and is not discussed in this entry.)

Historically, confirmation has been closely related to the problem of induction, the question of what to believe regarding the future in the face of knowledge that is restricted to the past and present. [David Hume](#) gives the classic formulation of the problem of the justification of induction in *A Treatise of Human Nature*:

Let men be once fully persuaded of these two principles, *that there is nothing in any object, consider'd in itself, which can afford us a reason for drawing a conclusion beyond it; and, that even after the observation of the frequent or constant conjunction of objects, we have no reason to draw any inference concerning any object beyond those of which we have had experience;* (Hume 1739/2000, book 1, part 3, section 12)

The reason is that any such inference beyond those objects of which we had experience needs to be justified – and, according to Hume, this is not possible.

In order to justify induction one has to provide a deductively valid or an inductively strong argument to the effect that our inductively strong arguments will continue to lead us to true conclusions (most of the time) in the future. (An argument consists of a set of premises P_1, \dots, P_n and a conclusion C . Such an argument is deductively valid just in case the truth of the premises guarantees the truth of the conclusion. There is no standard definition of an inductively strong argument, but the idea is that the premises speak in favor of or support the conclusion.) But there is no deductively valid argument whose premises are restricted to the past and present and whose conclusion is about the future – and all our knowledge is about the past and present. On the other hand, an inductively strong argument presumably has to be inductively strong in the very sense of our inductive practices – and thus begs the question. For more see the introductory Skyrms (2000), the intermediate Hacking (2001), and the advanced Howson (2000a).

Neglecting the background information B , as we will mostly do in the following, we can state the link between induction and confirmation as follows. The conclusion H of an inductively strong argument with premise E is confirmed by E . If r quantifies the strength of the inductive argument in question, the degree of confirmation of H by E is equal to r . Let us then start the discussion of

confirmation by the first serious attempts to define the notion, and to develop a corresponding logic of confirmation.

2. Hempel and the Logic of Confirmation

a. The Ravens Paradox

According to the *Nicod criterion* of confirmation (Hempel 1945), universal generalizations of the form “All *F*s are *G*s,” in symbols $\Box x(Fx \rightarrow Gx)$, are confirmed by their “instances” “This particular object *a* is both *F* and *G*,” $Fa \wedge Ga$. (It would be more appropriate to call $Fa \rightarrow Ga$ rather than $Fa \wedge Ga$ an instance of $\Box x(Fx \rightarrow Gx)$.) The universal generalization “All ravens are black” is thus said to be confirmed by its instance “*a* is a black raven.” As “*a* is a non-black non-raven” is an instance of “All non-black things are non-ravens,” the Nicod criterion says that “*a* is a non-black non-raven” confirms “All non-black things are non-ravens.” (It is sometimes said that a black raven confirms the ravens hypothesis “All ravens are black.” In this case, confirmation is a relation between a non-linguistic entity – viz. a black raven – and a hypothesis. I decided to construe confirmation as a relation between, among other things, evidential propositions and hypotheses, and so we have to state the above in a clumsier way.)

One of Hempel’s conditions of adequacy for any relation of confirmation is the *equivalence condition*. It says that logically equivalent hypotheses are confirmed by the same evidential propositions. “All ravens are black” is logically equivalent to “All non-black things are non-ravens.” Therefore a non-black non-raven like a white shoe or a red herring can be used to confirm the ravens-hypothesis “All ravens are black.” Surely, this is absurd – and this is known as the ravens paradox.

Even worse, “All ravens are black,” $\Box x(Rx \rightarrow Bx)$, is logically equivalent to “All things that are green or not green are not ravens or black,” $\Box x(Gx \vee \neg Gx \rightarrow \neg Rx \vee Bx)$. “*a* is green or not green, and *a* is not raven or black” is an instance of this hypothesis. Furthermore, it is logically equivalent to “*a* is not a raven or *a* is black.” As everything is green or not green, we get the similarly paradoxical result that an object which is not a raven or which is black – anything but a non-black raven which could be used to falsify the ravens hypothesis is such an object – can be used to confirm the ravens hypothesis that all ravens are black.

Hempel (1945), who discussed these cases of the ravens, concluded that non-black non-ravens (as well as any other object that is not a raven or black) can indeed be used to confirm the ravens hypothesis. He attributed the paradoxical character of this alleged paradox to the psychological fact that we assume there to be far more non-black objects than ravens. However, the notion of confirmation he was explicating was supposed to presuppose no background knowledge whatsoever. An example by Good (1967) shows that such an unrelativized notion of confirmation is not useful (see Hempel 1967, Good 1968).

Others have been led to the rejection of the Nicod criterion. Howson (2000b, 113) considers the hypothesis “Everybody in the room leaves with somebody else’s hat,” which he attributes to Rosenkrantz (1981). If the background contains the information that there are only three

individuals a, b, c in the room, then the evidence consisting of the two instances “ a leaves with b ’s hat” and “ b leaves with a ’s hat” falsifies rather than confirms the hypothesis. Besides pointing to the role played by the background information in this example, Hempel would presumably have stressed that the Nicod criterion has to be restricted to universal generalization in one variable only. Already in his (1945, 13: fn. 1) he notes that $R(a, b) \wedge \neg R(a, b)$ falsifies $\Box x \Box y (\neg [R(x, y) \wedge R(y, x)] \rightarrow R(x, y) \wedge \neg R(x, y))$, which is equivalent to $\Box x \Box y R(x, y)$, although it satisfies both the antecedent and the consequent of the universal generalization (cf. also Carnap 1950/1962, 469f).

b. The Logic of Confirmation

After discussing the ravens, Hempel (1945) considers the following conditions of adequacy for any relation of confirmation:

1. Entailment Condition: If an evidential proposition E logically implies some hypothesis H , then E confirms H .
2. Special Consequence Condition: If an evidential proposition E confirms some hypothesis H , and if H logically implies some hypothesis H' , then E also confirms H' .
3. Special Consistency Condition: If an evidential proposition E confirms some hypothesis H , and if H is not compatible with some hypothesis H' , then E does not confirm H' .
4. Converse Consequence Condition: If an evidential proposition E confirms some hypothesis H , and if H is logically implied by some hypothesis H' , then E also confirms H' .

(The equivalence condition mentioned above follows from 2 as well as from 4). Hempel then shows that any relation of confirmation satisfying 1, 2, and 4 is trivial in the sense that every evidential proposition E confirms every hypothesis H . This is easily seen as follows. As E logically implies itself, E confirms E according to the entailment condition. The conjunction of E and H , $E \wedge H$, logically implies E , and so the converse consequence condition entails that E confirms $E \wedge H$. But $E \wedge H$ logically implies H ; thus E confirms H by the special consequence condition. In fact, it suffices that confirmation satisfies 1 and 4 in order to be trivial: E logically implies and, by 1, confirms the disjunction of E and H , $E \vee H$. As H logically implies $E \vee H$, E confirms H by 4.

Hempel (1945) rejects the converse consequence condition as the culprit rendering trivial any relation of confirmation satisfying 1-4. The latter condition has nevertheless gained popularity in the philosophy of science – partly because it seems to be at the core of the account of confirmation we will discuss next.

3. Popper’s Falsificationism and Hypothetico-Deductive Confirmation

a. Popper’s Falsificationism

Although Popper was an opponent of any kind of induction, his falsificationism gave rise to a

qualitative account of confirmation. Popper started by observing that many scientific hypotheses have the form of universal generalizations, say “All metals conduct electricity.” Now there can be no amount of observational data that would *verify* a universal generalization. After all, the next piece of metal could be such that it does not conduct electricity. In order to verify this hypothesis we would have to investigate all pieces of metal there are – and even if there were only finitely many such pieces, we would never know this (unless there were only finitely many space-time regions we would have to search). However, Popper’s basic insight is that these universal generalization can easily be *falsified*. We only need to find a piece of metal that does not conduct electricity in order to know that our hypothesis is false (supposing we can check this). Popper then generalized this. He suggested that all science should put forth bold hypotheses, which are then severely tested (where bold means to have a high degree of falsifiability, in other words, to have many observational consequences). As long as these hypotheses survive their tests, scientists should stick to them. However, once they are falsified, they should be put aside if there are competing hypotheses that remain unfalsified.

This is not the place to list the numerous problems of Popper’s falsificationism. Suffice it to say that there are many scientific hypotheses that are neither verifiable nor falsifiable (for example, “Each planet has a moon”), and that falsifying instances are often taken to be indicators of errors that lie elsewhere, say errors of measurement or errors in auxiliary hypotheses. As Duhem and Quine noted, confirmation is holistic in the sense that it is always a whole battery of hypotheses that is put to test, and the arrow of error usually does not point to a single hypothesis (Duhem 1906/1974, Quine 1953).

According to Popper’s falsificationism (see Popper 1935/1994) the hallmark of *scientific* (rather than meaningful, as in the early days of logical positivism) hypotheses is that they are falsifiable: scientific hypotheses must have consequences whose truth or falsity can in principle (and with a grain of salt) be ascertained by observation (with a grain of salt, because for Popper there is always an element of conventionalism in stipulating the basis of science). If there are no conditions under which a given hypothesis is false, this hypothesis is not scientific (though it may very well be meaningful).

b. Hypothetico-Deductive Confirmation

The *hypothetico-deductive* notion of confirmation says that an evidential proposition E confirms a hypothesis H relative to background information B if and only if the conjunction of H and B , $H \wedge B$, logically implies E in some suitable way (which depends on the particular version of hypothetico-deductivism under consideration). The intuition here is that scientific hypotheses are tested; and if a hypothesis H survives a severe test, then, intuitively, this is evidence in favor of H . Furthermore, scientific hypothesis are often used for predictions. If a hypothesis H correctly predicts some experimental outcome E by logically implying it, then, intuitively, this is again evidence for the truth of H . Both of these related aspects are covered by the above definition, if surviving a test is tantamount to entailing the correct outcome.

Note that hypothetico-deductive confirmation – henceforth HD-confirmation – satisfies Hempel’s converse consequence condition. Suppose an evidential proposition E HD-confirms some hypothesis H . This means that H logically implies E in some suitable way. Now any hypothesis H' which logically implies H also logically implies E . But this means – at least under most conditions fixing the “suitable way” of entailment – that E HD-confirms H .

Hypothetico-deductivism has run into serious difficulties. To mention just two, there is the problem of irrelevant conjunctions and the problem of irrelevant disjunctions. Suppose an evidential proposition E HD-confirms some hypothesis H . Then, by the converse consequence condition, E also HD-confirms $H \wedge H'$, for any hypothesis H' whatsoever. Assuming that the anomalous perihelion of Mercury confirms the general theory of relativity GTR (Earman 1992), it also confirms the conjunction of GTR and, say, that there is life on Mars – which seems to be wrong. Similarly, if E HD-confirms H , then $E \vee E'$ HD-confirms H , for any evidential proposition E' whatsoever. For instance, the disjunctive proposition of the anomalous perihelion of Mercury or Luca's living on the second floor HD-confirms GTR (Grimes 1990, Moretti 2004).

Another worry with HD-confirmation is that it is not clear how it should be applied to statistical hypotheses that do not strictly entail anything (see, however, Albert 1992). The treatment of statistical hypotheses is no problem for probabilistic theories of confirmation, which we will turn to now.

4. Inductive Logic

For overview articles see Fitelson (2005) and Hawthorne (2005).

a. Kolmogorov's Axiomatization

Before we turn to inductive logic, let us define the notion of probability as it was axiomatized by Kolmogorov (1933; 1956).

Let W be a non-empty set (of outcomes or possibilities), and let \mathbf{A} be a field over W , that is, a set of subsets of W that contains the whole set W and is closed under complementation (with respect to W) and finite unions. That is, \mathbf{A} is a field over W if and only if \mathbf{A} is a set of subsets of W such that

- (i) $W \in \mathbf{A}$
- (ii) if $A \in \mathbf{A}$, then $(W \setminus A) = -A \in \mathbf{A}$
- (iii) if $A \in \mathbf{A}$ and $B \in \mathbf{A}$, then $(A \cup B) \in \mathbf{A}$

If (iii) is strengthened to

- (iv) if $A_1 \in \mathbf{A}, \dots, A_n \in \mathbf{A}, \dots$, then $(A_1 \cup \dots \cup A_n \cup \dots) \in \mathbf{A}$,

so that \mathbf{A} is closed under countable (and not only finite) unions, \mathbf{A} is called a σ -field over W .

A function $Pr: \mathbf{A} \rightarrow \mathfrak{R}$ from the field \mathbf{A} over W into the real numbers \mathfrak{R} is a (*finitely additive*) *probability measure* on \mathbf{A} if and only if it is a non-negative, normalized, and (finitely) additive measure; that is, if and only if for all $A, B \in \mathbf{A}$

$$(K1) Pr(A) \geq 0$$

$$(K2) Pr(W) = 1$$

(K3) if $A \cap B = \emptyset$, then $Pr(A \cup B) = Pr(A) + Pr(B)$

The triple $\langle W, \mathbf{A}, Pr \rangle$ with W a non-empty set, \mathbf{A} a field over W , and Pr a probability measure on \mathbf{A} is called a (finitely additive) probability space. If \mathbf{A} is a σ -field over W and $Pr: \mathbf{A} \rightarrow \mathbb{R}$ additionally satisfies

(K4) if $A_1 \supseteq A_2 \supseteq \dots \supseteq A_n \dots$ is a decreasing sequence of elements of \mathbf{A} , i.e. $A_1 \in \mathbf{A}, \dots, A_n \in \mathbf{A}, \dots$, such that $A_1 \cap A_2 \cap \dots \cap A_n \cap \dots = \emptyset$, then $\lim_{n \rightarrow \infty} Pr(A_n) = 0$,

Pr is a σ -additive probability measure on \mathbf{A} and $\langle W, \mathbf{A}, Pr \rangle$ is a σ -additive probability space (Kolmogorov 1933; 1956, ch. 2). (K4) asserts that

$$\lim_{n \rightarrow \infty} Pr(A_n) = Pr(A_1 \cap A_2 \cap \dots \cap A_n \cap \dots) = Pr(\emptyset) = 0$$

for a decreasing sequence of elements of \mathbf{A} . Given (K1-3), (K4) is equivalent to

(K5) if $A_1 \in \mathbf{A}, \dots, A_n \in \mathbf{A}, \dots$, and if $A_i \cap A_j = \emptyset$ for all natural numbers i, j with $i \neq j$, then $Pr(A_1 \cup \dots \cup A_n \cup \dots) = Pr(A_1) + \dots + Pr(A_n) + \dots$

A probability measure $Pr: \mathbf{A} \rightarrow \mathbb{R}$ on \mathbf{A} is *regular* just in case $Pr(A) > 0$ for every non-empty $A \in \mathbf{A}$. Let $\langle W, \mathbf{A}, Pr \rangle$ be a probability space, and define \mathbf{A}^* to be the set of all $A \in \mathbf{A}$ that have positive probability according to Pr , that is, $\mathbf{A}^* = \{A \in \mathbf{A}: Pr(A) > 0\}$. The *conditional* probability measure

$Pr(\cdot | \cdot): \mathbf{A} \times \mathbf{A}^* \rightarrow \mathbb{R}$ on \mathbf{A} (based on the unconditional probability measure Pr) is defined for all $A \in \mathbf{A}$ and $B \in \mathbf{A}^*$ by the fraction

$$(K6) Pr(A|B) = Pr(A \cap B) / Pr(B)$$

(Kolmogorov 1933; 1956, ch. 1, §4). The domain of the second argument place of $Pr(\cdot | \cdot)$ has to be restricted to \mathbf{A}^* , since the fraction $Pr(A \cap B) / Pr(B)$ is not defined when $Pr(B) = 0$. Note that $Pr(\cdot | B): \mathbf{A} \rightarrow \mathbb{R}$ is a probability measure on \mathbf{A} , for every $B \in \mathbf{A}^*$.

Here are some immediate consequences of the Kolmogorov axioms and the definition of conditional probability. For every probability space $\langle W, \mathbf{A}, Pr \rangle$ and all $A, B \in \mathbf{A}$,

$$\text{Law of Negation: } Pr(\neg A) = 1 - Pr(A)$$

$$\text{Law of Conjunction: } Pr(A \cap B) = Pr(B) \cdot Pr(A|B) \text{ whenever } Pr(B) > 0$$

$$\text{Law of Disjunction: } Pr(A \cup B) = Pr(A) + Pr(B) - Pr(A \cap B)$$

$$\text{Law of Total Probability: } Pr(B) = \sum_i Pr(B|A_i) \cdot Pr(A_i),$$

where the A_i form a countable partition of W , i.e. A_1, \dots, A_n, \dots is a sequence of mutually exclusive

($A_i \cap A_j = \emptyset$ for all i, j with $i \neq j$) and jointly exhaustive ($A_1 \cup \dots \cup A_n \cup \dots = W$) elements of \mathbf{A} . A special case of the Law of Total Probability is

$$Pr(B) = Pr(B|A) \cdot Pr(A) + Pr(B|-A) \cdot Pr(-A).$$

Finally the definition of conditional probability is easily turned into

$$\text{Bayes's Theorem: } Pr(A|B) = Pr(B|A) \cdot Pr(A) / Pr(B)$$

$$= Pr(B|A) \cdot Pr(A) / [Pr(B|A) \cdot Pr(A) + Pr(B|-A) \cdot Pr(-A)]$$

$$= Pr(B|A) \cdot Pr(A) / \sum_i Pr(B|A_i) \cdot Pr(A_i),$$

where the A_i form a countable partition of W . The important role played by Bayes's Theorem (in combination with some principle linking objective chances and subjective probabilities) for confirmation will be discussed below. For more on Bayes's Theorem see Joyce (2003).

The names of the first three laws above already indicate that probability measures can also be defined on formal languages. Instead of defining probability on a field \mathbf{A} over some non-empty set W , we can take its domain to be a formal language \mathbf{L} , that is, a set of (possibly open) well-formed formulas that contains the tautological sentence \top (corresponding to the whole set W) and is closed under negation \neg (corresponding to complementation) and disjunction \vee (corresponding to finite union). That is, \mathbf{L} is a language if and only if \mathbf{L} is a set of well-formed formulas such that

(i) $\top \in \mathbf{L}$

(ii) if $\alpha \in \mathbf{L}$, then $\neg\alpha \in \mathbf{L}$

(iii) if $\alpha \in \mathbf{L}$ and $\beta \in \mathbf{L}$, then $(\alpha \vee \beta) \in \mathbf{L}$

If \mathbf{L} additionally satisfies

(iv) if $\alpha \in \mathbf{L}$, then $\exists x\alpha \in \mathbf{L}$,

\mathbf{L} is called a *quantificational* language.

A function $Pr: \mathbf{L} \rightarrow \mathfrak{R}$ from the language \mathbf{L} into the reals \mathfrak{R} is a *probability* on \mathbf{L} if and only if for all $\alpha, \beta \in \mathbf{L}$,

(L0) $Pr(\alpha) = Pr(\beta)$ if α is logically equivalent (in the sense of classical logic CL) to β

(L1) $Pr(\alpha) \geq 0$,

$$(L2) Pr(\tau) = 1,$$

$$(L3) Pr(\alpha \vee \beta) = Pr(\alpha) + Pr(\beta), \text{ if } \alpha \wedge \beta \text{ is logically inconsistent (in the sense of CL).}$$

(L0) is not necessary, if (L2) is strengthened to: (L2⁺) $Pr(\alpha) = 1$, if α is logically valid. If \mathbf{L} is a quantificational language with an individual constant “ a_i ” for each individual a_i in the envisioned countable domain, $i = 1, 2, \dots, n, \dots$, and $Pr: \mathbf{L} \rightarrow \mathfrak{R}$ additionally satisfies

$$(L4) \lim_{n \rightarrow \infty} Pr(\alpha[a_1/x] \wedge \dots \wedge \beta[a_n/x]) = Pr(\Box x \alpha),$$

Pr is called a *Gaifman-Snir* probability. Here “ $\alpha[a_i/x]$ ” results from “ $\alpha[x]$ ” by substituting the individual constant “ a_i ” for all occurrences of the individual variable “ x ” in “ α .” “ x ” in “ $\alpha[x]$ ” indicates that “ x ” occurs free in “ α ,” that is to say, “ x ” is not bound in “ α ” by a quantifier like it is in “ $\Box x \alpha$.”

Given (L0-3) and the restriction to countable domains, (L4) is equivalent to

$$(L5) \lim_{n \rightarrow \infty} Pr(\alpha[a_1/x] \vee \dots \vee \alpha[a_n/x]) = \sup \{Pr(\alpha[a_1/x] \vee \dots \vee \alpha[a_n/x]): n \in \mathbf{N}\} = Pr(\exists x \alpha),$$

where the equation on the right-hand side is the slightly more general definition adopted by Gaifman & Snir (1982, 501). A probability $Pr: \mathbf{L} \rightarrow \mathfrak{R}$ on \mathbf{L} is *regular* just in case $Pr(\alpha) > 0$ for every consistent $\alpha \in \mathbf{L}$. For $\mathbf{L}^* = \{\alpha \in \mathbf{L}: Pr(\alpha) > 0\}$ the conditional probability $Pr(\cdot | \cdot): \mathbf{L} \times \mathbf{L}^* \rightarrow \mathfrak{R}$ on \mathbf{L} (based on Pr) is defined for all $\alpha \in \mathbf{L}$ and all $\beta \in \mathbf{L}^*$ by the fraction

$$(L6) Pr(\alpha|\beta) = Pr(\alpha \wedge \beta)/Pr(\beta).$$

As before, $Pr(\cdot | \beta): \mathbf{L} \rightarrow \mathfrak{R}$ is a probability on \mathbf{L} , for every $\beta \in \mathbf{L}$.

Each probability Pr on a language \mathbf{L} induces a probability space $\langle W, \mathbf{A}, Pr^* \rangle$ with W the set *Mod* of all models for \mathbf{L} , \mathbf{A} the smallest σ -field containing the field $\{Mod(\alpha) \subseteq Mod: \alpha \in \mathbf{L}\}$, and Pr^* the unique σ -additive probability measure on \mathbf{A} such that $Pr^*(Mod(\alpha)) = Pr(\alpha)$ for all $\alpha \in \mathbf{L}$. (A model for a language \mathbf{L} with an individual constant for each individual in the envisioned domain can be represented by a function $w: \mathbf{L} \rightarrow \{0,1\}$ from \mathbf{L} into the set $\{0,1\}$ such that for all $\alpha, \beta \in \mathbf{L}$: $w(\neg\alpha) = 1 - w(\alpha)$, $w(\alpha \vee \beta) = \max\{w(\alpha), w(\beta)\}$, and $w(\exists x \alpha) = \max\{w(\alpha[a/x]): \text{“}a\text{” is an individual constant of } \mathbf{L}\}$.)

In conclusion, it is to be noted that some authors take conditional probability $Pr(\cdot \text{ given } -)$ as

primitive and define probability as $Pr(\cdot \text{ given } W)$ or $Pr(\cdot \text{ given } \tau)$ (see Hájek 2003b). For more on probability and its interpretations see Hájek (2003a), Hájek & Hall (2000), Fitelson & Hájek & Hall (2005).

b. Logical Probability and Degree of Confirmation

There has always been a close connection between probability and induction. Probability was thought to provide the basis for an inductive logic. Early proponents of a logical conception of probability include Keynes (1921/1973) and Jeffreys (1939/1967). However, by far the biggest effort to construct an inductive logic was undertaken by Carnap in his *Logical Foundations of Probability* (1950/1962). Carnap starts from a simple formal language with countably many individual constants (such as “Carl Gustav Hempel”) denoting individuals (viz. Carl Gustav Hempel) and finitely many monadic predicates (such as “is a great philosopher of science”) denoting properties (viz. being a great philosopher of science), but not relations (such as being a better philosopher of science than). Then he defines a state-description to be a complete description of each individual with respect to all the predicates. For instance, if the language contains three individual constants “ a ,” “ b ,” and “ c ” (denoting the individuals a , b , and c , respectively), and four monadic predicates “ P ,” “ Q ,” “ R ,” and “ S ” (denoting the properties P , Q , R , and S , respectively), then there are $2^3 \cdot 4$ state descriptions of the form:

$$\begin{aligned} &\pm Pa \wedge \pm Qa \wedge \pm Ra \wedge \pm Sa \wedge \pm Pb \wedge \pm Qb \wedge \pm Rb \wedge \\ &\pm Sb \wedge \pm Pc \wedge \pm Qc \wedge \pm Rc \wedge \pm Sc, \end{aligned}$$

where “ \pm ” indicates that the predicate in question is either unnegated as in “ Pa ” or negated as in “ $\neg Pa$.” That is, a state description determines for each individual constant “ a ” and each predicate “ P ” whether or not Pa . Based on the notion of a state description, Carnap then introduces the notion of a structure description, a maximal disjunction of state descriptions which can be obtained from each other by uniformly substituting individual constants for each other. In the above example there are, among others, the following structure descriptions:

$$(Pa \wedge Qa \wedge Ra \wedge Sa) \wedge (Pb \wedge Qb \wedge Rb \wedge Sb) \wedge (Pc \wedge Qc \wedge Rc \wedge Sc)$$

$$\begin{aligned} &((Pa \wedge Qa \wedge Ra \wedge Sa) \wedge (Pb \wedge Qb \wedge Rb \wedge \neg Sb) \wedge (Pc \wedge Qc \wedge \neg Rc \wedge Sc)) \vee ((Pb \wedge Qb \wedge Rb \wedge Sb) \wedge (Pa \wedge Qa \wedge \\ &Ra \wedge \neg Sa) \wedge (Pc \wedge Qc \wedge \neg Rc \wedge Sc)) \vee ((Pc \wedge Qc \wedge Rc \wedge Sc) \wedge (Pb \wedge Qb \wedge Rb \wedge \neg Sb) \wedge (Pa \wedge Qa \wedge \neg Ra \wedge Sa)) \\ &\vee ((Pa \wedge Qa \wedge Ra \wedge Sa) \wedge (Pc \wedge Qc \wedge Rc \wedge \neg Sc) \wedge (Pb \wedge Qb \wedge \neg Rb \wedge Sb)) \end{aligned}$$

So a structure description is a disjunction of one or more state descriptions. It says *how many* individuals satisfy the maximally consistent predicates (Carnap calls them Q -predicates) that can be formulated in the language. It may but need not say *which* individuals. The first structure description above says that all three individuals a , b , and c have the maximally consistent property $Px \wedge Qx \wedge Rx \wedge Sx$. The second structure description says that exactly one individual has the maximally consistent property $Px \wedge Qx \wedge Rx \wedge Sx$, exactly one individual has the maximally consistent property $Px \wedge Qx \wedge Rx \wedge \neg Sx$, and exactly one individual has the maximally consistent property $Px \wedge Qx \wedge \neg Rx \wedge Sx$. It does not say which of a , b , and c as the property in question.

Each function that assigns non-negative weights w_i to the state descriptions z_i whose sum $\sum_i w_i$ equals 1 induces a probability on the language in question. Carnap then argues – by postulating various principles of symmetry and invariance – that each of the finitely many structure (not state) descriptions s_j should be assigned the same weight v_j such that their sum $\sum_j v_j$ is equal to 1. This

weight v_j should then be divided equally among the state descriptions whose disjunction constitutes the structure description s_j . The probability so obtained is Carnap's favorite m^* , which, like any other probability, induces what Carnap calls a confirmation function (and we have called a conditional probability): $c^*(H, E) = m^*(H \wedge E)/m^*(E)$

(In case the language contains countably infinitely many individual constants, some structure descriptions are disjunctions of infinitely many state descriptions. These state descriptions cannot all get the same positive weight. Therefore Carnap considers the limit of the measures m_n^* for the languages L_n containing the first n individual constants in some enumeration of the individual constants, provided this limit exists.)

c^* allows learning from experience in the sense that

$$c^*(\text{the } n+1 \text{st individual is } P, k \text{ of the first } n \text{ individuals are } P) > c^*(\text{the } n+1 \text{st individual is } P, \tau) \\ = m^*(\text{the } n+1 \text{st individual is } P),$$

where τ is the tautological sentence. If we assigned equal weights to the state descriptions instead of the structure descriptions, no such learning would be possible. Let us check that c^* allows learning from experience for $n=2$ in a language with three individual constants "a," "b," and "c" and one predicate "P." There are eight state descriptions and four structure descriptions:

$z_1 = Pa \wedge Pb \wedge Pc$	$s_1 = Pa \wedge Pb \wedge Pc$: All three individuals are P .
$z_2 = Pa \wedge Pb \wedge \neg Pc$	$s_2 = (Pa \wedge Pb \wedge \neg Pc) \vee (Pa \wedge \neg Pb \wedge Pc) \vee (\neg Pa \wedge Pb \wedge Pc)$: Exactly two individuals are P .
$z_3 = Pa \wedge \neg Pb \wedge Pc$	
$z_4 = Pa \wedge \neg Pb \wedge \neg Pc$	$s_3 = (Pa \wedge \neg Pb \wedge \neg Pc) \vee (\neg Pa \wedge Pb \wedge \neg Pc) \vee (\neg Pa \wedge \neg Pb \wedge Pc)$: Exactly one individual is P .
$z_5 = \neg Pa \wedge Pb \wedge Pc$	
$z_6 = \neg Pa \wedge Pb \wedge \neg Pc$	
$z_7 = \neg Pa \wedge \neg Pb \wedge Pc$	
$z_8 = \neg Pa \wedge \neg Pb \wedge \neg Pc$	$s_4 = \neg Pa \wedge \neg Pb \wedge \neg Pc$: None of the three individuals is P .

Each structure description s_1 - s_4 gets weight $v_j = 1/4$ ($j = 1, \dots, 4$).

$$s_1 = z_1: v_1 = m^*(Pa \wedge Pb \wedge Pc) = 1/4$$

$$s_2 = z_2 \vee z_3 \vee z_5: v_2 = m^*((Pa \wedge Pb \wedge \neg Pc) \vee (Pa \wedge \neg Pb \wedge Pc) \vee (\neg Pa \wedge Pb \wedge Pc)) = 1/4$$

$$s_3 = z_4 \vee z_6 \vee z_7: v_3 = m^*((Pa \wedge \neg Pb \wedge \neg Pc) \vee (\neg Pa \wedge Pb \wedge \neg Pc) \vee (\neg Pa \wedge \neg Pb \wedge Pc)) = 1/4$$

$$s_4 = z_8: v_4 = m^*(\neg Pa \wedge \neg Pb \wedge \neg Pc) = 1/4$$

These weights are equally divided among the state descriptions z_1 - z_8 .

$$z_1: w_1 = m^*(Pa \wedge Pb \wedge Pc) = 1/4 \quad z_5: w_5 = m^*(\neg Pa \wedge Pb \wedge Pc) = 1/12$$

$$z_2: w_2 = m^*(Pa \wedge Pb \wedge \neg Pc) = 1/12 \quad z_6: w_6 = m^*(\neg Pa \wedge Pb \wedge \neg Pc) = 1/12$$

$$z_3: w_3 = m^*(Pa \wedge \neg Pb \wedge Pc) = 1/12 \quad z_7: w_7 = m^*(\neg Pa \wedge \neg Pb \wedge Pc) = 1/12$$

$$z_4: w_4 = m^*(Pa \wedge \neg Pb \wedge \neg Pc) = 1/12 \quad z_8: w_8 = m^*(\neg Pa \wedge \neg Pb \wedge \neg Pc) = 1/4$$

Let us now compute the values of the confirmation function c^* .

$$\begin{aligned} c^*(\text{the 3rd individual is } P, 2 \text{ of the first 2 individuals are } P) &= \\ &= m^*(\text{the 3rd individual is } P \wedge \text{the first 2 individuals are } P) / m^*(\text{the first 2 individuals are } P) \\ &= m^*(\text{the first 3 individuals are } P) / m^*(\text{the first 2 individuals are } P) \\ &= m^*(Pa \wedge Pb \wedge Pc) / m^*(Pa \wedge Pb) \\ &= (1/4) / (1/4 + 1/12) \\ &= 3/4 \\ &> 1/2 = m^*(Pc) = c^*(\text{the 3rd individual is } P) \end{aligned}$$

The general formula is (Carnap 1950/1962, 568)

$$\begin{aligned} c^*(\text{the } n+1 \text{st individual is } P, k \text{ of the first } n \text{ individuals are } P) &= \\ &= (k + \omega) / (n + \kappa) \\ &= (k + (\omega/\kappa)\epsilon\kappa) / (n + \kappa), \end{aligned}$$

where ω is the “logical width” of the predicate “ P ” (Carnap 1950/1962, 127), that is, the number of maximally consistent properties or Q -predicates whose disjunction is logically equivalent to “ P ”

($\omega = 1$ in our example: “ P ”). $\kappa = 2^\pi$ is the total number of Q -predicates ($\kappa = 2^1 = 2$ in our

example: “ P ” and “ $\neg P$ ”) with π being the number of primitive predicates ($\pi = 1$ in our example:

“ P ”). This formula is dependent on the logical factor ω/κ of the “relative width” of the predicate “ P ,” and the empirical factor k/n of the relative frequency of P s.

Later on, Carnap (1952) generalizes this to a whole continuum of confirmation functions c_λ where the parameter λ is inversely proportional to the impact of evidence. λ specifies how the

confirmation function c_λ weighs between the logical factor ω/κ and the empirical factor k/n . For

$\lambda = \infty$, c_λ is independent of the empirical factor k/n : c_λ (the $n+1$ st individual is P , k of the

first n individuals are P) = $\frac{k}{n}$ (Carnap 1952, §13). For $\lambda=0$, c_λ is independent of the logical factor $\frac{k}{n}$: c_λ (the $n+1$ st individual is P , k of the first n individuals are P) = k/n and thus coincides with what is known as the straight rule (Carnap 1952, §14). c^* is the special case with $\lambda=0$ (Carnap 1952, §15). The general formula is (Carnap 1952, §9)

$$c_\lambda \text{ (the } n+1 \text{st individual is } P, k \text{ of the first } n \text{ individuals are } P) = (k + \lambda/n)/(n + \lambda).$$

In his (1963) Carnap slightly modifies the set up and considers *families* of monadic predicates $\{P_1, \dots, P_p\}$ like the family of color predicates $\{\text{red, green, blue}\}$. For a given family $\{P_1, \dots, P_p\}$ and each individual constant " a " there is exactly one predicate " P_j " such that $P_j a$. Families thus generalize $\{P, \neg P\}$ and correspond to random variables. Given his axioms (including *A15*) Carnap (1963, 976) shows that for each family $\{P_1, \dots, P_p\}$, $p \geq 2$,

$$c_\lambda \text{ (the } n+1 \text{st individual is } P_j, k \text{ of the first } n \text{ individuals are } P_j) = (k + \lambda/p)/(n + \lambda).$$

One of the peculiar features of Carnap's systems is that universal generalizations get degrees of confirmation (alias conditional probability) 0. Hintikka (1966) further elaborates Carnap's project in this respect. For a neo-Carnapian approach see Maher (2004a).

Of more interest to us is Carnap's discussion of "the controversial problem of the *justification of induction*" (1963, 978, emphasis in the original). For Carnap, the justification of induction boils down to justifying the axioms specifying a set of confirmation functions. The "reasons are based upon our intuitive judgments concerning inductive validity". Therefore "[i]t is impossible to give a purely deductive justification of induction," and these "reasons are a priori" (Carnap 1963, 978). So according to Carnap, induction is justified by appeals to intuition about inductive validity. We will see below that Goodman, who is otherwise very skeptical about the prospects of Carnap's project, shares this view of the justification of induction. In fact, the view also seems to be widely accepted among current Bayesian confirmation theorists and their desideratum/explicatum approach (see Fitelson 2001 for an example). [According to Carnap (1952, ch. I), an explication is "the transformation of an inexact, prescientific concept, the *explicandum*, into a new exact concept, the *explicatum*." (Carnap 1952, 3) The desideratum/explicatum approach consists in stating various "intuitively plausible desiderata" the explicatum is supposed to satisfy. Proposals for explicata that do not satisfy these desiderata are rejected. This appeal to intuitions is fine as long as we are doing conceptual analysis. However, contemporary confirmation theorists also sell their accounts as normative theories. Normative theories are not justified by appeal to intuitions, though. They are justified relative to a goal by showing that the norms in question further the goal at issue. See section 7.]

First, however, we will have a look at what Carnap has to say about Hempel's conditions of adequacy.

c. Absolute and Incremental Confirmation

As we saw in the preceding section, one of Carnap's goals was to define a *quantitative* notion of confirmation, explicated by a confirmation function in the manner indicated above. It is important to note that this quantitative concept of confirmation is a relation between two propositions H and E (three, if we include the background information B), a number r , and a confirmation function c . In chapters VI and VII of his (1950/1962) Carnap discusses comparative and qualitative concepts of confirmation. The explicans for qualitative confirmation he offers is that of positive probabilistic relevance in the sense of some logical probability m . That is, E qualitatively confirms H in the sense of some logical measure m just in case E is positively relevant to H in the sense of m , that is,

$$m(H \wedge E) > m(H) \cdot m(E).$$

If both $m(H)$ and $m(E)$ are positive – which is the case whenever both H and E are not logically false, because Carnap assumes m to be regular – this is equivalently expressed by the following inequality:

$$c(H, E) > c(H, \tau) = m(H)$$

So provided both H and E have positive probability, E confirms H if and only if E raises the conditional probability (degree of confirmation in the sense of c) of H . Let us call this concept *incremental* confirmation. Again, note that qualitative confirmation is a relation between two propositions H and E , and a conditional probability or confirmation function c . Incremental confirmation, or positive probabilistic relevance, is a qualitative notion, which says *whether* E raises the conditional probability (degree of confirmation in the sense of c) of H . Its natural quantitative counterpart measures *how much* E raises the conditional probability of H . This measure may take several forms which will be discussed below.

Incremental confirmation is different from the concept of *absolute* confirmation on which it is based. The quantitative explication of absolute confirmation is given by one of Carnap's confirmation functions c . The qualitative counterpart is to say that E absolutely confirms H in the sense of c if and only if the degree of absolute confirmation of H by E is sufficiently high, $c(H, E) > r$. So Carnap, who offers degree of absolute confirmation $c(H, E)$ as explication for the quantitative notion of confirmation of H by E , and who offers incremental confirmation or positive probabilistic relevance between E and H as explication of the qualitative notion of confirmation, is, to say the least, not fully consistent in his terminology. He switches between absolute confirmation (for the quantitative notion) and incremental confirmation (for the qualitative notion). This is particularly peculiar, because Carnap (1950/1962, §87) is the locus classicus for the discussion of Hempel's conditions of adequacy mentioned in section 2b.

d. Carnap's Analysis of Hempel's Conditions

In analyzing the special consequence condition, Carnap argues that

Hempel has in mind as explicandum the following relation: "the degree of confirmation of H by E is greater than r , where r is a fixed value, perhaps 0 or 1/2 (Carnap 1962, 475; notation adapted);

that is, the *qualitative* concept of *absolute* confirmation. Similarly when discussing the special

consistency condition:

Hempel regards it as a great advantage of any explicatum satisfying [a more general form of the special consistency condition 3] “that it sets a limit, so to speak, to the strength of the hypotheses which can be confirmed by given evidence” ... This argument does not seem to have any plausibility for *our* explicandum, (Carnap 1962, 477; emphasis in original)

which is the *qualitative* concept of *incremental* confirmation,

[b]ut it is plausible for the second explicandum mentioned earlier: the degree of [absolute] confirmation exceeding a fixed value r . Therefore we may perhaps assume that Hempel’s acceptance of [a more general form of 3] is due again to an inadvertent shift to the second explicandum. (Carnap 1962, 477-478)

Carnap’s analysis can be summarized as follows. In presenting his first three conditions of adequacy, Hempel was mixing up two distinct concepts of confirmation, two distinct explicanda in Carnap’s terminology, viz.

- (i) the qualitative concept of *incremental* confirmation (positive probabilistic relevance) according to which E confirms H if and only if E (has non-zero probability and) increases the degree of absolute confirmation (conditional probability) of H , and
- (ii) the qualitative concept of *absolute* confirmation according to which E confirms H if and only if the degree of absolute confirmation (conditional probability) of H by E is greater than some value r .

Hempel’s second and third condition, 2 and 3, respectively, hold true for the second explicandum (for $r \geq 1/2$), but they do not hold true for the first explicandum. On the other hand, Hempel’s first condition holds true for the first explicandum, but it does so only in a *qualified* form (Carnap 1950/1962, 473) – namely only if E is not assigned probability 0, and H is not already assigned probability 1.

This, however, means that, according to Carnap’s analysis, Hempel first had in mind the explicandum of incremental confirmation for the entailment condition. Then he had in mind the explicandum of absolute confirmation for the special consequence and the special consistency conditions 2 and 3, respectively. And then, when Hempel presented the converse consequence condition, he got completely confused and had in mind still another explicandum or concept of confirmation (neither the first nor the second explicandum satisfies the converse consequence condition). This is not a very charitable analysis. It is not a good one either, because the qualitative concept of absolute confirmation, which Hempel is said to have had in mind for 2 and 3, also satisfies 1 – and it does so *without* the second qualification that H be assigned a probability smaller than 1. So there is no need to accuse Hempel of mixing up two concepts of confirmation. Indeed, the analysis is bad, because Carnap’s reading of Hempel also leaves open the question of what the third explicandum for the converse consequence condition might have been. For a different analysis of Hempel’s conditions and a corresponding logic of confirmation see Huber (2007a), respectively.

5. The New Riddle of Induction and the Demise of the

Syntactic Approach

According to Goodman (1983, ch. III), the problem of justifying induction boils down to *defining* valid inductive rules, and thus to a definition of confirmation. The reason is that an inductive inference is justified by conformity to an inductive rule, and inductive rules are justified by their conformity to accepted inductive practices. One does not have to follow Goodman in this respect, however, in order to appreciate his insight that whether a hypothesis is confirmed by a piece of evidence depends on features other than their syntactical form.

In his (1946) he asks us to suppose a marble has been drawn from a certain bowl on each of the ninety-nine days up to and including VE day, and that each marble drawn was red. Our evidence can be described by the conjunction “Marble 1 is red and ... and marble 99 is red,” in symbols: $Ra_1 \wedge \dots \wedge Ra_{99}$. Whatever the details of our theory of confirmation, this evidence will confirm the hypothesis “Marble 100 is red,” R_{100} . Now consider the predicate $S =$ “is drawn by VE day and is red, or is drawn after VE day and is not red.” In terms of S rather than R our evidence is described by the conjunction “Marble 1 is drawn by VE day and is red or it is drawn after VE day and is not red, and ..., and marble 99 is drawn by VE day and is red or it is drawn after VE day and is not red,” $Sa_1 \wedge \dots \wedge Sa_{99}$. If our theory of confirmation relies solely on syntactical features of the evidence and the hypothesis, our evidence will confirm the conclusion “Marble 100 is drawn by VE and is red, or it is drawn after VE day and is not red,” S_{100} . But we know that the next marble will be drawn after VE day. Given this, S_{100} is logically equivalent to the negation of R_{100} . So one and the same piece of evidence can be used to confirm a hypothesis and its negation, which is certainly absurd.

One might object to this example that the two formulations do not describe one and the same piece of evidence after all. The first formulation in terms of R should be the conjunction “Marble 1 is drawn by VE day and is red, and ..., and marble 99 is drawn by VE day and is red,” $(Da_1 \wedge Ra_1) \wedge \dots \wedge (Da_{99} \wedge Ra_{99})$. The second formulation in terms of S should be “Marble 1 is drawn by VE day and it is drawn by VE day and red or drawn after VE and not red, and ..., and marble 99 is drawn by VE day and it is drawn by VE day and red or drawn after VE day and not red,” $(Da_1 \wedge Sa_1) \wedge \dots \wedge (Da_{99} \wedge Sa_{99})$. Now the two formulations really describe one and the same piece of evidence in the sense of being logically equivalent. But then the problem is whether any interesting statement can ever be confirmed. The syntactical form of the evidence now seems to confirm $Da_{100} \wedge Ra_{100}$, equivalently $Da_{100} \wedge Sa_{100}$. But we know that the next marble is drawn after VE day; that is, we know $\neg Da_{100}$. That the future resembles the past in all respects is thus false. That it resembles the past in some respects is trivial. The new riddle of induction is the question in which respects the future resembles the past, and in which it does not.

It has been suggested that the puzzling character of Goodman’s example is due to its mentioning a particular point of time, viz. VE day. A related reaction has been that gerrymandered predicates, whether or not they involve a particular point of time, cannot be used in inductive inferences. But there are plenty of similar examples (Stalker 1994), and it is commonly agreed that Goodman has succeeded in showing that a purely syntactical definition of (degree of) confirmation won’t do. Goodman himself sought to solve his new riddle of induction by distinguishing between “projectible” predicates such as “red” and unprojectible predicates such as “is drawn by VE day and is red, or is drawn after VE day and is not red.” The projectibility of a predicate is in turn determined by its entrenchment in natural language. This comes very close to saying that the projectible predicates are the ones that we do in fact project (that is, use in inductive inferences).

(Quine's 1969 "natural kinds" are special cases of what can be described by projectible predicates.)

6. Bayesian Confirmation Theory

Bayesian confirmation theory is by far the most popular and elaborated theory of confirmation. It has its origins in Rudolf Carnap's work on inductive logic (Carnap 1950/1962), but relieves itself from defining confirmation in terms of *logical* probability. More or less any subjective degree of belief function satisfying the Kolmogorov axioms is considered to be an admissible probability measure.

a. Subjective Probability and the Dutch Book Argument

In Bayesian confirmation theory, a probability measure on a field of propositions is usually interpreted as an agent's degree of belief function. There is disagreement as to how broad the class of admissible probability measures is to be construed. Some objective Bayesians such as the early Carnap insist that the class consist of a single logical probability measure, whereas subjective Bayesians admit any probability measure. Most Bayesians will be somewhere in the middle of this spectrum when it comes to the question which particular degree of belief functions it is reasonable to adopt in a particular situation. But they will agree that from a purely logical point of view any (regular) probability measure is acceptable. The standard argument for this position is the Dutch Book Argument.

The Dutch Book Argument starts with the assumption that there is a link between subjective degrees of belief and betting ratios. It is further assumed that it is *pragmatically* defective to accept a series of bets which guarantees a sure loss, that is, a Dutch Book. By appealing to the Dutch Book Theorem that an agent's betting ratios satisfy the probability axioms just in case they do not make the agent vulnerable to such a Dutch Book, it is inferred that it is *epistemically* defective to have degrees of belief that violate the probability axioms. The strength of this inference is, of course, dependent on the link between degrees of belief and betting ratios. If this link is identity – as it is when one defines degrees of belief as betting ratios – the distinction between pragmatic and epistemic defectiveness disappears, and the Dutch Book Argument is a deductively valid argument. But this comes at the cost of rendering the link between degrees of belief and betting ratios implausible. If the link is weaker than identity – as it is when degrees of belief are only measured by betting ratios – the Dutch Book Argument is not deductively valid anymore, but it has more plausible assumptions.

The pragmatic nature of the Dutch Book Argument has led to so called *depragmatized* versions. A depragmatized Dutch Book Argument starts with a link between degrees of belief and *fair* betting ratios, and it assumes that it is *epistemically* defective to *consider* a series of bets that guarantees a sure loss *as fair*. Using the depragmatized Dutch Book Theorem that an agent's fair betting ratios obey the probability calculus if and only if the agent never considers a Dutch Book as fair, it is then inferred that it is epistemically defective to have degrees of belief that do not obey the probability calculus. The thesis that an agent's degree of belief function should obey the probability calculus is called probabilism. For more on the Dutch Book Argument see Hájek (2005). For a different justification of probabilism in terms of the accuracy of degrees of belief see Joyce (1998).

b. Confirmation Measures

Let \mathbf{A} be a field of propositions over some set of possibilities W , let H, E, B be propositions from \mathbf{A} , and let Pr be a probability measure on \mathbf{A} . We already know that H is incrementally confirmed

by E relative to B in the sense of Pr if and only if $\text{Pr}(H \circ E|B) > \text{Pr}(H|B) \cdot \text{Pr}(E|B)$, and that this is a relation between three propositions and a probability space whose field contains the propositions. The central notion in Bayesian confirmation theory is that of a confirmation measure. A real valued function $c: \mathbf{P} \rightarrow \mathfrak{R}$ from the set \mathbf{P} of all probability spaces $\langle W, \mathbf{A}, \text{Pr} \rangle$ into the reals \mathfrak{R} is a *confirmation measure* if and only if for every probability space $\langle W, \mathbf{A}, \text{Pr} \rangle$ and all $H, E, B \in \mathbf{A}$:

$$c(H, E, B) > 0 \Leftrightarrow \text{Pr}(H \circ E|B) > \text{Pr}(H|B) \cdot \text{Pr}(E|B)$$

$$c(H, E, B) = 0 \Leftrightarrow \text{Pr}(H \circ E|B) = \text{Pr}(H|B) \cdot \text{Pr}(E|B)$$

$$c(H, E, B) < 0 \Leftrightarrow \text{Pr}(H \circ E|B) < \text{Pr}(H|B) \cdot \text{Pr}(E|B)$$

The six most popular confirmation measures are (what I now call) the Carnap measure c (Carnap 1962), the distance measure d (Earman 1992), the log-likelihood or Good-Fitelson measure l (Fitelson 1999 and Good 1983), the log-ratio or Milne measure r (Milne 1996), the Joyce-Christensen measure s (Christensen 1999, Joyce 1999, ch. 6), and the relative distance measure z (Crupi & Tentori & Gonzalez 2007).

$$c(H, E, B) = \text{Pr}(H \circ E|B) - \text{Pr}(H|B) \cdot \text{Pr}(E|B)$$

$$d(H, E, B) = \text{Pr}(H|E \circ B) - \text{Pr}(H|B)$$

$$l(H, E, B) = \log [\text{Pr}(E|H \circ B) / \text{Pr}(E|-H \circ B)]$$

$$r(H, E, B) = \log [\text{Pr}(H|E \circ B) / \text{Pr}(H|B)]$$

$$s(H, E, B) = \text{Pr}(H|E \circ B) - \text{Pr}(H|-E \circ B)$$

$$z(H, E, B) = [\text{Pr}(H|E \circ B) - \text{Pr}(H|B)] / \text{Pr}(-H|B) \text{ if } \text{Pr}(H|E \circ B) \geq \text{Pr}(H|B)$$

$$= [\text{Pr}(H|E \circ B) - \text{Pr}(H|B)] / \text{Pr}(H|B) \text{ if } \text{Pr}(H|E \circ B) < \text{Pr}(H|B)$$

(Mathematically speaking, there are uncountably many confirmation measures.) For an overview article, see Eells (2005). Book length expositions are Earman (1992) and Howson & Urbach (1989/2005).

c. Some Success Stories

Bayesian confirmation theory captures the insights of Popper's falsificationism and hypothetico-deductive confirmation. Suppose evidence E falsifies hypothesis H relative to background

information B in the sense that $B \circ H \circ E = \emptyset$. Then $Pr(E \circ H|B) = 0$, and so $Pr(E \circ H|B) = 0 < Pr(H|B) \cdot Pr(E|B)$, provided both $Pr(H|B)$ and $Pr(E|B)$ are positive. So as long as H is not already known to be false (in the sense of having probability 0 conditional on B) and E is a possible outcome (one with positive probability conditional on B), falsifying E incrementally disconfirms H relative to B in the sense of Pr .

Remember, E HD-confirms H relative to B if and only if the conjunction of H and B logically implies E (in some suitable way). In this case $Pr(E \circ H|B) = Pr(H|B)$, provided $Pr(B) > 0$. Hence as long as $Pr(E|B) < 1$, we have

$$Pr(E \circ H|B) > Pr(H|B) \cdot Pr(E|B),$$

which means that E incrementally confirms H relative to B in the sense of Pr (Kuipers 2000).

If the conjunction of H and B logically implies E , but E is already known to be true in the sense of having probability 1 conditional on B , E does not incrementally confirm H relative to B in the sense of Pr . In fact, no E which receives probability 1 conditional on B can incrementally confirm any H whatsoever. This is the so called problem of old evidence (Glymour 1980). It is a special case of a more general phenomenon. The following is true for many confirmation measures (d , l , and r , but not s). If H is positively relevant to E given B , the degree to which E incrementally confirms H relative to B is greater, the smaller the probability of E given B . Similarly, if H is negatively relevant for E given B , the degree to which E disconfirms H relative to B is greater, the smaller the probability of E given B (Huber 2005a). If $Pr(E|B) = 1$ we have the problem of old evidence. If $Pr(E|B) = 0$ we have the above mentioned problem that E cannot disconfirm hypotheses it falsifies.

Some people simply deny that the problem of old evidence is a problem. Bayesian confirmation theory, it is said, does not explicate whether and how much E confirms H relative to B . It explicates whether E is *additional* evidence for H relative to B , and how much *additional* confirmation E provides for H relative to B . If E already has probability 1 conditional on B , it is part of the background knowledge, and so does not provide any additional evidence for H . More generally, the more we already believe in E , the less additional (dis)confirmation this provides for positively (negatively) relevant H . This reply does not work in case E is a falsifier of H with probability 0 conditional on B , for in this case $Pr(H|E \circ B)$ is not defined. It also does not agree with the fact that the problem of old evidence is taken seriously in the literature on Bayesian confirmation theory (Earman 1992, ch. 5). An alternative view (Joyce 1999, ch. 6) sees several different but equally legitimate concepts of confirmation at work. The intuition behind one concept is the reason for the implausibility of the explication of another.

In contrast to hypothetico-deductivism, Bayesian confirmation theory has no problem with assigning degrees of incremental confirmation to statistical hypotheses. Such alternative statistical hypotheses H_1, \dots, H_n, \dots are taken to specify the probability of an outcome E . The probabilities $Pr(E|H_1), \dots, Pr(E|H_n), \dots$ are called the likelihoods of the hypotheses H_i . Together with their prior probabilities $Pr(H_i)$ the likelihoods determine the posterior probabilities of the H_i via Bayes's Theorem:

$$\Pr(H_i|E) = \Pr(E|H_i) \cdot \Pr(H_i) / [\sum_j \Pr(E|H_j) \cdot \Pr(H_j) + \Pr(E|H) \cdot \Pr(H)]$$

The so called “catchall” hypothesis H is the negation of the disjunction or union of all the alternative hypotheses H_i , and so it is equivalent to $\neg(H_1 \cup \dots \cup H_n \cup \dots)$. It is important to note the implicit use of something like the principal principle (Lewis 1980) in such an application of Bayes’s Theorem. The probability measure \Pr figuring in the above equation is an agent’s degree of belief function. The statistical hypotheses H_i specify the objective chance of the outcome E as $\text{Ch}_i(E)$. Without a principle linking objective chances to subjective degrees of belief, nothing guarantees that the agent’s conditional degree of belief in E given H_i , $\Pr(E|H_i)$, is equal to the chance of E as specified by H_i , $\text{Ch}_i(E)$. The principal principle says that an agent’s conditional degree of belief in a proposition A given the information that the chance of A is equal to r (and no further inadmissible information) should be r , $\Pr(A|\text{Ch}(A) = r) = r$. For more on the principal principle see Hall (1994), Lewis (1994), Thau (1994), as well as Vranas (2004a). Spohn shows that the principal principle is a special case of the reflection principle (van Fraassen 1984; 1995). The latter principle says that an agent’s current conditional degree of belief in A given that her future belief in A equals r should be r ,

$$\Pr_{\text{now}}(A|\Pr_{\text{later}}(A) = r) = r \quad \text{provided } \Pr_{\text{now}}(\Pr_{\text{later}}(A)=r) > 0.$$

Bayesian confirmation theory can also handle the ravens paradox. As we have seen, Hempel thought that “ a is neither black nor a raven” confirms “All ravens are black” relative to no or tautological background information. He attributed the unintuitive character of this claim to a conflation of it and the claim that “ a is neither black nor a raven” confirms “All ravens are black” relative to our actual background knowledge A – and the fact that A contains the information that there are more non-black objects than ravens. The latter information is reflected in our degree of belief function \Pr by the inequality

$$\Pr(\neg Ba|A) > \Pr(Ra|A).$$

If we further assume that the probabilities of finding a non-black object as well as finding a raven are independent of whether or not all ravens are black,

$$\Pr(\neg Ba|\Box x(Rx \rightarrow Bx) \wedge A) = \Pr(\neg Ba|A),$$

$$\Pr(Ra|\Box x(Rx \rightarrow Bx) \wedge A) = \Pr(Ra|A),$$

we can infer (when we assume all probabilities to be defined) that

$$\Pr(\Box x(Rx \rightarrow Bx)|Ra \wedge Ba \wedge A) > \Pr(\Box x(Rx \rightarrow Bx)|\neg Ra \wedge \neg Ba \wedge A) > \Pr(\Box x(Rx \rightarrow Bx)|A).$$

So Hempel’s intuitions are vindicated by Bayesian confirmation theory to the extent that the above independence assumptions are plausible (or there are weaker assumptions entailing a similar result), and to the extent he also took non-black non-ravens to confirm the ravens hypothesis relative to our actual background knowledge. For more, see Vranas (2004b).

Let us finally consider the problem of irrelevant conjunction in Bayesian confirmation theory. HD-confirmation satisfies the converse consequence condition, and so has the undesirable feature that

E confirms $H \wedge H'$ relative to B whenever E confirms H relative to B , for any H' whatsoever. This is not true for incremental confirmation. Even if $\Pr(E \wedge H|B) > \Pr(E|B) \cdot \Pr(H|B)$, it need not be the case that $\Pr(E \wedge H \wedge H'|B) > \Pr(E|B) \cdot \Pr(H \wedge H'|B)$. However, the following special case is also true for incremental confirmation.

If $H \wedge B$ logically implies E , then E incrementally confirms $H \wedge H'$ relative to B , for any H' whatsoever (whenever the relevant probabilities are defined).

In the spirit of the last paragraph, one can, however, show that $H \wedge H'$ is less confirmed by E relative to B than H alone (in the sense of the distance measure d and the Good-Fitelson measure l) if H' is an irrelevant conjunct to H given B with respect to E in the sense that

$$\Pr(E|H \wedge H' \wedge B) = \Pr(E|H \wedge B)$$

(Hawthorne & Fitelson 2004). If $H \wedge B$ logically implies E , then every H' such that $\Pr(H \wedge H' \wedge B) > 0$ is irrelevant in this sense. For more see Fitelson (2002), Hawthorne & Fitelson (2004), Maher (2004b).

7. Taking Stock

Let us grant that Bayesian confirmation theory adequately explicates the concept of confirmation. If so, then this is the concept scientists use when they say that the anomalous perihelion of Mercury confirms the general theory of relativity. It is also the concept more ordinary epistemic agents use when they say that, relative to what they have experienced so far, the dark clouds on the sky are evidence for rain. The question remains what happened to Hume's problem of the *justification* of induction. We know – by definition – that the conclusion of an inductively strong argument is well-confirmed by its premises. But does that also justify our acceptance of that conclusion? Don't we first have to justify our definition of confirmation before we can use it to justify our inductive inferences?

It seems we would have to, but, as Hume argued, such a justification of induction is not possible. All we could hope for is an adequate description of our inductive practices. As we have seen, Goodman took the task of adequately describing induction as being tantamount to its justification (Goodman 1983, ch. III, ascribes a similar view to Hume, which is somehow peculiar, because Hume argued that a justification of induction is *impossible*). In doing so he appealed to deductive logic, which he claimed to be justified by its conformity to accepted practices of deductive reasoning. But that is not so. Deductive logic is *not* justified because it adequately describes our practices of deductive reasoning – it doesn't. The rules of deductive logic are justified *relative* to the goal of truth preservation in all possible worlds. The reasons are that (i) in going from the premises of a deductively valid argument to its conclusion, truth is preserved in all possible worlds (this is known as soundness); and that (ii) any argument with that property is a deductively valid argument (this is known as completeness). Similarly for the rules of nonmonotonic logic, which are justified *relative* to the goal of truth preservation in all "normal" worlds (for normality see e.g. Koons 2005). The reason is that all and only nonmonotonically valid inferences are such that truth is preserved in all normal worlds when one jumps from the premises to the conclusion (Kraus & Lehmann & Magidor 1990, for a survey see Makinson 1994). More generally, the justification of a canon of normative principles – such as the rules of deductive logic, the rules of nonmonotonic

logic, or the rules of inductive logic – are only justified *relative to a certain goal* when one can show that adhering to these normative principles in some sense furthers the goal in question.

Similarly to Goodman, Carnap sought to justify the principles of his inductive logic by appeals to intuition (cf. the quote in section 4b). Contemporary Bayesian confirmation theorists with their desideratum/explicatum approach follow Carnap and Goodman at least insofar as they apparently do not see the need for justifying their accounts of confirmation by more than appeals to intuition. These are supposed to show that their definitions of confirmation are adequate. But the alleged impossibility of justifying induction does not entail that its adequate description or explication in form of a particular theory of confirmation is sufficient to justify inductive inferences based on that theory. Moreover, as noted by Reichenbach (1938; 1940), a justification of induction is not impossible after all. Hume was right in claiming that there is no deductively valid argument with knowable premises and the conclusion that inductively strong arguments will always lead us to true conclusions. But that is not the only conclusion that would justify induction. Reichenbach was mainly interested in the limiting relative frequencies of particular outcomes in various sequences of events. He could show that a particular inductive rule – the straight rule that conjectures that the limiting relative frequency is equal to the observed relative frequency – will lead us to the true limiting relative frequency, if any inductive rule does. However, the straight rule is not the only rule with this property. Therefore its justification relative to the goal of discovering limiting relative frequencies is at least incomplete. If we want to keep the analogy to deductive logic, we can put things as follows: Reichenbach was able to establish the soundness, but not the completeness, of his inductive logic (that is, the straight rule) with respect to the goal of eventually arriving at the true limiting relative frequency. (Reichenbach himself provides an example that proves the incompleteness of the straight rule with respect to this goal.)

While soundness in this sense is not sufficient for a justification of the straight rule, such results provide more reasons than appeals to intuition. They are necessary conditions for the justification of a normative rule of inference relative to a particular goal of inquiry. A similar view about the justification of induction is held by formal learning theory. Here one considers the objective *reliability* with which a particular method (such as the straight rule or a particular confirmation measure) finds out the correct answer to a given question. The use of a method to answer a question is only justified when the method reliably answers the question, if any method does. As different questions differ in their complexity, there are different senses of reliability. A method may correctly answer a question after finitely many steps and with a sign that the question is answered correctly – as when we answer the question whether the first observed raven is black by saying “yes” if it is, and “no” otherwise. Or it may answer the question after finitely many steps and with a sign that it has done so when the answer is “yes,” but not when the answer is “no” – as when we answer the question whether there exists a black raven by saying “yes” when we first observe a black raven, and by saying “no” otherwise. Or it may stabilize to the correct answer in the sense that the method conjectures the right answer after finitely many steps and continues to do so forever without necessarily giving a sign that it has arrived at the correct answer – as when we answer the question whether the limiting relative frequency of black ravens among all ravens is greater than .5 by saying “yes” as long as the observed relative frequency is greater than .5, and by saying “no” otherwise (under the assumption that this limit exists). And so on. This provides a classification of all problems in terms of their complexity. The use of a particular method for answering a question of a certain complexity is only justified if the method reliably answers the question in the sense of reliability determined by the complexity of the question. A discussion of Bayesian confirmation theory from the point of view of formal learning theory can be found in Kelly & Glymour (2004). Schulte (2002) gives an introduction to the main philosophical ideas of

formal learning theory. A technically advanced book length exposition is Kelly (1996). The general idea is the same as before. A rule is justified relative to a certain goal to the extent that the rule furthers that goal.

So can we justify particular inductive rules in the form of confirmation measures along these lines? We had better, for otherwise there might be inductive rules that would reliably lead us to the correct answer about a question where our inductive rules won't (cf. Putnam 1963a; see also his 1963b). Before answering this question, let us first be clear which goal confirmation is supposed to further. In other words, why should we accept well-confirmed hypotheses rather than any other hypotheses? A natural answer is that science and our more ordinary epistemic enterprises aim at *true* hypotheses. The justification for confirmation would then be that we should accept well-confirmed hypotheses, because we are in some sense guaranteed to arrive at true hypotheses if (and only if) we stick to well-confirmed hypotheses. Something along these lines is true for absolute confirmation according to which degree of confirmation is equal to probability conditional on the data. More precisely, the Gaifman and Snir convergence theorem (Gaifman & Snir 1982) says that for almost every world or model w for the underlying language – that is, all worlds w except, possibly, for those in a set of measure 0 (in the sense of the measure Pr^* on the σ -field \mathbf{A} from section 4a) – the probability of a hypothesis conditional on the first n data sentences from w converges to its truth value in w (1 for true, 0 for false). It is assumed here that the set of all data sentences separates the set of all worlds (in the sense that for any two distinct worlds there is a data sentence which is true in the one and false in the other world). If we accept a hypothesis as true as soon as its probability is greater than .5 (or any other positive threshold value < 1), and reject it as false otherwise, we are guaranteed to almost surely arrive at true hypotheses after finitely many steps. That does not mean that no other method can do equally well. But it is more than to simply appeal to our intuitions, and a necessary condition for the justification of absolute confirmation relative to the goal of truth. See also Earman (1992, ch. 9) and Juhl (1997).

A more limited result is true for incremental confirmation. Based on the Gaifman and Snir convergence theorem one can show for every confirmation measure c and almost all worlds w that there is an n such that for all later m : the conjunction of the first m data sentences confirms hypotheses that are true in w to a non-negative degree, and it confirms hypotheses that are false in w to a non-positive degree (the set of all data sentences is again assumed to separate the set of all worlds). Even if this more limited result were a satisfying justification for the claim that incremental confirmation furthers the goal of truth, the question remains why one has to go to *incremental* confirmation in order to arrive at true theories. It also remains unclear what *degrees* of incremental confirmation are supposed to indicate, for it is completely irrelevant for the above result whether a positive degree of confirmation is high or low – all that matters is that it is positive. This is in contrast to absolute confirmation. There a high number represents a high probability – that is, a high probability of being true – which almost surely converges to the truth value itself. To make these vague remarks more vivid, let us consider an example.

Suppose my 35 year old friend is pregnant and I am curious as to who is the father. I know that it is either the 35 year old Alberto or the 55 year old Ben or the 55 year old Cesar. My initial degree of belief function Pr is such that

$$Pr(A) = .9, Pr(B) = Pr(C) = .05, Pr(A \wedge B) = Pr(A \wedge C) = Pr(B \wedge C) = 0,$$

$$Pr(A \vee B) = Pr(A \vee C) = .95, Pr(B \vee C) = .1, Pr(A \vee B \vee C) = 1,$$

$$Pr(A \wedge G) = .4, Pr(B \wedge G) = .03, Pr(C \wedge G) = .03, Pr(G) = .46,$$

where A is the proposition that Alberto is the father, and similarly for B and C . G is the proposition that the father has grey hair. [More precisely, the probability space is $\langle L, Pr \rangle$ with L the propositional language over the set of propositional variables $\{A, B, C, G\}$ and Pr such that $Pr(A \wedge G) = .4, Pr(B \wedge G) = .03, Pr(C \wedge G) = .03, Pr(A \wedge \neg G) = .5, Pr(B \wedge \neg G) = .02, Pr(C \wedge \neg G) = .02, Pr(A \wedge B) = Pr(A \wedge C) = Pr(B \wedge C) = Pr(\neg A \wedge \neg B \wedge \neg C) = 0$.] This is a fairly reasonable degree of belief function. Most men at the age of 55 I know have grey hair. Less than 50% of the men of age 35 I know have grey hair. And I tend to use the principal principle whenever I can (assuming a close connection between objective chances and relative frequencies). Now suppose I learn that the father has grey hair. My new degrees of belief are

$$Pr(A|G) = 40/46, Pr(B|G) = 3/46, Pr(C|G) = 3/46,$$

$$Pr(A \vee B|G) = Pr(A \vee C|G) = 43/46, Pr(B \vee C|G) = 6/46, Pr(A \vee B \vee C|G) = 1.$$

G incrementally confirms $B, C, B \vee C, A \vee C, B \vee C$, it neither incrementally confirms nor incrementally disconfirms $A \vee B \vee C$, and it incrementally disconfirms A .

However, my degree of belief in A is still more than thirteen times my degree of belief in B and my degree of belief in C . And whether I have to bet on these propositions or whether I am just curious as to who is the father of my friend's baby, all I care about after having received evidence G will be my new degrees of belief in the various answers – and my utilities, including my desire to answer the question. I will be willing to bet on A at less favorable odds than on either B or C or even their disjunction; and should my friend tell me she is going to marry the father of her baby – she assuming that I know who it is – I would buy my wedding present on the assumption that she is going to marry Alberto (unless, of course, I can ask her first). In this situation, incremental confirmation and degrees of incremental confirmation are at best misleading.

[What is important is a way of updating my old degree of belief function by the incoming evidence. The above example assumes evidence to come in the form of a proposition that I become certain of. In this case, probabilism says I should update my degree of belief function by Strict Conditionalization:

If Pr is your subjective probability at time t , and between t and t' you learn E and no logically stronger proposition in the sense that your new degree of belief in E is 1, then your new subjective probability at time t' should be $Pr(\cdot|E)$.

As Jeffrey (1983) observes, we usually do not learn by becoming certain of a proposition. Evidence often merely changes our degrees of belief in various propositions. Jeffrey Conditionalization is a more general update rule than Strict Conditionalization:

If Pr is your subjective probability at time t , and between t and t' your degrees of belief in the countable partition $\{E_1, \dots, E_n, \dots\}$ change from $Pr(E_i)$ to $p_i \in [0,1]$ (with $Pr(E_i) = p_i$ for $Pr(E_i) \in \{0,1\}$), and your positive degrees of belief do not change on any superset thereof, then your new subjective probability at time t' should be Pr^* , where for all A , $Pr^*(A) = \sum_i Pr(A|E_i) \in p_i$.

For evidential input of the above form, Jeffrey Conditionalization turns regular probability measures into regular probability measures, provided no contingent evidential proposition receives an extreme value $p \in \{0,1\}$. Radical probabilism (Jeffrey 2004) urges you not to assign such extreme values, and to have a regular initial degree of belief function – that is, whenever you can (but you can't always). Field (1978) proposes an update rule for evidence of a different format.

This is also the place to mention different formal frameworks besides probability theory. For an overview, see Huber (2008a).]

More generally, degrees of belief are important to us, because together with our desires they determine which acts it is rational for us to take. The usual recommendation according to rational choice theory for choosing one's acts is to maximize one's expected utility (the mathematical representation of one's desires), that is, the quantity

$$EU(a) = \sum_{s \in S} su(a(s)) \cdot Pr(s).$$

Here S is an exclusive and exhaustive set of states, u is the agent's utility function over the set of outcomes $a(s)$ which are the results of an act a in a state s (acts are identified with functions from states s to outcomes), and Pr is the agent's probability measure on a field over S (Savage 1972). From this decision-theoretic point of view all we need – besides our utilities – are our degrees of belief encoded in Pr . Degrees of confirmation encoding how much one proposition increases the probability of another are of no use here.

In the above examples, I only consider the propositions A, B, C , because they are sufficiently *informative* to answer my question. If truth were the only thing I am interested in, I would be happy with the tautological answer that *somebody* is the father of my friend's baby, $A \vee B \vee C$. But I am not. The reason is that I want to know what is going on out there – not only in the sense of having true beliefs, but also in the sense of having informative beliefs. In terms of decision theory, my decisions do not only depend on my degrees of belief – they also depend on my utilities. This is the idea behind the plausibility-informativeness theory (Huber 2008b), according to which epistemic utilities reduce to informativeness values. If we take as our epistemic utilities in the above example the informativeness values of the various answers (with positive probability) to our question, we get

$$I(A) = I(B) = I(C) = 1, I(A \vee B) = I(A \vee C) \approx 40/83, I(B \vee C) = 60/83, I(A \vee B \vee C) = 0,$$

where the question "Who is the father of my friend's baby?" is represented by the partition $Q = \{A, B, C\}$ and the informativeness values of the various answers are calculated according to

$$I(A) = 1 - [1 - \sum_i Pr^*(X_i|A)]/[1 - \sum_i Pr^*(X_i)],$$

a measure proposed by Hilpinen (1970). Contrary to what Hilpinen (1970, 112) claims, $I(A)$ does *not* increase with the logical strength of A . The probability Pr^* is the posterior degree of belief function from our example, $Pr(\cdot|G)$. If we insert these values into the expected utility formula,

$$EU(a) = \sum_{s \in \mathcal{S}} su(a(s)) \in Pr^*(s) = \sum_{X \in \mathcal{X}} Qu(a(X)) \in Pr^*(X) = \sum_{X \in \mathcal{X}} QI(X) \in Pr^*(X),$$

we get the result that the act of *accepting* A as answer to our question maximizes our expected epistemic utility.

Not all is lost, however. The distance measure d turns out to measure the expected utility of accepting H when utility is identified with informativeness measured according to a measure proposed by Carnap & Bar-Hillel (1953) (one can think of this measure as measuring how much an answer informs about the most difficult question, viz. which world is the actual one?). Similarly, the Joyce-Christensen measure s turns out to measure the expected utility of accepting H when utility is identified with informativeness about the data measured according to a proposal by Hempel & Oppenheim (1948). So far, this is only interesting. It gets important by noting that d and s can also be justified relative to the goal of *informative truth* – and not just by appealing to our intuitions about maximizing expected utility. When based on a regular probability, there almost surely is an n such that for all later m : relative to the conjunction of the first m data sentences, contingently true hypotheses get a positive value and contingently false hypotheses get a negative value. Moreover, within the true hypotheses, logically stronger hypotheses get a higher value than logically weaker hypotheses. The logically strongest true hypothesis (the complete true theory about the world w) gets the highest value, followed by all logically weaker true hypotheses all the way down to the logically weakest true hypothesis, the tautology, which is sent to 0. Similarly within the false hypotheses: the logically strongest false hypothesis, the contradiction, is sent to 0, followed by all logically weaker false hypotheses all the way down to the logically weakest false hypothesis (the negation of the complete theory about w). As informativeness increases with logical strength, we can put this as follows (assuming that the underlying probability measure is regular): d and s do not only distinguish between true and false theories, as do all confirmation measures (as well as all conditional probabilities). They additionally distinguish between informative and uninformative true theories, as well as between informative and uninformative false theories. In this sense, they reveal the following structure of almost every world w [$w(p) = w(q) = 1$ in the toy example]:

	informative and contingently true in w	$p \wedge q$
> 0	contingently true in w	$p, q, p \leftrightarrow q$
	uninformative and contingently true in w	$p \vee q, \neg p \vee q, p \vee \neg q$
$= 0$	logically determined	$p \vee \neg p, p \wedge \neg p$
	informative and contingently false in w	$\neg p \wedge \neg q, p \wedge \neg q, \neg p \wedge q$
< 0	contingently false in w	$\neg p, \neg q, p \leftrightarrow \neg q$
	uninformative and contingently false in w	$\neg p \vee \neg q$

This result is also true for the Carnap measure c , but it does not extend to all confirmation measures. It is false for the Milne measure r , which does not distinguish between informative and uninformative false theories. And it is false for the Good-Fitelson measure l , which distinguishes neither between informative and uninformative true theories nor between informative and uninformative false theories. For more see Huber (2005b).

The reason c , d , and s have this property of distinguishing between informative and uninformative truth and falsehood is that they are probabilistic assessment functions in the sense of the plausibility-informativeness theory (Huber 2008b) – and the above result is true for all probabilistic

assessment functions (not only those that can be expressed as expected utilities). The plausibility-informativeness theory agrees with traditional philosophy that truth is an epistemic goal. Its distinguishing thesis is that there is a second epistemic goal besides truth, namely, informativeness, which has to be taken into account when we evaluate hypotheses. Like confirmation theory, the plausibility-informativeness theory assigns numbers to hypotheses in the light of evidence. But unlike confirmation theory, it does not appeal to intuitions when it comes to the question why one is justified in accepting hypotheses with high assessment values. The plausibility-informativeness theory answers this question by showing that accepting hypotheses according to the recommendation of an assessment function almost surely leads one to (the most) informative (among all) true hypotheses (against, this can be seen as a soundness result). (The corresponding completeness result that *only* acceptances according to the recommendations of assessment functions almost surely lead to informative true hypotheses does not hold. For a discussion of this, see Huber 2008b, sec. 6.2.)

It is idle to speculate what Hume would have said to all this ado. Suffice it to note that his problem would not have got off the ground without our desire for informativeness.

8. References and Further Reading

Albert, Max (1992), "Die Falsifikation Statistischer Hypothesen." *Journal for General Philosophy of Science* 23, 1-32.

Alchourrón, Carlos E. & Gärdenfors, Peter & Makinson, David (1985), "On the Logic of Theory Change: Partial Meet Contraction and Revision Functions." *Journal of Symbolic Logic* 50, 510-530.

Carnap, Rudolf (1950/1962), *Logical Foundations of Probability*. 2nd ed. Chicago: University of Chicago Press.

Carnap, Rudolf (1952), *The Continuum of Inductive Methods*. Chicago: University of Chicago Press.

Carnap, Rudolf (1963), "Replies and Systematic Expositions. Probability and Induction." In P.A. Schilpp (ed.), *The Philosophy of Rudolf Carnap*. La Salle, IL: Open Court, 966-998.

Carnap, Rudolf & Bar-Hillel, Yehoshua (1953), *An Outline of a Theory of Semantic Information*. Technical Report 247. Research Laboratory of Electronics, MIT. Reprinted in Y. Bar-Hillel (1964), *Language and Information. Selected Essays on Their Theory and Application*. Reading, MA: Addison-Wesley, 221-274.

Christensen, David (1999), "Measuring Confirmation." *Journal of Philosophy* 96, 437-461.

Crupi, Vincenzo and Tentori, Katya, and Gonzalez, Michel (2007), On Bayesian Measures of Evidential Support: Theoretical and Empirical Issues. *Philosophy of Science* 74, 229-252.

Duhem, Pierre (1906/1974), *The Aim and Structure of Physical Theory*. New York: Atheneum.

Earman, John (1992), *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory*.

Cambridge, MA: MIT Press.

Eells, Ellery (2005), "Confirmation Theory." In J. Pfeifer & S. Sarkar (eds.), *The Philosophy of Science. An Encyclopedia*. Oxford: Routledge.

Field, Hartry (1978), "A Note on Jeffrey Conditionalization." *Philosophy of Science* 45, 361-367.

Fitelson, Branden (1999), "The Plurality of Bayesian Measures of Confirmation and the Problem of Measure Sensitivity." *Philosophy of Science* 66 (Proceedings), S362-S378.

Fitelson, Branden (2001), *Studies in Bayesian Confirmation Theory*. PhD Dissertation. Madison, WI: University of Wisconsin-Madison.

Fitelson, Branden (2002), "Putting the Irrelevance Back Into the Problem of Irrelevant Conjunction." *Philosophy of Science* 69, 611-622.

Fitelson, Branden (2005), "Inductive Logic." In J. Pfeifer & S. Sarkar (eds.), *The Philosophy of Science. An Encyclopedia*. Oxford: Routledge.

Fitelson, Branden & Hájek, Alan & Hall, Ned (2005), "Probability." In J. Pfeifer & S. Sarkar (eds.), *The Philosophy of Science. An Encyclopedia*. Oxford: Routledge.

Gaifman, Haim & Snir, Marc (1982), "Probabilities over Rich Languages, Testing, and Randomness." *Journal of Symbolic Logic* 47, 495-548.

Gärdenfors, Peter (1988), *Knowledge in Flux. Modeling the Dynamics of Epistemic States*. Cambridge, MA: MIT Press.

Gärdenfors, Peter & Rott, Hans (1995), "Belief Revision." In D.M. Gabbay & C.J. Hogger & J.A. Robinson (eds.), *Handbook of Logic in Artificial Intelligence and Logic Programming. Vol. 4. Epistemic and Temporal Reasoning*. Oxford: Clarendon Press, 35-132.

Glymour, Clark (1980), *Theory and Evidence*. Princeton: Princeton University Press.

Good, Irving John (1967), "The White Shoe is a Red Herring." *British Journal for the Philosophy of Science* 17, 322.

Good, Irving John (1968), "The White Shoe qua Herring is Pink." *British Journal for the Philosophy of Science* 19, 156-157.

Good, Irving John (1983), *Good Thinking: The Foundations of Probability and Its Applications*. Minneapolis: University of Minnesota Press.

Goodman, Nelson (1946), "A Query on Confirmation." *Journal of Philosophy* 43, 383-385.

Goodman, Nelson (1983), *Fact, Fiction, and Forecast*. 4th ed. Cambridge, MA: Harvard University Press.

Grimes, Thomas R. (1990), "Truth, Content, and the Hypothetico-Deductive Method." *Philosophy*

of Science 57, 514-522.

Hacking, Ian (2001), *An Introduction to Probability and Inductive Logic*. Cambridge: Cambridge University Press.

Hájek, Alan (2003a), "Interpretations of Probability." In E.N. Zalta (ed.), *Stanford Encyclopedia of Philosophy*.

Hájek, Alan (2003b), "What Conditional Probability Could Not Be." *Synthese* 137, 273-323.

Hájek, Alan (2005), "Scotching Dutch Books?" *Philosophical Perspectives* 19 (Epistemology), 139-151.

Hájek, Alan & Hall, Ned (2000), "Induction and Probability." In P. Machamer & M. Silberstein (eds.), *The Blackwell Guide to the Philosophy of Science*. Oxford: Blackwell, 149-172.

Hall, Ned (1994), "Correcting the Guide to Objective Chance." *Mind* 103, 505-518.

Hawthorne, James (2005), "Inductive Logic." In E.N. Zalta (ed.), *Stanford Encyclopedia of Philosophy*.

Hawthorne, James & Fitelson, Branden (2004), "Re-solving *Irrelevant Conjunction* with Probabilistic Independence." *Philosophy of Science* 71, 505-514.

Hempel, Carl Gustav (1945), "Studies in the Logic of Confirmation." *Mind* 54, 1-26, 97-121.

Hempel, Carl Gustav (1962), "Deductive-Nomological vs. Statistical Explanation." In H. Feigl & G. Maxwell (eds.), *Scientific Explanation, Space and Time. Minnesota Studies in the Philosophy of Science* 3. Minneapolis: University of Minnesota Press, 98-169.

Hempel, Carl Gustav (1967), "The White Shoe: No Red Herring." *British Journal for the Philosophy of Science* 18, 239-240.

Hempel, Carl Gustav & Oppenheim, Paul (1948), "Studies in the Logic of Explanation." *Philosophy of Science* 15, 135-175.

Hilpinen, Risto (1970), "On the Information Provided by Observations." In J. Hintikka & P. Suppes (eds.), *Information and Inference*. Dordrecht: D. Reidel, 97-122.

Hintikka, Jaakko (1966), "A Tw-Dimensional Continuum of Inductive Methods." In J. Hintikka & P. Suppes (eds.), *Aspects of Inductive Logic*. Amsterdam: North-Holland, 113-132.

Hitchcock, Christopher R. (2001), "The Intransitivity of Causation Revealed in Graphs and Equations." *Journal of Philosophy* 98, 273-299.

Howson, Colin (2000a), *Hume's Problem: Induction and the Justification of Belief*. Oxford: Oxford University Press.

Howson, Colin (2000b), "Evidence and Confirmation." In W.H. Newton-Smith (ed.), *A*

Companion to the Philosophy of Science. Oxford: Blackwell, 108-116.

Howson, Colin & Urbach, Peter (1989/2005), *Scientific Reasoning: The Bayesian Approach*. 3rd ed. La Salle, IL: Open Court.

Huber, Franz (2005a), "Subjective Probabilities as Basis for Scientific Reasoning?" *British Journal for the Philosophy of Science* 56, 101-116.

Huber, Franz (2005b), "What Is the Point of Confirmation?" *Philosophy of Science* 75, 1146-1159.

Huber, Franz (2008a) "Formal Epistemology." In E. N. Zalta (ed.), *Stanford Encyclopedia of Philosophy*.

Huber, Franz (2008b), "Assessing Theories, Bayes Style." *Synthese* 161, 89-118.

Hume, David (1739/2000), *A Treatise of Human Nature*. Ed. by D.F. Norton & M.J. Norton. Oxford: Oxford University Press.

Jeffrey, Richard C. (1965/1983), *The Logic of Decision*. 2nd ed. Chicago: University of Chicago Press.

Jeffrey, Richard C. (2004), *Subjective Probability: The Real Thing*. Cambridge: Cambridge University Press.

Jeffreys, Harold (1939/1967), *Theory of Probability*. 3rd ed. Oxford: Clarendon Press.

Joyce, James F. (1998), "A Non-Pragmatic Vindication of Probabilism." *Philosophy of Science* 65, 575-603.

Joyce, James F. (1999), *The Foundations of Causal Decision Theory*. Cambridge: Cambridge University Press.

Joyce, James M. (2003), "Bayes's Theorem." In E.N. Zalta (ed.), *Stanford Encyclopedia of Philosophy*.

Juhl, Cory (1997), "Objectively Reliable Subjective Probabilities." *Synthese* 109, 293-309.

Kelly, Kevin T. (1996), *The Logic of Reliable Inquiry*. Oxford: Oxford University Press.

Kelly, Kevin T. & Glymour, Clark (2004), "Why Probability does not Capture the Logic of Scientific Justification." In C. Hitchcock (ed.), *Contemporary Debates in the Philosophy of Science*. Oxford: Blackwell, 94-114.

Keynes, John Maynard (1921/1973), *A Treatise on Probability. The Collected Writings of John Maynard Keynes*. Vol. III. New York: St. Martin's Press.

Kolmogoroff, Andrej N. (1933), *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Berlin: Springer.

Kolmogorov, Andrej N. (1956), *Foundations of the Theory of Probability*, 2nd ed. New York:

Chelsea Publishing Company.

Koons, Robert (2005), "Defeasible Reasoning." In E.N. Zalta (ed.), *Stanford Encyclopedia of Philosophy*.

Kraus, Sarit & Lehmann, Daniel & Magidor, Menachem (1990), "Nonmonotonic Reasoning, Preferential Models, and Cumulative Logics." *Artificial Intelligence* 40, 167-207.

Kuipers, Theo A.F. (2000), *From Instrumentalism to Constructive Realism. On Some Relations between Confirmation, Empirical Progress, and Truth Approximation*. Dordrecht: Kluwer.

Kyburg, Henry E. Jr. (1961), *Probability and the Logic of Rational Belief*. Middletown, CT: Wesleyan University Press.

Lewis, David (1980), "A Subjectivist's Guide to Objective Chance." In R.C. Jeffrey (ed.), *Studies in Inductive Logic and Probability*. Vol. II. Berkeley: University of California Press, 263-293. Reprinted in D. Lewis (1986), *Philosophical Papers*. Vol. II. Oxford: Oxford University Press, 83-113.

Lewis, David (1994), "Humean Supervenience Debugged." *Mind* 103, 473-490.

Maher, Patrick (1999), "Inductive Logic and the Ravens Paradox." *Philosophy of Science* 66, 50-70.

Maher, Patrick (2004a), "Probability Captures the Logic of Scientific Confirmation." In C. Hitchcock (ed.), *Contemporary Debates in Philosophy of Science*. Oxford: Blackwell, 69-93.

Maher, Patrick (2004b), "Bayesianism and Irrelevant Conjunction." *Philosophy of Science* 71, 515-520.

Makinson, David (1994), "General Patterns in Nonmonotonic Logic." In D.M. Gabbay & C.J. Hogger & J.A. Robinson (eds.), *Handbook of Logic in Artificial Intelligence and Logic Programming*. Vol. 3. *Nonmonotonic Reasoning and Uncertain Reasoning*. Oxford: Clarendon Press, 35-110.

Milne, Peter (1996), " $\log[P(h|eb)/P(h/b)]$ is the One True Measure of Confirmation." *Philosophy of Science* 63, 21-26.

Moretti, Luca (2004), "Grimes on the Tacking by Disjunction Problem." *Disputatio* 17, 16-20.

Pearl, Judea (2000), *Causality: Models, Reasoning, and Inference*. Cambridge: Cambridge University Press.

Popper, Karl R. (1935/1994), *Logik der Forschung*. Tübingen: J.C.B. Mohr.

Putnam, Hilary (1963a), "Degree of Confirmation and Inductive Logic." P.A. Schilpp (ed.), *The Philosophy of Rudolf Carnap*. La Salle, IL: Open Court, 761-784. Reprinted in H. Putnam (1975/1979), *Mathematics, Matter and Method*. 2nd ed. Cambridge: Cambridge University Press,

270-292.

Putnam, Hilary (1963b), "Probability and Confirmation." *The Voice of America, Forum Philosophy of Science 10*, U.S. Information Agency. Reprinted in H. Putnam (1975/1979), *Mathematics, Matter and Method*. 2nd ed. Cambridge: Cambridge University Press, 293-304.

Quine, Willard Van Orman (1953), "Two Dogmas of Empiricism." *The Philosophical Review 60*, 20-43.

Quine, Willard van Orman (1969), "Natural Kinds." In N. Rescher et.al. (eds.), *Essays in Honor of Carl G. Hempel*. Dordrecht: Reidel, 5-23.

Reichenbach, Hans (1938), *Experience and Prediction. An Analysis of the Foundations and the Structure of Knowledge*. Chicago: University of Chicago Press.

Reichenbach, Hans (1940), "On the Justification of Induction." *Journal of Philosophy 37*, 97-103.

Rosenkrantz, Roger (1981), *Foundations and Applications of Inductive Probability*. New York: Ridgeview.

Roush, Sherrilyn (2005), "Problem of Induction." In J. Pfeifer & S. Sarkar (eds.), *The Philosophy of Science. An Encyclopedia*. Oxford: Routledge.

Savage, Leonard J. (1954/1972), *The Foundations of Statistics*. 2nd ed. New York: Dover.

Schulte, Oliver (2002), "Formal Learning Theory." In E.N. Zalta (ed.), *Stanford Encyclopedia of Philosophy*.

Skyrms, Brian (2000), *Choice and Chance. An Introduction to Inductive Logic*. 4th ed. Belmont, CA: Wadsworth Thomson Learning.

Spohn, Wolfgang (1988), "Ordinal Conditional Functions: A Dynamic Theory of Epistemic States." In W.L. Harper & B. Skyrms (eds.), *Causation in Decision, Belief Change, and Statistics II*. Dordrecht: Kluwer, 105-134.

Stalker, Douglas F. (ed.) (1994), *Grue! The New Riddle of Induction*. Chicago: Open Court.

Thau, Michael (1994), "Undermining and Admissibility." *Mind 103*, 491-504.

van Fraassen, Bas C. (1984), "Belief and the Will." *Journal of Philosophy 81*, 235-256.

van Fraassen, Bas C. (1995), "Belief and the Problem of Ulysses and the Sirens." *Philosophical Studies 77*, 7-37.

Vineberg, Susan (2005), "Dutch Book Argument." In J. Pfeifer & S. Sarkar (eds.), *The Philosophy of Science. An Encyclopedia*. Oxford: Routledge.

Vranas, Peter B.M. (2004a), "Have Your Cake and Eat It Too: The Old Principal Principle

Reconciled with the New.” *Philosophy and Phenomenological Research* 69, 368-382.

Vranas, Peter B.M. (2004b), “Hempel’s Raven Paradox: A Lacuna in the Standard Bayesian Solution.” *British Journal for the Philosophy of Science* 55, 545-560.

Woodward, James F. (2003), *Making Things Happen. A Theory of Causal Explanation*. Oxford: Oxford University Press.

Author Information:

Franz Huber

Email: franz@caltech.edu

California Institute of Technology

© 2008