*Erschienen in: Computational models of rationality : essays dedicated to Gabriele Kern-Isberner on the occasion of her 60th birthday / Beierle, Christoph; Brewka, Gerhard; Thimm, Matthias (Hrsg.). - London : College Publications, 2016. - (Tributes ; 29). - S. 112-131. - ISBN 978-1-84890-198-8*

# ENUMERATIVE INDUCTION[1]

*Wolfgang Spohn*

*Fachbereich Philosophie*
*Universität Konstanz*
*78457 Konstanz*
*Germany*

**Abstract:** This is a philosophical paper about enumerative induction, a basic rule of inductive or defeasible reasoning. It shows that it is almost satisfactorily accounted for in probability theory and even more adequately in ranking theory. As such it certainly deals with the foundations of formal rationality and should be of interest not only to philosophers, but also to the Artificial Intelligence community, which is deeply engaged in inductive reasoning as well.

## 1. Introduction

*Enumerative induction* is a kind of inference: from "the 1st $F$ is $G$, the 2nd $F$ is $G$, …, the $n$-th $F$ is $G$" or from "all observed $F$'s are $G$" infer "all $F$'s are $G$" or even "it's a law that all $F$'s are $G$". This is obviously not a deductively valid inference. Rather, it is a defeasible or nonmonotonic inference: add the premise "the $n$+1st or the next observed $F$ is not $G$", and you would reject the conclusion. So, is it an inductively valid inference? Hard to say; we have no good criterion of inductive validity. For some it is the most basic inductive inference of all. For some it is much too primitive and not anything we can rely on in our inductive practice. Still others think that it leads into contradictions straightaway.

In any case, it is a most suggestive and most venerable inductive inference rule. Since mankind can think, it uses this rule. We generalize all the time, we wouldn't survive without doing so, and this finds its immediate expression in this rule. As indicated, the rule is contested and has been much discussed. The present situation, though, is strangely silenced. I guess, most philosophers think that this rule has finally found its Bayesian home. And even though this home may not be fully comfortable, they feel we need no longer be concerned.

[1] I dedicate this paper to Gabriele Kern-Isberner. I am not well suited to grasp all the merits of her work. However, from my perspective I am very happy and grateful for the brilliant uses she has made of ranking theory invented by me in 1983, for instance in her Habilitationsschrift (1999), in Kern-Isberner (2004), and many papers thereafter.

Well, we shouldn't take this to be the end of the story. There is at least an afterword to it. As this paper will explain, the ranking-theoretic home is even more adequate. We will find that within this framework the above inferences hold good, even with the stronger nomological conclusion and without danger of contradiction. The afterword will not be particularly complicated. In my view, it's just a matter of an adequate conceptualization – which ranking theory is able to provide.

The plan of the paper is as follows: First, it should, and does, deliver a bit of background. Section 2 will give a very brief historical sketch of enumerative induction. This history is interwoven with the history of inductive skepticism, as indicated in section 3. Section 4 will explain how enumerative induction finally arrived at its Bayesian home, not without frictions, though. Whenever there is a probabilistic story, there is also an analogous ranking-theoretic story. This holds in the present case, too. So, section 5 will tell that analogous story, in which the Bayesian frictions will simply vanish. Section 6 concludes with two observations concerning Goodman's new riddle of induction and concerning the alleged apriority of the uniformity of nature.

The Bayesians already discovered that their (slightly distorted) version of enumerative induction is entailed by symmetry considerations (and minor additional premises). This entailment will stand out even more clearly in the ranking-theoretic story. Thus, enumerative induction is not a basic inductive inference, as it has seemed through centuries, but is derived from even more basic features of our inductive constitution.

## 2. A Few Historical Remarks About Enumerative Induction

Human thinking generalizes. And as soon as reflection sets in, one wonders what one is really doing there. So it is not surprising that this generalizing inference was an important topic already in early Indian philosophy. How early is hard to say, because there is often a big time gap between thinking and writing down in Indian philosophy. In any case, there was a debate how many positive instances were required for drawing this inference, and there was a doctrine that one positive instance, one $F$ that is a $G$, may suffice, at least when $F$ and $G$ stand for the right kind of universals. There was also a clear awareness of the danger of overgeneralization. The generalizations need to be hedged by something like varying the conditions and noting the ensuing differences. Only then one can hope to specify the right kind of $F$ (which may, of course, be logically complex taking account of all the conjectured necessary and sufficient conditions).[2]

[2] Cf., e.g., Smart (1967, pp. 164ff.)

The inference was also a concern in Greek philosophy. Aristotle discussed it in his *Topoi* and his *Analytica Priora* under the label *epagoge*. Being aware that it is not cogent, but only more or less persuasive, he tried to find conditions of admissibility. The inference was also an issue between the Epicureans and the Stoics. The more empirically inclined Epicureans tried to spell out conditions under which the inference acquires more certainty, while the more rationalistic Stoics emphasized its ineliminable uncertainty. It's no surprise that already the ancient Skeptics complained that the inference lacks any justification. We owe, by the way, the term *inductio* to Cicero who thus translated the Aristotelian term.[3]

The rediscovery of Greek philosophy in the Middle Ages naturally led to a rich discussion of enumerative induction in Scholastic philosophy.[4] Still, all those treatments appear quite academic. The topic acquired real methodological importance only with the rise of empirical science in the 16th century, and, moreover, theoretical wit with the rise of what can be properly called probability theory in the 17th century. And, no doubt, the herald of the new inductive methods was Francis Bacon.

Bacon, particularly in his (1620), developed quite sophisticated canons of the scientific method or of inductive reasoning, which served as a sort of blueprint for John Stuart Mill's much more elaborate theory in his (1843, Book III, chs. 8 – 10) more than 200 years later. One might say that Bacon's ideas were guided by trust and mistrust in enumerative induction alike. Unguarded generalization would be childish. Hence, he developed what he called the Table of Affirmation, the Table of Negation, and most importantly, the Table of Comparison, which Mill turned into this Method of Concomitant Variations. These refinements are required, since inductive generalizations are counteracted by eliminations through counter-instances. And this initiates a search for ever more detailed and accurate generalizations.[5]

In modern terms, not available at those times, one might say that the conclusion of enumerative induction is only a ceteris paribus law: ceteris paribus, all $F$'s are $G$. And then Bacon's and Mill's sophisticated methodology might be understood as trying to spell out how to substantiate this sweeping reference to ceteris paribus conditions. Not that this remark would clarify much; all the present treatments of ceteris paribus laws are just as tentative as those historical inductive canons.[6] It is noteworthy, though, that the problems addressed now and then are the same.

---

[3] Cf. Ruzicka (1976, pp. 323-325).

[4] Cf. Ruzicka (1976, pp. 326f.).

[5] Cf. also Cranston (1967, pp. 238f.).

[6] See Spohn (2014) which contains a very brief critical overview of those treatments as well as my explication of ceteris paribus conditions in ranking-theoretic terms, i.e., in terms to be applied here to enu-

The advent of probability theory definitely widened the perspective. Probabilistic inference was inductive inference *par excellence*. And the probabilistic generalization of enumerative induction was straightforward. It is called the *statistical inference* or the *straight rule* and says: if $m$ of the $n$ observed $F$'s are $G$, then infer that the (statistical) probability of an $F$ being $G$ is $m/n$. Surely, this is beset with all the problems of enumerative induction and more.

Still, the widening of the perspective was most important. The old writings appeared to conceive of induction only in the form of enumerative induction. So-called eliminative induction is also a very old idea, certainly to be found, e.g., already in Bacon's and definitely in Mill's methods, and gained central importance in Popper's (1934) philosophy of science. However, it is a deductive inference, a virtue Popper has continuously emphasized, with the merely negative conclusion that some generalizations are false. Thus, it was only in the 18th century, after the advent of probability theory, that methodologists started clearly discerning various forms of inductive inference. Bayesian hegemony, though, the idea that all forms of inductive reasoning can be reduced to probabilistic reasoning was a much later thought, taking shape only in the late 20th century.

At the same time, the multitude of inductive inferences casted doubt on the role of enumerative induction. Maybe it is just too primitive, and there are superior inductive methods? For instance, Peirce started conceiving of abduction; and nowadays, many philosophers of science favor the idea that the inference to the best explanation (IBE = abduction) is such a superior inductive inference, which is characteristic of the modern sciences.[7] Clarity about the nature of IBE is quite disproportionate, though, to this emphasis. Thus, calling enumerative induction primitive is at least ambiguous. It may mean: not sufficiently elaborated to be useful. But it may still mean: *basic* – so that it must be accounted for, before one can hope to do justice to more sophisticated forms of inductive inference.

Another issue came to the fore in the vigorous debate between so-called inductivists and deductivists in the 19th century, apparently not only between philosophers, but also within the scientific disciplines themselves. Inductivists recommended enumerative induction or some of its sophistications as a method of generating scientific hypotheses. Deductivists took this to be absurd; finding hypotheses is a matter of informed imagination, in the first place, and of checking then whether the deductive consequences are the desired ones. Newton could never have arrived at his laws by enumerative induction!

---

merative induction as well. The paper also contains a simple learning algorithm for ceteris paribus conditions, which I did not relate to the historical canons. Maybe one should.

[7] See, e.g., Lipton (1991).

Following Popper (1934), it was concluded in the 20th century that this debate was situated mainly within the context of discovery, which is the wrong context, anyway. The appropriate context rather is the context of justification. Induction should not be taken as a heuristic method; it is a matter of justifying or confirming the consequent by the premises. In fact, I perceive here an ambiguity in the notion of inference. Inference as a process might lead us to novel insights, whereas inference as a mere relation or connection between sentences or propositions invites us to assess the quality of that connection. I am entirely on the side of the latter interpretation. One must be aware, however, that it is quite a step to conceive of inference as confirmation.

## 3. Brief Remarks on Inductive Skepticism

The previous point reminds us of the fact that inductive inference was accompanied by inductive skepticism at all times. I have already mentioned the ancient Skeptics. The undeniable point is, simply, that the conclusion of an inductive inference might be false while its premises are true. This is indeed a defining characteristic of inductive or ampliative inference! So, what good reason is there to believe in the conclusion?

In our modern times it was David Hume who developed inductive skepticism in so masterly a manner, in his *Treatise* (1739) as well as in his *Enquiry* (1748). His basic point again was that inductive inference cannot be deductively cogent. However, he made very clear that there is also no way around this basic point. There is also no inductive justification of inductive inference. Such justification would just use and hence presuppose the forms of inference to be justified. Causal inference, another main target of Hume, is no better off than inductive inference in general.

A probabilistic interpretation of inductive inference doesn't help, either. We might say that an inductive conclusion is at least very probable. But in which sense? If "very probable" means "high relative frequency of truth", i.e., if "all observed *F*'s are *G*", is to entail at least "most *F*'s are *G*", this entailment is as unjustified as before. If, however, "very probable" signifies only our high confidence, nothing is gained; it is just the justification of this high confidence which is at issue.[8]

Hume finally observed that all would be fine if we could presuppose the uniformity of nature as a most general law. Yet, how could we ever assume such a law? We could arrive at it only by a higher-order inductive generalization. This is no solution, no way to ground our inductive inferences. And so Hume acquiesced in his inductive skepti-

[8] Cf. Salmon (1966, pp. 5ff. and pp. 48ff.)

cism. Inductive inferences are nothing but our habits of thought. We pursue the habits we have. What else could we do? But don't ask for justification.[9]

Kant tried to do better by trying to establish the uniformity of nature as an a priori principle of thought, without which any kind of experience would be impossible. In his terms this took the form of a general law of causality.[10] Let's not deepen the issue now. Surely, though, one might say that Hume convinced more philosophers than Kant did.

Later on it became clear that even the law of the uniformity of nature wouldn't help. This was the radicalization of Hume's inductive skepticism by Goodman (1946). He invented unnatural generalizations like "all emeralds are grue". His point then was that there are countless diverging generalizations that all agree with the observed facts. Suppose that $G$ and $G'$ disagree only for unobserved $F$'s and that all observed $F$'s were $G$ as well as $G'$. Hence, by enumerative induction, all $F$'s are $G$ as well as $G'$. But this can't be, provided there are $F$'s as yet unobserved. So, which of the two generalizations should we prefer? It won't do to call $G$ natural and $G'$ unnatural. That's precisely the issue. Thus the uniformity of nature may take countless shapes, and we have apparently no criterion for conjecturing rather this than that shape. In other words, not only are all inductive inferences deductively invalid, they all seem equally bad.[11]

Goodman's solution of his 'new riddle of induction' was roughly the same as Hume's. He also referred to the habits of thought or rather to the entrenched social practices. The riddle provoked an intense discussion. However, I haven't seen anything moving essentially beyond the original skeptical solution.[12]

What's the present status of inductive skepticism? My attitude – which seems widespread, though I am perhaps overoptimistic – is this: Basically, one has to accept Hume's skeptical solution. Just referring to the habits of thought, to social practice, etc. is, however, too psychologistic, too empirically minded, more defeatist than necessary. There are quite a number of normative principles or rationality postulates guiding our inductive behavior. And those principles have strong normative foundations; at least they allow for reasonable and sophisticated discussion. They need not uniquely deter-

---

[9] This is the quite common psychologistic interpretation of Hume. I don't think that it is really fair to Hume. I rather see the exercise of reason in his ‚associations' and ‚habits of thought', an exercise that is clearly rationally reconstructible.

[10] This is the famous second analogy in Kant (1781/87, B232), which states the a priori principle: "All alterations take place in conformity with the law of connection of cause and effect". This entails the uniformity of nature: "That which follows or happens must follow according to a universal rule from that which was contained in the previous state" (B245).

[11] Quine (1960, §17) and elsewhere made the insightful remark that all language use builds on a two-fold induction concerning nature as well as our own behavior or concerning the reference of signs and the signs themselves. Apply skepticism to this double induction, and you end up with meaning skepticism à la Kripke (1982).

[12] Freitag (2015), though, is one of the cutest recent contributions.

mine our inductive behavior. But they provide severe constraints, and how far-reaching and consequential they are is an open, fruitful, and constructive issue. I will return to this below.

## 4. The Present Status of Enumerative Induction

If this is really the present representative attitude concerning (skepticism about) inductive inference in general, what does this specifically entail for enumerative induction? What is its present status?

The first observation is a bit surprising, I find: The last 40 years have seen an explosion of formal inductive theorizing. For centuries probability theory was the only game in town, and now there are default logic, belief revision theory, Dempster-Shafer belief functions, possibility theory, formal learning theory, ranking theory, indeed all kinds of defeasible or non-monotonic reasoning.[13] However, to the best of my knowledge these theories hardly took a stance towards enumerative induction; this was simply no topic for them. This is a blunt contrast to the fact that in all the past centuries inductive inference was mainly about enumerative induction.

Well, there are a few exceptions. Ranking theory is one; I'll come to this. Pollock (1990) just accepted the statistical inference – as mentioned, the straightforward probabilistic generalization of enumerative induction – among his representative collection of basic defeasible inference rules without deepening the insight in this inference. Baumgartner (2009) promotes a research tradition proposing algorithms for projecting causal structures or deterministic generalizations from singular data; this may indeed be called sophisticated enumerative induction.[14] And most importantly, formal learning theory as initiated by Kevin Kelly is foremost a most elaborate account of enumerative induction; this is still an active research program[15], which I cannot further discuss here.

These exceptions are definitely worth attending, but they still seem minority projects. Therefore it is entirely appropriate, I think, to say that the present status of enumerative induction is represented by its Bayesian account. What is this account? Let me first introduce some terminology so that we can be a bit more precise:

Let us replace the language of first-order logic informally used so far by the set-theoretic language of variables. Let $X$ be some generic variable taking values in some range $V$. And Let $X_1$, $X_2$, … be an infinite series of possible realizations of $X$. Thus $X_1$,

---

[13] Halpern (2003) gives a beautiful systematization of a broad spectrum of theories, and Huber, Schmidt-Petri (2009) provide a very useful anthology.

[14] Again, this should, but hasn't been compared with the algorithm proposed in Spohn (2014).

[15] Cf, e.g., Kelly (2008).

$X_2$, … is an infinite series of variables all taking values in $V$. For instance, we could consider an infinite series of $F$'s and let $X_n$ take value 1 or 0 according to whether or not the $n$-th $F$ is $G$. Or we could consider an infinite series of objects or experiments and let $X_n$ take values 1, …, 4 according to whether the $n$-th object has, or the $n$-th experiment results in, $F$ & $G$, $F$ & non-$G$, non-$F$ & $G$, or non-$F$ & non-$G$. The generalization that all $F$'s are $G$ then translates into the assertion that none of the variables $X_n$ takes value 2. Or $X_n$ could represent which value the $n$-th particle takes in some state space $S$ or which trajectory in $S^T$ it takes through $S$ during the times in $T$. So, we observe the behavior of the first $n$ variables $X_1, \ldots, X_n$, and the issue of enumerative induction is what to infer from this concerning the behavior of the further variables $X_{n+1}, X_{n+2}, \ldots$

Now, a possible world, or a possible course of events, as far as it can be represented by the variables, is just a sequence of values $v_1, v_2, \ldots$ in $V$ which the variables $X_1, X_2, \ldots$ might take. For convenience I shall assume that $V$ is finite (but not a singleton, of course). Let **N** be the set of non-negative integers, $\mathbf{N}^+$ the set of positive integers, and $\mathbf{N}^\infty = \mathbf{N} \cup \{\infty\}$. So $W = V^{\mathbf{N}^+}$ may be taken as the set of possible worlds or possible courses of events. Even if $V$ is finite, $W$ is very rich, indeed uncountable. $W$ may also be taken as the domain of the variables $X_1, X_2, \ldots$. Then, for any $\mathbf{v} = (v_1, v_2, \ldots) \in W$, we may define $X_n(\mathbf{v}) = v_n$. Let $\{X_n = v\} = \{\mathbf{v} \in W \mid X_n(\mathbf{v}) = v\}$ and $\{X_n \in U\} = \{\mathbf{v} \in W \mid X_n(\mathbf{v}) \in U\}$, respectively, represent the proposition that $X_n$ takes the value $v$ or some value in $U \subseteq V$. All these *atomic* propositions generate a σ-algebra $\mathcal{A}$ of propositions over $W$. Our study of enumerative induction must hence focus on the general propositions $G_U = \bigcap_{n \in \mathbf{N}^+} \{X_n \in U\}$ that all variables take values in $U$.

This is all the algebraic material we need. So, what is the Bayesian account of enumerative induction? As a first step this means to study the issue entirely in terms of subjective probabilities, i.e., in terms of a (σ-additive) probability measure $P$ on $\mathcal{A}$ and its rational behavior. I have already mentioned a second step: Within the context of justification, inductive inference should be interpreted as confirmation. A third step then is Carnap's explication of confirmation as probabilistic positive relevance.[16] Thus, enumerative induction turns into the following claim: The proposition $\{X_1 \in U\} \cap \ldots \cap \{X_n \in U\}$ is positively relevant to the proposition $G_U$.

At this point, an obstacle emerged: the problem of the null confirmation of laws. Under the assumptions of Carnap's so-called λ-continuum of inductive methods (which includes symmetry as introduced below) each infinite generalization $G_U$ (for $U \subset V$) provably receives probability 0.[17] Hence, no proposition can be positively relevant to any contingent infinite generalization.

---

[16] See Carnap (1950/62, chs. VI and VII).

[17] CF. Carnap (1950/62, § 110F).

Hintikka (1966) ingeniously circumvented this problem with his so-called two-dimensional continuum of inductive methods. Here, universal generalizations receive a positive a priori probability and can thus be confirmed by positive instances.[18] Somehow, though, Hintikka's ideas have not been well received. I am wondering why; I don't know of any telling refutation. Perhaps my impression was a shared one, namely that Hintikka's two-dimensional continuum was designed ad hoc in order to yield the intended results. Probabilities are anchored in reality in relative frequencies. So, inductive probabilities are made for somehow estimating or approaching relative frequencies. Then, however, there does not seem to be any good reason for favoring extreme relative frequencies, 1 and 0, in such a way that they get a positive a priori weight, whereas otherwise only intervals of relative frequencies get positive a priori weight. Granted, strict laws are peculiar. However, I think the specific characteristics of strict laws are better captured in the picture developed below.

Perhaps, though, it was the strong conception of inductive logic in general which fell out of favor, and with it Hintikka's proposal. Be this as it may, the main probabilistic line was simply Carnap's: If the infinite generalization has probability 0, who cares about the infinite generalization? What counts are the cases within our life span or, simply, the next instance. Thus, in a fourth step, Carnap transformed enumerative induction into his principle of positive instantial relevance: $\{X_1 \in U\} \cap \ldots \cap \{X_n \in U\}$ is positively relevant to $\{X_{n+1} \in U\}$. Or, more generally, however $X_1, \ldots, X_n$ realize, $\{X_{n+1} \in U\}$ is positively relevant to $\{X_{n+2} \in U\}$. Quite some transformation!

Let's be a bit more exact and distinguish various notions of instantial relevance. For any possible world $\mathbf{v} = (v_1, v_2, \ldots)$ let's abbreviate $\{X_1 = v_1\} \cap \ldots \cap \{X_n = v_n\}$ by $E_n(\mathbf{v})$. Then $P$ satisfies $\text{PIR}_\text{n}$ (the *principle of positive instantial relevance, nonconditional version*) iff for any possible evidence $E_n(\mathbf{v})$ and any non-empty set of values $U \subset V$

$$P(\{X_{n+2} \in U\} \mid \{X_{n+1} \in U\} \cap E_n(\mathbf{v})) > P(\{X_{n+2} \in U\} \mid E_n(\mathbf{v})).$$

And $P$ satisfies $\text{PIR}_\text{c}$ (the *principle of positive instantial relevance, conditional version*) iff for any such $E_n$ and $U$

$$P(\{X_{n+2} \in U\} \mid \{X_{n+1} \in U\} \cap E_n(\mathbf{v})) > P(\{X_{n+2} \in U\} \mid \{X_{n+1} \notin U\} \cap E_n(\mathbf{v})).$$

Clearly, in the probabilistic case $\text{PIR}_\text{n}$ and $\text{PIR}_\text{c}$ are equivalent. Still, I have introduced the distinction, because it will make a difference in the ranking-theoretic case. Moreover, let's say that $P$ satisfies $\text{NNIR}_\text{n}$ (the *principle of non-negative instantial relevance, nonconditional version*) iff $P$ satisfies the inequality for $\text{PIR}_\text{n}$ when $>$ is replaced by $\geq$.

[18] See also Kuipers (1978) for an elaborate study of Hintikka's theory.

And similarly, $P$ satisfies $NNIR_c$ (the *principle of non-negative instantial relevance, conditional version*) iff $P$ satisfies the inequality for $PIR_c$ when > is replaced by ≥. Finally, $P$ satisfies $IR_n$ or, respectively, $IR_c$ (*instantial relevance*) iff > is replaced by ≠ in the relevant inequalities.

So, the upshot of the Bayesian transformation so far is that enumerative induction is explicated as the probabilistic PIR (= $PIR_n$ or $PIR_c$). The crux of the Bayesian account lies now in the fifth and final step: Enumerative induction need not be axiomatically assumed as a basic inductive rule, as it seemed all the centuries before. Rather, it is entailed and hence justified by more basic assumptions. The crucial assumption is *symmetry* or *exchangeability*: $P$ is *symmetric* (with respect to $X_1, X_2, \ldots$) iff for any $n \in \mathbf{N}^+$, any permutation $\pi$ of $\{1, \ldots, n\}$, and any (not necessarily different) values $v_1, \ldots v_n$ in $V$

$$P(\{X_{\pi(1)} = v_1\} \cap \ldots \cap \{X_{\pi(n)} = v_n\}) = P(\{X_1 = v_1\} \cap \ldots \cap \{X_n = v_n\}).$$

According to symmetry, what counts probabilistically is only how many variables take which value. Which specific variables do so, does not make any probabilistic difference. This is extremely plausible when we miss any special information about specific variables (or the objects or experiments they represent).[19] So I need not repeat here the overwhelming credibility of this assumption.

The first result now is that symmetry entails NNIR (= $NNIR_n$ or $NNIR_c$). So, we may secondly conclude that symmetry plus IR (= $IR_n$ or $IR_c$) entail PIR, i.e., the Bayesian version of enumerative induction. Why assume IR? This is overwhelmingly plausible as well; without IR we could not learn anything at all from our observations. However, we could also argue that IR is entailed by the so-called Reichenbach axiom, which says this: For any possible world $\mathbf{v} = (v_1, v_2, \ldots)$ and any value $v \in V$ let $rf(\mathbf{v}, v)$ denote the relative frequency with which $v$ occurs among the first $n$ values $v_1, \ldots v_n$. Then

$$\lim_{n\to\infty} [P(\{X_{n+1} = v\} \mid E_n(\mathbf{v})) - rf(\mathbf{v}, v)] = 0.$$

In other words, the probability that the next variable will take the value $v$ converges to the relative frequency of $v$ with increasing evidence. Again such limiting learning behavior seems reasonably required. For all these results see, e.g., Humburg (1971).

As explained there, these results ultimately ground in de Finetti's fundamental representation theorem from (1937), which says that any symmetric probability measure over $\mathcal{A}$ is a *unique* mixture of *Bernoulli measures* over $\mathcal{A}$, according to which all variables $X_n$

[19] This kind of symmetry is also assumed in all conceptions of inductive logic. Above, though, when saying that inductive logic fell out of favor I referred to stronger conceptions which assume symmetry not only with respect to objects, but also with respect to predicates, i.e., in our terminology, with respect to the possible values in $V$. This is indeed very questionable, but not relevant here.

have the same distribution and each variable $X_n$ is probabilistically independent of all the other variables. We might conceive of such a Bernoulli measure as a statistical hypothesis about the objective probabilities governing $X_1$, $X_2$, … In this interpretation, which is not de Finetti's, the representation theorem says that our subjective assessment $P$, if symmetric, uniquely corresponds to a second order distribution over (= mixture of) possible statistical hypotheses. The final twist then is that we can verify from the beginning whether our $P$ has the appropriate limiting behavior as required by the Reichenbach axiom. That is, if the so-called carrier of that second-order distribution (= the smallest topologically closed set having measure 1) corresponding to $P$ is the space of all Bernoulli measures over $\mathcal{A}$, then $P$ satisfies the Reichenbach axiom.[20]

All in all, this is a beautiful justificatory story for the Bayesian version of enumerative induction and thus indeed huge progress. In view of this story I am no longer impressed by Hume's skepticism. We have now found good reasons why we must assume PIR, and we may do without an argument to the effect that higher subjective probability somehow entails higher relative frequency of truth. And we should follow the good reasons. Indeed, we even know that our probabilities will converge to the relative frequency of truth, although we must grant that at no point can we be sure where the point of convergence comes to lie.

In this way, the topic seems to have come to a rest, and we may be content with the Bayesian account of enumerative induction. Really?

## 5. A More Adequate Ranking-Theoretic Account

I don't think that we should be fully satisfied. We need not put up with all the Bayesian transformations. In particular, with the step to PIR, positive *instantial* relevance, we have lost all reference to generality so characteristic of enumerative induction. Indeed, enumerative induction is originally foreign to probabilistic epistemology, and not only because it is much older. Its Bayesian naturalization is only due to the fact that confirmation theory took an exclusively probabilistic shape. This was not always so. Hempel (1945), the modern classic of confirmation theory, started a research program of so-called qualitative confirmation theory. However, this program turned out to be infeasible. At least, this was the conclusion beautifully summarized in Niiniluoto (1972). Since, we are left with a probabilistic confirmation theory.

I think that the abandoning of Hempel's program was premature. Confirmation theory can also be developed within ranking theory, and this comes much closer to

[20] For all this, see again Humburg (1971).

Hempel's original intentions of a qualitative confirmation theory. There is no place here to develop this claim on a larger scale. Let us see, though, how enumerative induction fares within ranking-theoretic confirmation theory. So, we still conceive of inductive inference as confirmation, and we may still stick to Carnap's explication of confirmation as positive relevance. However, we now replace probabilistic by ranking-theoretic positive relevance. How does this work in detail?

We deal with the same atomic propositions as before. However, we now assume $\mathcal{A}$ to be the complete algebra generated by these atomic propositions. Next, instead of the probability measure $P$ we consider a negative ranking function $\kappa$ on $\mathcal{A}$. $\kappa$ is *negative ranking function* on $\mathcal{A}$ iff $\kappa$ is a function from $\mathcal{A}$ into $\mathbf{N}^{\infty}$ such that for all $A, B \in \mathcal{A}$: (a) $\kappa(W) = 0$ and $\kappa(\emptyset) = \infty$, (b) $\kappa(A \cup B) = \min \{\kappa(A), \kappa(B)\}$ (*minimitivity*). Moreover, let's assume that $\kappa$ is *completely minimitive*, i.e., that for all $\mathcal{B} \subseteq \mathcal{A}$ $\kappa(\bigcup \mathcal{B}) = \min_{B \in \mathcal{B}} \kappa(B)$.[21]

If the algebra $\mathcal{A}$ were finite, there would be no point in considering complete minimitivity. However, we attend to generalizations, which are infinite Boolean combinations of atomic propositions (and we need not restrict ourselves to countable combinations, as we actually do). Hence, we must take a stance on how a ranking function should behave vis à vis such infinite propositions. I think it is reasonable to assume complete minimitivity, as I have more extensively defended in Spohn (2012, pp. 73f.).[22]

The standard interpretation of a negative ranking function $\kappa$ is as degrees of disbelief (where disbelieving means taking to be false). This is why those functions are called negative. They don't take negative values, but their positive values express negative facts. Hence, I disbelieve $A$ iff $\kappa(A) > 0$, and I believe $A$ iff $\kappa(\bar{A}) > 0$.[23] Still, my belief can be more or less firm, as expressed by the ranks. It is convenient to define two-sided ranks. The *two-sided ranking function* $\tau$ belonging to the negative ranking function $\kappa$ is defined by $\tau(A) = \kappa(\bar{A}) - \kappa(A)$. Thus, $A$ is taken to be true or false or neither according to whether $\tau(A) > 0$ or $\tau(A) < 0$ or $= 0$.

*Conditional negative ranks* are defined as $\kappa(B \mid A) = \kappa(A \cap B) - \kappa(A)$, provided $\kappa(A) < \infty$. Thereby, minimitivity can be expressed as saying $\min \{\kappa(A \mid A \cup B), \kappa(B \mid A \cup B)\} = 0$. This means that given the disjunction you can't take both disjuncts to be false; this

---

[21] This definition goes back to my Habilitationsschrift Spohn (1983, sect. 5.3.) Its first appearance in English is in Spohn (1988), where negative ranking functions were still called ordinal conditional functions. Theory and applications of these functions are comprehensively presented in Spohn (2012).

[22] There would have been no point in considering $\sigma$-minimitivity, since it is equivalent to the apparently stronger complete minimitivity. And then it also fits better to build ranking functions on complete instead of $\sigma$-algebras. For all these niceties see Spohn (2012, pp. 72ff.).

[23] One may also define a stricter notion of belief by saying that $A$ is believed iff $\kappa(\bar{A}) > z$ for some threshold $z > 0$. This well accounts for the vagueness of the notion of belief (or disbelief). However, whatever the threshold, belief is always consistent and deductively closed. See Spohn (2012, pp. 76f.) for details. Here, we may well neglect this point.

is obviously rationally mandated. In my view, this clearly extends to infinite disjunctions, i.e., to complete minimitivity. Finally, conditional two-sided ranks are defined as $\tau(B \mid A) = \kappa(\bar{B} \mid A) - \kappa(B \mid A)$. For all fuller explanations of the basics of ranking theory I must refer to Spohn (2012, ch. 5).

Now we are prepared to study enumerative induction in ranking-theoretic terms.[24] Again, I assume that we only deal with symmetric ranking functions, for the same overwhelming reasons as in the probabilistic case. κ is *symmetric* iff κ satisfies the above symmetry condition with κ replacing $P$, i.e., if for any $n \in \mathbf{N}^+$, any permutation π of $\{1, \ldots, n\}$, and any values $v_1, \ldots v_n$ in $V$

$$\kappa(\{X_{\pi(1)} = v_1\} \cap \ldots \cap \{X_{\pi(n)} = v_n\}) = \kappa(\{X_1 = v_1\} \cap \ldots \cap \{X_n = v_n\}).$$

Then we have a first nice surprise: the credibility of a generalization is the very *same* as that of its instances. More formally, if κ is symmetric, then $\tau(G_U) = \tau(X_n \in U)$ for all $n \in \mathbf{N}^+$. This is a direct consequence of complete minimitivity. Or more generally: if $G_{>n,U} = \bigcap_{k>n} \{X_k \in U\}$ is the relevant generalization restricted to the future and if $E_n(\mathbf{v})$ is any evidence about the first $n$ variables, then $\tau(G_{>n,U} \mid E_n(\mathbf{v})) = \tau(X_k \in U \mid E_n(\mathbf{v}))$ for all $k > n$.[25] Of course, this holds not only for the ranks specified, but also for rank comparisons as required for confirmatory relations. Sloppily stated, this means that generalization is automatically built in into symmetric ranking functions as a consequence of the basic rationality postulates of ranking theory. Hence, we need not despair of the 'null confirmation' of laws, we need not take Carnap's escape route to instantial relevance, and we need not choose ad hoc measures à la Hintikka (1966). I take this to be a first important advantage of the ranking-theoretic account.

The next surprise, however, is less pleasant. We can transfer all the above notions of (positive, non-negative) instantial relevance in the nonconditional or the conditional version to ranking theory simply by substituting the two-sided ranking function τ (not the negative ranking function κ) for $P$ in the defining conditions above. (Relevance is more succinctly expressible in terms of τ.) So, we know what $\text{PIR}_n$, $\text{PIR}_c$, $\text{NNIR}_n$, and $\text{NNIR}_c$ mean in ranking theoretic terms. Let us also assume that our ranking function κ is *regular* in the sense that $\kappa(E_n(\mathbf{v})) < \infty$ for all $n \in \mathbf{N}^+$ and $\mathbf{v} \in W$. So, ranks conditional on $E_n(\mathbf{v})$ are always defined. Only then are we guaranteed to be able to learn from any kind of evidence. Then we have first to observe that, in contrast to the probabilistic

[24] The basic reason why the ranking-theoretic story will be similar to the Bayesian story is quite obvious from the axioms, according to which the minimum, the sum, and the difference of ranks, respectively, roughly correspond to the sum, the product, and the quotient of probabilities. For the precise formal relation between ranks and probabilities see Theorem 10.1 in Spohn (2012, pp. 203f.) which explains the formal similarities as well as the subtle formal differences between the two theories.

[25] For the simple proof see Spohn (2012, p. 282).

case, the conditional and nonconditional versions of these notions are not equivalent. Rather, $PIR_n$ entails $PIR_c$, and $NNIR_c$ entails $NNIR_n$, but the reverse entailments do not hold. So, we have to make a choice. As I have argued extensively in Spohn (2012, pp. 106f.), confirmation (or the reason relation, as I call it there) is more adequately captured by the conditional versions. So, is $PIR_c$ ranking-theoretically entailed by symmetry, as it is probabilistically? If so, the above extension to future generalizations $G_{>n,U}$ would already conclude our business?

On the contrary; this is the unpleasant surprise. We have the following theorem: There is no regular symmetric ranking function on $\mathcal{A}$ satisfying $PIR_c$.[26] We might settle for the weaker $NNIR_c$, which is satisfiable by regular symmetric ranking functions. However, this doesn't look attractive. $NNIR_c$ looks too weak; instantial irrelevance most of the time wasn't what we expected to get. Moreover, in contrast to the probabilistic case, $NNIR_c$ is not entailed by ranking-theoretic symmetry; we would have to additionally stipulate it.

Something seems to have gone badly wrong. The solution I have pursued in Spohn (2012, sect. 12.4 – 5) is to sharply distinguish between generalizations and laws. Generalizations or regularities are simply propositions, members of the propositional algebra, of the form $G_U$ or $G_{>n,U}$. Laws, by contrast, are something entirely different. The issue is embedded in a large issue of philosophy of science. There are mere regularities like "all gold spheres are smaller than one mile in diameter" (which may well be true) and true laws like "all uranium spheres are smaller than one mile in diameter".[27] What distinguishes lawlike sentences from mere generalizations (whether true or false)?[28] This has proved to be a remarkably recalcitrant problem. Early attempts at this distinction all failed. Thus, it has become apparent that lawlikeness or nomicity is connected with explanatory force, with entailing counterfactual conditionals, and with inductive behavior – three very soft topics in philosophy of science. It is a huge task to sort out all these connections. For us, only the last point is relevant. Somehow, it seems that enumerative induction doesn't extend to all generalizations whatsoever, but applies only to laws or potential laws. This is certainly what all philosophers from earlier centuries would have said who were not aware of the intricacies of that distinction.

So, it seems that we have made a mistake so far by trying to ranking-theoretically reconstruct enumerative induction for generalizations. We should restrict it to laws. But what are laws? We can't evade engaging into this distinction. However, here I have to

---

[26] This is Theorem 12.9 of Spohn (2012, p. 283). It is not trivial at all. For a proof see there p. 298.

[27] The example is from van Fraassen (1989, p. 27), who attributes it to the philosophy of science folklore of the 1960s.

[28] For an in-depth discussion of the issue see, e.g., Lange (2000).

cut a long story short.[29] The fundamental point is that laws are not propositions at all, contrary to what the mainstream concerning the issue has taken for granted! Laws do not make assertions and do not have truth conditions. This seems to be a bizarre claim; no wonder that the mainstream has not taken it seriously. Still, the claim has respectable philosophical precedence starting with Ramsey (1929).[30]

So, if laws are not propositions, what are they? The general, though obscure slogan is: laws rather are inference tickets. (And here you may again read "inference" as "confirmation"!). What is my ranking-theoretic translation of that slogan and the Ramsey quote? Basically, the idea is very simple: We mentioned above that in the probabilistic case statistical laws or hypotheses for $\mathcal{A}$ are represented by Bernoulli measures for $\mathcal{A}$. Each single case is characterized by a certain (objective) probability distribution, and this is turned into a law in the independent, identically distributed repetition of the siingle case as represented by the corresponding Bernoulli measure. So, correspondingly, what is a deterministic law? Just the corresponding ranking-theoretic notion. That is, I explicate that the (negative) ranking-function $\lambda$ is a *subjective law* for $\mathcal{A}$ iff all variables $X_n$ have the same distribution (in terms of ranks) and each variable $X_n$ is ranking-theoretically independent of all the other variables according to $\lambda$. (The latter boils down to instantial irrelevance: for all $n$ and $\mathbf{v}$, given $E_n(\mathbf{v})$ $X_{n+1}$ is irrelevant to $X_{n+2}$.) The analogy to statistical laws is charming; for a philosophical defense of this explication I have to refer to Spohn (2012, sect. 12.4).[31]

Subjective laws are related to generalizations. Let $\lambda$ be such a law and define $U = \{v \mid \lambda(X_n = v) = 0\}$. Then $\lambda$ contains the belief in the generalization $G_U$ and in no stronger generalization; that is, $U$ is the largest subset of $V$ such that $\lambda(\bar{G}_U) > 0$. But of course this belief may be realized in ranking functions in many different ways. Hence, laws represent a very specific attitude towards generalizations.

I have used a disturbing term by calling such a $\lambda$ a *subjective* law. This is owed to the fact that ranking functions still represent only the epistemic state of some epistemic subject. So, such a law $\lambda$ is just an epistemic attitude, too. There are ways, though, to objectivize this notion and to turn some subjective laws into objective ones. However, this remark can only be clarified by engaging into what I call the objectivization theory for

---

[29] The longer story is told in Spohn (2012, sect. 12.4).

[30] My key witness are the following quotes from Ramsey (1929): "Many sentences express cognitive attitudes without being propositions; and the difference between saying yes or no to them is not the difference between saying yes or no to a proposition" (pp. 135f.). "Laws are not either" (namely propositions) (p. 150). Rather "the general belief consists in (a) A general enunciation, (b) A habit of singular belief" (p. 136).

[31] Clearly, though, the independent, identically distributed repetition may be taken as a mathematical explication of Ramsey's 'habit of singular belief'.

ranking functions, which I develop in Spohn (2012, ch. 15). Here, we better leave this issue aside and go along with the notion of a subjective law.

Now we can proceed as in the probabilistic case by transferring de Finetti's representation theorem to ranking theory and getting all its benefits. Let me mention first the positive results and then the caveats.[32] We can indeed prove that each regular symmetric ranking function is a unique mixture of subjective laws. Well, roughly; for the caveats see below. That is, if $\Lambda$ is the set of regular subjective laws for $\mathcal{A}$ and $\rho$ is any ranking function for $\Lambda$, then the mixture $\kappa$ of $\Lambda$ by $\rho$ is a regular symmetric ranking function for $\mathcal{A}$, where this mixture is defined by $\kappa(A) = \min \{\lambda(A) + \rho(\lambda) \mid \lambda \in \Lambda\}$ for all $A \in \mathcal{A}$. I call $\rho$ an *impact function*, since it tells which impact each subjective law has on the mixture. And reversely – that's the difficult part – if $\kappa$ is a regular symmetric ranking function for $\mathcal{A}$, then there is a unique impact function $\rho$ such that $\kappa$ is the mixture of $\Lambda$ by $\rho$.

The next point in my dialectics is this: So far, we have notions of relevance and confirmation only for propositions; evidence may or may not confirm generalizations. However, if laws are not propositions, these notions cannot be applied to laws. The impact functions fill the gap: Let our initial symmetric $\kappa$ be represented by the impact function $\rho$. We learn by observing the first $n$ variables and conditionalizing our $\kappa$ on those observations. The conditionalized $\kappa$ – let's call it $\kappa'$ – will no longer be symmetric concerning the first $n$ variables, but it will stay symmetric with respect to all future variables $X_{n+1}$, $X_{n+2}$, … The point now is that $\kappa'$ will be represented by a different impact function $\rho'$. Thus, the impact of a subjective law $\lambda$ will change or possibly stay the same from $\rho$ to $\rho'$, and accordingly we can say that $\lambda$ has been confirmed or disconfirmed or neither.

More precisely, we find this: Let $\kappa_n$ be the ranking function reached after observing the first $n$ variables. Now we observe $X_{n+1}$ to take the value $v$ and thus move to $\kappa_{n+1}$. Let the corresponding impact functions be $\rho_n$ and $\rho_{n+1}$. Then for any subjective law $\lambda$ we have:

$$\rho_{n+1}(\lambda) - \rho_n(\lambda) = \lambda(X_{n+1} = v) - \kappa_n(X_{n+1} = v).$$ [33]

That is, the impact of $\lambda$ is upgraded or downgraded precisely to the extent to which the credibility it gives to the observed value of $X_{n+1}$ deviates from the credibility this value has according to $\kappa_n$.

This is my ranking-theoretic account of enumerative induction. It does not refer to the next single case, as did the Bayesian account. It even does not refer to generaliza-

[32] The full story is mathematically involved. Together with all the proofs it is presented in Spohn (2012, sect. 12.5).

[33] This is Theorem 12.22 of Spohn (2012, p. 299).

tions as such. It does refer specifically to (subjective) laws, as originally intended, and indeed in an even quantitatively appealing way. In this way it improves upon the Bayesian account. And it preserves the virtues of the Bayesian account, by not just postulating enumerative induction as a basic inductive inference rule, but by essentially deriving it from the more basic epistemic rationality postulates of regularity and symmetry. In this way, enumerative induction seems to have finally found an even more comfortable home.

Now to the caveats: So far, I have not told the full truth about my ranking-theoretic duplication of de Finetti's representation theorem. Things are more complicated. A first qualification is mathematically interesting, but philosophically neutral, as far as I see. It is not true that each regular symmetric κ is the mixture result of exactly one impact function ρ. We must more cautiously define the notion of what I call a *minimal* mixture, and this in fact turns out to be unique. The second qualification is less harmless. Regularity and symmetry alone do not suffice for representation. We also have to assume that κ is *concave* in a mathematically well-specifiable sense. The problem here is that concavity is required for mathematical reasons, although I can so far not present good philosophical or normative reasons why our ranking function should be concave. I am optimistic; but the case is mathematically intricate. This is presently just an open flank of my argument. In this respect, the Bayesian account does better, since it relies only on well-justifiable features of our subjective probabilities.[34] So, to be precise, the ranking-theoretic representation theorem says: A ranking function κ for $\mathcal{A}$ is regular, symmetric, and concave if and only if there is a unique impact function ρ for Λ such that κ is the minimal mixture of Λ by ρ.[35] This displays the two qualifications.

## 6. Two Concluding Observations

The ranking-theoretic representation theorem and the confirmation of laws entailed by it have various consequences. Let me finally mention two of them.

First, we may observe that Goodman's new riddle of induction partially evaporates. Our above equation for $\rho_{n+1}(\lambda) - \rho_n(\lambda)$ entails that each subjective law λ for which the observed $X_{n+1}$ taking the value $v$ is a positive instance, i.e., for which $\lambda(X_{n+1} = v) = 0$, is thereby confirmed. So, if $\{X_{n+1} = v\}$ represents that the n+1st emerald is green, crazy hypotheses like "all emeralds are grue" are thereby just as well confirmed as plausible laws like "all emeralds are green". We may well grant this; there is no reason to be scared by this observation. Indeed, there is no contradiction; in the positive relevance

[34] For a formal explanation of these differences see footnote 24.

[35] Cf. Theorems 12.14 and 12.18 in Spohn (2012, p. 293 and p. 296).

sense incompatible laws may be simultaneously confirmed. It is only that not all laws can be confirmed. There also are many potential laws that get disconfirmed, natural ones like "all emeralds are blue" and crazy ones like "all emeralds are bleen". If all goes well, "all emeralds are grue" will get disconfirmed soon enough. But, of course, we never know.

By no means may we conclude that all the incompatible hypotheses would be equally good. On the contrary, the a priori impact of the crazy hypotheses is so terribly low that it will stay low even after many confirmations. However, it is not infinitely low, and this is only reasonable. Why should we not be able to discover in the end that there is a hidden connection between the color of an emerald and the time of its first observation (or some other apparently entirely unrelated feature)? However, why should the a priori impact of gruesome hypotheses be very low? That's the snag. I have not given any reason for this. Impacts are just subjectively chosen, and our choice is clear, whereas the defenders of the grue hypothesis choose entirely different initial impacts. This is why I have said that Goodman's new riddle evaporates only partially. We might have expected a rationalization of our a priori impacts, which, however, I cannot provide. Still, my point stands. There is not any paradox in the fact that many incompatible hypotheses are simultaneously confirmed.[36]

The same point could have been made within a Bayesian framework, however less comfortably. In a Bayesian framework we would have to give non-zero initial weights to all the crazy hypotheses (in order to possibly confirm them), and it matters to our probabilities for the factual propositions which weights we choose. This is the so-called problem of the catch-all hypothesis which some take to be a severe objection to Bayesianism.[37] By contrast, in the ranking-theoretic framework very low impacts of crazy hypotheses need not surface in the ranks for factual propositions. This is what I have called the innocence of the worst explanation in Spohn (2012, sect. 14.15). Hence, my partial response to Goodman's new riddle is more easily maintained in ranking-theoretic terms.

The second and final observation concerns the so-called law of the uniformity of nature: Just as in the probabilistic case we may, somewhat sloppily, say that the ranking-theoretic representation theorem shows that our symmetric (and concave) epistemic state is always a mixture of (subjective) laws. We are bound to think in laws! Now, this overstates the case a little bit. Formally, $\lambda_0$ defined by $\lambda_0(A) = 0$ for all $A \in \mathcal{A}$ is also a subjective law. However, it rather represents the belief in complete lawlessness instead

---

[36] I neglect here the extension of Goodman's new riddle to meaning skepticism according to which it is unclear whether "all emeralds are green" really means this or actually means "all emeralds are grue".

[37] See, e.g., Earman (1992, sect. 7.2).

of belief in a specific law. (Recall my above remark about which generalizations are believed to hold in subjective laws; $\lambda_0$ only contains the belief in the empty or tautological generalization.) Still, all the other subjective laws have substantial content.

The more precise statement now is this. Let the regular, symmetric and concave $\kappa$ be a mixture of $\Lambda$ by the impact function $\rho$. Does the mixture contain $\lambda_0$? Maybe. Define sup $\kappa$ = sup $\{\kappa(A) \mid A \in \mathcal{A}\}$. sup $\kappa$ may be 0, positive, but finite, or $\infty$. Note that regularity requires only that $\kappa(E_n(\mathbf{v})) < \infty$ for all $n$ and $\mathbf{v}$. But this is compatible with the $\kappa(E_n(\mathbf{v}))$ having no upper bound and even compatible with $\kappa(A) = \infty$ for some infinitary proposition $A$. So sup $\kappa = \infty$ is a possibility as well. Now the formal observation is that $\rho(\lambda_0)$ = sup $\kappa$.[38] What does this mean?

sup $\kappa = 0$ means that $\kappa(A) = 0$ for all $A \in \mathcal{A}$. Such a $\kappa$ is incapable of any inductive inference, and we may well exclude it as a totally unreasonable epistemic state. sup $\kappa$ = $\infty$ means that $\rho(\lambda_0) = \infty$, i.e., a maximal denial of $\lambda_0$; and it entails that $\lambda_0$ remains maximally excluded after arbitrary amounts of evidence. (Infinite ranks cannot change.) So, whatever the evidence, we stick to the firm belief that some substantial law will hold. This may well be called an unrevisably a priori belief in lawfulness or the uniformity of nature. The third case is that sup $\kappa = s$ for some finite $s > 0$. This means that $\rho(\lambda_0) = s >$ 0, and since $\rho$ is a negative ranking function, this means disbelief in $\lambda_0$. Again, this may be called an a priori belief in the uniformity of nature. But it is not unrevisable; it may be defeated. We may (but need not) receive bewildering evidence that ever more upgrades the impact of $\lambda_0$, possibly up to 0. Then we would have reached a state of total perplexity in which we do not dare projecting any substantial law into the future. The last two cases may be actually hard to decide; sup $\kappa$ may be finite, but so large, and hence the impact of $\lambda_0$ so low, that we never actually reach the point of despair.

In any case, this observation shows that the belief in the uniformity of nature need not be as unjustified as Hume has thought. It is rationally required (as much as regularity, symmetry and concavity are rationally required). So, Kant may have been right with his a priori claims. At the same time, though, the remarks show that we have to differentiate. Belief in lawfulness or the uniformity of nature may be unrevisably or defeasibly a priori. And Kant was certainly not aware of the latter possibility.

[38] See Spohn (2012, p. 301).

## References

Bacon, F. (1620), *Novum Organum*, London.

Baumgartner, M. (2009), "Uncovering Deterministic Causal Structure: A Boolean Approach", *Synthese* 170, 71-96.

Carnap, Rudolf (1950/62), *The Logical Foundations of Probability*, Chicago: Chicago University Press, 2nd ed. 1962.

Cranston, M. (1967), "Bacon, Francis", in: P. Edwards (ed.), *The Encyclopedia of Philosophy, Vol. 1*, New York: Macmillan, pp. 235-240.

de Finetti, B. (1937), "La Prévision: Ses Lois Logiques, Ses Sources Subjectives", *Annales de l'Institut Henri Poincaré* 7. Engl. translation: "Foresight: Its Logical Laws, Its Subjective Sources", in: H.E. Kyburg jr., H.E. Smokler (eds.), *Studies in Subjective Probability*, New York: John Wiley & Sons 1964, pp. 93-158.

Earman, J. (1992), *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory*, Cambridge, Mass.: MIT Press.

Freitag, W. (2015), "I bet you'll solve Goodman's Riddle", *Philosophical Quarterly* 65, 254-267.

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

Halpern, J. Y. (2003), *Reasoning about Uncertainty*, Cambridge, Mass.: MIT Press.

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

Hintikka, J. (1966), "A Two-Dimensional Continuum of Inductive Methods", in: J. Hintikka, P. Suppes (eds.), *Aspects of Inductive Logic*, Amsterdam: North-Holland, pp. 113-132.

Huber, Franz, Christoph Schmidt-Petri (eds.) (2009), *Degrees of Belief. An Anthology*, Oxford: Oxford University Press.

Humburg, J. (1971), "The Principle of Instantial Relevance", in: R. Carnap, R. C. Jeffrey (eds.), *Studies in Inductive Logic and Probability, Volume I*, Berkeley: University of California Press, pp. 225-233.

Hume, David (1739), *A Treatise Concerning Human Nature, Book I*, London.

Hume, David (1748), *An Enquiry into Human Understanding*, London.

Kant, I. (1781/87), *Kritik der reinen Vernunft*; engl. translation: *Critique of Pure Reason*, London: Macmillan 1929.

Kelly, Kevin T. (2008), "Ockham's Razor, Truth, and Information", in: P. Adriaans, J. van Benthem (eds.), *Handbook of the Philosophy of Science, Vol. 8, Philosophy of Information*, Amsterdam: Elsevier, pp. 321-359.

Kern-Isberner, Gabriele (1999), *A Unifying Framework for Symbolic and Numerical Approaches to Nonmonotonic Reasoning and Belief Revision*. Fachbereich Informatik der FernUniversität Hagen, Habilitationsschrift.

Kern-Isberner, Gabriele (2004), "A Thorough Axiomatization of a Principle of Conditional Preservation in Belief Revision", *Annals of Mathematics and Artificial Intelligence* 40, 127-164.

Kripke, S. A. (1982), *Wittgenstein on Rules and Private Language*, Oxford, Blackwell.

Kuipers, T. A. F. (1978), *Studies in Inductive Logic and Rational Expectation*, Dordrecht: Reidel.

Lange, M. (2000), *Natural Laws in Scientific Practice*, Oxford: Oxford University Press.

Lipton, P. (1991), *Inference to the Best Explanation*, London: Routledge, 2nd ed. 2004.

Mill, J. St. (1843), *A System of Logic*, London.

Niiniluoto, I. (1972), "Inductive Systematization: Definition and a Critical Survey", *Synthese* 25, 25-81.

Pollock, J. L. (1990), *Nomic Probability and the Foundations of Induction*, Oxford: Oxford University Press.

Popper, K. R. (1934), *Logik der Forschung*, Wien: Springer; engl. translation: *The Logic of Scientific Discovery*, London: Hutchinson, 1959.

Quine, W. V. O. (1960), *Word and Object*, Cambridge, Mass: MIT Press.

Ramsey, F. P. (1929), “General Propositions and Causality”, in: F. P. Ramsey, *Foundations. Essays in Philosophy, Logic, Mathematics and Economics*, ed. by D. H. Mellor, London: Routledge & Kegan Paul 1978, pp. 133-151.

Ruzicka, R. (1976), “Der Induktionsbegriff von Aristoteles bis Galilei”, in: K. Gründer, *Historisches Wörterbuch der Philosophie, Band 4*, Basel: Schwabe, pp. 323-329.

Salmon, W. C. (1966), *The Foundations of Scientific Inference*, Pittsburgh: University Press.

Smart, N. (1967), “Indian Philosophy”, in: P. Edwards (ed.), *The Encyclopedia of Philosophy, Vol. 4*, New York: Macmillan, pp. 155-169.

Spohn, Wolfgang (1983), *Eine Theorie der Kausalität*, unpublished Habilitationsschrift, Universität München, pdf-version at: http://www.uni-konstanz.de/FuF/Philo/Philosophie/philosophie/files/habilitation.pdf

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,* vol.II, Dordrecht; Kluwer, pp. 105-134.

Spohn, W. (2012), *The Laws of Belief. Ranking Theory and Its Philosophical Applications*, Oxford: Oxford University Press.

Spohn, W. (2014), “The Epistemic Account of Ceteris Paribus Conditions”, *European Journal for the Philosophy of Science* 4, 385-408.

van Fraassen, Bas C. (1989), *Laws and Symmetry*, Oxford University Press, Oxford.