First publ. in: Philosophy of Science 72 (2005), 1, pp. 164-187

ENUMERATIVE INDUCTION AND LAWLIKENESS

Wolfgang Spohn
Fachbereich Philosophie
Universität Konstanz
D-78457 Konstanz

Abstract: The paper is based on ranking theory, a theory of degrees of disbelief (and hence belief). On this basis, it explains enumerative induction, the confirmation of a law by its positive instances, which may indeed take various schemes. It gives a ranking theoretic explication of a possible law or a nomological hypothesis. It proves, then, that such schemes of enumerative induction uniquely correspond to mixtures of such nomological hypotheses. Thus, it shows that de Finetti's probabilistic representation theorems may be transformed into an account of confirmation of possible laws and that enumerative induction is equivalent to such an account. The paper concludes with some remarks about the apriority of lawfulness or the uniformity of nature.

1. Introduction*

Enumerative induction says that a law is confirmed by its positive instances or may be inductively inferred from them (in the absence of negative instances). It is, for sure, the most venerable and primitive of all inductive rules. But it has a bad press. It is very crude; science does not seem to proceed with such simple rules. Goodman's new riddle of induction has shown that enumerative induction is inconsistent, if generally applied; and it seems impossible to say what the appropriate restrictions are. On the face of it, it is a rule of qualitative confirmation theory; but philosophers have dispaired of constructing such a theory.

The rule has finally found a Bayesian home. It is true, though, that at least within inductive logic as developed by Carnap (1971/80) nothing can confirm a law because each law has probability 0 (if its domain of quantification is infinite). The natural idea was then to turn enumerative induction into the Principle of Positive Instantial Relevance according to which each positive instance confirms that the next instance is also positive. This seems reasonable, and accepted. So, why bother any longer?

---

* I am indebted to two anonymous referees whose rich remarks led to numerous improvements and clarifications of this paper.

Well, “primitive” is ambiguous. It may indeed mean “not workable”. But it also means “basic”. If we do not fully understand the basic things, how can we ever hope to come to terms with the more complicated things? So whoever is concerned with inductive, plausible, or uncertain reasoning should be concerned to understand such a primitive rule as enumerative induction. The goal of the paper is to enhance this understanding. The way to reach the goal is to bring enumerative induction home from quantitative to qualitative confirmation theory, and the reason why this is feasible is that in the meantime we have a fully general qualitative confirmation theory at our disposal, namely ranking theory. This needs a little explanation.

Traditionally, confirmation theory is a field within philosophy of science. Its quantitative or probabilistic version, i.e. Bayesianism, has been a major option from the beginning. In the 50’s and 60’s we also saw forceful attempts to construct a qualitative confirmation theory. However, the project was abandoned in the 70’s, for reasons nicely summarized in Niiniluoto (1972). Thus, at least within philosophy of science Bayesianism had won the day. However, logicians and computer scientists were very active since around 1975 in producing alternatives, though rarely under the labels ‘induction’ or ‘confirmation’ (see, e.g., the many theories collected in Gabbay, Smets 1998-2000, in particular vol. 1 and 3) and hence scarcely noticed in epistemology and philosophy of science. The multiplicity of proposals developed there is quite confusing. Still, I believe that ranking theory as developed by me in Spohn (1983, sect. 5.3, and 1988), though under a different name, is the most suitable qualitative account of induction or confirmation.

This introduction is not the place for extensively arguing the case; the old reasons given in Spohn (1988) still apply. Let me state only the most important point. The central notion in this connection is the notion of *conditional belief*. In order to say whether some evidence would qualitatively confirm some hypothesis we have, to put it vaguely, to look at whether the hypothesis would be believed given the evidence and not given the evidence. If we want to give an account of induction, we have to give an account of belief change; so I have argued in Spohn (2000). And belief change best works by conditionalization rules that essentially refer to conditional beliefs, just as probabilistic conditionalization rules refer to conditional probabilities. We do need an adequate notion of conditional belief.

Hence, we should look at the various attempts to explain it. Belief revision theory (cf., e.g., Gärdenfors 1988) makes a plausible proposal: B is believed given A in a certain belief state iff B is believed in the revision of that state by A. But as I have argued in Spohn (1988), belief revision theory, as it is presented up to date, is defective and the proposal therefore inadequate. One might say that B is believed given A iff $P(B \mid A) = 1$, but this proposal is incomplete, because in standard probability theory this conditional probability is undefined if $P(A) = 0$. One might insist on the proposal by interpreting *P* as a Popper measure that fills this incompleteness by taking conditional probability as an undefined primitive. However, as shown in Spohn (1986), this idea is defective in just the way belief revision theory is. And so on. In the end, I claim, one must turn to ranking theory that offers the most adequate account of conditional belief.

Here, I simply want to proceed on the basis of this scarcely redeemed claim. The point of the introductory remarks was only to suggest that the most promising way to study enumerative induction is in terms of ranking functions. This is what I want to do here. Hence, the plan of the paper is this. In section 2 I shall introduce the theory of ranking functions as far as we need it here. Section 3 will then apply ranking theory to enumerative induction which, as we shall see, may realize in a variety of schemes. This will turn out to be a brief and rather boring exercise; the insights come later. In section 4, I shall propose a ranking theoretical explication of what a possible law or a nomological hypothesis is. In section 5, we shall be able to show that there is a one-one-correspondence between schemes of enumerative induction as found in section 3 and mixtures of nomological hypotheses as explained in section 4. Thus, our ranking theoretic analysis will result in transferring de Finetti's deep account of the confirmation of statistical hypotheses to the deterministic or qualitative realm. Section 6 will conclude with some remarks on the defeasible or unrevisable apriority of lawfulness or the uniformity of nature.

2. Ranking Functions

Let us start with a set *W* of possible worlds, small worlds in the sense of Savage rather than maximally large worlds in the sense of Lewis. Each subset of *W* is a truth condition or *proposition*. Hence, the set of propositions forms a complete Boolean algebra. I shall outright assume propositions to be the objects of doxastic attitudes, thereby taking these attitudes to be intensional. We know well that this is problematic, that the so-called propositional attitudes are presumably hyperintensional. But we scarcely know what to do about the problem. Hence, my assumption is just to signal that I do not want to worry here about these kinds of problems.[1]

Moreover, I assume that there is a distinguished class of (logically independent) *atomic propositions*. The paradigmatic atomic proposition states that a certain object has a certain property. Finally, I shall assume that the complete algebra of

---

[1] The *locus classicus* concerning (hyper-)intensionality is Carnap (1947); cf., in particular §§11-15. He there proposed to solve the problem of hyperintensionality with his notion of intensional structure. Quine responded by directly taking sentences as objects of belief. And till today the issue has remained obscure and undecided.

propositions is generated by the atomic propositions. Thus, each possible world is tantamount to a maximally consistent and possibly infinite conjunction of atomic propositions. A proposition is called *molecular* iff it is a member of the Boolean algebra generated by the atomic propositions, i.e., iff it is generated from the atomic propositions by *finitely* many Boolean operations.[2]

This is all we need to introduce our basic notion:

Definition 1: κ is a *ranking function* (for W) iff κ is a function from W into the set of extended non-negative integers $\mathbf{N}^+ = \mathbf{N} \cup \{\infty\}$[3] such that $\kappa(w) = 0$ for some $w \in W$. For each proposition $A \subseteq W$ the *rank* $\kappa(A)$ of A is defined by $\kappa(A) = \min \{\kappa(w) \mid w \in A\}$ and $\kappa(\emptyset) = \infty$. For $A, B \subseteq W$ the (*conditional*) *rank* $\kappa(B \mid A)$ of B *given* A is defined by $\kappa(B \mid A) = \kappa(A \cap B) - \kappa(A)$.

Since singletons of worlds are propositions as well, the point and the set function are interdefinable. The point function is simpler, but auxiliary, the set function is the one to be interpreted as a doxastic state.

Indeed, ranks are best interpreted as *degrees of disbelief*. $\kappa(A) = 0$ says that A is not disbelieved at all; $\kappa(A) = 1$ says that A is disbelieved (and hence $\overline{A}$ believed) to degree 1; etc. Note that $\kappa(A) = 0$ does not say that A is believed; this is rather expressed by $\kappa(\overline{A}) > 0$, i.e., that non-A is disbelieved (to some degree). The clause that $\kappa(w) = 0$ for some $w \in W$ is thus a *consistency* requirement. It guarantees that at least some proposition, and in particular W itself, is not disbelieved (and hence that some proposition, e.g. $\emptyset$, is not believed). This entails the

law of negation: for each $A \subseteq W$, either $\kappa(A) = 0$ or $\kappa(\overline{A}) = 0$ or both.

The set $C_\kappa = \{w \mid \kappa(w) = 0\}$ is called the *core* of κ (or of the doxastic state represented by κ). C_κ is the strongest proposition believed (to be true) in κ. In fact, a proposition is believed in κ if and only if it is a superset of C_κ. Hence, the set of beliefs is *deductively closed* according to this representation.[4]

These observations make clear the following essential point: On the one hand, the degrees of disbelief are the basic notion. On the other hand, these degrees also

---

[2] Cf. also Carnap (1971) who proceeds with a similar algebraic framework.

[3] This is a deviation from the definition I have given in earlier papers. It will be explained below.

[4] Consistency and deductive closure are standard in doxastic logic; they have been often attacked and equally often defended. The issue of logical omniscience is indeed highly problematic and closely related to the issue of hyperintensionality of propositional attitudes already mentioned. We have, however, decided the issue already by assuming propositions as objects of doxastic attitude; under this assumption consistency and deductive closure are quite trivial rationality requirements.

contain an all-or-nothing notion of disbelief (and thus belief): disbelief *is* disbelief to some positive degree. If we would confine ourselves to a static perspective, this all-or-nothing notion, which I sometimes called *plain* (dis-)belief and which is well studied in doxastic logic, would be good enough. However, in order to define an adequate notion of conditional belief and thus to account for the dynamics of the all-or-nothing notion, we have to introduce the degrees. I emphasize this point because it marks an important advantage of ranking over probability theory. The latter cannot offer an adequate notion of plain belief, and hence those raised in probabilistic thinking tend to find the notion disreputable. But, intuitively, we have the notion, and it is basic to large parts of epistemology. Ranking theory satisfies both needs here, the one for the all-or-nothing notion and the other for the degrees.

There are two laws for the distribution of degrees of disbelief: the

law of conjunction: $\kappa(A \cap B) = \kappa(A) + \kappa(B \mid A)$.

That is, the degree of disbelief in A and the degree of disbelief in B given A add up to the degree of disbelief in A-and-B; this appears highly intuitive. With Definition 1 we may say conversely that this is precisely how conditional degrees of disbelief are to be understood. And there is the

law of disjunction: $\kappa(A \cup B) = \min\{\kappa(A), \kappa(B)\}$.

That is, the degree of disbelief in a disjunction is the minimum of the degrees of the disjuncts. Given the definition of conditional ranks, this law is nothing but a conditional consistency requirement; if it would not hold the inconsistency could arise that both $\kappa(A \mid A \cup B)$, $\kappa(B \mid A \cup B) > 0$, i.e., that both A and B are disbelieved given A-or-B.

According to Definition 1, the law of disjunction indeed extends to disjunctions of arbitrary cardinality. I find this reasonable, since an inconsistency is to be avoided in any case, be it finitely or infinitely generated. Note that this entails that each countable set of ranks must have a minimum (not only an infimum) and that the range of a ranking function must therefore be well-ordered. Hence, the range $\mathbf{N}^+$ is a natural choice. This point will become important later on.[5]

---

[5] It is obvious that one has various options at this point. For instance, in Spohn (1988) I still took the range to consist of arbitrary ordinal numbers, but the advantages of this generality did not make up for the complications. By contrast, Hild (t.a., sect. 3.2) does not extend the law of disjunction to the infinite case and is thus free to adopt non-negative reals as values.

It is also obvious that the issue about infinite disjunctions is closely related to the discussion of the Limit Assumption in Lewis (1973, sect. 1.4). Without this assumption, it may happen that "if A

I immediately add

Definition 2: A ranking function is *regular* iff all consistent molecular propositions have finite ranks.

In the sequel we shall consider only regular ranking functions. In earlier papers I have assumed a stronger form of regularity by outright defining a ranking function to be a function from W into **N** so that only ∅ receives infinite rank. If all propositions are molecular, there is no difference. In this paper, however, we want to consider possibly infinite and thus non-molecular generalizations, and then this stronger form of regularity is not feasible. Whence the present weaker assumption.

There is no need here to develop ranking theory more extensively. A general remark may be more helpful: ranking theory works in almost perfect parallel to probability theory. Take any probabilistic theorem, replace probabilities by ranks, the sum of probabilities by the minimum of ranks, the product of probabilities by the sum of ranks, and the quotient of probabilities by the difference of ranks, and you are almost guaranteed to arrive at a ranking theorem. Additivity of probabilities thus translates into the law of disjunction for ranks. The probabilistic law of multiplication translates into the above law of conjunction. It is easy to prove the ranking analogue to the formula of total probability, the

formula of the total rank: $\kappa(A) = \min_{i \leq n} [\kappa(A \mid B_i) + \kappa(B_i)]$,

which says for a partition $\{B_1,...,B_n\}$ of W how to compute the rank of some proposition A from the rank of A given various hypotheses B_i and the ranks of the hypotheses B_i themselves. One may continue with a ranking version of Bayes' theorem.[6] One can even develop the whole theory of Bayesian nets in ranking terms.[7] And so on.

The general reason is that ranks may roughly be interpreted as orders of magnitude of (infinitesimal) probabilities. Consider a non-standard probability measure taking non-standard reals as values. The logarithm of a product of such probabilities is the sum of the logarithms of the factors, w.r.t. any base. And the order of magnitude (= the logarithm in round figures) of a sum of such probabilities is the minimum of the orders of magnitude of its terms, at least w.r.t. an infinitesimal

were the case, then B_i would be the case" is true for infinitely many B that are jointly unsatisfiable. Lewis finds reason to accept this situation. I prefer to accept the Limit Assumption instead.

6 This point is strongly developed in Hild (t.a.).

7 This was my original motivation. The basis of this theory, namely the so-called graphoid axioms of conditional independence, are proved for ranks in Spohn (1983, sect. 5.3) and (1988, sect. 6).

base. This perspective explains the translatability. However, I should emphasize that the translation is only an excellent rule of thumb, but not perfectly reliable, as we shall see later on (cf. also Spohn 1994). The matter is not fully cleared up.

It is still annoying, perhaps, that belief is not characterized in a positive way. But there is remedy.

Definition 3: β is the *belief function* associated with κ (and thus a *belief function*) iff β is the function assigning integers to propositions such that $\beta(A) = \kappa(\overline{A}) - \kappa(A)$ for each $A \subseteq W$. Similarly, $\beta(B \mid A) = \kappa(\overline{B} \mid A) - \kappa(B \mid A)$.

Recall that at least one of the terms $\kappa(\overline{A})$ and $\kappa(A)$ must be 0. Hence, $\beta(A) > 0$, < 0, or $= 0$ iff, respectively, A is believed, disbelieved, or neither; and A is the more strongly believed, the larger $\beta(A)$. Thus, belief functions may appear to be more natural. But their formal behavior is more awkward. I shall use both notions.

Since this is an essay about confirmation theory, we must ask: what is confirmation with respect to ranking functions? The same as elsewhere, namely *positive relevance*.

Definition 4: A *confirms* or is a *reason for B* relative to κ iff A is positively relevant to B, i.e., iff $\beta(B \mid A) > \beta(B \mid \overline{A})$, i.e., iff $\kappa(\overline{B} \mid A) > \kappa(\overline{B} \mid \overline{A})$ or $\kappa(B \mid A) < \kappa(B \mid \overline{A})$ or both.[8]

There is an issue here whether the condition should require $\beta(B \mid A) > \beta(B)$ or only $\beta(B \mid A) > \beta(B \mid \overline{A})$, as stated. In the corresponding probabilistic case, the two conditions are equivalent if all three terms are defined, but the first condition is a bit more general, since it may be defined while the second is not. That is why the first is often preferred. In the ranking case, however, all three terms are always defined, and the second condition may be satisfied while the first is not. In that case the second condition on which my definition is based seems to be more adequate.[9]

---

[8] I believe that if epistemologists talk of justification and warrant, they ought to refer basically to this relation of A being a reason for B; cf. Spohn (2001). That's, however, a remark about a different context.

[9] A relevant argument is provided by the so-called problem of old evidence. The problem is that after having accepted the evidence it can no longer be confirmatory. However, this is so only on the basis of the first condition. According to the second condition, learning about A can never change what is confirmed by A, and hence the problem does not arise. This point, or its probabilistic analogue, is made by Joyce (1999, sect. 6.4) by using Popper measures, relative to which the second condition is defined even if $P(\overline{A}) = 0$. However, cf. my skeptical remark about Popper measures in section 1.

Let me close my presentation of ranking theory with formally introducing a point that will receive great importance later on: Ranking functions can be mixed, just as probability measures can. For instance, if κ_1 and κ_2 are two ranking functions for W and if κ^* is defined by

$$\kappa^*(A) = \min \{\kappa_1(A), \kappa_2(A) + n\} \text{ for some } n \in \mathbf{N}^+ \text{ and all } A \subseteq W,$$

then κ^* is again a ranking function for W. Or more generally:

Definition 5: Let K be a set of ranking functions for W and ρ a ranking function for K. Then κ^* defined by

$$\kappa^*(A) = \min \{\kappa(A) + \rho(\kappa) \mid \kappa \in K\} \text{ for all } A \subseteq W$$

is (obviously) a ranking function for W and is called the *mixture* of K by ρ.

Note the similarity of this definition with the formula of the total rank; the various κ take here the role of the various hypotheses B_i in that formula.

3. Symmetry and Non-negative Instantial Relevance

Now we are well prepared turn to our proper topic, enumerative induction. Let us start with simplifying the propositional structure as far as our topic allows: by considering an infinite series of objects and just one property P. So, each object can either have or lack P, and there are just two universal generalizations: "all objects are P", and "all objects are not P". Concerning the objects this is all the generality we need; concerning the properties we proceed minimally. This will facilitate our business. It will be clear, though tedious to prove, that the results below generalize to any finite number of properties. So, the results are considerably stronger than they appear. However, I don't know how things stand with an infinity of properties that may be generated, e.g., by a real-valued magnitude.

This simplification allows us to represent each possible world by a sequence $\boldsymbol{z} = (z_1, z_2, ...)$ of 1's and 0's, where $z_n = 1$ or 0 means, respectively, that the n-th object has or lacks P. $\{\boldsymbol{x}$ takes $z_{i_1}, ..., z_{i_n}\}$ is short for the proposition $\{\boldsymbol{x} \mid x_{i_j} = z_{i_j}$ for $j = 1,...,n\}$.

The most basic assumption ranking functions will be supposed to satisfy is *symmetry*. This means that ranking functions should be able to distinguish different objects only with respect to the properties considered, in our case P and non-P. Let us make this precise in

Definition 6: κ is *symmetric* iff for any sequences $\boldsymbol{y}$ and $\boldsymbol{z}$ and any permutation π of $\mathbf{N}$ $\kappa(\boldsymbol{x}$ takes $y_1, ..., y_n) = \kappa(\boldsymbol{x}$ takes $z_{\pi(1)}, ..., z_{\pi(n)})$ if $y_i = z_{\pi(i)}$ for $i = 1,...,n$.

Regular symmetric ranking functions take a particularly simple form, as stated in the obvious

Theorem 1: For each regular symmetric κ there is a *representative function* f from $\mathbf{N} \times \mathbf{N}$ into $\mathbf{N}$ such that $\kappa(\boldsymbol{x}$ takes $z_1,..., z_{m+n}) = f(m,n)$ if $\sum_{i=1}^{m+n} z_i = m$, i.e., if exactly m of the first $m+n$ objects have P and the others lack P. This function satisfies $f(0,0) = 0$ and the minimum property $f(m,n) = \min\,[f(m+1,n), f(m,n+1)]$ (for a proof apply the law of disjunction to the fact that $\{\boldsymbol{x}$ takes $z_1,..., z_{m+n}\} = \{\boldsymbol{x}$ takes $z_1,..., z_{m+n+1}$ and $z_{m+n+1} = 1\} \cup \{\boldsymbol{x}$ takes $z_1,..., z_{m+n+1}$ and $z_{m+n+1} = 0\}$). Conversely, any function f from $\mathbf{N} \times \mathbf{N}$ into $\mathbf{N}$ with these two properties represents a regular symmetric ranking function.

This entails that f can be visualized as in infinite triangle of non-negative integers

$$
\begin{array}{ccccc}
 & & f(0,0) & & \\
 & f(1,0) & & f(0,1) & \\
f(2,0) & & f(1,1) & & f(0,2) \\
\ldots & & \ldots & & \ldots
\end{array}
$$

If a *path* in such a triangle is any sequence which starts at any point $f(m,n)$ and in which each member is succeeded by its left or right neighbor immediately below, then the minimum property entails that each such path is non-decreasing and that whenever a path increases by going left any path going right at this point does not increase, and vice versa.

Symmetry has a long and venerable history. Indeed, van Fraassen (1989) even went so far to argue that lawlikeness is a confused idea we should dispense with and that symmetry takes the key role in scientific reasoning in its place. This paper will in fact confirm van Fraassen's view, with the minor divergence that lawlikeness need not be dispensed with, but will receive an appropriate account through the notion of symmetry. In any case, we shall pursue our investigation of enumerative induction only in terms of symmetric (and regular) ranking functions.

The first noteworthy observation in this pursuit is that given symmetry there is no difference between belief in the next instance and belief in the universal generalization about all further instances. Suppose that after some evidence concerning the first n objects you believe that the n+1st object will have P; that is, your κ is such that $\kappa(\boldsymbol{x}$ takes $z_{n+1} = 0) = s$ for some $s > 0$. Because of symmetry you then believe that any further object will have P, that is, $\kappa(\boldsymbol{x}$ takes $z_{n+k} = 0) = s > 0$ for any $k \geq 1$. And because of the infinite variant of the law of disjunction this entails that you believe in *all* further objects having *P with the same strength*; that is, your disbelief that *some* future object lacks P is as strong as your disbelief that a specific future object lacks P; i.e., $\kappa(\bigcup_{k \geq 1} \boldsymbol{x}$ takes $z_{n+k} = 0) = s > 0$. If this sounds counterintuitive[10], we have to return arguing about the law of disjunction and the conditional consistency it reflects. However, don't be confused; your disbelief that *all* further objects lack P may still be much stronger or even infinite.

This means that as far as positive confirmation is concerned, i.e., confirmation that generates or strengthens belief instead of merely diminishing disbelief, there is no difference between the next or any other positive instance and the universal generalization about all further instances. Hence, Carnap's problem of the null confirmation of universal generalizations disappears in the ranking theoretic context, and the recourse to instantial relevance which was only a substitute in the Bayesian framework is fully legitimate here.[11]

Instantial relevance can take a stronger and a weaker form. The *principle of positive instantial relevance* (*PIR*) says that, given any evidence concerning the first n objects, the n+1st object having or lacking P confirms, respectively, the n+2nd object having or lacking P. The weaker *principle of non-negative instantial relevance* (*NNIR*) requires only that the contrary is not confirmed. Hence, let us state

Definition 7: A regular symmetric ranking function κ *satisfies PIR* iff $\beta(\boldsymbol{x}$ takes $z_{n+1} \mid \boldsymbol{x}$ takes $z_1, ..., z_n) < \beta(\boldsymbol{x}$ takes $z_{n+2} \mid \boldsymbol{x}$ takes $z_1, ..., z_{n+1})$ whenever $z_{n+1} = z_{n+2}$, i.e., iff for the relevant representative function f and all $m, n \in \mathbf{N}$ $f(m+2,n) - f(m+1,n+1) < f(m+1,n) - f(m,n+1) < f(m+1,n+1) - f(m,n+2)$. κ *satisfies NNIR* iff the weak inequalities hold instead.

---

[10] It is not unlikely, though, that your intuitions are probabilistically trained, and then it is difficult to tell apart the intuitions and the training.

[11] Within the probabilistic context, the strongest proposal to overcome Carnap's problem of the null confirmation of universal gneralizations is the K-dimensional system of Hintikka and Niiniluoto (1976). It would be interesting to compare it with the ranking theoretic approach.

PIR may look like the correct formalization of enumerative induction; alas, we have

Theorem 2: There is no regular symmetric ranking function satisfying PIR.

Proof: Let us try to satisfy PIR by an appropriate representative function f. So we start with $f(0,0) = 0$ and, without loss of generality, $f(1,0) = 0$ and $f(0,1) = r \geq 0$. This entails $f(2,0) = 0$. Hence, if we set $f(1,1) = r$, we already violate PIR. So, we must choose $f(1,1) = s > r$ and $f(0,2) = r$. This in turn entails $f(3,0) = 0$ and $f(0,3) = r$. But we cannot complete, then, the fourth line of our triangle: we must set $f(2,1)$ or $f(1,2) = s$, and both choices violate PIR. □

This failure should not come as a surprise. If we try to satisfy PIR with respect to the positive instances and increase the disbelief in a negative instance with increasing positive evidence, we cannot at the same time satisfy PIR with respect to the negative instances. For, many negative instances are then just as disbelieved as a single one, and hence the negative instances cannot be positively relevant to further negative instances. We cannot have it both ways.

Hence, we are forced to settle for the weaker NNIR. It is easily seen to be consistent. Within a probabilistic setting non-negative instantial relevance is in fact entailed by symmetry. (Cf. Humburg 1971.) Thus it is worth noting that this is not the case here; it is obvious that there are symmetric ranking functions violating NNIR (because there are representative functions violating the additional condition of definition 7).

Where do we stand? If we want to account for enumerative induction within the ranking theoretic setting, we have to accept the second best explication, i.e., NNIR. We should also keep in mind that, within this setting, positively confirming the next instance is tantamount to confirming the corresponding generalization. Thus, we may preliminarily conclude that each symmetric ranking function satisfying NNIR is a way to realize enumerative induction, there being indeed an infinity of such ways.

Still, the preliminary conclusion does not look right. There is a definite loss in the retreat from PIR to NNIR. Even partial instantial *ir*relevance does not really seem compatible with enumerative induction; it is strange that the confirming effect of a positive instance must fail at least sometimes. This is a negative illusion, though. In section 5 all doubts dissolve. We shall find that NNIR is exactly right and that, contrary to appearance, positive relevance can be fully reestablished.

Our investigation has remained superficial so far. The topic gains depth only when we remind ourselves of the fact that enumerative induction was never taken to apply to all universal generalizations whatsoever, but rather only to laws or potential laws; at most with respect to laws it may claim to be a reasonable rule of inductive inference. Where is this crucial point reflected in our ranking theoretic explication? Well, it *is* reflected, but not at all in an obvious way. In order to uncover it, we have to think about what lawlikeness may mean in ranking theoretic terms.

4. Laws

In our simple setting we considered just two universal generalizations: G_1 = (1,1,...) and G_0 = (0,0,...). What could it mean to treat G_1, say, as a law and not as an accidental generalization? I think, quite unoriginally, that this shows in our inductive behavior. To believe in G_1 as a law is, first, to believe in G_1, as expressed by $\kappa(\overline{G_1}) > 0$. But, as we already know, the belief in G_1 can be realized in many different ways; this belief alone does not fix the inductive relations between the various instances. Which forms may they take? Well, if you learn about positive instances of G_1, you do not change your beliefs about the further instances according to κ, since you expected them to be positive, anyway. Crucial differences emerge only when we look at how you respond to negative instances according to the various attitudes. Let me focus for a while on two particular responses, which I call the 'persistent' and the 'shaky' attitude:

If you have the *persistent* attitude, your belief in further positive instances is unaffected by negative ones, i.e., $\kappa(\boldsymbol{x} \text{ takes } z_{n+1} = 0) = \kappa(\boldsymbol{x} \text{ takes } z_{n+1} = 0 \mid \boldsymbol{x} \text{ takes } z_1 = ... = z_n = 0)$. If, by contrast, you have the *shaky* attitude, your belief in further positive instances is destroyed by a negative instance, i.e., $\kappa(\boldsymbol{x} \text{ takes } z_2 = 0 \mid \boldsymbol{x} \text{ takes } z_1 = 0) = 0$, and, due to symmetry, also by several negative instances.

The difference is, I find, characteristic of the distinction between lawlike and accidental generalizations. Let us look at two famous examples. First the coins:

(1) All Euro coins are round.
(2) All of the coins in my pocket today are made of silver.

It seems intuitively clear to me that we have the persistent attitude towards (1) and the shaky one towards (2). If we come across a cornered Euro coin, we wonder what might have happened to it, but our confidence that the next coin will be round again is not shattered. If, however, I find a copper coin in my pocket, my expectations concerning the further coins simply collapse; if (2) has proved wrong in one case, it may prove wrong in any case.

Or look at the metal cubes, which are often thought to be the toughest example, because they display no perspicuous syntactic or semantic difference:

(3) All solid uranium cubes are smaller than one cubic mile.
(4) All solid gold cubes are smaller than one cubic mile.

What I said about (1) and (2) applies here as well, I find. If we bump into a gold cube this large, we are surprised – and start thinking there might well be further ones. If we stumble upon a uranium cube of this size, we are surprised again. But we find our reasons for thinking that such a cube cannot exist unafflicted and will instead start investigating this extraordinary case (if it obtains for long enough). As far as I see, the difference between the shaky and the persistent attitude applies as well to the other examples prominent in the literature.[12]

I am well aware that this sounds at best partially convincing. I am deliberately painting black and white here in order to elaborate the opposition between the persistent and the shaky attitude. Obviously, one would be prepared to say how one would respond in such cases only if they would be described in much more detail, especially concerning the evidence which led one to believe in the relevant generalizations in the first place. So, there is also a lot of grey.

There are at least two different kinds of grey. First, there is a broad range of attitudes between the two extremes I have described. Being shaky means to be *very* shaky; the belief in further positive instances may instead fade more slowly. And being persistent means to be *strictly* persistent; the belief in further positive instances may instead fade so lately that we never come to the point of testing it. Second, if confronted with such cases, we would in a sense widen our perspective. Take the uranium cube again. If we would really bump into such a large uranium cube, we would not simply mumble "impossible!" and stick to the belief that there will be no further exceptions. Rather, we would say that our original law was qualified by a *ceteris paribus* clause, anyway, and that a thorough investigation of the case will allow us to get clearer about normal and exceptional conditions. However, as fascinating as it is, the issue of *ceteris paribus* laws is certainly beyond the sope of this paper.[13]

---

[12] Cf., e.g., the overview in Lange 2000, pp. 11f. As Köhler (t.a.) pointed out to me, Bode's law of the logarithmic distribution of the planets in the solar system aptly illustrates my dichotomy. This law appeared to be accidental, and one counter-instance would have destroyed the confidence in it. Only recently it has acquired lawlike status via very sophisticated considerations, and the discovery of an anomaly would not impair this status.

[13] But I am convinced that ranking theory helps understanding this bewildering issue. At least I have argued so in Spohn (2002).

There are now two ways to respond. One may either say there is too much grey not decomposable into black and white. Or one may say that there is an important insight in my black and white distinction which opens a fruitful way to analyse the shades of grey. I hope I have given at least some plausibility to proceeding on the second response.

If this is the right way to see the matter, treating a generalization strictly as a law is really to take the strictly persistent attitude towards it. This conclusion leads us to a further consequence, namely that the characteristic of lawlikeness is not something to be found in the propositional content of the generalization; it rather lies in our inductive attitude towards it or its instantiations. This consequence will be of crucial importance in the sequel.

The account given so far is obviously very close to the old idea that laws are not general statements, but rather inference rules or inference licenses. The idea goes back at least to Ramsey (1929) who stated it very clearly: "Many sentences express cognitive attitudes without being propositions; and the difference between saying yes or no to them is not the difference beween saying yes or no to a proposition" (pp. 135f.). And "... 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). The idea has become quite popular among philosophers.

From a purely logical point of view, however, it is hard to see the difference between accepting the generalization as an axiom and accepting the corresponding inference rule for each instantiation. The only difference is that the rule is logically weaker; the rule is made admissible by the axiom, but the axiom cannot be inferred with the help of the rule. What else beside this unproductive logical point could be meant by the slogan "laws are inference rules" has been little explained.

Still, one might say that the inference-license perspective emphasizes the single case. This emphasis has now been stripped of its merely rhetorical character; it is reflected, I think, in my central notion of persistence and thus finds a precise induction-theoretic basis. In this perspective, the mark of laws is not their universality which breaks down with a single counter-instance, but rather their operation in each single case, which is not impaired by exceptions. Here, my account meets with Cartwright (1989) and her continuous efforts to explain that physical laws are deceptive and that we should rather attend to the single case and to the capacities (co-)operating in it. In Spohn (2002) the point is argued a bit more extensively.

So much for some striking agreements. The most obvious disagreement is with Popper, of course. There is no doubt about how much philosophy of science owes to Popper. In view of this, my account is really ironic, since its conclusion is, in a way, that the mark of laws is their *not* being falsifiable by negative instances; only accidental generalizations are subject to such falsification. To be a bit more precise: Of course, any generalization is falsified by a single counter-instance. But falsified generalizations are to be rejected according to Popper. By contrast I have argued that the belief in the further instances is shattered by the falsifying instances only in the case of accidental generalizations, but not in the case of laws. No doubt, the idea that the belief in laws is not given up so easily is familiar at least since Kuhn (1962), and already Popper (1934, ch. IV, §22) has insisted that the falsification of laws proceeds by more specialized counter-laws rather than by mere counter-instances. Here, however, the point is boiled down to its induction-theoretic essence.

5. Laws and Enumerative Induction

There is a striking and severe tension between sections 3 and 4. We saw that, given symmetry, PIR is not feasible. Hence, we retreated to NNIR as a preliminary explication of enumerative induction. Then we noticed that enumerative induction applies only to laws. Finally, I have proposed an explication of laws according to which instances are *independent* of each other; this is what persistence amounts to. Thus we arrived at complete instantial *irrelevance* which is rather a caricature of NNIR and not in agreement with enumerative induction at all. Something must have gone badly wrong.

No, there is only a subtle confusion. Belief in a law is more than belief in a proposition. It is a certain doxastic attitude, and that attitude as such is characterized by the independence in question: if I would have just this attitude, just the belief in a strict law and no further belief, my κ would exhibit this persistence or independence. Enumerative induction, by contrast, is not about what the belief in a law *is*, but about how we may acquire or *confirm* this belief. The two inductive attitudes involved may be easily confused, but the confusion cannot be identified as long as one thinks that belief in a law is just belief in a proposition.

However, what could it mean to confirm a law if it does not mean to confirm a proposition? My definition of confirmation in section 2 applies only to the latter. Hence, the talk of the confirmation of laws, i.e., of a second-order inductive attitude towards a first-order inductive attitude, is so far mere metaphorics. Can we do better?

Yes, we can. There is fortunately clear precedent in the literature. Given the close similarity between probability and ranking theory, one might notice that *a law* as I conceived it is nothing but *a sequence of independent, identically distributed random variables* translated into ranking terms. It thus becomes obvious that de Finetti (1937) addresses exactly our problem in the probabilistic context. In his celebrated theorems de Finetti showed that there is a one-one correspondence between symmetric probability measures for an infinite sequence of random variables and mixtures of Bernoulli measures according to which the variables are independent and identically distributed; and he showed that the mixture focusses more and more on a single Bernoulli measure as evidence accumulates. He thus showed to the objectivist that subjective symmetric measures provide everything he wants: beliefs about statistical hypotheses that converge toward the true one with increasing evidence.

De Finetti's issue between objectivism and subjectivism is not my concern. Ranking functions are thoroughly epistemological and have as such no objective interpretation.[14] Still, we can immediately translate de Finetti's theory into an account of the confirmation of laws as conceived here. The basic construction is, I find, illuminating, despite its formalistic appearance.

Let us return to our simple one-property frame. We had two universal generalizations G_1 and G_0. But there are infinitely many persistent, lawlike attitudes. If we define for all $r,s \in \mathbf{N}^+$

$$\lambda_{-r}(\boldsymbol{x} \text{ takes } z_1, ..., z_n) = r \cdot \sum_{i=1}^{n} z_i \text{, and } \lambda_s(\boldsymbol{x} \text{ takes } z_1, ..., z_n) = s \cdot (n - \sum_{i=1}^{n} z_i) \text{ ,}$$

then $\Lambda = \{\lambda_t \mid t \in \mathbf{Z}^+)$ includes all and only the persistent attitudes (where $\mathbf{Z}^+ = \mathbf{Z} \cup \{\infty, -\infty\}$). Λ contains precisely the 'Bernoullian' ranking functions which are symmetric and according to which each instance is independent from all others. For $t > 0$ λ_t believes in G_1 and disbelieves in each negative instance with rank t. For $t < 0$ it is just the other way around; such a λ_t believes in G_0 and disbelieves in each positive instance with rank t. In short, each λ_t counts the number of counter-instances within $\{\boldsymbol{x}$ takes $z_1, ..., z_n\}$ to the generalization it believes in and multiplies it by t (or $-t$).

What then is the difference between, e.g., λ_1 and λ_2? There is none in content and none in persistence. The only difference lies in the disbelief in negative instances; λ_2 is firmer a law, one might say, than λ_1. Rather for technical reasons we have to include λ_∞ and $\lambda_{-\infty}$. λ_0, finally, does not represent a law at all. It rather represents *lawlessness*, indeed complete agnosticism; nothing (except the tautology) is believed in λ_0. Its special role will be discussed in the final section.

Now, believing in laws, confirming and falsifying laws, etc. are doxastic attitudes towards laws, which will here be modelled, of course, by a ranking function

[14] But see Spohn (1993), where I tried to reduce the tension between my ranking theoretic and hence subjective explication of causation and the hardly deniable view that causation is an objective relation in the world.

ρ over the set Λ of possible laws. If the possible laws are possible first-order attitudes, then ρ is a second-order attitude, which, however, induces a first-order attitude. What, according to ρ, is the rank of a proposition $A \subseteq W$, i.e. the degree of disbelief in A? It is the minimum of all the disbeliefs in A according to the possible laws in Λ weighed by the disbelief in the laws according to ρ; that is, the first-order attitude induced by ρ is just the mixture of Λ by ρ as defined in definition 5.

Are we talking about a specific second-order attitude ρ? No, you may have any ρ you like. The following considerations are perfectly general in this respect. Let us call ρ *proper*, though, iff at most one of $\rho(\lambda_\infty)$ and $\rho(\lambda_{-\infty})$ is finite. Now we can start translating de Finetti's theorems.

First, we have:

Theorem 3: For each proper ρ over Λ, the mixture of Λ by ρ is a regular symmetric ranking function satisfying NNIR.

Proof: Regularity and symmetry are obvious since all λ_s are regular and symmetric. The proof of NNIR is essentially a tedious exercise. And since ρ is to be proper, the mixture is regular. □

Second, we have: For each regular symmetric ranking function κ satisfying NNIR there is a proper ranking function ρ over Λ such that κ is the mixture of Λ by ρ. We may indeed strengthen the claim. Suppose we mix, e.g., λ_1 and λ_2 by some ρ with $\rho(\lambda_1) = \rho(\lambda_2) = 0$. Then λ_2 is obviously a redundant component of the mixture; it never determines the result of the mixture, i.e., the relevant minimum. Because of such redundant components mixtures are never unique.[15] Uniqueness can be achieved only with minimal mixtures, as we might call them. However, we must be careful in catching the right kind of minimality. The point of the following definition will become fully clear only with theorem 5 below.

Definition 8: λ_s is a *redundant component* of the mixture of Λ by ρ w.r.t. a proposition A iff there is no proposition B such that $\min \{\lambda_t(A \cap B) + \rho(\lambda_t) \mid t \in \mathbf{Z}^+\} < \min \{\lambda_t(A \cap B) + \rho(\lambda_t) \mid t \in \mathbf{Z}^+ - \{s\}\}$, i.e., iff λ_s does not determine the value of the mixture for any $A \cap B$. λ_s is a *strongly redundant component* of the mixture of Λ by ρ iff λ_s is a redundant component of the mixture w.r.t. to all $A_{m,n}$ ($m,n \geq 0$), where $A_{m,n}$ is the proposition that (in some order) m of the first $m+n$ objects have P

[15] This is a noticeable difference to probabilistic mixtures where every ingredient with positive weight contributes to the mixture, however slightly.

and the other n objects lack P. Finally, the mixture of Λ by ρ is called *minimal* iff for all its strongly redundant components λ_s $\rho(\lambda_s) = \infty$.

Hence, in a minimal mixture all strongly redundant components get weight ∞ and cannot enter the mixture at all. The strengthened claim then is:

Theorem 4: For each regular symmetric ranking function κ satisfying NNIR there is a unique ρ over Λ such that κ is the minimal mixture of Λ by ρ.

Proof: Let κ be a regular symmetric function satisfying NNIR, let f be its representative function forming an infinite triangle of non-negative integers, and let c = sup f be the supremum of f, which may be finite or infinite. Let us focus on *simple* paths starting at the boundary of the triangle and making no turns. These paths take two forms. For each $m \geq 0$ there is the *right* path $f(m,0), f(m,1), f(m,2)$, ... starting at the left and going always right, and for each $n \geq 0$ there is the *left* path $f(0,n), f(1,n)$, $f(2,n)$, ... starting at the right and going always left. We know that the simple paths are non-decreasing (like all others). NNIR entails, moreover, that the simple paths do not *accelerate*; whenever i, j, k are three consecutive members of such a path, then $k - j \leq j - i$.

Each simple path either goes to infinity or reaches a maximum and then remains constant. Let us define a_m to be the supremum of the m-th right path $f(m,0), f(m,1)$, ... and b_n to be the maximum of the n-th left path $f(0,n), f(1,n)$, ... $(m,n \geq 0)$. Again, both sequences $\boldsymbol{a} = (a_0, a_1, ...)$ and $\boldsymbol{b} = (b_0, b_1, ...)$ must be non-decreasing and, due to NNIR, also non-accelerating. Either $a_0 = 0$ or $b_0 = 0$ or both, and $c = \sup \boldsymbol{a} = \sup \boldsymbol{b}$.

With the help of the two sequences $\boldsymbol{a}$ and $\boldsymbol{b}$ we can construct now the relevant minimal mixture ρ. If $a_1 - a_0 := r$, we set $\rho(\lambda_{-r}) = a_0$; and if a_m is any point at which $\boldsymbol{a}$ decelerates, i.e., such that $a_m - a_{m-1} > a_{m+1} - a_m := r$, we set $\rho(\lambda_{-r}) = a_m - mr$. Similarly, if $b_1 - b_0 := s$, we set $\rho(\lambda_s) = b_0$; and if b_n is any point at which $\boldsymbol{b}$ decelerates, i.e., such that $b_n - b_{n-1} > b_{n+1} - b_n := s$, we set $\rho(\lambda_s) = b_n - ns$. If for any $t \in \mathbf{Z}^+$ $\rho(\lambda_t)$ is not thereby defined, we set $\rho(\lambda_t) = \infty$. This completes the construction of ρ. Note, in particular, that this entails $\rho(\lambda_0) = c$. Hence, the lawless λ_0 is a relevant component of the mixture only if c is finite.

Since either $a_0 = 0$ or $b_0 = 0$, there is some $t \in \mathbf{Z}^+$ with $\rho(\lambda_t) = 0$. Since at least one of a_1 and b_1 is finite, either $\rho(\lambda_\infty) = \infty$ or $\rho(\lambda_{-\infty}) = \infty$ or both. Hence, ρ is a proper ranking function over Λ.

The mixture of Λ by ρ indeed generates the representative function f: For all $m,n \geq 0$ we have either $f(m,n) = a_m$ or $f(m,n) = b_n$, since either $f(m,n+1) = f(m,n)$ or $f(m+1,n) = f(m,n)$, and thus either the right or the left simple path through $f(m,n)$

does not increase after $f(m,n)$. Now suppose $f(m,n) = b_n \leq a_m$, and let us check whether our mixture yields the same result:

As above, let $A_{m,n}$ be the proposition that (in some order) m of the first $m+n$ objects have P and the other n objects lack P. Hence,

$$f(m,n) = \kappa(A_{m,n}) = \min\,\{\lambda_t(A_{m,n}) + \rho(\lambda_t) \mid t \in \mathbf{Z}^+\}$$
$$= \min_{r,s \geq 0}\ [mr + \rho(\lambda_{-r}),\ ns + \rho(\lambda_s)]$$

How to calculate this minimum? Let b_{n^*} be the largest point before b_n where $\boldsymbol{b}$ decelerates and let $s^* = b_{n^*+1} - b_{n^*}$. Hence, $b_{n^*} = b_n - s^*(n - n^*)$. What about $ns^* + \rho(\lambda_{s^*})$? We have:

$$ns^* + \rho(\lambda_{s^*}) = ns^* + b_{n^*} - n^*s^* \text{ (according to the definition of } \rho)$$
$$= ns^* + b_n - s^*(n - n^*) - n^*s^* = b_n.$$

Moreover, it is clear from the construction that $ns + \rho(\lambda_s) \geq ns^* + \rho(\lambda_{s^*})$ for all other $s \in \mathbf{N}^+$. The same reasoning shows that $mr + \rho(\lambda_{-r}) \geq a_m \geq b_n$ for all $r \in \mathbf{N}^+$. Thus, indeed, $f(m,n) = b_n$ according to the mixture.

Of course, if $f(m,n) = a_m \leq b_n$, the corresponding argument holds.

Some λ_t ($t \in \mathbf{Z}^+$) receiving finite rank by ρ may be a redundant component of the mixture of Λ by ρ w.r.t. $A_{0,0}$ (= the tautology); this always happens when two successive members a_m and a_{m+1} of $\boldsymbol{a}$ or b_n and b_{n+1} of $\boldsymbol{b}$ are points of deceleration. But none of them is strongly redundant, and the mixture is indeed minimal in the sense of definition 8. This, however, will become clear only with the next theorem. It will also be obvious, then, that the ρ we have constructed is unique, i.e., provides the only minimal mixture generating the representative function f. □

The final step in our translation of de Finetti is to inquire how the mixture is changed by evidence. This can be directly read off from the results above. Suppose that we collect the evidence $A_{m,n}$ that m of the first $m+n$ objects have and the other n objects lack P. If we start with the regular symmetric κ with representative function f, what is then the a posteriori ranking function $\kappa_{m,n}$ on the space of possibilities for the infinitely many remaining objects? Well, we learn by conditionalization; hence, for any proposition B within this space $\kappa_{m,n}(B) = \kappa(B \mid A_{m,n})$. The representative function $f_{m,n}$ of $\kappa_{m,n}$ is thus given by $f_{m,n}(p,q) = f(m+p, n+q) - f(m,n)$.

Now, suppose that κ is the minimal mixture of Λ by ρ. What is then the unique $\rho_{m,n}$ so that $\kappa_{m,n}$ is the minimal mixture of Λ by $\rho_{m,n}$? We know that f is the result of the mixture by ρ, i.e.,

$$f(m,n) = \min_{r,s\geq 0} [\lambda_{-r}(A_{m,n}) + \rho(\lambda_{-r}), \lambda_s(A_{m,n}) + \rho(\lambda_s)]$$
$$= \min_{r,s\geq 0} [\rho(\lambda_{-r}) + mr, \rho(\lambda_s) + ns].$$

Thus, we have for all $p,q \in \mathbf{N}$:

$$f_{m,n}(p,q) = f(m+p, n+q) - f(m,n)$$
$$= \min_{r,s\geq 0} [\rho(\lambda_s) + (n+q)s , \rho(\lambda_{-r}) + (m+p)r] - f(m,n)$$
$$= \min_{r,s\geq 0} [\rho(\lambda_s) + ns - f(m,n) + qs , \rho(\lambda_{-r}) + mr - f(m,n) + pr].$$

This already suggests how to define $\rho_{m,n}$. However, $\rho_{m,n}$ has to be a minimal mixture, and therefore we still need to eliminate some of the components originally having finite rank. For this purpose, let a_{m^*} be the largest member of $\boldsymbol{a}$ up to a_m where $\boldsymbol{a}$ decelerates and b_{n^*} the largest member of $\boldsymbol{b}$ up b_n where $\boldsymbol{b}$ decelerates (thus, possibly $a_{m^*} = a_m$ and $b_{n^*} = b_n$), and let $r^* = a_{m^*+1} - a_{m^*}$ and $s^* = b_{n^*+1} - b_{n^*}$. Now we can state

Theorem 5: Define for $r,s \in \mathbf{N}^+$:

$\rho_{m,n}(\lambda_{-r}) = \rho(\lambda_{-r}) + mr - f(m,n)$ for $r \leq r^*$ and
$\rho_{m,n}(\lambda_s) = \rho(\lambda_s) + ns - f(m,n)$ for $s \leq s^*$;

and if $r > r^*$ and $s > s^*$, then $\rho_{m,n}(\lambda_{-r}) = \rho_{m,n}(\lambda_s) = \infty$. Then $\kappa_{m,n}$ is the minimal mixture of Λ by $\rho_{m,n}$.

Proof: It is obvious from the construction for theorem 4 that λ_{-r} and λ_s are strongly redundant components of $\rho_{m,n}$ for $r > r^*$ and $s > s^*$. Thus the minimality of the mixture of Λ by $\rho_{m,n}$ carries over to $\rho_{m,n}$. Therefore, the above calculations already prove that $f_{m,n}$ is generated by $\rho_{m,n}$. □

The point of defining minimality as we did in definition 8 now becomes clear. As mentioned, some components of the mixture of Λ by ρ may be initially redundant, i.e., w.r.t. to $A_{0,0}$. Still, they may become non-redundant after conditionalization by $A_{m,n}$. Hence, they have to be included already in the original mixture. Otherwise, we could not have obtained $\rho_{m,n}$ from ρ so easily as in theorem 5.

The theorem has three important consequences. First, it helps to reestablish positive instantial relevance. Suppose, we find the $m+n+1$st object to have P; thus, our evidence increases from $A_{m,n}$ to $A_{m+1,n}$. How does the mixture change from $\rho_{m,n}$ to $\rho_{m+1,n}$? Insofar $\rho_{m+1,n}$ is finite we have for $r,s \geq 1$:

$$\rho_{m+1,n}(\lambda_{-r}) = \rho(\lambda_{-r}) + (m+1)r - f(m+1,n) \text{ and } \rho_{m+1,n}(\lambda_s) = \rho(\lambda_s) + ns - f(m+1,n).$$

Hence, in any case $\rho_{m+1,n}(\lambda_{-r}) - \rho_{m+1,n}(\lambda_s) = r + \rho_{m,n}(\lambda_{-r}) - \rho_{m,n}(\lambda_s)$. That is, each λ_{-r} as opposed to any of the λ_s is more disbelieved in $\rho_{m+1,n}$ than in $\rho_{m,n}$ (by r ranks). In other words, the additional positive instance is positively relevant to the positive lawlike attitudes. So, on the level of the second-order attitudes we indeed have exceptionless positive instantial relevance, which is blurred, though, by the mixture and thus weakens to NNIR on the level of first-order attitudes. Theorem 2 has shown that this weakening is unavoidable, but now we see that it is only an artifact of the mixture.

This observation teaches us, secondly, that as more and more positive instances accumulate and $m - n$ diverges to infinity, $\rho_{m,n}(\lambda_{-r}) - \rho_{m,n}(\lambda_s)$ ($r,s \geq 1$) diverges to infinity as well, i.e., the disbelief in the negative lawlike attitudes heads for infinite firmness. This parallels de Finetti's observation in the probabilistic case.

So, all in all, we have seen that de Finetti's account of the confirmation of statistical hypotheses may be perfectly translated into ranking theoretic terms, thus deepening our understanding of enumerative induction and lawlikeness.

There is still a third lesson, which has in fact no probabilistic analogue. It thus goes a little step beyond de Finetti and deserves a brief concluding section of its own.

6. The Apriority of Lawfulness

This lesson concerns the special role of λ_0. We noticed already that λ_0 is total agnosticism expressing lawlessness instead of lawfulness. Now, we either have $\rho(\lambda_0) = \infty$, which entails $\rho_{m+n}(\lambda_0) = \infty$ for all $m,n \in \mathbf{N}$. Then ρ embodies the maximally firm belief that some law or other will obtain. This belief would indeed be invariable, not refutable even by very long sequences of apparent random behavior of the instances with respect to P. This does not appear reasonable.

The alternative is that we give $\rho(\lambda_0)$ some finite value; hence, $\rho_{m,n}(\lambda_0) = \rho(\lambda_0) - f(m,n)$. This entails that with each unexpected realization of an instance λ_0 gets less disbelieved. After too many disappointments we shall eventually have lost our belief in lawfulness and any belief about the behavior of new objects concerning P, the belief in lawlessness being the only remaining option. This may also sound implausible. However, $\rho(\lambda_0)$ may be very large so that the agnostic state is in fact never reached.

The more relevant observation, though, is that the whole story I have told about the single property P can be generalized to any finite number of properties P_1, ..., P_m in a straightforward way. We can define Carnap's Q-predicates, i.e., the atoms of the Boolean algebra of properties generated by P_1, ..., P_m; for each Q-predicate Q_k we can consider the generalization "there is no Q_k" and the corresponding laws, i.e., persistent attitudes; and then all the theorems of section 4 continue to hold. So, what we would really do if lawlessness with respect to P threatens is to try to correlate P with some other properties and to pursue the investigation within a larger space of properties.

Within such a larger space also more complex forms of laws become available going beyond persistent attitudes towards "there is no Q_k". As already mentioned, the ranking theoretic framework in particular allows of an analysis of ceteris paribus laws (cf. Spohn 2002, sect. 4). So, there are rich prospects of generalization. I don't know, though, whether and how the de Finettian story I have told concerning simple laws (about P or the Q_k) carries over to such more complex laws. And I don't know of any working account of conceptual change answering the threat of lawlessness within any given set of properties or conceptual framework. So, there is still a lot to do as well.

However, let me finally emphasize what my brief discussion of λ_0 means in more traditional terms. Kant tried to overcome Hume's objectivity skepticism generally with his transcendental logic and its synthetic principles a priori and Hume's inductive skepticism particularly with his a priori principle of causality. This principle ascertained rather only the rule- or law-guidedness of everything happening and was thus as well called the principle of uniformity of nature (cf., e.g., Salmon 1966, pp.40ff.). As was often observed, this principle did not offer any constructive solution of the problem of induction, since it does not give any direction as to specific causal laws or specific inductive inferences. Still, it provided, if a priori true, an abstract guarantee that our inductive efforts are not futile in principle. *Is* it a priori true?

Nowadays, two notions of apriority are usually distinguished. A proposition is *unrevisably a priori* if it must be believed and cannot be given up under any evidential circumstances. This is certainly the notion which Kant used, though did not express it in this way, and which Quine attacked when attacking analyticity. By contrast, a proposition is *defeasibly a priori* if it is to be believed initially, prior to any experience (and may be given up later on). The prior probabilities discussed by Bayesians are a paradigm of defeasible apriority because they are, of course, expected to change.

Now, our initial ranking function is some regular symmetric κ satisfying NNIR. Via theorem 4, κ uniquely corresponds to some ranking function ρ over Λ. The belief in lawfulness, then, is the same as the disbelief in lawlessness, i.e. $\rho(\lambda_0) > 0$. We saw that this is an extremely reasonable assumption. And we now see that it is tantamount to the defeasible apriority of lawfulness: we must start believing in the uniformity of nature.

The unrevisable apriority of lawfulness, however, is expressed by the stronger condition $\rho(\lambda_0) = \infty$. We also saw that this condition does not appear reasonable, at least if one relates it to the property *P* or, more generally, to any fixed set of properties. Still, it may be unrevisably a priori that there is *some* set of properties with respect to which nature is uniform. I am not prepared to decide whether or not the unrevisable apriority of lawfulness is defensible in this sense. But I think the issue is more clearly arguable on the basis provided here.

References

Carnap, R. (1947), *Meaning and Necessity*, Chicago: University Press, 2nd ed. 1956.

Carnap, R. (1971/80), “A Basic System of Inductive Logic”, Part I in: R. Carnap, R.C. Jeffrey (eds.), *Studies in Inductive Logic and Probability, Vol. I*, Berkeley: University of California Press, 1971, pp. 33-165; Part II in: R.C. Jeffrey (ed.), *Studies in Inductive Logic and Probability, Vol. II*, Berkeley: University of California Press, 1980, pp. 7-155

Cartwright, N. (1989), *Nature’s Capacities and Their Measurement*, Oxford: Clarendon Press.

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.

Gabbay, D.M., P. Smets (eds.) (1998-2000), *Handbook of Defeasible Reasoning and Uncertainty Management Systems*, Vol. 1-5, Dordrecht: Kluwer.

Gärdenfors, P. (1988), *Knowledge in Flux*, Cambridge, Mass.: MIT Press.

Hild, M. (t.a.), *Introduction to Induction. On the First Principles of Reasoning*.

Hintikka, J., Niiniluoto, I. (1976), “An Axiomatic Foundation for the Logic of Inductive Generalization”, in: M. Przelecki, K. Szaniawski, R. Wójcicki (eds.), *Formal Methods in the Methodology of Empirical Sciences*, Dordrecht: Reidel, pp. 57-81.

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

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

Köhler, E. (t.a.), “Comments on Wolfgang Spohn’s Paper”, in: M.C. Galavotti, F. Stadler (eds.), *Induction and Deduction in the Sciences*, Dordrecht: Kluwer.

Kuhn. T.S. (1962), *The Structure of Scientific Revolutions*, Chicago: The University of Chicago Press.

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

Lewis, D.K. (1973), *Counterfactuals*, Oxford: Blackwell.

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

Popper, K.R. (1934), *Logik der Forschung*, Wien: Springer.

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.

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

Spohn, W. (1983), *Eine Theorie der Kausalität*, unpublished Habilitationsschrift, München.

Spohn, W. (1986), “The Representation of Popper Measures”, *Topoi* 5, 69-74.

Spohn, W. (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. (1993), “Causal Laws are Objectifications of Inductive Schemes”, in: J.-P. Dubucs (ed.), *Philosophy of Probability*, Dordrecht: Kluwer, pp. 223-252.

Spohn, W. (1994), “On the Properties of Conditional Independence”, in: P. Humphreys (ed.), *Patrick Suppes: Scientific Philosopher, Vol. 1, Probability and Probabilistic Causality*, Dordrecht: Kluwer, pp. 173-194.

Spohn, W. (2000), “Wo stehen wir heute mit dem Problem der Induktion?”, in: R. Enskat (ed.), *Erfahrung und Urteilskraft*, Würzburg: Königshausen & Naumann, pp. 151-164.

Spohn, W. (2001), “Vier Begründungsbegriffe”, in: T. Grundmann (ed.), *Erkenntnistheorie. Positionen zwischen Tradition und Gegenwart*, Paderborn: Mentis, pp. 33-52.

Spohn, W. (2002), „Laws, Ceteris Paribus Conditions, and the Dynamics of Belief“, *Erkenntnis* 57, 373-394.

van Fraassen, B.C. (1989), *Laws and Symmetry*, Oxford: Clarendon Press.