ENUMERATIVE INDUCTION

Wolfgang Spohn
Fachbereich Philosophie
Universität Konstanz
D-78457 Konstanz

Abstract: The paper explains enumerative induction, the confirmation of a law by its positive instances, in ranking theoretic terms. It gives a ranking theoretic explication of a possible law or a nomological hypothesis. And it shows, finally, that such schemes of enumerative induction uniquely correspond to mixtures of such nomological hypotheses. Thus the paper shows that de Finetti's probabilistic representation theorems may be directly transformed into an account of confirmation of possible laws and that enumerative induction is equivalent to such an account.

1. Introduction

Enumerative induction, or Nicod's rule, is an old philosopher's hat. It says that a law is confirmed by its positive instances. It is the oldest and most primitive of all inductive rules. And it has had a bad press. It seems much too primitive. Even if Popper's polemics against all kinds of inductivism is exaggerated, it has made all too clear that science does not proceed with such simple rules. Goodman's new riddle of induction with the grue emeralds has shown that enumerative induction is inconsistent, if generally applied; but it is hard 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 Carnap's inductive logic – within the variants proposed by Hintikka it would be different – nothing can confirm a law because each law has probability 0 (if its domain of quantification is infinite). But the natural idea was then to turn Nicod's rule into the Principle of Positive Instantial Relevance according to which each positive instance confirms that the next instance is also positive. This seems to be reasonable, and accepted. So, why bother any longer?

Well, the Bayesian home is not entirely comfortable as the point about the null confirmation of laws indicates. But the main point is that "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 or uncertain reasoning should be concerned to understand such a primitive rule as enumerative induction. I promise that by the end of this paper our understanding will be enhanced.

My promise is grounded in the fact that it is no longer true that there is no general qualitative confirmation theory. The reasons why the project had been abandoned in the 70's are nicely summarized in Niiniluoto (1972). So, Bayesianism had won the day. However, logicians and computer scientists were very active since around 1975 in producing alternatives, though not under the labels 'induction' or 'confirmation'. Therefore these activities were little recognized in epistemology and philosophy of science. However, it is precisely in this area where we find a full qualitative account of induction or confirmation. Let me explain.

What should we expect an account of induction to achieve? I take the view that it is equivalent to a theory of belief revision or, more generally, to an account of the dynamics of doxastic states. This is why the topic is so inexhaustible. Everybody, from the neurophysiologist to the historian of ideas, can contribute to it, and one can deal with it from a descriptive as well as from a normative perspective.

Philosophers would like to come up with a very general normative account. Bayesianism provides such an account that is almost complete. There, rational doxastic states are described by probability measures, and their rational dynamics is described by conditionalization rules. However, in order to do justice to the qualitative aspects we should have an account of doxastic states that represents belief or acceptance-as-true. Bayesianism fails here. So, belief revision theory (cf., e.g., Gärdenfors 1988) and other theories were devised to fill the gap. Unfortunately, the dynamics it provides turned out to be incomplete as well (cf. Spohn 1988, sect. 3). There have been several attempts to plug the holes, but I am still convinced that ranking theory, proposed in Spohn (1983, sect. 5.3, and 1988), though under a different name, offers essentially the most adequate account for a full dynamics of belief or acceptance-as-true.[1] Thus, what I want to do here is simply to study enumerative induction in terms of ranking functions.

The plan of the paper is this. In section 2 I shall introduce the theory of ranking functions as far we need it here. Section 3 will then apply ranking theory to enumerative induction; this will indeed turn out to be a brief and 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, finally, 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 trans-

---

[1] One may think that the incompleteness of Bayesianism is filled by Popper measures. However, if one combines the lessons of Spohn (1986) and (1988), it is clear that Popper measures are just as incomplete as is AGM belief revision theory. So, even the Bayesian has reason to buy into ranking theory as suggested in Spohn (1988), sect. 7.

ferring de Finetti's deep account of the confirmation of statistical hypotheses to the deterministic realm.

2. Ranking Functions

Let us start with a set W of possible worlds, small rather than large worlds. Each subset of W is a truth condition or *proposition*. Hence, the set of propositions forms a complete Boolean algebra. I assume propositions to be the objects of doxastic attitudes. We know well that this is problematic, and we scarcely know what to do about the problem. Hence, my assumption is just an act of front alignment.

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 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.

This is all we need to introduce our basic notion: κ is a *ranking function* (for W) iff κ is a function from W into the set of extended non-negative integers $\mathbf{N}^+ = \mathbf{N} \cup \{\infty\}$ 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 *grades of disbelief.* $\kappa(A) = 0$ says that A is not disbelieved at all. It 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. 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. Indeed, 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.[2]

[2] Consistency and deductive closure have been often attacked and equally often defended. The issue of logical omniscience is indeed highly problematic. But we have already decided the issue by taking propositions as objects of doxastic attitude. I don't see viable alternatives.

There are two laws for the distribution of grades of disbelief. The *law of conjunction*: $\kappa(A \cap B) = \kappa(A) + \kappa(B \mid A)$, i.e., the grade of disbelief in A and the grade of disbelief in B given A add up to the grade of disbelief in A-and-B. And the *law of disjunction*: $\kappa(A \cup B) = \min\{\kappa(A), \kappa(B)\}$, i.e., the grade of disbelief in a disjunction is the minimum of the grades of the disjuncts. The latter is just a conditional consistency requirement; if the law would not hold the inconsistency could arise that $\kappa(A \mid A \cup B), \kappa(B \mid A \cup B) > 0$, i.e., both A and B are disbelieved given A-or-B.

According to the above definition, 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 has a minimum and thus that the range of a ranking function is well-ordered. Hence, the range $\mathbf{N}^+$ is a natural choice.[3]

A ranking function is called *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 function from W into $\mathbf{N}$ so that only $\emptyset$ receives infinite rank. Since we want to consider here possibly infinite universal generalizations, 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. For instance, you thus get a ranking version of Bayes' theorem. Or you can develop the whole theory of Bayesian nets in ranking terms. And so on. The general reason is that one can roughly interpret ranks as the orders of magnitude of (infinitesimal) probabilities.

The parallel extends to the laws of doxastic change, i.e., to rules of conditionalization. Thus, it is at least plausible that ranking theory provides a complete dynamics of doxastic states (as may be shown in detail; cf. Spohn, 1988, sect. 5). This is why qualitative inductive matters are best based on them.

Above, I claimed that a full dynamics of belief is tantamount to an account of induction and confirmation. So, what is confirmation with respect to ranking func-

---

[3] 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 that assumption, it may happen that "if A were the case, then B_i would be the case" is true for infinitely many B_i which are jointly unsatisfiable. Lewis finds reason to accept this situation. I prefer to accept the Limit Assumption instead.

tions? The same as elsewhere, namely *positive relevance*: *A confirms* or is a *reason for B* relative to κ iff $\kappa(\overline{B} \mid A) > \kappa(\overline{B} \mid \overline{A})$ or $\kappa(B \mid A) < \kappa(B \mid \overline{A})$ or both.

A final point that will prove relevant 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, if K is 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 a ranking function for W. The function κ^* may be called the *mixture* of K by ρ.

This is all the material we shall need. So let us turn to our proper topic, enumerative induction.

3. Symmetry and Non-negative Instantial Relevance

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 we need no more generality, with the properties we proceed minimally. This will facilitate our business. But it will turn out to be an easy exercise to generalize the results below 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.

Hence, we can represent each possible world by a sequence $\mathbf{a} = (a_1, a_2, ...)$ of 1's and 0's, where $a_n = 1$ or 0 means, respectively, that the n-th object has or lacks P. $\{\mathbf{x}$ takes $a_1, ..., a_n\}$ is short for the proposition $\{\mathbf{x} \mid x_i = a_i \text{ for } i = 1,...,n\}$.

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 therefore define: κ is *symmetric* iff for any sequences $\mathbf{a}$ and $\mathbf{b}$ and any permutation π of $\mathbf{N}$ $\kappa(\mathbf{x}$ takes $a_1, ..., a_n) = \kappa(\mathbf{x}$ takes $b_{\pi(1)}, ..., b_{\pi(n)})$ if $a_i = b_{\pi(i)}$ for $i = 1,...,n$.

Regular symmetric ranking functions take a particularly simple form. For each such function κ there is a *representative function* f from $\mathbf{N} \times \mathbf{N}$ into $\mathbf{N}$ such that $\kappa(\mathbf{x}$

takes $a_1, ..., a_{m+n}) = f(m,n)$ whenever $\sum_{i=1}^{m+n} a_i = m$, i.e., exactly m of the first $m+n$ objects have P. Moreover, we have $f(0,0) = 0$ and the minimum property $f(m,n) = \min[f(m+1,n), f(m,n+1)]$. Indeed, any such function f represents a regular symmetric ranking function. This entails that f can be visualized as in infinite triangle:

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

If a *path* in such a triangle is a sequence which starts with $f(0,0)$ and in which each member is succeeded by its left or right neighbor below, then the minimum property entails that each such path is non-decreasing and 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 confirmation concerning the next instance and confirmation concerning the universal generalization. This is so because, given symmetry and any evidence concerning the first n objects, the rank of the $n+1$st object having (or lacking) P is the same as the rank of any further object having (or lacking) P and thus indeed the same as the rank of *all* further objects having (or lacking) P. Hence, Carnap's problem of the null confirmation of universal generalizations does not turn up in our context, and the recourse to instantial relevance which was only a substitute in the Bayesian framework is here fully legitimate.

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 $n+1$st object having or lacking P confirms, respectively, the $n+2$nd 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 define that a regular symmetric ranking function κ *satisfies PIR* iff {x takes a_{n+1}} confirms {x takes a_{n+2}} given {x takes $a_1, ..., a_n$} whenever $a_{n+1} = a_{n+2}$, i.e., iff $f(m+2,n) - f(m+1,n+1) < f(m+1,n) - f(m,n+1) < f(m+1,n+1) - f(m,n+2)$; and κ *satisfies NNIR* iff {x takes a_{n+1}} does not disconfirm {x takes a_{n+2}} given {x takes a_1, ..., a_n}, i.e., iff the weak inequalities hold instead.

PIR may look like the correct formalization of enumerative induction; alas, it is *inconsistent*. *Proof*: Let us try to satisfy PIR. So, we start with $f(0,0) = 0$ and, say, $f(1,0) = 0$ 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 then we can't complete the fourth line of our triangle. We must set $f(2,1)$ or $f(1,2) = s$, but both choices violate PIR.

Hence, we must settle for the weaker NNIR. It is easily seen to be consistent. In fact, within a probabilistic setting non-negative instantial relevance is entailed by symmetry.[4] Thus it is interesting to note that this is not the case here; it is straightforward to construct symmetric ranking functions violating NNIR.

My preliminary conclusion is that NNIR is the best approximation to enumerative induction within the ranking theoretic setting. It is all the better, since instantial relevance is here tantamount to relevance for the corresponding generalization. So, each symmetric ranking function satisfying NNIR is one way to realize enumerative induction, and there is indeed an infinity of such ways. Still, one should grant that 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. In the final section, though, we shall find a way to fully reestablish positive relevance.

What does this teach us so far? Not so much. 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 a bit about what lawlikeness may mean in ranking theoretic terms.

4. Laws

In our simple setting we had 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 instead as an accidental generalization? I think, as many before me since Ramsey, 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, this belief in G_1 can be realized in many different ways. Let me focus for a while on two particular ways, which I call the 'persistent' and the 'shaky' attitude. If you learn about positive instances of G_1, you do not change your beliefs according to κ, since you expected them to be posi-

[4] Cf. Humburg (1971).

tive, anyway. The crucial difference emerges when we look at how you respond to negative instances according to the various attitudes:

If you have the *persistent* attitude, your belief in further instantiations is unaffected by negative instances, i.e., $\kappa(x_{n+1} = 0) = \kappa(x_{n+1} = 0 \mid x_1 = ... = x_n = 0)$. If, by contrast, you have the *shaky* attitude, your belief in further instantiations is destroyed by a negative instance, i.e., $\kappa(x_2 = 0 \mid x_1 = 0) = 0$, 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 attitude 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:

(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.[5] However, intuitions may not always be certain. But it is clear, of course, that I am describing two extremes here. 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 late that we never come to the point of testing it. But treating a generalization strictly as a law is, if I am right, to take the strictly persistent attitude towards it. The conclusion, hence, is that the characteristic of lawlikeness does not lie in the propositional

[5] Cf., e.g., the overview in Lange 2000, pp. 11f.

content of the generalization, but in our inductive attitude towards it or, rather, its instantiations.

This account is perhaps closest 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 states 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 had many followers since.

From a purely logical point of view, however, it was always difficult 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" was always hard to explain. Still, one might say that the inference-license perspective puts more emphasis on what to do in the single case. This emphasis is not mere rhetorics; 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 one counter-instance, but rather their operation in each single case, which is not impaired by exceptions.

So much for some important agreements. The most obvious disagreement is with Popper, of course. Given how much we have learned from Popper about philosophy of science, my account is really ironic, since it concludes in a way that it is the mark of laws that they are *not* falsifiable by negative instances; it is only the accidental generalizations that are so falsifiable. Of course, the idea that the belief in laws is not given up so easily is familiar at least since Kuhn's days, and even Popper (1934, ch. IV, §22) insisted from the outset that falsifications of laws proceed by more specialized counter-laws rather than simply by counter-instances. But here we have stripped down the point to its induction-theoretic bones.[6]

5. Laws and Enumerative Induction

There seems to be a severe tension between section 3 and 4. We saw that, given symmtery, PIR is not feasible. So we retreated to NNIR as an explication of enumerative induction. Then we noticed that enumerative induction applies only to

[6] The case is more extensively argued in Spohn (2002).

laws. Finally, I have proposed an explication of laws according to which instances are independent of each other. 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 be wrong here.

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 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 then to confirm a law if not to confirm a proposition? Indeed, my above definition of confirmation applies only to the latter, and to talk of the confirmation of laws, i.e., of a second-order inductive attitude towards a first-order inductive attitude, is at best metaphorical. Can we do better?

Yes, 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 famous 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 concentrates 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, i.e., 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.[7] Still, we can immediately transform de Finetti's theory into an account of the confirmation of laws as conceived here. Despite its artificial and formalistic appearance, the basic construction is, I find, illuminating.

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 each $s \in \mathbf{N}$

[7] But see Spohn (1993).

$$\lambda_s(a_1, ..., a_n) = s \cdot (n - \sum_{i=1}^{n} a_i) \text{ and } \lambda_{-s}(a_1, ..., a_n) = s \cdot \sum_{i=1}^{n} a_i,$$

then $\Lambda = \{\lambda_s \mid s \in \mathbf{Z})$ exhausts all the persistent attitudes. Λ contains precisely the symmetric ranking functions according to which each instance is independent from all others. For $s > 0$ λ_s believes in G_1 and disbelieves in each negative instance with rank s. For $s < 0$ it is just the other way around; then λ_s believes in G_0 and disbelieves in each positive instance with rank s. So, what 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, then λ_1. λ_0 has a special role. It does not represent a law at all. It rather represents lawlessness, indeed complete agnosticism; nothing (except the tautology) is believed in λ_0.

Having a second-order attitude towards all these first-order attitudes means in our context having a ranking function ρ over Λ. And the resulting first-order attitude is just the mixture of Λ by ρ as defined in section 2. Now we can start translating de Finetti's theorems.

First, we have: *For each* 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. And the proof of NNIR is no more than a tedious exercise.

Second, we have: For each regular symmetric ranking function κ satisfying NNIR there is a 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. Hence, let us define that the mixture of Λ by ρ is *minimal* iff we have $\rho(\lambda_s) < \infty$ only if λ_s is a non-redundant component of the mixture, i.e., only if there is a proposition A such that min $\{\lambda_r(A) + \rho(\lambda_r) \mid r \in \mathbf{Z}\} <$ min $\{\lambda_r(A) + \rho(\lambda_r) \mid r \in \mathbf{Z} - \{s\}\}$, the first minimum thus being achieved only with λ_s. Hence, in a minimal mixture all redundant components get weight ∞ and cannot enter the mixture at all.

Then, the strengthened claim is: *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, and let f be its representative function forming an infinite triangle of non-negative integers. Now, let's focus on *simple* paths in this triangle taking at most one turn, i.e., taking either the *left-turn* form $f(0,0), ..., f(0,n), f(1,n), f(2,n), ...$ $(n \geq 0)$ or the *right-turn* form $f(0,0), ..., f(m,0), f(m,1), f(m,2), ...$ $(m \geq 0)$. We know already that each such path is non-decreasing. NNIR entails, moreover, that the simple paths don't *accelerate*;

i.e., if $r_0, r_1, \ldots$ is such a path, then $r_{n+1} - r_n \leq r_n - r_{n-1}$, provided that r_{n-1}, r_n, r_{n+1} do not form the turn. Hence, each such path either goes to infinity or reaches a maximum and then remains constant. So, let us define a_n to be the maximum of the n-th left-turn path $f(0,0), \ldots, f(0,n), f(1,n), f(2,n), \ldots$, and b_m to be the maximum of the m-th right-turn path $f(0,0), \ldots, f(m,0), f(m,1), f(m,2), \ldots$ $(m,n \geq 0)$. Again, both sequences $a_0, a_1, \ldots$ and $b_0, b_1, \ldots$ must be non-decreasing and non-accelerating, either $a_0 = 0$ or $b_0 = 0$ or both, and $\max\{a_i\} = \max\{b_i\} := c$, which may be either finite or infinite.

With the help of the two sequences we can construct now the relevant minimal mixture ρ. Take $a_0, a_1, \ldots$ first, and let a_n be any point at which the sequence flattens, i.e., such that $a_n - a_{n-1} > a_{n+1} - a_n := s$. Then we set $\rho(\lambda_s) = a_n - ns$. Similarly, if b_m is a point at which $b_0, b_1, \ldots$ flattens, i.e., such that $b_m - b_{m-1} > b_{m+1} - b_m := r$, then we set $\rho(\lambda_{-r}) = b_m - mr$. If for any $t \in \mathbf{Z}$ $\rho(\lambda_t)$ is not defined thereby, then we set $\rho(\lambda_t) = \infty$. 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.

Clearly, $\rho(\lambda_t) = 0$ for some $t \in \mathbf{Z}$, since either $a_0 = 0$ or $b_0 = 0$. Hence, ρ is a ranking function over Λ. Moreover, it is not difficult to verify that the mixture thus defined generates exactly the representative function f; hence, κ *is* the mixture of Λ by ρ. Moreover, this mixture is minimal and unique; any other ρ over Λ would either generate not a minimal mixture or a different one; this may again be easily inferred from the construction. *End of proof.*

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 then given simply 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.,

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

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

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

This suggests that ρ^* as defined by

$$\rho^*(\lambda_s) = \rho(\lambda_s) + ns - f(m,n) \text{ and } \rho^*(\lambda_{-r}) = \rho(\lambda_{-r}) + mr - f(m,n)$$

is the $\rho_{m,n}$ we are looking for. However, ρ^* may take negative values and it need not generate a minimal mixture. Hence, we define for all $s \in \mathbf{Z}$:

$\rho_{m,n}(\lambda_s) = \rho^*(\lambda_s)$, if $\rho^*(\lambda_s) \geq 0$ and there is no s' with $|s'| < |s|$ such that $\rho^*(\lambda_{s'}) \leq \rho^*(\lambda_s)$; $\rho_{m,n}(\lambda_s) = \infty$, otherwise.

It is then easy to verify that $f_{m,n}$ is indeed generated by $\rho_{m,n}$ and hence that $\kappa_{m,n}$ *is the minimal mixture of* Λ by $\rho_{m,n}$.

This observation has three immediate 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 $r,s \geq 1$:

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

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, the λ_{-r} as opposed to the λ_s are 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 positive instantial relevance that may be blurred in the mixed first-order attitude satisfying only NNIR.

This 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 belief that the negative lawlike attitudes are incorrect heads for infinite firmness. This parallels de Finetti's observation in the probabilistic case.

Finally, we should take a look at the special role of λ_0, which has indeed no probabilistic analogue. 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 and invariable belief that some law or other will obtain. This does not appear very 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*.

This may also sound implausible. However, $\rho(\lambda_0)$ may be very large so that the agnostic state is in fact never reached. More importantly, the whole story I have told about the single property *P* can be generalized to any finite number of properties in a straightforward way. 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 all these considerations within a larger space of properties.

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

References

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.

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

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

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.

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.

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. (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.