Laws:
A Ranking Theoretic Account

Wolfgang Spohn

FEW 2008
Madison, WI, May 15-18

Preview

This talk may be understood:

- as an application of ranking theory to philosophy of science, or
- as a proposal for understanding what a strict or deterministic law is, or, still more specifically,
- as a justification of enumerative induction and an explanation why it works only for laws.

And it proceeds by

- an introduction,
- a brief list of the required essentials of ranking theory,
- a lay-out of the formal set-up of our specific discussion,
- a partially surprising discussion of the confirmation of generalizations,
- an account of what laws are (pro Ramsey and contra Popper),
- A de Finettian account of the confirmation of such laws.

The General Philosophical Strategy

Trying to understand natural necessity (such as nomic necessity, causal necessity, probability, counterfactuality, and determination, dispositions, powers, forces, and capacities)

not via Humean Supervenience,

but via Humean Projection

as covertly epistemological notions.

Here, I shall talk only about nomic necessity or lawlike-ness.

The Problem of Lawlikeness

Is to say what (with the exception of their truth) is characteristic of laws so that one can distinguish:

(1a) All (past, present, or future) humans on earth are less than 150 years old.

(1b) All persons (presently or ever) in this room are less than 80 years old.

(2a) All uranium spheres have a diameter of less than one mile.

(2b) All gold spheres have a diameter of less than one mile.

(3a) All apples in mystery bowl are green.

(3b) All apples in my bowl are green.

Six Characteristic Features

(i) Laws are completely general [not true],

(ii) Laws refer to all possibilities, not only to actuality [too circular],

(iii) Laws are systematic [of doubtful use],

(iv) Laws have explanatory force,

(v) Laws support counterfactuals,

(vi) Laws are projectible to and confirmed by their instances.

The Further Strategy

The last three features are closely connected.

We may either account for them via some underlying feature [hopeless, as the history of inductive skepticism shows],

or explicate them directly, perhaps by taking one of the three as basic.

Leaving aside a long argument, I take projectibility and confirmability as my starting point.

And I shall base my inquiry on ranking, not probability theory as the pertinent acount of confirmation.

Ranking Functions I

Let W be a possibility space and $\mathcal{A}$ be a complete algebra of propositions over W.

κ is a (*negative*) *ranking function* over $\mathcal{A}$ iff κ is a function from $\mathcal{A}$ into $\mathbf{N}^+ = \mathbf{N} \cup \{\infty\}$ such that

(a) $\kappa(W) = 0$ and $\kappa(\emptyset) = \infty$,

(b) for any $\mathcal{B} \subseteq \mathcal{A}$ $\kappa(\bigcup \mathcal{B}) = \min \{\kappa(A) \mid A \in \mathcal{B}\}$ [the law of (infinite) disjunction],

It follows: (c) $\kappa(A) = 0$ or $\kappa(\neg A) = 0$ [the law of negation].

Ranking Functions II

Ranking functions are gradings of disbelief.

$\kappa(A) > 0$ says that A is disbelieved (to some degree),

$\kappa(A) = 0$ says that A is not disbelieved,

$\kappa(\neg A) > 0$ says that A is believed.

Axioms (a) and (b) entail that belief is consistent and deductively closed.

One might also define a *positive ranking function* π by $\pi(A) = \kappa(\neg A)$ expressing belief directly or a *two-sided ranking function* τ by $\tau(A) = \pi(A) - \kappa(A)$ expressing both at once, so that $\tau(A) > 0$, < 0, or $= 0$, according to whether A is believed, disbelieved or neutral.

Negative ranking functions are the theoretically most fruitful version.

Ranking Functions III

If $\kappa(A) < \infty$, the *conditional rank of B given A* is defined as $\kappa(B \mid A) = \kappa(A \cap B) - \kappa(A)$. This is equivalent to $\kappa(A \cap B) = \kappa(A) + \kappa(B \mid A)$ [the law of conjunction].

Given this definition, the law of (infinite) disjunction asserts nothing but conditional consistency.

There is a pervasive analogy between probability and ranking theory: Replace the sum of probabilities by the minimum of ranks and the product and quotient of probabilities by the sum and the difference of ranks, and any probabilistic theorem is almost guaranteed to turn into a ranking theorem.

Ranking Functions IV

A is a *reason for*, or *confirms*, or is *positively relevant*, to B wrt κ or τ iff $\tau(B \mid A) > \tau(B \mid \neg A)$.

A is a *reason against*, or *disconfirms*, or is *negatively relevant*, to B wrt κ or τ iff $\tau(B \mid A) < \tau(B \mid \neg A)$.

A is *irrelevant to*, or *independent of*, B wrt κ or τ iff $\tau(B \mid A) = \tau(B \mid \neg A)$.

Positive ranking relevance behaves very much like positive probabilistic relevance. It is reflexive, symmetric, but not transitive, and it embraces deductive reasons.

The Formal Set-up I

Let $\mathcal{A}$ be generated by a sequence of variables X_1, X_2, … that are all repetitions of the first variable X_1, and that are all functions from the possibility space into some finite range I.

E.g.: X_1 = Raven #1 is black or not, X_2 = Raven #2 is black or not, etc.

Or: X_1 = trajectory through state space in trial #1, X_2 = trajectory through state space in trial #2, etc.

Thus, the possibility space W may be conceived to consist of all sequences $\mathbf{w} = (w_1, w_2, \ldots)$, where each $w_n \in I$, so that $X_n(\mathbf{w}) = w_n$.

The Formal Set-up II

We are interested at least in generalizations.

For $J \subseteq I$ let $G_J = \{\mathbf{w} \in W \mid \text{for all } n\ w_n \in J\}$ saying that all objects or trials take values in J,
and $G^J = \{\mathbf{w} \in W \mid \text{for all } n\ w_n \notin J\}$ saying that no object or trial takes a value in J.

The generalizations $G^i = G^{\{i\}}$ ($i \in I$) can be taken as basic since $G^J = \bigcap_{i \in J} G^i$.

We shall see that such generalizations or regularities may not be taken as laws.

The Formal Set-up III

We shall need a way of denoting (sequences of) singular facts.

If $i \in I$ and $J \subseteq I$, let us use $\{w_n = i\}$ and $\{w_n \in J\}$ as short for $\{\mathbf{w} \in W \mid w_n = i\}$ and $\{\mathbf{w} \in W \mid w_n \in J\}$, i.e., for the proposition that the variable X_n takes the value i or some value in J.

Finally, let $\mathcal{A}_n$ be the complete algebra of propositions only about the first n objects or trials, i.e., generated by $X_1, \ldots, X_n$ or by propositions of the form $\{w_1 = x_1, \ldots, w_n = x_n\}$.

The Confirmation of Generalizations I

The task to study the inductive behavior of laws probably means to study how such generalizations are confirmed by singular facts wrt to a ranking function κ. I shall assume that:

(a) κ is *regular*, i.e., that for all n and all (molecular) non-empty $A \in \mathcal{A}_n$ $\kappa(A) < \infty$,

(b) κ is *symmetric*, i.e., that κ assigns the same rank to all permutations of a given sequence of results.

For a symmetric ranking function it matters only how many times the values in I are realized. If $E_\mathbf{n}$ says that the value i realizes with the absolute frequency n_i ($i \in I$) in the first n trials, we have $\kappa(E_\mathbf{n}) = \kappa(A)$ for each $\emptyset \neq A \subseteq E_\mathbf{n}$.

The Confirmation of Generalizations II

There is no problem of null confirmation of infinite generalizations for symmetric ranking functions.
A generalization is just as firmly believed as each of its instances.

In other words: For $J \subseteq I$,
$\kappa(w_1 \in J) = \kappa(w_n \in J)$ for all n,
and since $\neg G^J = \bigcup_{n \in \mathbf{N}} \{w_n \in J\}$
$\kappa(\neg G^J) = \kappa(w_1 \in J)$ due to the law of infinite disjunction.

The Confirmation of Generalizations III

In particular, our task is to study the confirmation of laws through their instances and thus perhaps first the same for generalizations. So we define:

(Regular symmetric) κ *satisfies PIR$_c$* (i.e., *positive instantial relevance in the conditional sense*) iff for the associated τ, for all $E \in \mathcal{A}_n$ ($n \geq 0$), and all $i \in I$

$$\tau(\{w_{n+2}=i\} \mid E \cap \{w_{n+1}=i\}) > \tau(\{w_{n+2}=i\} \mid E \cap \{w_{n+1} \neq i\}).$$

κ *satisfies PIR$_n$* (i.e., *PIR* in the *non-conditional sense*) iff for all $E \in \mathcal{A}_n$ and all $i \in I$

$$\tau(\{w_{n+2} = i\} \mid E \cap \{w_{n+1} = i\}) > \tau(\{w_{n+2} = i\} \mid E).$$

The Confirmation of Generalizations IV

κ *satisfies*, respectively, *$NNIR_c$* or *$NNIR_n$* (*non-negative instantial relevance in the conditional* and *the non-conditional sense*) iff the weak inequality ≥ holds in *PIR_c* or *PIR_n* instead of the strict >.

Theorem: PIR_n entails PIR_c, $NNIR_c$ entails $NNIR_n$, and none of the reverse entailments holds.

Theorem: Symmetry implies NNIR in the probabilistic context, but not the ranking context.

Theorem: There is no regular symmetric ranking function for A satisfying PIR_c.

The Confirmation of Generalizations V

What does this negative result teach us about lawlikeness? Since the confirmation of the next instances is tantamount to the confirmation of generalizations (concerning all future instances), we must conclude that all generalizations G_J (concerning future instances) can be confirmed by at most finitely many positive instances. Hence, it seems we have reached a dead end; we cannot use enumerative induction, i.e., confirmability by positive instances for characterizing strict laws.

We may settle for the weaker NNIR_c, if this is the only ranking-theoretically feasible way. This seems unsatisfactory, though, since it leaves arbitrary room for instantial irrelevance.

Laws and Persistence I

So far, we inquired the confirmability of generalizations in general, not that of laws in particular; but so far we have no hint to the latter.

Also, the inquiry may have seemed misguided insofar as it studied PIR and NNIR under any evidence whatsoever, even evidence that has (repeatedly) falsified the relevant generalization. However, this is not as absurd as it seems.

Projectibility of a law is usually equated with its confirmability; we project its past success onto the future. This we do with lawlike, not with accidental generalizations. But we may give projectibility a stronger reading, not as the extension of past observations to future cases, but simply as the continuous application of the law to future cases, whatever the past. Let me explain:

Laws and Persistence II

Let us ask: What does it mean to believe in a law? Hard to say as long as we don't know what a law is. However, we know what it is to believe in a generalization G_J : namely, $\kappa(\neg G_J) > 0$.

However, this belief can be realized in many ways, even for symmetric κ; the inductive relations among the various instances can take many forms. Which form should they take? It is clear that as long as one observes only positive instances the belief in G_J is maintained or even strengthened; this is what NNIR_c implies.

What happens, though, when one observes negative instances? If the negative instances get believed, then G_J gets disbelieved according to any κ; this is a trivial matter of deductive logic. However, with regard only to the future instances, anything might happen according to κ, even given symmetry and NNIR_c. Let us look at two paradigmatic (extreme) responses to negative evidence.

Laws and Persistence III

If you have the (very) *persistent* attitude, your belief in further po-sitive instances is unaffected by observing negative instances, i.e.,

$\tau(\{w_{n+1} \in J\} \mid \{w_1, \ldots, w_n \notin J\}) = \tau(\{w_{n+1} \in J\}) > 0.$

If, by contrast, you have the (very) *shaky* attitude, your belief in further positive instances is destroyed by the first negative instance (and due to $NNIR_c$ also by several negative instances), i.e.,

$\tau(\{w_{n+1} \in J\} \mid \{w_1, \ldots, w_{n-1} \in J\} \cap \{w_n \notin J\}) \leq 0.$

I want to suggest that the different attitudes are distinctive of treating generalizations as lawlike or accidental. (Look at our initial problematic examples.)

Laws and Persistence IV

With this suggestion I am painting black and white. I admit there is also a lot of grey. Still, I insist on my ideal types and would like to offer four considerations.

(1) Intuitions may be probabilistically biased and thus induce reluctance towards my suggestion.

(2) The ideal types delimit a broad range of less extreme attitudes.

(3) When confronted with apparent counter-instances we never simply accept them shaking our generalization or persistently write them off.

Laws and Persistence V

(4) (and most importantly): I only want to claim that, if you would firmly believe in the law and nothing else, then you would have the persistent attitude.

The last point also dissolves the complaint that the persistent attitude is obviously unable to learn. The doxastic attitude that expresses the belief in a law need not display learning ability by itself. What is rationally required is only that one is not tied to that doxastic attitude, but is able to change it in response to evidence.

Laws and Persistence VI

However, if the belief in a law is a certain kind of doxastic attitude, we so far have no account of the change and of the confirmation or disconfirmation of such an attitude i.e., of the belief in a law.

Thus, our first-order account of the confirmation of propositions needs to be complemented by a so far missing second-order account of the confirmation of such first-order attitudes. Such an account will be sketched in the last part of the talk.

Laws and Persistence VII

Ramsey (1929) says: "Many sentences express cognitive attitudes without being propositions; and the difference between saying yes or no to them is not the difference between saying yes or no to a proposition" (pp. 135f.). And "… *laws are not either*" (namely pro-positions) (p. 150). Rather: "The general belief con-sists in (*a*) A general enunciation, (*b*) A habit of singular belief" (p. 136).

This is the familiar view that laws are inference rules or inference licences. That may be either trivial or obscure, but it might be said to emphasize the single case, as does my notion of persistence.

Laws and Persistence VIII

Given how much philosophy of science owes to Popper, my account is really ironic, since it concludes that it is the mark of laws that they are *not* falsifiable by negative instances (this is the persistent attitude); only accidental generalizations are so falsifiable (this is the shaky attitude).

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 insisted from the outset that the falsification of laws proceeds by more specialized counter-laws rather than simply by counter-instances. However, I have not seen the point elsewhere being so radically stripped to its induction-theoretic bones.

Laws and Persistence IX

So, my core claim is that the belief in a strict law consists in the persistent attitude. The latter in turn is characterized by the ranking-theoretic *independence* of the instantiations of the law, i.e., of the variables X_1, X_2, … In other words, if ξ is any ranking function for I, we may define λ_ξ as the independent and identically distributed infinite repetition of ξ.

Does the core claim help us to say what a law *is*? So far, I was careful only to analyze the *belief* in a law. The point is that the belief in a law turned out, following Ramsey, to be not (merely) the belief in something; "belief in a law" is a not further parsable phrase.

Laws and Persistence X

I propose to call the λ_ξ's *subjective laws*. This is sufficiently artificial a term to indicate the rhetorical move explicitly. And there is a twofold justification behind that move.

(1) We shall be able to fully explain the phrase "belief in a law" in a parsable manner, as belief in something, namely exactly in such λ_ξ's; this has to do with the distinction of first-order and second-order attitudes already indicated.

Laws and Persistence XI

(2) The adjective “subjective” is not merely to signify that subjective laws are still something doxastic, i.e., entertained by doxastic subjects. Subjective laws are so far nothing that could be called true or false.

How such subjective laws can be objectivized so as to be called true or false and what that could mean at all is another story that I shall attempt to tell at the First Formal Epistemology Festival on Conditionals and Ranking Functions on July 28-30 in Konstanz (where you are cordially invited).

The Confirmation of Laws I

What could it mean to confirm subjective laws? Fortunately, there is clear precedent provided by Bruno de Finetti's philosophy of probability. Given the close similarity between probability and ranking theory, we may attempt to translate de Finetti's account of statistical laws. That attempt indeed works.

First, we can note that what I called a subjective law is nothing but the ranking-theoretic analogue to a Bernoulli measure.

De Finetti famously proved: each symmetric probability measure for the infinite sequence of random variables is a unique mixture of Bernoulli measures.

The Confirmation of Laws II

Thus, if the symmetric measure expresses your subjective probabilities and the Bernoulli measures represent hypotheses about objective probabilities, your subjective opinion is a unique mixture of objective statistical hypotheses, the weights of the mixture representing your credence in these hypotheses.

The mixture changes through evidence that favors the hypotheses close to the observed relative frequencies and disfavors the other ones.

In fact, if the evidence converges to a certain limit of relative frequencies the mixture converges to the Bernoulli measure taking these limiting relative frequencies as objective probabilities.

The Confirmation of Laws III

This works in the very same way in ranking theory:

We start from the set Λ of possible subjective laws for $\mathcal{A}$ (first-order attitudes), and we consider a complete negative ranking function ρ for Λ (a second-order attitude) that is to represent our (dis-)belief in the laws.

We can mix the subjective laws in Λ by ρ and receive *the mixture* κ defined by:

$$\kappa(A) = \min\{\lambda(A) + \rho(\lambda) \mid \lambda \in \Lambda\} \text{ for all } A \in \mathcal{A}.$$

Let us call ρ an *impact function*; the subjective law λ has *impact* $\rho(\lambda)$ on the mixture κ.

The Confirmation of Laws IV

Theorem: A ranking function κ for $\mathcal{A}$ is regular, symmetric, and concave if and only if it is the mixture of the set Λ of subjective laws by some proper impact function ρ.

Theorem: For each regular symmetric concave ranking function κ for $\mathcal{A}$ there is a unique impact function ρ for Λ such that κ is the minimal mixture of Λ by ρ.

Theorem: A regular symmetric concave ranking function satisfies $NNIR_c$.

Theorem: If $I = \{1, 2\}$, i.e., if the X_i are only two-valued variables, then $NNIR_c$ entails concavity.

The Confirmation of Laws V

Consequence: We can thereby fully restore positive instantial relevance with respect to the subjective laws instead of the next instances.

If X_{n+1} taking value i is more disbelieved according to λ than in $\kappa_{\mathbf{n}}$, i.e. κ given evidence $E_{\mathbf{n}}$, then the additional observation of X_{n+1} taking value i decreases the impact of λ proportionally, i.e., λ gets proportionally more *dis*believed.

If that observation is exactly as unexpected according to λ as according to $\kappa_{\mathbf{n}}$, the impact of λ does not change.

If that observation was less disbelieved according to λ than in $\kappa_{\mathbf{n}}$, then λ gets proportionally confirmed, i.e., less disbelieved.

This is how we intuitively expect the confirmation of laws to behave.

The Confirmation of Laws VI

Consequence: Precisely those subjective laws that do not exclude the values infinitely often realized do not drop out of the mixture. All the other ones get disbelieved with ever greater firmness diverging to infinity.

The Long and the Short of All This

Take everything you know about statistical laws, translate it from probability into ranking theory, and you learn everything you wanted to know about strict or deterministic laws!

Thanks so much for your attention!