Notes for Week 5 of *Confirmation*

09/26/07

Branden Fitelson

1 Preliminary Remarks About Nicod, Probability, and Inductive Logic

While Nicod intends to be carrying forward the work of Keynes on universal induction, there is a noticeable shift from Keynes to Nicod. Keynes was primarily talking about induction as an *epistemological* problem. Nicod, on the other hand, seems to prefer to think of induction in "*logical*" terms. [Historical note: the word "confirmation" is used for the first time here by Nicod in connection with a "logical" conception of induction. This sets the stage for the subsequent work of Hempel, Carnap, and others.] Unfortunately, Nicod makes several significant logical errors during the course of setting up his essay on "the logical problem of induction". As a result, several of his main lines of argument are frustrated, and various confusions arise.

Before delving into some of the details of Nicod's essay, I want to discuss a few of these confusions. First, an important distinction that arises (in rather inchoate/confused form) in Nicod's discussion is that between *probability logic* (PL) and *probabilistic inductive logic* (PIL). (PL) is *deductive*: it involves deductive entailment relations between certain kinds of probability statements, and it is typically used to provide a theoretical account of how uncertainty in the premises of various sorts of arguments "propagate through" to their conclusions. (PIL), on the other hand, is *inductive*: it involves genuine probabilistic relations between premises and conclusions, where these relations are themselves *logical* in nature (they are *analogous* to deductive relations). When Nicod sets up the framework of his discussion, he begins with a sort of "preamble":

> We first regard inference as the perception of a connection between the premises and conclusion which asserts that the conclusion is true if the premises are true. This connection is implication, and we shall say that an inference grounded in it is a *certain inference*. But there are weaker connections which are also the basis of inferences. They have not until recently received any universal name. Let us call them with Mr. Keynes *relations of probability*. The presence of one of these relations among the group A of propositions and the proposition B indicates that in the absence of any other information, if A is true, B is probable to a degree p. A is still a group of premises, B is still a conclusion, and the perception of such a relation between A and B is still an inference: let us call this second kind of inference *probable inference*.

Here, Nicod is postulating *logical* probability relations (analogous to entailments), the "perception" of which constitute inductive *inference* (presumably, an *epistemological* relation). He seems to think this is what Keynes has in mind. I don't think so. As I explained last week, I think Keynes's probabilities are (all) *epistemic*, *not* logical.[1] Be that as it may, Nicod is clearly interpreting Keynesian *conditional probability* statements $\ulcorner \Pr(B \mid A) = p \urcorner$ as *conditionals* with probabilistic consequents: $\ulcorner$if A is true, B is probable to a degree $p\urcorner$. This is also unfortunate, since this has $\ulcorner \Pr(B \mid B) = 1 \urcorner$ coming out as $\ulcorner$if B is true, B is probable to degree $1\urcorner$, which is crazy, since not all true propositions have probability 1 (in *any* sense of "probability"), but all true propositions *do* confer conditional probability 1 on themselves (on any reasonable theory of conditional probability, including Keynes's). Continuing on with his "preamble", Nicod says the following:

> First of all, probable inference says that in the absence of any other information, the truth of its premises renders its conclusion probable to a degree p. It then says *less* than what is claimed by certain inference, as we have just called it; but it says it with as complete certainty. It does not conclude in a *probable manner*, but it reaches a *probable conclusion* in as certain a manner as the other reaches its conclusion.

Thus, Nicod is clearly talking about (PL) *not* (PIL) here. He makes his meaning even more clear in the following final section of the "preamble", where he more precisely distinguishes certain *vs* probable inference:

> We have so far considered only premises that are certain. But any inference which yields something, starting from premises taken as certain, still yields somethlng starting from premises taken only as probable, and this holds for both types of inferences, certain and probable. We can even assert that starting from premises which, taken together have a probability p, a certain inference will confer on its conclusion the same probability p; and a probable inference, which would confer on its conclusion the probability q if its premises were certain, will confer on its conclusion the probability pq.

Here's what Nicod has in mind. On the one hand, we can represent what Nicod calls "certain inference" as the following deductive argument pattern, involving probabilistic premises and a probabilistic conclusion:

[1] This is the only sensible reading of Keynes, in light his insistence that his inductive methods are applicable even to *arithmetic*. When we get to Carnap, we'll see a precise and explicit account of logical probabilities, and how they relate to epistemic probabilities.

$$\Pr(A \mid \top) = p. \left[\begin{array}{l}\text{“starting from premises which,}\\ \text{taken together have a probability } p\text{”}\end{array}\right]$$

$$\Pr(B \mid A) = 1. \text{ [“a certain inference”]}$$

$$\therefore \Pr(B \mid \top) = p. \text{ [“will confer on its conclusion the same probability } p\text{”]}$$

This isn't quite right. If the conclusion were $\Pr(B \mid \top) \geq p$ instead of $\Pr(B \mid \top) = p$, then this would be a deductively valid argument, and a theorem of (PL). It is supposed to undergird Nicod's claim that "starting from premises which, taken together have a probability p, a certain inference will confer on its conclusion the same probability p." The real underlying (PL)-moral here is that certain inferences cannot have conclusions that are any *less* probable than their (least probable) premises. On the other hand, (merely) *probable* inferences generally *are* apt to have conclusions that are less probable than their (least probable) premises. Nicod claims that the following is a theorem of (PL), which characterizes "probable inference":

$$\Pr(A \mid \top) = p. \left[\begin{array}{l}\text{“starting from premises which,}\\ \text{taken together have a probability } p\text{”}\end{array}\right]$$

$$\Pr(B \mid A) = q. \left[\begin{array}{l}\text{“a probable inference, which would confer on its con-}\\ \text{clusion the probability } q \text{ if its premises were certain”}\end{array}\right]$$

$$\therefore \Pr(B \mid \top) = pq. \text{ [“will confer on its conclusion the probability } pq\text{”]}$$

Unfortunately, as Hailperin (and Braithwaite) points out, this is *not* a deductively valid argument/theorem of (PL) either. Again, the conclusion of this argument should be $\Pr(B \mid \top) \geq pq$, not $\Pr(B \mid \top) = pq$.[2] In any event, Nicod is close to articulating another theorem of (PL) here. And, Nicod seems to think that such theorems characterize both certain and probable inference. While this is not really what Keynes had in mind (nor is it what Carnap had in mind), it is a stance that other commentators have subsequently taken in the literature since Nicod. For instance, Popper (and in some moods Adams and Hailperin) had views similar to Nicod's. Such "probability logicians" see all inference as — at bottom — *deductive*. For them, the conclusions of "good inductive arguments" are *probability statements* which are *entailed by* other probability statements. I won't be discussing (PL) much more after this week, but this was a good opportunity to bring out this important distinction between (PL) and (PIL). We will be talking a lot more about (PIL) in the coming weeks.

2 Nicod on Primary *vs* Secondary Inductions

For Nicod, a *primary induction* is one "whose premises do not derive their certainty or probability from any induction", whereas a *secondary induction* "has among its premises the conclusion of another induction." Appealing to the (PL)-style arguments described above (as he does throughout the essay), Nicod argues that:

> the probability conferred by an inference, of any sort, upon its conclusion is at most equal to that of the least probable of its premises. The probability supplied by any induction whatsoever cannot exceed the highest probability that a primary induction can yield. That is why primary induction should be analyzed before any other. For it is not only the logical foundation of induction, but it also marks the limit of all inductive assurance.

Here, again, Nicod seems to misunderstand the significance of the (PL)-theorems discussed above. What follows from them is that the *lower bound* of the probability conferred by an inference, of any sort, upon its conclusion is at most equal to that of the least probable of its premises. Nonetheless, with this correction, there is a kernel of truth in what Nicod is saying here. Since a secondary induction will already contain the conclusion of a primary induction as a premise, there is a sense in which primary inductions *can* constrain the lower bound on the probability of the conclusion of a secondary induction. But, this constraint will have no force, unless the conclusion of the primary induction in question is less probable than the other premises of the secondary induction in question. So, it seems untrue that primary inductions will (always) constitute the sort of "limit of all inductive assurance" that Nicod seems to have in mind here. Moreover, as Braithwaite points out, Nicod is hung up on the *posterior probability* of the conclusion, rather than the *degree of relevance* that the premises have for the conclusion. Recall that Keynes (especially in his discussion

[2] In the special case where $\Pr(A \mid \top) = p = 1$, this argument *is* valid. This explains why we *can* safely interpret "a probable inference, which would confer on its conclusion the probability q if its premises were certain" as $\Pr(B \mid A) = q$ in its second premise.

of Humean "circularity", which involved a case of secondary induction in Nicod's sense) emphasized the *relevance* of premises to conclusion, rather than *posterior probability* conferred on conclusion by premises.

3 Nicod on the Asymmetry Between Confirmation and Invalidation

Nicod notes an asymmetry between confirmation and invalidation:

> A favourable case increases more or less the probability of a law, whereas a contrary case annihilates it entirely. Confirmation supplies only a probability; invalidation on the contrary, creates a certainty. Confirmation is only favourable, while invalidation is fatal.

Here, Nicod is echoing part of what Keynes says. Recall that Keynes was committed to both of the following claims (where $K_\top$ contains whatever presuppositions are required for inductive methods to "work"):

- $\Pr[(\forall x)(Fx \supset Gx) \mid Fa \,\&\, Ga \,\&\, K_\top] > \Pr[(\forall x)(Fx \supset Gx) \mid K_\top] > 0$
- $\Pr[(\forall x)(Fx \supset Gx) \mid Fa \,\&\, {\sim}Ga \,\&\, K_\top] = 0$

Nicod is also accepting these (with respect to "logical probability", I'd say, as opposed to Keynes's $\Pr_\phi$), and he is calling the first *confirmation*, and the second *invalidation*. There is clearly an asymmetry here. And, the asymmetry can be put in another way: The degree to which $Fa \,\&\, {\sim}Ga$ disconfirms $(\forall x)(Fx \supset Gx)$ is greater than the degree to which $Fa \,\&\, Ga$ confirms $(\forall x)(Fx \supset Gx)$. The idea here is that $Fa \,\&\, {\sim}Ga$ *refutes* $(\forall x)(Fx \supset Gx)$, whereas $Fa \& Ga$ is merely positively relevant to $(\forall x)(Fx \supset Gx)$. As I mentioned last time, Keynes also talks about ${\sim}Fa \,\&\, Ga$ and ${\sim}Fa \,\&\, {\sim}Ga$ as well. He thinks the latter confirms $(\forall x)(Fx \supset Gx)$, and he's unsure about the former. Interestingly, Nicod never discusses these other cases at all (more on that next time in Hempel). Moreover, as I mentioned last time, it is rather unclear as to why anyone should accept the first claim about "positive instances" confirming universal "laws". After all, $(\forall x)(Fx \supset Gx) \nvDash Fa \,\&\, Ga$, so why think such "instances" *must* raise the probability of their corresponding universal generalizations?

As I pointed out last time, it *is* true that $(\forall x)(Fx \supset Gx) \vDash Fa \supset Ga$. As a result, we *do* have the following as a *theorem* of probability calculus (assuming non-extreme probabilities for the "laws" and "instances"):

- $\Pr[(\forall x)(Fx \supset Gx) \mid Fa \supset Ga \,\&\, K_\top] > \Pr[(\forall x)(Fx \supset Gx) \mid K_\top] > 0$

With this in mind, let $\mathcal{H} \overset{\text{def}}{=} (\forall x)(Fx \supset Gx)$, and let $\mathcal{E}_\supset \overset{\text{def}}{=} Fa \supset Ga$. Then, we can state the following asymmetry *desideratum* that any probabilistic relevance measure $\mathfrak{c}$ of degree of confirmation should satisfy:

- $\mathfrak{c}({\sim}\mathcal{H}, {\sim}\mathcal{E}_\supset \mid K_\top) > \mathfrak{c}(\mathcal{H}, \mathcal{E}_\supset \mid K_\top) > 0$

It seems to me that this is the correct asymmetry claim that should come out of Nicodian considerations (indeed, we might even make the stronger claim that $\mathfrak{c}({\sim}\mathcal{H}, {\sim}\mathcal{E}_\supset \mid K_\top)$ should take a *maximal* value in some more absolute sense). We'll return to this issue later when we talk about Hempel, Carnap, and Goodman.

Putting this error to one side, Nicod thinks that this sort of asymmetry has a tendency to "lure" people into thinking *all* induction is grounded in invalidation.[3] He thinks Keynes has fallen prey to this affliction himself, in his treatment of Analogy. Keynes thinks all universal induction relies to some extent on Analogy in addition to Pure Induction (which Nicod calls "induction by *confirmation*"). Nicod reads Keynes's treatment of Analogy as resting asymmetrically on *invalidation*. I think this is somewhat unfair. It is true that Keynes talks about getting "as close to perfect analogy as possible" by "reducing the comprehensiveness of those resemblances ϕ_1 between the instances which our generalization disregards". And, it is also true that in cases of imperfect analogy and incomplete knowledge of analogies and disanalogies, this will involve some amount of invalidation. But, in the limiting case of perfect analogy itself (and complete knowledge of all analogies and disanalogies), Keynes has to be committed to the basic claim that $Fa \,\&\, Ga$ — *in the absence of any other knowledge whatsoever* — is favorable to $(\forall x)(Fx \supset Gx)$, relative to $K_\top$. Thus, at bottom, I think Keynes is relying almost as heavily on confirmation as he is on invalidation. In this sense, Keynes is no "invalidationist". Moreover, Nicod's arguments in his critical discussion of Keynes on Analogy all trade on his (mis)understanding of the (PL)-style theorems we discussed above. So, it seems they don't have much probative value. For this reason, I will focus mainly on Nicod's objections to Keynes on Pure Induction, which have to do with the conditions under which multiple observations of "instances" will drive the posterior probability of a universal claim to one. Here, Nicod has many interesting things to say.

[3] Karl Popper comes to mind as a philosopher who was, indeed, obsessed with invalidation, and who became quite skeptical about confirmation playing any role whatsoever in induction. David Miller continues to defend a version of Popperian "falsificationism".

4 Nicod on Pure Induction

4.1 Keynes on Pure Induction

Before getting into Nicod's worries about Keynes's remarks on Pure Induction, it is helpful to summarize what Keynes himself says. Let $\mathcal{E}_i$ be the i^{th} "positive instance" of the form $Fa_i \,\&\, Ga_i$ (for some object a_i) observed by an agent ϕ. Keynes appeals to the following assumptions in his proof of the "limit theorem":

(1) For all i, $\Pr_\phi(\mathcal{E}_i \mid \mathcal{H} \,\&\, K_\top) = 1$.

(2) For all i and j, $\Pr_\phi(\mathcal{E}_i \mid K_\top) = \Pr_\phi(\mathcal{E}_j \mid K_\top)$.

(3) For all $i > 1$, $\Pr_\phi(\mathcal{E}_i \mid \&_{j=1}^{i-1} \mathcal{E}_j \,\&\, K_\top) < 1$.

(4) $\Pr_\phi(\mathcal{H} \mid K_\top) \in (0, 1)$.

(5) $\lim_{n\to\infty} \Pr_\phi(\&_{i=1}^{n} \mathcal{E}_i \mid {\sim}\mathcal{H} \,\&\, K_\top) = 0$.

Specifically, Keynes claims to have proven the following limit claim from these five assumptions:

$$\lim_{n\to\infty} \Pr_\phi(\mathcal{H} \mid \&_{i=1}^{n} \mathcal{E}_i \,\&\, K_\top) = 1$$

Nicod's worries are mainly about the status and plausibility of (4) and (5), and whether the limit theorem really does follow from (1)–(5). Before getting to Nicod's worries, I want to talk a bit more about (1). Why does Keynes assume this? I've already asked this question twice before. Keynes says a few things about (1). Here's the odd footnote that comes right after the initial statement of (1) as an assumption (our notation):

> In the most general sense we can regard any proposition as the generalisation of all the propositions which follow from it. For if p is any proposition, and we put $Fx \overset{\text{def}}{=} \ulcorner x$ can be inferred from $p\urcorner$ and $Gx \overset{\text{def}}{=} \ulcorner x \urcorner$ [or, perhaps, $Gx \overset{\text{def}}{=} \ulcorner x$ is true$\urcorner$?], then $(\forall x)(Fx \supset Gx)$ [every statement that can be inferred from p is true?] is logically equivalent to p. Since Pure Induction consists in finding as many instances of a generalisation as posslble, it is, in the widest sense, the process of strengthening the probability of any proposition by adducing numerous instances of known truths which follow from it.

This is the best translation I could give of this footnote into modern logical-ese, and it is *bizarre* to say the least. Here, Keynes seems to be using monadic predicate logic meta-meta-linguistically (the variables seem to range over some class of object-language statements, and F seems to involve the metalinguistic relation of entailment). Even if we could make sense of this spaghetti bowl, it does not address the question as to *why* anyone should believe (1). Later, Keynes summarizes the limit-theorem he has in mind, as follows:

> We have shown that if each of the instances necessarily follows from the generalisation, then each additional instance increases the probability of the generalisation, so long as the new instance could not have been predicted with certainty from a knowledge of the former instances.

This quote also clearly indicates that Keynes is falsely assuming that the $\mathcal{E}_i$'s *necessarily follow from* $\mathcal{H}$. They don't. As I have discussed above, what follows are the *conditional* (*viz.*, *proper*) instances $Fa_i \supset Ga_i$. So, it seems to me that this entire argument rests on a mistake [unless $K_\top$ is assumed to contain the antecedents of the $\mathcal{E}_\supset$ conditionals, which would not be consistent with what Keynes (or Nicod) are after].

4.2 Nicod on Keynes on Pure Induction

Nicod's worries about Keynes's argument have mainly to do with presuppositions of "*determinism*" that Nicod thinks are required for Keynes's argument to go through. Keynes seems to deny such presuppositions:

> The common notion, that each successive verification of a doubtful principle strengthens it, is formally proved without any appeal to conceptions of law or of causality.

However, as Braithwaite points out, Keynes does assume determinism "in the establishment of the *a priori* probability of any law by the Principle of Limited Variety (though Mr. Keynes does not make this as clear in his exposition as might be desired)". But, that is a much different kind of determinism (at the level of groups of properties, not instances), and this isn't something that Nicod seems to be terribly worried about.

Nicod does seem to be worried that, if the probability of the law is to be driven to *certainty* by verification, then its prior probability must also be 1, *if* a statement about its prior probability [(4)] is also to be a premise in the argument. This worry (again) seems to stem from Nicod's (mis)application of the (PL)-theorems above. One thing Nicod is right about here is that we are now in the realm of deductive arguments with probabilistic premises and conclusions. But, Nicod is confused if he thinks that no argument with the conclusion Keynes wants can be made cogently without already assuming that $\mathcal{H}$ has prior probability 1. For one thing, the conclusion of the argument is not a simple probability claim. It is a limit claim. Moreover, even if it were a probability claim, it would be a claim that conditionalizes on an infinite conjunction, which takes us well beyond the resources of simple probability logic. In any event, I will not dwell on this worry further here.

Some form "determinism" seems to be at work here when Keynes supposes that $\Pr_{\phi}(\mathcal{E}_i \mid K_{\top}) = 1$ *if* $K_{\top}$ includes a previous instance $\mathcal{E}_j$ that is "identical with" $\mathcal{E}_i$. But, given Keynes's other characterizations of this assumption (and looking at his proof), all he really *needs* to assume here is that future (or present) instances cannot (reasonably) be predicted with certainty on the basis of past instances. And, this seems fine to me, since we never actually get (at any finite stage) to be certain that $\mathcal{H}$ is true on the basis of our observations up to the present time. And, even if we were certain that $\mathcal{H}$ was true at some finite stage, $\mathcal{E}_i$ *still* wouldn't follow from it. [But, since we're presupposing (1) here, I guess I can see why this worry arises.]

Nicod next discusses preconditions (4) and (5). Keynes claims that (4) and (5) follow from his assumption of Absolute Finiteness (AF): that all the "apparent qualities" in the universe are determined by a finite number of "generator properties". Nicod agrees that (4) follows from (AF) — properly understood — but he suggests in a footnote (278) the following amendment to (AF), in order to ensure that precondition (4) does follow:

> To speak in all rigour, it would be necessary to assume not only that the number of groups of connected characters is some finite number x, but also that there is a finite probability that x is less than a given number — than a billion, for instance. For if all the finite numbers have the same chances of being x, it is infinitely more probable that x *is higher or lower than any assigned number*, and hence not finite.

Braithwaite voices a similar complaint:

> Mr. Keynes has been guilty of a fallacy similar to that of the confusion of Convergence with Uniform Convergence in the theory of Infinite Series [or of $(\forall x)(\exists y)Rxy$ with $(\exists y)(\forall x)Rxy$]. What is required to give an initial probability to any generalisation is not that the number of the groups is finite, but that it is less than some number given in advance. And when we remember that Mr. Keynes's Principle has to have some *a priori* probability, we see that this apparently slight change makes an enormous difference.

The problem which emerges is that, while it might be reasonable (*a priori*) to assume that the number of groups is finite, it seems unreasonable (*a priori*) to assume it is less than some number given in advance. The reason why (4) is said to follow from this stronger assumption is, according to Nicod and Braithwaite, that (given said assumption) "there is a finite chance that both [F and G] are determined by the same set of generator properties". If there is a finite probability that F and G are determined by the same set of generator properties, then, one supposes, this implies that there is a finite probability that F and G are coextensional, and hence that $(\forall x)(Fx \equiv Gx)$. And, since $(\forall x)(Fx \equiv Gx) \vDash \mathcal{H}$, this ensures a finite probability of $\mathcal{H}$ as well. I'm still fuzzy on the precise connection between the metaphysics and epistemology here. I suppose they are assuming that a *Principle of Indifference* applies to the underlying metaphysical states. But, why accept this? Do we have epistemic access to these "underlying states"? Perhaps this isn't as much a problem for Nicod, who seems to be working with *logical* probabilities? We'll return to these issues again, below.

As for (5), Nicod has a lot of interesting remarks about it. First, Nicod confirms my suspicion that some *Principle of Indifference* is at work here as well. Nicod gives an argument for (5), which runs (roughly) as:

(a) If the law $\mathcal{H}$ is false, it must be false in at least one instance.

(b) The number of "non-identical" instances is some pre-specified, finite number N.

(c) There is no reason to suppose that anyone of these N will be the next instance rather than any other.

(d) So, presumably by some sort of *Principle of Indifference*, the probability that the next instance will invalidate the law is always greater than or equal to $\frac{1}{N}$, for some pre-specified finite N.

(e) Therefore, the probability that the next instance will confirm (indeed, be entailed by!) the law is always less than or equal to $1 - \frac{1}{N} < 1$, which is sufficient to ensure that $\lim_{n\to\infty} \Pr_{\phi}(\&_{i=1}^{n} \mathcal{E}_i \mid {\sim}\mathcal{H} \,\&\, K_{\top}) = 0$.

Premise (d)/(e) is the one Nicod and Braithwaite call into question. According to Braithwaite, (d) implies that

> the number of numerically distinct instances arising out of each group of generator properties is equal, and that the fact that we continue not meeting with one of these instances does not diminish the probability that shall meet it in the next example.

Both Nicod and Braithwaite reject this assumption. But, Braithwaite suggests a weaker alternative to (e):

(e′) The probability that the next (n^{th}) instance will confirm the law (conditional on the falsity of the law and its satisfaction in the previous $n-1$ instances) does not tend to one more rapidly than $\frac{1}{n}$.

Braithwaite thinks (e′) suffices for the desired limit theorem, and that it is "not unplausible", or, at least, that it is more plausible than (e). As we will see, below, Hosiasson-Lindenbaum proves that something even stronger than (e′) follows from Keynes's (AF) assumption. Indeed, she precisely clarifies what Nicod is right (and wrong) about here. This turns out to involve some rather sophisticated logic and probability theory.

We should also wonder about (b). It seems to imply that there are only N *particulars* in the universe. [Or, at least, that there are N *discernible* particulars.] This is a strong assumption to make *a priori* (Carnap makes a formally similar assumption, for similar technical probabilistic reasons, as we'll see later). More to the point, if the requisite probabilistic constraints here are *epistemic*, then I don't see why they need to rely on any metaphysical assumptions. Presumably, all that is required is that it would be *reasonable* for an epistemic agent's credences to satisfy the constraints. And, that could be true even if these (or other) metaphysical assumptions about the "underlying realm" of properties *turned out to be false*. For instance, one might argue for assumptions like those required by the theorem on the grounds that it is reasonable to assign non-extreme probabilities to claims that are not self-evident *a priori* truths (or falsehoods). On the other hand, if the constraints here are meant to be "logical" in nature, then I'm not sure *what* considerations are relevant for constraining such "logical" probabilities (we'll talk more about this when we get to Carnap).

4.3 Hosiasson-Lindenbaum on the Keynes-Nicod Dispute Concerning Pure Induction

In her classic paper "On Confirmation", Janina Hosiasson-Lindenbaum (JHL) brings refreshing rigor to the otherwise not fully rigorous dispute between Keynes and Nicod. JHL gives a precise formal axiomatization of conditional probability (which she calls "confirmation" — note this usage differs from Nicod's probabilistic *relevance* usage — we'll see *that* ambiguity again in Carnap), and she uses it to prove the following theorem:

Theorem (JHL). Assume $\sim\mathcal{H}$ is equivalent to a disjunction $\bigvee_{i=1}^{k}\mathcal{H}_i$ of k mutually exclusive alternatives to $\mathcal{H}$ (*i.e.*, where each $\mathcal{H}_i$ is incompatible with $\mathcal{H}$). ***If*** the following four conditions are satisfied:

(i) For every i ($1 \leq i \leq k$), $\Pr(\mathcal{H}_i \mid \sim\mathcal{H} \,\&\, K_\top) > 0$.

- $*$ None of the $\mathcal{H}_i$'s is "excluded *a priori*."

(ii) For every i ($1 \leq i \leq k$) and every s ($1 \leq s \leq n$),

$$\Pr(\mathcal{E}_s \mid \mathcal{H}_i \,\&\, K_\top) = \Pr(\mathcal{E}_s \mid \mathcal{H}_i \,\&\, \&_{j=1}^{s-1}\mathcal{E}_j \,\&\, K_\top) = \text{constant (depending } \textit{only} \text{ on } i) > 0.$$

- $*$ The $\mathcal{E}_s$'s are "facts which possess constant finite probabilities with respect to each $\mathcal{H}_i$, and which are independent of the previously observed facts $\&_{j=1}^{s-1}\mathcal{E}_j$, *given* each of the $\mathcal{H}_i$'s."

(iii) For every s ($1 \leq s \leq n$), there exists an $i \neq j$ such that: $\Pr(\mathcal{E}_s \mid \mathcal{H}_i \,\&\, K_\top) \neq \Pr(\mathcal{E}_s \mid \mathcal{H}_j \,\&\, K_\top)$. [JHL claims that this follows from (AF); but Keynes, Nicod, and Braithwaite all seem to have missed it.]

(iv) For each s ($1 \leq s \leq n$), $\Pr(\mathcal{E}_s \mid \mathcal{H} \,\&\, K_\top) = 1$. [The usual assumption about "positive instances".]

then, the following three conclusions follow:

(v) $\Pr(\mathcal{E}_n \mid \&_{i=1}^{n-1}\mathcal{E}_i \,\&\, \sim\mathcal{H} \,\&\, K_\top) > \Pr(\mathcal{E}_{n-1} \mid \&_{i=1}^{n-2}\mathcal{E}_i \,\&\, \sim\mathcal{H} \,\&\, K_\top)$.

- $*$ "given $\sim\mathcal{H}$, the probability that the next observed instance will be a positive instance of $\mathcal{H}$ increases with the number of previously observed positive instances of $\mathcal{H}$."

(vi) $\lim_{n\to\infty}\Pr(\mathcal{E}_n \mid \&_{i=1}^{n-1}\mathcal{E}_i \,\&\, \sim\mathcal{H} \,\&\, K_\top) = 1$ *only if* $\Pr(\mathcal{E}_s \mid \mathcal{H}_t \,\&\, K_\top) = 1$ for some s and t.

(vii) $\lim_{n\to\infty}\Pr_{\phi}(\mathcal{H} \mid \&_{i=1}^{n}\mathcal{E}_i \,\&\, K_\top) = 1$.

So, Keynes was right about his limit theorem (vii). On the other hand, Nicod (and Braithwaite) was right that (v) also follows from Keynes's presuppositions. However, because Keynes's (AF) also entails that, for each s and t, $\Pr(\mathcal{E}_s \mid \mathcal{H}_t \,\&\, K_\top) < 1$, (vi) implies that $\lim_{n\to\infty}\Pr(\mathcal{E}_n \mid \&_{i=1}^{n-1}\mathcal{E}_i \,\&\, \sim\mathcal{H} \,\&\, K_\top) \neq 1$, *contrary* to what Nicod suggests in his essay. Thus, Nicod was incorrect about (vi). In these ways, JHL's paper definitively resolves these nuances of the Keynes-Nicod dispute concerning Pure Induction, using only assumptions that all parties seem to agree upon. JHL's paper is also important because it contains a crucial result concerning Hempel's paradox of confirmation. We'll get to that in our next set of readings on Hempel (in two weeks).