A Correction Rule for Inductive Methods

Ruurik Holm, Helsinki

I will discuss the problem of choosing the correct inductive method from Carnap's (1952) continuum. My proposal is to use a correction rule to adjust the method according to obtained evidence. I will discuss a minimum requirement such a rule has to satisfy, especially from a consturctive point of view. The question of refuting inductive scepticism by means of a correction rule is assessed.

Carnap (e.g. 1950, 564; 1952, 38) regards the extreme method $\lambda=\infty$ as seemingly inappropriate for sound scientific reasoning on the grounds that it gives no consideration to experience in making expectations or estimations. However, the argument is based on the presupposition that inductive reasoning is sound. A strict anti-inductivist holds that $\lambda=\infty$ is the right choice precisely because it gives no regard to experience. Therefore, the question remains how to reject $\lambda=\infty$ in Carnap's continuum without making the inductivist presupposition.

My aim in this paper is to argue for a correction rule for adjusting inductive methods, including $\lambda=\infty$, to overcome this difficulty which can be considered as one formulation of the problem of induction. First, however, I will discuss what kind of minimum requirement any correction rule must have, especially from the point of view of constructive semantics. Then I will proceed to present a concrete example of such a rule. In the final section, I will discuss if using the rule presented in section 2 satisfying the requirement presented in section 1 is more rational than adhering to $\lambda=\infty$.

1. The Minimum Requirement

My first proposal for a minimum requirement for any acceptable correction rule is that the method it yields should be the same in the limit as the method corresponding to the limit of the degree of order of the sequence, always when one of the limits exist.

However, in constructive mathematics, limits can only exist for sequences that have been constructively given. Let us consider the state description obtained by tossing an indestructible coin with an unknown bias. It is clear that a particular infinite sequence of tosses cannot be given as a computable or constructive function – otherwise the result of each toss would be fixed in advance. A classical formalization is available: one can define the *i*'th toss to be the value $t(i)$ of some function $t: N \rightarrow \{Heads, Tails\}$, even without being able to give a computation rule for t.

Martin-Löf's (1990) nonstandard type theory provides a constructive semantics for sequences that are not given by a computable rule. However, there remains another problem with respect to the constructive interpretation of the minimum requirement as phrased above.

A sequence can constructively converge towards a limit only if one can compute the terms of the sequence up to infinity since otherwise it is impossible to know the value of the limit and hence also impossible to consider that the sequence has a limit in the constructive sense. On the other hand, the motivation behind the correction rule is that it would eventually guide us towards the optimum method, even if it is not knowable what the optimum method is. The problem is how to formulate this idea in constructive semantics.

In nonstandard type theory, one can give a constructive formulation of the minimum requirement. Denote the sequence of x tosses by $w(x)$, the degree of order of the sequence $w(y)$ by $d_o(w(y))$ and the correction rule applied to $w(y)$ by $Corr(\lambda_a, w(y))$, where λ_a is the initial method. The function δ: /→/ gives the corresponding degree of order for each real-valued method (hence, when $\lambda\rightarrow\infty$, $\delta(\lambda)\rightarrow 0.5$ in the case of the coin tossing example since then the maximum degree of disorder is 0.5). By Δ_ω we denote any infinite subsequence (in the nonstandard sense) of the set natural numbers $\{0,1,2,\ldots\}$ that consists of consecutive numbers, and a is an arbitrary real number:

$$(\forall y \in \Delta_\omega)(d_o(w(y)) \approx a) \Leftrightarrow (\forall y \in \Delta_\omega)(\delta(Corr(\lambda_a, w(y))) \approx a).$$

The above formula says that for any (in the nonstandard sense) infinite interval, the real degree of order remains within the same boundaries as the degree of order obtained by *Corr*. The standard interpretation of the formula is that there is a minimum length such that for all intervals with at least that length, if the degree of order is within some distance ε from a in that interval, also the degree of order obtained by using the correction rule is within ε from a, and vice versa. In other words, it follows from the minimum requirement that for all sufficiently long intervals, the results of the correction rule have the same boundaries as the degree of order.

The motivation for using nonstandard type theory here arises from the infinitistic property of the nonstandard number ω of being provably bigger than any standard natural number. The meaning of expressions dealing with such an infinitely long interval is constructively explained by reference to multiple finite intervals so that the minimum requirement is not in contradiction with any proposition obtained by negating it and replacing the infinite interval Δ_ω with an arbitrary finite interval.

Since it is now clear that the minimum requirement makes sense also constructively, I can proceed to present a concrete example of a correction rule which fulfils the requirement.

2. Formulating the Rule

It is obvious that the correction rule must have its first effects on the initial method when some finite amount of data is received. One cannot change the method unless one begins at some finite point.

Let the parameter c denote a positive real number which expresses how cautiously the method is changed according to the observed degree of order; a big value for c means moderate changes. The variable w_x denotes a sequence of x first tosses.

The following formulas give an inductive definition of a simple but still adequate correction rule G, where the induction variable denotes the number of performed tosses:

$$\begin{cases} G(\lambda_a, c, 0, w_0) = \lambda_a, \\ G(\lambda_a, c, x+1, w_{x+1}) = \delta^{-1}(\delta(G(\lambda_a, c, x, w_x)) + \frac{1}{c}(d_0(w_x) - \delta(G(\lambda_a, c, x, w_x)))). \end{cases}$$

The definition can also be applied when $\lambda_a \rightarrow \infty$.

The idea of the rule is to take the difference of the observed degree of order and the one implied by the currently used method as the basis of correction. This number is then multiplied by the caution factor 1/c to acquire the magnitude of the adjustment. The final result, which is a number denoting an inductive method, is obtained by application of the inverse function of δ.

3. The Problem of Induction

The problem of induction formulated in the context of inductive probability concerns the justification of increasing or decreasing the probability value of a hypothesis on the basis of evidence.

Strict anti-inductivism represented by the choice of $\lambda=\infty$ means that one will never adopt the optimum method when the order in the state description under examination is higher than the theoretical minimum. The question is if this kind of complete rejection of inductive inference is justified, i.e., whether it is justified to hold that past observations should not influence the probabilities of future observations even though it follows that one makes less than optimal estimates about the relative frequency of properties except in the case of minimum order, the error being bigger the higher the degree of order of the state description in question is.

Hans Reichenbach's (1949) vindication of induction was based on the idea that one should adopt a method of which it is known that it will lead to successful approximations about the relative frequencies of properties in an infinite domain, provided that such success is possible, i.e., that the limits of relative frequencies exist when the domain size tends to infinity.

A related idea can be applied here. The correction rule will eventually lead to closer and closer approximations of the optimum method when the sample size increases, but it does not exclude the possibility of *a priori* considerations about the optimum method. The problem with both Reichenbach's straight rule and the correction rule is that no particular body of obtained data can really justify inductive predictions concerning future observations. In the correction rule approach, this means that no particular sample can justify the shift from the anti-inductivism toward moderate inductivism.

However, even if one holds that inductive reasoning is not justified, one must admit that it is possible that the method $\lambda=\infty$ is not optimal any longer when more information is obtained. There is always some (logical) probability that future will manifest order, even according to the anti-inductivist method, because it assigns positive probabilities to ordered state descriptions. If one does not change the method on the basis of evidence at all when one has set $\lambda=\infty$ in the beginning, one thus runs the risk of making an infinite number of inaccurate predictions, while the employment of an adequate correction rule would eliminate this risk. Hence, there seems to be a pragmatic argument suggesting the utilization of a correction rule.

On the other hand, if the unknown optimum method corresponding to the unknown state description is $\lambda=\infty$, every correction rule satisfying the minimum requirement will necessarily agree with any approximation of the real degree of order. Hence, there is no risk of making an infinite number of inaccurate predictions. The risk one takes is only that the first observed terms of the sequence are misleading with respect to relative frequencies in the full sequence, guiding the adjustment of the method away from the optimum value. There is thus a possibility of an arbitrarily high but still finite number of inaccurate predictions.

To add up, the correction rule never leads to an infinite number of inaccurate predictions, whereas for $\lambda=\infty$ this possibility exists. I hold this to be an argument for not adhering to $\lambda=\infty$, which does not contain the inductivist presupposition. What remains to be discussed in this context is the role of probabilities, namely those assigned by $\lambda=\infty$ to state descriptions where $\lambda=\infty$ is not the optimum method as contrasted to those assigned to state descriptions where the method is indeed optimal.

Literature

Carnap, R. 1950 *Logical Foundations of Probability*, Chicago: The University of Chicago Press.

Carnap, R. 1952 *The Continuum of Inductive Methods*, Chicago: The University of Chicago Press.

Martin-Löf, P. 1990 "Mathematics of Infinity", in P. Martin-Löf and G. Mints (eds.), *COLOG-88, Lecture Notes in Computer Science* 417, Berlin: Springer,146-197.

Reichenbach, H. 1949 *The Theory of Probability*, Berkeley: University of California Press.