Comment: Causality, Complexity and Determinism

Patrick Suppes

As always, Jack Good provides enough topics for a dozen commentators. I have restricted what I have to say to three issues that are of central interest to philosophers and statisticians. These are: probabilistic causality, randomness and complexity, and determinism. The topics I have selected for comment do not reflect how much I agree with Good's general philosophical position, especially his emphasis on a Bayes/non-Bayes compromise and on the essential role of intuitive judgment in statistical practice.

PROBABILISTIC CAUSALITY

Good and I are not far apart on the theory of probabilistic causality. Certainly his early work (1961/62) was a direct influence on my own major effort in this area (1970). As he rightly remarks, mine is primarily a qualitative theory, and what he has offered is a more quantitative one. It is exactly this issue that I want to focus on here. For many kinds of foundational matters we certainly want a general quantitative theory. A good example would be the fundamental theory of measurement of extensive quantities, as applied, for example, to length or mass. We expect this fundamental theory to work without modification across a wide range of applications and in the context of many different theories. Certainly across the range of classical physical theories we expect the measurement of mass or length to remain unchanged. Given these expectations, supported by a wealth of experience, it is certainly appropriate to develop in complete detail the general theory of extensive measurement, with standard quantitative results.

It is much less clear to me that this is the case with the more general notion of cause. Good's attempt to develop a quantitative theory of causality is certainly the most interesting and sustained effort that has been made, as far as I know. On the other hand, it also seems to me that it has had surprisingly few detailed applications. It seems to me that there is a clear reason for this absence of a wide range of applications, in spite of its considerable intrinsic interest. The reason is that the quantitative theory of causality is not fixed or invariant across contexts, but varies considerably from one application to another. In contrast, the ordinal or qualitative theory, reflecting the broad intuitions of common sense, should remain invariant across quantitative variation.

Let me illustrate the contrast I have in mind with a simple comparison of two learning models or, more generally, models with some sort of feedback. In the first case, we have a simple linear response model with discrete trials. On each trial the probability of response R_i—for simplicity I shall restrict myself to two responses, $i = 1, 2$—, is a linear function of the probability of such a response on the previous trial and the reinforcement or feedback on the previous trial. Thinking in terms of the feedback causally influencing response, let C_i, $i = 1, 2$, be the reinforcement, and let x_n be the sequence of responses and feedbacks through trial n. We may write this simple linear model with a single learning parameter θ as follows:

$$P(R_{i,n+1} \mid C_{j,n}, R_{k,n}, x_{n-1}) = (1-\theta)P(R_{i,n} \mid x_{n-1}) + \theta\delta_{ij},$$

where δ_{ij} is the Dirac delta function.

Consider in contrast Luce's beta model (1959) where, for the purposes at hand, there are two beta parameters with β_i being the applicable parameter when the feedback at the end of the trial is C_i. In terms of this notation, then, we may write the beta model as

$$P(R_{1,n+1} \mid C_{i,n}R_{j,n}x_{n-1}) = \frac{P(R_{1,n} \mid x_{n-1})}{(1-\beta_i)P(R_{1,n} \mid x_{n-1}) + \beta_i}.$$

The single important difference between the linear learning model and Luce's beta model that I want to concentrate on here is that in the case of the beta model the operators commute, that is, the probability of a response on trial $n + 1$ is just a function of the initial probability together with the number of causes of each type that have occurred in the first n trials. In contrast, the linear model is completely dependent on the order in which the causes occur, that is, the causal operators, to put it in such language, do not commute.

What happens in these two cases is that we get two different quantitative causal theories for exactly the same phenomena. It seems to me that there is no reason whatsoever to think that a general quantitative theory of causality should discriminate between two such models. It is very much a matter of detailed empirical theory and experiment in particular

Patrick Suppes is Lucie Stern Professor of Philosophy, Stanford University, Stanford, California 94305.