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.

domains. It is for this reason that I am skeptical that we will see a successful general quantitative theory of probabilistic causality, in spite of Jack Good's interesting and original efforts.

As Good points out in the present article and elsewhere, there may be a use for a quantitative theory of causality in situations in which we have no other theory. This is in the same spirit with which one applies linear regression models in the absence of any fundamental theory of the phenomena in question. I have no quarrel with such an approach, but I retain my skepticism that a quantitative theory of causality can compare in any sense with the full criteria we expect a quantitative theory to satisfy, as in the case of the measurement of mass, length, temperature, etc., or as in the grander examples of classical mechanics and other quantitative physical theories.

RANDOMNESS AND COMPLEXITY

In the article published here, Good makes some remarks about randomness, but he has a lot more to say about it in his volume *Good Thinking* (1983f), to which he refers repeatedly in this article. On page 88 of *Good Thinking*, Good raises the problem of accepting the enciphering of a secret message when the randomly enciphered message "by extraordinarily bad luck" came out exactly the same as the original one. He poses the question "would the enciphering agent be prepared to use it." There are two questions that I want to address to him in the context of his thoughts about randomness as expressed especially in Chapter 8 of *Good Thinking*.

The first is the importance of making a distinction between random processes or procedures on the one hand and random results on the other. It seems to me that in much of the discussion of randomness there is not always a clear distinction between these two ideas. By random *procedure* I mean, of course, a procedure that we characterize as having appropriate randomness characteristics. In contrast, by random *result* I mean the actual sequence obtained. Here we apply the type of test advocated by Kolmogorov and others to decide whether the actual sequence is random, with the basic criterion being that there is no essentially shorter way of describing the sequence than by reproducing the sequence itself. The Kolmogorov criterion does not consider at all the method by which the sequence is generated—it concentrates entirely on the features of the sequence itself. It seems to me that this distinction is not often enough brought to the fore in statistical discussions of the place of randomness in experimentation.

My second observation is related to this and is a second query to Jack Good. If we replace random procedures by random results, why not go a step further and replace random results by *complex* results. Then we could have a table of complex sequences with some threshold of complexity being satisfied by the table. We simply draw any sequence of a given length from the table, or, as a still different approach, we could use some standard method of generating random numbers, but we throw out any constructed sequence whose complexity is below an agreed upon complexity threshold. It is evident enough that this idea of a complexity threshold is robust. The exact value chosen is not important. There will remain for any experiments of any size a more than adequate number of complex sequences.

The tension between randomness and complexity is apparent. A *sampling procedure* is random. Often, any sequence, simple or complex, is as likely as any other. But the *result* of using the random procedure is a given sequence whose complexity can be measured. My suggestion is that we move from procedures to results and from randomness to complexity as the essential measure. I think that what I am suggesting here is sympathetic to what Good has had to say about randomness in different places, especially in the chapter quoted, but I am not entirely certain.

DETERMINISM

Perhaps my most direct disagreement with Good is in his drawing a contrast between determinism and probability. For example, he says early on in the present article "Physical probability too is a useful fiction even if the world is deterministic $\cdots$." What I want to say is in disagreement with this, but even here I suspect that the spirit of what I have to say is not something that Jack will find particularly unsettling. It is a point though that I think philosophically is worth some clarification.

That we are already not far apart is evident from the following quotation (1983f, pp. 91–92).

> In the discussion some one raised the question of whether the physical universe might be "infinitely random" in the sense that an infinite amount of information would be required to make an accurate prediction of a closed system. The following observation is relevant. Consider a model of classical statistical mechanics in which the various particles are assumed to be perfectly elastic spheres. Then, to predict a time T ahead, to a given tolerance, I think we would need to know the position and velocities of the particles to a number of decimal places proportional to T. There must be some degree of accuracy that is physically impossible to measure, and in this sense classical statistical mechanics provides indeterminism arising out of determinism.

The particular example I want to cite, an important one in the theory of classical mechanics, goes beyond Good's example in the following sense. Random sequences, under any definition that is used, are generated by simple models of classical mechanics. The generation of these random sequences does not in any way depend upon ignorance of initial conditions and measurement tolerances, but is simply a fact independent of measurement about the mechanical systems themselves. The existence of such systems shows, it seems to me very clearly, that there is no opposition between completely deterministic systems and random systems.

The example I use is a special case of the three-body problem, certainly the most extensively studied problem in the history of mechanics. Our special case is this. There are two particles of equal mass moving according to Newton's inverse-square law of gravitation in elliptic orbits relative to their common center of mass, which is at rest. The third particle has a nearly negligible mass, so it does not affect the motion of the other two particles, but they affect its motion. This third particle is moving along a line perpendicular to the plane of motion of the first two particles and intersecting the plane at the center of their mass. From symmetry considerations, we can see that the third particle will not move off this line. The restricted problem is to describe the motion of the third particle. The analysis of this easily described situation is quite complicated and technical, but some of the results are simple to state in informal terms and directly relevant to my focus on determinism and randomness. Near the escape velocity for the third particle—the velocity at which it leaves and does not periodically return, the periodic motion is very irregular. In particular, the following remarkable theorem can be proved. Let t_1, t_2, $\cdots$ be the time at which the particle intersects the plane of motion of the other two particles. Let s_k be the largest integer equal to or less than the difference between t_{k+1} and t_k times a given constant. Variation in the s_k's obviously measures the irregularity in the periodic motion. The theorem in the version given by Moser (1973), due to the Russian mathematicians Sitnikov (1960) and Alekseev (1968a, b; 1969a, b), is this:

THEOREM. *Given that the eccentricity of the elliptic orbit is positive but not too large, there exists an integer, say* α, *such that any infinite sequence of terms* s_k *with* $s_k \geq \alpha$ *corresponds to a solution of the deterministic differential equation governing the motion of the third particle.*

The correspondence between a solution of the differential equation and a sequence of integers is the source of the term *symbolic dynamics*. The idea of such a correspondence originated with G. D. Birkhoff in the 1930s.

A corollary about random sequences immediately follows. Let s be any random sequence of heads and tails—for this purpose we can use any of the several variant definitions—Church, Kolmogorov, Martin-Löf, etc. We pick two integers greater than α to represent the random sequence—the lesser of the two representing heads, say, and the other tails. We then have:

COROLLARY. *Any random sequence of heads and tails corresponds to a solution of the deterministic differential equation governing the motion of the third particle.*

In other words, for each random sequence there exists a set of initial conditions that determines the corresponding solution. Notice that in essential ways the motion of the particle is completely unpredictable even though deterministic. This is a consequence at once of the associated sequence being random. This example demonstrates the startling fact that the same phenomena can be both deterministic and random. The underlying explanation is the extraordinary instability of the deterministic phenomena.

From all the recent work on chaotic systems in classical mechanics, it is clear that there is nothing isolated or unusual about this example. There are certainly other deterministic physical systems of a simple nature that will generate similar random phenomena. The classical philosophical dichotomy between determinism and randomness is a mistaken one.

ADDITIONAL REFERENCES

ALEKSEEV, V. M. (1968a). Quasirandom dynamical systems. I. Quasirandom diffeomorphisms. *Math. USSR-Sb.* **5** 73–128.

ALEKSEEV, V. M. (1968b). Quasirandom dynamical systems. II. One-dimensional nonlinear oscillations in a field with periodic perturbation. *Math. USSR-Sb.* **6** 505–560.

ALEKSEEV, V. M. (1969a). Quasirandom dynamical systems. III. Quasirandom oscillations of one-dimensional oscillators. *Math. USSR-Sb.* **7** 1–43.

ALEKSEEV, V. M. (1969b). Quasirandom dynamical systems. *Mat. Zametki* **6** 489–498.

LUCE, R. D. (1959). *Individual Choice Behavior: A Theoretical Analysis.* Wiley, New York.

MOSER, J. (1973). *Stable and Random Motions in Dynamical Systems with Special Emphasis on Celestial Mechanics.* Herman Weyl Lectures, Institute for Advanced Study, Princeton Univ. Press, Princeton, N. J.

SITNIKOV, K. A. (1960). Existence of oscillating motions for the three-body problem. *Dokl. Akad. Nauk SSSR* **133** 303–306.