Discussion of Causal Diagrams for Empirical Research by J. Pearl

Stephen E. Fienberg; Clark Glymour; Peter Spirtes

Biometrika, Vol. 82, No. 4 (Dec., 1995), 690-692.

Stable URL:
http://links.jstor.org/sici?sici=0006-3444%28199512%2982%3A4%3C690%3ADOCDFE%3E2.0.CO%3B2-0

Biometrika is currently published by Biometrika Trust.

This raises the question as to how we can use scientific understanding and empirical data to construct the requisite causal model. By saying little about this specification problem, Pearl is in danger of being misunderstood to say that it is not important. To build either a distributional or a counterfactual causal model, we need to assess evidence on how interventions affect the system, and what remains unchanged. This will typically require a major scientific undertaking. Given this structure, distributional aspects can, in principle, be estimated from suitable empirical data, if only these are available, and we can then apply the manipulations described by Pearl to address problems of the 'effects of causes'. But much more would be needed to address 'causes of effects', since counterfactual probabilities are, almost by definition, inaccessible to direct empirical study. Empirical data can be used to place bounds on these (Balke & Pearl, 1994), but these will usually only be useful when they essentially determine the functions in (3). And, for this, it will be necessary to conduct studies in which the variables ε_i are explicitly identified and observed. Thus the whole mechanism needs to be broken down into essentially deterministic sub-mechanisms, with randomness arising solely from incomplete observation. In most branches of science such a goal is quite unattainable.

I emphasise the distinction drawn above, between inference about 'effects of causes' and 'causes of effects', because it might be tempting to try to extend Pearl's analysis, particularly in its formulation (3), to the latter problem. For both problems serious difficulties attend the initial model specification, but these are many orders of magnitude greater for 'causes of effects', and the inferences drawn will be very highly sensitive to the specification.

On a different point, I am intrigued by possible connexions between Pearl's clear distinction between conditioning and intervening, and the prequential framework of Dawid (1984, 1991), especially as elaborated by Vovk (1993). Suppose A plays a series of games, involving coins, dice, roulette wheels, etc. At any point, the game chosen may depend on the observed history. We could model this dependence probabilistically, or leave it unspecified. Now suppose we are informed of the sequence of games actually played, and want to say something about their outcomes. In a fully probabilised model, we could condition on the games played, but this would involve unpleasant analysis, and be sensitive to assumptions. Alternatively, and seemingly very reasonably, we can use the 'prequential model', which treats the games as having been fixed in advance. This is obtained from a fully specified model, with its natural temporally defined causal model, by 'setting' the games, rather than conditioning on them.

[*Received May* 1995]

Discussion of 'Causal diagrams for empirical research' by J. Pearl

BY STEPHEN E. FIENBERG

Department of Statistics, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213-3890, U.S.A.

CLARK GLYMOUR AND PETER SPIRTES

Department of Philosophy, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213-3890, U.S.A.

In recent years we have investigated the use of directed graphical models (Spirtes, 1995; Spirtes, Glymour & Scheines, 1993) in order to analyse predictions about interventions that follow from causal hypotheses. We therefore welcome Pearl's development and exposition. Our goal here is to indicate some other virtues of the directed graph approach, and compare it to alternative formalisations.

Directed graph models have a dual role, explicitly representing substantive hypotheses about influence and implicitly representing hypotheses about conditional independence. We can connect the two dimensions, one causal and the other stochastic, by explicit mathematical axioms. For example, the causal Markov axiom requires that, in the graph, each variable be independent of its nondescendants conditional on its set of parents. The formalism allows one to hold causal hypotheses fixed while varying the axiomatic connexions to probabilistic constraints. In this way, one can prove the correctness of computable conditions for prediction, for the statistical equivalence of models, and for the possibility or impossibility of asymptotically correct model search, all under alternative axioms and under a variety of circumstances relevant to causal inference, including the presence of latent variables, sample selection bias, mixtures of causal structures, feedback, etc. Thus it is possible to derive Pearl's Theorem 3, and other results in his paper, from the Markov condition alone, provided one treats a manipulation as conditionalisation on a 'policy' variable appropriately related to the variable manipulated. Further, two extensions of Theorem 3 follow fairly directly. First, if the sufficient conditions in Theorem 3 for the equalities of probabilities are violated, distributions satisfying the Markov condition exist for which the equalities do not hold. Secondly, if the Markov condition entails all conditional independencies holding in a distribution, an axiom sometimes called 'faithfulness', the conditions of Theorem 3 are also necessary for the equalities given there.

The graphical formalism captures many of the essential features common to statistical models that sometimes accompany causal or constitutive hypotheses, including linear and nonlinear regression, factor analysis, and both recursive and nonrecursive structural equation models. In many cases, these models are representable as graphical models with additional distribution assumptions. In some cases, the graphical formalism provides an alternative parametrisation of subsets of the distributions associated with a family of models, as, for example, for the graphical subset of distributions from the log-linear parametrisation of the multinomial family (Bishop, Fienberg & Holland, 1975; Whittaker, 1990). Directed graphs also offer an explicit representation of the connexion between causal hypotheses and independence and conditional independence hypotheses in experimental design, and, under various axioms, permit the mathematical investigation of relations between experimental and nonexperimental designs.

Rubin (1974), Rosenbaum & Rubin (1983), Holland (1988) and Pratt & Schlaifer (1988) have provided an important alternative treatment of the prediction of the results of interventions from partial causal knowledge. As Pearl notes, their approach, which involves conditional independence of measured and 'counterfactual' variables, gives results in agreement with the directed graphical approach under an assumption they refer to as 'strong ignorability'. For example, a result given without proof by Pratt & Schlaifer provides a 'sufficient and almost necessary' condition for the equality of the probability of Y when X is manipulated, and the conditional probability of the counterfactual of Y on X. A direct analogue of their claim of sufficiency is provable from the Markov condition and necessity follows from the faithfulness condition, which is true with probability 1 for natural measures on linear and multinomial parameters. This offers a reasonable reconstruction of what they may have meant by 'almost necessary'. The Rubin approach to prediction has some advantages over directed graph approaches, for example in the representation of circumstances in which features of units influence other units. The disadvantages of the framework stem from the necessity of formulating hypotheses explicitly in terms of the conditional independence of actual and counterfactual variables rather than in terms of variables directly influencing others. In our experience, even experts have difficulty reliably judging the conditional independence relations that do or do not follow from assumptions. For example, we have heard many statistically trained people deny, before doing the calculation, that the normality and independence of X, Y and e, coupled with the linear equation $Z = aX + bY + e$, entail that X and Y are dependent conditional on Z. For the same reason, the Rubin framework may make more difficult mathematical proofs of results about invariance, equivalence, search, etc.

There are at least two other alternative approaches to the graphical formalism: Robins' (1986)

G-computation algorithm for calculating the effects of interventions under causal hypotheses expressed as event trees, an extension of the Rubin approach; and Glenn Shafer's (1996) more recent and somewhat different tree structure approach. Where both are applicable, they seem to give the same results as do procedures Pearl describes for computing on directed graphs. An advantage of the directed graph formalism is the naturalness of the representation of influence. Questions regarding the relative power of these alternative approaches are as follows.

(i) Is the graphical approach applicable to cases where the alternatives are not, particularly when there are structures in which it is not assumed that every variable either influences or is influenced by every other?

(ii) Is the graphical approach faster in some instances, because the directed graphs can encode independencies in their structure while event trees cannot?

(iii) Can the alternatives, like the graphical procedure, be extended to cases in which the distribution forced on the manipulated variable is continuous?

As far as we can tell, none of the approaches to date has been able to cope with causal language associated with explanatory variables in proportional hazards models, where the nonlinear structure does not lend itself naturally to conditional independence representations.

[*Received April* 1995]

Discussion of 'Causal diagrams for empirical research' by J. Pearl

BY DAVID FREEDMAN

Department of Statistics, University of California, Berkeley, California 94720, U.S.A.

Causal inference with nonexperimental data seems unjustifiable to many statisticians. For others, the trick can be done almost on a routine basis, with the help of regression and its allied techniques, like path analysis or simultaneous-equation models. However, typical regression studies are problematic, because inferences are conditional on unvalidated, even unarticulated, assumptions: for discussion and reviews of the literature, see Freedman (1991, 1995).

Deriving causation from association by regression depends on stochastic assumptions of the familiar kind, and on less familiar causal assumptions. Building on earlier work by Holland (1988) and Robins (1989) among others, Pearl develops a graphical language in which the causal assumptions are relatively easy to state. His formulation is both natural and interesting. It captures reasonably well one intuition behind regression analysis: causal inferences can be drawn from associational data if you are observing the results of a controlled experiment run by Nature, and the causal ordering of the variables is known. When these assumptions hold, there is identifiability theory that gives an intriguing description of permissible inferences.

Following Holland (1988), I state the causal assumptions along with statistical assumptions that, taken together, justify inference in conventional path models. There is an observational study with n subjects, $i = 1, \ldots, n$. The data will be analysed by regression. There are three measured variables, X, Y, Z. The path diagram has arrows from X to Y; then, from X and Y to Z. The diagram is interpreted as a set of assumptions about causal structure: the data result from coupling together two thought experiments, as specified below. Statistical analysis proceeds from the assumption that subjects are independent and identically distributed in certain respects. That is the basis for estimating regression functions, an issue Pearl does not address; customary tests of significance would follow too.

Random variables are represented in the usual way on a sample space Ω. With notation like Holland's, $Y_{i,x}(\omega)$ represents the Y-value for subject i at $\omega \in \Omega$, if you set the X-value to x. The