examine the issue of spurious causation. Since spurious causation is typically defined as a case in which certain marginal dependencies vanish upon conditioning, the results are relevant to literature in graphical modeling that equates the absence of causation with conditional independence. The idea behind causation in distribution is to examine the distribution of the response $\underline{Y}_x$ when every element of the population has the same value $\underline{x}$ on the causal vector $(\underline{X})$ and to compare the distributions as $\underline{x}$ varies. If the distributions do not change as $\underline{x}$ varies, one says $\underline{X}$ does not cause $\underline{Y}$ in distribution and otherwise one says $\underline{X}$ causes $\underline{Y}$ in distribution. For a conditioning set $\underline{X}_{R^*}$, I show (1) $\underline{X} \perp\!\!\!\perp \underline{Y} \mid \underline{X}_{R^*}$ does not imply $\underline{X}$ does not cause $\underline{Y}$ in distribution, and (2) $\underline{X}$ does not cause $\underline{Y}$ in distribution, does not imply $\underline{X} \perp\!\!\!\perp \underline{Y} \mid \underline{X}_{R^*}$. For example, if $\underline{X}_{R^*}$ is prior to variable X, and X prior to variable Y, with no variables intervening between X and Y, the results state that X may (or may not) "directly influence" Y (using the sense of directly influence in the graphical modelling literature), but X may not (may) cause Y in distribution. Note also there is no path connecting X to Y in this example. This should suggest that causal inferences based on the usual conditional independence relations do not generally sustain a manipulative account of the causal relation. Sobel (1992) also gives necessary and sufficient conditions for equivalence of conditional independence and causation in distribution.

The foregoing suggests more cautious use of the term "causation" in future work. Not surprisingly, I do not like the terms "causal network" and "influence diagrams"; is not influence just another synonym for causation? The terms employed by Spiegelhalter et al. (directed graphical model, belief networks) seem preferable. Finally, I want to briefly take up the term "irrelevance," sometimes defined via structures that satisfy the axioms of generalized conditional independence (Smith, 1988). (Smith uses the term "uninformative" and is always careful to mention the conditioning set.) From my view, scientists often allow the connotative aspects of words to creep into their use of technical terms, and this can be detrimental. Thus, one might want to choose terms whose connotative aspects are in accord, as much as possible, with the technical definition. In that vein, relevance seems to encompass many things, including causation; for example, the phrase "causally irrelevant" describes one form of irrelevance. Even leaving aside causation, adding information to the conditioning set of marginalizing over this set can make "irrelevant" variables become "relevant"; should these variables have been called irrelevant to begin with?

Comment

Joe Whittaker

It gave me great pleasure to read these articles. Here we have two papers on the application of conditional independence: one to the specification of a graphical model for assessing association in multivariate responses and the other to message passing on a directed graph, in a paper which expertly summarises the probabilistic view of dealing with uncertainty in expert systems. Right at the outset, let me state my own belief that it is not so much the graphic display but the notion of conditional dependence and independence and the idea of a ternary relationship that X_1 affects (or is irrelevant to) X_2 in the presence of X_3, which constitutes the fundamental contribution of graphical models to statistical analysis.

I particularly want to focus on the Cox and Wermuth (CW) paper, which I believe raises some unresolved issues, and discuss three topics in more detail: the value of a graphical representation, the distinction between multivariate and "block" regression and the role of the Schur complement as a partial variance.

Joe Whittaker is Senior Lecturer, Mathematics Department, Lancaster University, LA1 4YF, United Kingdom.

VALUE OF A GRAPHICAL REPRESENTATION

Few practising statisticians can be unaware of the immediate and powerful impact of visual display in conveying the results of a statistical analysis to a consulting client. A tremendous selling point of graphical models is the graph: a fact which is well known to statistical researchers in related areas such as path analysis, causal modelling, factor analysis and structural equation modelling. The same lesson can be learnt from the recently expanding field of neural networks, where statisticians [for instance, Ripley (1993) and Cheng and Titterington (1993)] are discovering that neuroscientists and computer scientists have been busy proposing neural network formulations of nonlinear statistical classification methods. While perhaps not