A Note On “Causality: Models, Reasoning, and Inference” by Judea Pearl

Dr. Alex Liu, August 2005 ***

Executive Summary

The note offers a reflective analysis of Judea Pearl’s Causality: Models, Reasoning, and Inference, highlighting its significance and contributions to the field of empirical research methods. Pearl's work is lauded for introducing new tools and concepts for causal analysis, notably graphical tools, a nonparametric extension of SEM, discussions on causality, and algorithms for discovering causal structures. The author argues that while Pearl aimed to influence the statistics community, his contributions are more directly applicable to empirical researchers seeking to deepen their understanding of causality beyond traditional statistical measures. The note acknowledges the debate around Pearl's assumptions and the practical application of his theories, contrasting his work with other scholars like D.A. Freedman and the SGS team. Despite criticisms and the acknowledged limitations of Pearl’s algorithms in practical scenarios, the note concludes that Pearl and SGS’s methodologies, supplemented with subject knowledge and additional statistical methods, can significantly improve the practice of inferring causality from statistical evidence.


Although it seems the book introduces a new paradigm for conducting causal research, to many empirical scholars, including myself, the main purpose of this book is to:

1) Develop graphical tools to assist in causal analysis.

2) Develop a nonlinear and nonparametric extension of SEM (Structural Equation Modeling).

3) Discuss causality.

4) Develop an algorithm that uses partial correlations to discover causal structure under certain assumptions.

Despite this, the book is a must-read for those involved in empirical research methods. The author expended significant effort to convince the statistics community of the value of his ideas, which I believe is a misplaced effort. His work is more beneficial for those using statistics for empirical research than for statisticians themselves.

Experts in research methods often assert that "research methods are not equivalent to statistics." Research methods encompass statistics and more. Pearl's work aims to formalize this "more" and provide explicit tools for its application. In essence, Pearl's contributions aid in processing statistical results for causal analysis rather than in enhancing statistical analysis itself.

In traditional empirical analysis, at least as taught in mainstream methods, the "more" for causal analysis stipulates that variable A is a cause of variable B if:

1) A and B are correlated.
2) The association arises because A causes B, not vice versa, due to temporal, logical, or theoretical reasons.

3) The association between A and B is not spurious.

I find at least three aspects of Pearl's work particularly noteworthy and applicable to empirical research projects:

1) His explicit definition of the "more."
2) His formal representation of these concepts.
3) His development of rules and tools for managing these concepts.

After fully understanding these aspects, I believe we can utilize Pearl's work to enhance our causal analysis in empirical research.

According to Pearl, statistics encompasses mean, variance, correlation, regression, dependence, conditional independence, association, likelihood, collapsibility, risk ratio, odds ratio, marginalization, conditionalization, "controlling for," etc. In contrast, causal analysis involves randomization, influence, effect, confounding, "holding constant," disturbance, spurious correlation, instrumental variables, intervention, explanation, attribution, etc. The latter minus the former represents the "more."

Professor Pearl's formal language for representing causal analysis includes both structural equation models (linear, nonlinear, and nonparametric) and graphical diagrams. He uses "do(x)" to denote intervention. Many methodologists agree that Pearl's work has provided unprecedented formalization of concepts like spuriousness and confounding.

His proposed rules include criteria for selecting covariates for adjustment, intervention calculus, and counterfactual analysis. Professor Pearl also introduced the IC* algorithm for discovering causal structures.

These contributions are significant milestones in Pearl's work. However, this is just the beginning. Generally, the book raises more questions than it answers, with many gaps that need bridging for effective causal analysis. For instance, indirect effects are not as thoroughly covered as direct and total effects, and estimating the strength of causal influence is also overlooked.

D.A. Freedman of UC Berkeley offers a differing perspective from Pearl's (Freedman, 2004). Freedman and other scholars have criticized Pearl for not conducting modeling or empirical work, instead approaching causation from mathematical or philosophical angles. While this criticism may not wholly diminish the value of theoretical discussions, Freedman argues that Pearl's work relies on numerous unrealistic assumptions, difficult to verify in applied research. Freedman points out that Pearl acknowledges some of these assumptions, like on page 83 of his book, but does not clarify all of them.

Published in 1993 (with a second edition in 2000 by MIT Press), Causation, Prediction, and Search by Spirtes, Glymour, and Scheines (SGS) is also recommended reading. They developed software for their algorithms and applied it to extensive real research. The dialogue between SGS and Freedman delves into whether the transition from statistical evidence to causal inference can be automated without subject knowledge.

However, both the algorithms developed by Pearl and SGS have limitations. Professor
Freedman claims they are based on false assumptions. Many scholars, including myself, have tested these algorithms on empirical data, often finding they lead to inconclusive or erroneous results. Nonetheless, the concepts presented in these algorithms, combined with subject knowledge and other statistical methods like SEM, can be instrumental in generating hypotheses and testing fitted models. Professor Bill Shipley's work is exemplary in this regard (Shipley, 2000).

In conclusion, inferring causality from statistical evidence like correlation undoubtedly requires subject knowledge, additional statistical methods, and diligent effort. Yet, the contributions of Pearl and SGS can significantly enhance current practices.

**Reference**


*** This note was written when Alex worked at IBM Research from Dec 2004 to April 2005, then was modified in August 2005, and edited for wording later with an executive summary added. The author benefited from discussion on this matter with Dr. Sunil Noronha and Joseph Kramer of IBM Research.***
Navigating the Complexity of Causality: Why the AI-Human Collaborative Approach Is Our Best Path Forward | by Dr. Alex Liu | Medium

RM4Es - Data - Equation - Estimation - Evaluation - Explanation/Execution (researchmethods.org)