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

By Dr. Alex Liu

August 2005 ***

This is a note on my reading Judea Pearl’s book “Causality: Models, Reasoning, and Inference” 1999 Cambridge University Press.

Even it sounds like the book is creating a NEW paradigm of conducting causal research, to many empirical scholars including me; the main purpose of this book is to:

1) Develop graphical tools in assisting causal analysis
2) Develop a non-linear and non-parametric extension of SEM
3) Discuss about causality
4) Develop an algorithm using partial correlations to discover causal structure under certain assumptions

However, all the above has already made this book a must read for people in empirical research methods. The author made a lot of effort to convince the statistics community for the acceptance of his ideas. I think that is a wrong approach. His work is more useful to people using statistics for empirical research, than to statisticians.

Experts of research methods often say that “research methods do not equal to statistics”. Research methods equal statistics plus something else. Pearl’s work is to formalize this “something else” and provide tools to work on them explicitly. In other words, Pearl’s work can help us processing statistical results for causal analysis, but not much to improve statistical analysis.

In traditional empirical analysis, at least in the mainstream methods teaching, this “something else” for causal analysis is that variable A is a cause of variable B, if:

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

It seems to me that at least three parts of Pearl work are worth studying and even being applied to some empirical research projects.

(1) His work of explicitly defining the “something else”
(2) His work of formally representing them
(3) His work of developing rules and tools for us to handle them

After gaining a full understanding of the above three items, I think that we can use Pearl’s work to assist our causal analysis in empirical research.

According to Pearl, statistics deals with mean, variance, correlation, regression, dependence, conditional independence, association, likelihood, collapsibility, risk ratio, odd ratio, marginalization, conditionalization, “controlling for”, … While causal analysis deals with randomization, influence, effect, confounding, “holding constant”, disturbance, spurious correlation, instrumental variables, intervention, explanation, attribution, … The second part minus the first part is the “something else”.

Professor Pearl’s language to formally represent causal analysis and its components include both structural equation models (linear, nonlinear and nonparametric) and graphical diagrams. Pearl uses do(x) to represent intervention. As many methodologists will agree, with Pearl’s work, method concepts like spuriousness and confounding, are much better formalized than ever before.

His proposed rules include criterion to select covariates for adjustment, intervention calculus, and counterfactual analysis. Professor Pearl also proposed IC* algorithm to discover causal structures.

These are good contributions made by Pearl’s work. But, this is just a beginning. In general, I think there are more questions than answers in this book. There are also many missing links we need to bridge, in order to conduct a good causal analysis. For example, indirect effects are not covered as much as the direct effects and total effects. How to estimate the strength of a causal influence is also left out.

D.A. Freedman of UC Berkeley takes a different view than that of Pearl (Freedman 2004). Many scholars including Freedman mentioned that Pearl did not do any modeling or empirical work, but just talked causation mathematically or philosophically, that may not be a fair comment as theoretical discussion along can be very valuable. Due to this, Freedman claims that Pearl’s work is based on many assumptions that are unrealistic and difficulty to confirm in applied research. Freedman claims that Pearl acknowledged some of these assumptions like in page 83 of his book, but did not make all them clear.

Published in 1993 (2nd edition in 2000 by MIT Press), the book Causation, Prediction and Search by Spirtes, Glymour, and Scheines (SGS) is worth reading as they actually developed a software for their developed algorithms and applied to a lot of real research. Between SGS and Freedman, there are also many dialogues in discussing whether the work from statistical evidence to causal inference can be automated without any needs for subject knowledge.

Actually, both the algorithms developed by Pearl and SGS do not work well. Professor Freedman of UC Berkeley claims these algorithms do not work as they are based on false assumptions. As I know, quite many scholars including myself tried these algorithms on some empirical data, and found these algorithms often lead us to nowhere or to some errors. However, many ideas presented in these algorithms can be used, in combination with subject knowledge and other statistical methods like structural equation modeling method, to aid us in generating hypotheses and also in testing fitted models. Professor Bill Shipley has some good work along this line (Shipley 2000).

In general, I believe to successfully infer causality from statistical evidence like correlation does require some subject knowledge, additional statistical methods and hard work. But, the work of Pearl and SGS can help to improve the current practice greatly.

***Reference***

Freedman D. A. 2004 Statistical Models for Causation, UCB Statistics Technical Report No. 651 www.stat.berkeley.edu/-census/651.pdf

Freedman D. A. 1998 Are There Algorithms That Discover Causal Structure, UCB Statistics Technical Report No. 514

Halpern, Joseph Y. and Judea Pearl 2001 Cause and Explanations: A Structural-Model Approach, in Proceedings of the Seventeenth Conference on Uncertainy inartificial Intelligence, San Francisco, CA: Morgan Kaufmann

Heckerman, David and Ross Shachter 1995 Decision-theoretic foundation for causal reasoning, Journal of Artificial Intelligence Research 3: 405-430

Pearl, Judea 2003 Statistics and Causal Inference: A Review, Test (2003) Vol. 12, No. 2: 281-345

Newberg, Leland Gerson 2003 Review of Causality Econometric Theory, 19, 2003: 675-685

Shipley, Bill 2000 Cause and Correlation in Biology: A User's Guide to Path Analysis, Structural Equations and Causal Inference. Cambridge University Press

Spirtes, P., Glymour, C., and Scheines, R. 1993 Causation, Prediction, and Search. Springer Lecture Notes in Statistics, no. 81, Springer-Verlag, New York. 2nd edn. 2000, MIT Press

***

**This note was written when Alex worked in IBM Research from Dec 2004 to April 2005, then was modified in August 2005. The author benefited from discussion on this matter with Dr. Sunil Noronha and Joseph Kramer of IBM Research.**

**For further work of Dr. Alex Liu on this subject, please visit below for his book ~ *From Model Building to Model Mapping*:**

**http://www.researchmethods.org/modeling-mapping.htm**

**Or visit below for the RM software where causality reasoning and techniques have been incorporated.**

**http://www.researchmethods.org/rmplatforms.htm**