J. R. Statist. Soc. B (2003)
65, *Part* 4, *pp.* 937–941

A general condition for avoiding effect reversal after marginalization

D. R. Cox

Nuffield College, Oxford, UK

and Nanny Wermuth

University of Mainz, Germany

[Received November 2002. Revised February 2003]

Summary. The paper examines the effect of marginalizing over a possibly unobserved background variable on the conditional relation between a response and an explanatory variable. In particular it is shown that some conclusions derived from least squares regression theory apply in general to testing independence for arbitrary distributions. It is also shown that the general condition of independence of the explanatory variable and the background ensures that monotonicity of dependence is preserved after marginalization. Relations with effect reversal and with collapsibility are sketched.

Keywords: Causality; Collapsibility; Graphs; Randomization; Yule–Simpson paradox

1. Introduction

In many contexts the following issue arises. We have, in the simplest case, three random variables, Y, X and W, and wish to study the dependence of Y on X given W, i.e. we are interested in the conditional density $f_{Y|X,W}(y|x,w)$, in particular as a function of x for arbitrary w. We may, however, marginalize over W, i.e. consider $f_{Y|X}(y|x)$, either for simplification or because W is not observed. For example, in a randomized trial, X is the treatment and W a background variable. In an observational study of the effect of an exposure X on a response Y appropriate adjustment for potential confounders W would commonly be made. If in fact W is not observed, the possible effect on the form of the conditional density must be considered. In many applications there would be other observed variables that are conditioned on throughout but there is no need to show these explicitly in our notation.

2. Linear system

We denote the least squares linear regression coefficient of Y on X adjusting for W by $\beta_{YX.W}$ and the total regression coefficient obtained by marginalizing over W by β_{YX}. Then (Cochran, 1938)

$$\beta_{YX} = \beta_{YX.W} + \beta_{YW.X}\beta_{XW},$$

the two terms on the right-hand side corresponding to the paths from X to Y in Fig. 1(a). It follows that $\beta_{YX} = \beta_{YX.W}$, i.e. the linear dependence of Y on X is unaltered in slope, if and only

Address for correspondence: David Cox, Nuffield College, Oxford, OX1 1NF, UK.
E-mail: david.cox@nuf.ox.ac.uk

 1369–7412/03/65937

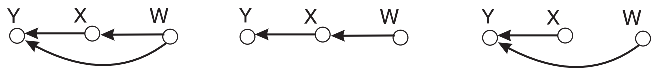

Fig. 1. Simple graphs of relations between response Y, explanatory variable X, in the presence of possibly unobserved variable W (other background variables held fixed are not shown): (a) general relation (two paths from W to Y); (b) one path absent, $Y \perp\!\!\!\perp W|X$; (c) a different path absent, $X \perp\!\!\!\perp W$, as would occur if X randomized treatment

if either $\beta_{YW.X} = 0$ or $\beta_{XW} = 0$. For multivariate Gaussian distributions this implies that

$$Y \perp\!\!\!\perp W|X \quad \text{or} \quad X \perp\!\!\!\perp W. \tag{1}$$

These two cases are represented in Figs 1(b) and 1(c).

In particular, if X represents a randomized treatment the second condition is satisfied by design. In observational investigations and situations in which W is not observed, $X \perp\!\!\!\perp W$ would be an assumption, needing support by subject-matter knowledge and evidence from other studies and possibly investigation by sensitivity analysis (Rosenbaum (2002), chapter 4).

3. General distributions

We now consider corresponding results for arbitrary densities. The general dependences in Fig. 1(a) are represented via the recursive factorization of the joint density of Y, X and W in the form

$$f_{Y,X,W}(y,x,w) = f_W(w)\, f_{X|W}(x|w)\, f_{Y|X,W}(y|x,w).$$

Conditional independences, in particular those shown in Figs 1(b) and 1(c), are represented by missing edges in the graph, as follows.

Given the first condition in expression (1), i.e. the absence of a direct edge from W to Y, it follows that the conditional density of Y given X and W, namely $f_{Y|X,W}(y|x,w)$, does not depend on w. That is, $f_{Y|X,W}(y|x,w)$ depends only on x and hence is the same as $f_{Y|X}(y|x)$ obtained by multiplying by $f_W(w)$ and integrating over w.

The second condition in expression (1), i.e. $X \perp\!\!\!\perp W$, is often the more interesting. In this case, however, $f_{Y|X,W} \neq f_{Y|X}$, in general, in a slightly condensed notation. Importantly, however, if in addition $Y \perp\!\!\!\perp X|W$ then X is independent of both Y and W, so no apparent association in the distribution of Y and X is induced, a familiar advantage of randomization.

In the non-null case, i.e. where there is dependence between Y and X given W, the following important qualitative conclusion holds, namely that monotonicity of dependence is maintained.

The dependence of a random variable V on another random variable U is called stochastically increasing if $P(V > v|U = u)$ is increasing in u for all v, i.e. if V is continuous the partial derivative of the conditional distribution function of V given $U = u$, namely $G(v|u)$, satisfies

$$\partial G(v|u)/\partial u \leqslant 0 \tag{2}$$

for all v and u with strict inequality in a region of positive probability. If U is discrete, partial differentiation is replaced by differencing between adjacent levels. Now suppose that Y given X and W is stochastically increasing in X for all w, so that $\partial F(y|x,w)/\partial x \leqslant 0$ for all y, x and w, where $F(y|x,w)$ is the conditional distribution function of Y given $X = x$ and $W = w$. In general the distribution function of Y given $X = x$ marginalized over W is

$$F(y|x) = \int F(y|x,w)\, f(w|x)\, \mathrm{d}w,$$

where again we have condensed the notation by omitting the suffix from $f_{W|X}(w|x)$. On differentiating with respect to x, we have that for regular distributions

$$\frac{\partial F(y|x)}{\partial x} = \int \left\{ \frac{\partial F(y|x,w)}{\partial x} f(w|x) + F(y|x,w) \frac{\partial f(w|x)}{\partial x} \right\} \mathrm{d}w. \tag{3}$$

Now if $X \perp\!\!\!\perp W$ the second term in the integral is 0 and the conclusion that $\partial F(y|x)/\partial x \leqslant 0$ follows immediately, so by inequality (2) Y remains stochastically increasing in x after marginalization.

If X is discrete then, as noted above, the same argument applies using differencing instead of differentiation. If the levels of X are qualitative and if it is possible to order the levels of X so that Y is stochastically increasing in this ordering in its conditional distribution given X and W, then we have shown that the same property is retained after marginalization over W. There are minor changes if W is discrete. A referee has pointed out that a particularly simple direct proof of monotonicity is available if Y is binary.

Because $f(w|x)$ is a density in w and thus has total integral 1, we have, provided that W is not independent of X, that

$$\int \frac{\partial f(w|x)}{\partial x} \mathrm{d}w = 0,$$

so that in different parts of its range $\partial f(w|x)/\partial x$ takes different signs. This implies that in some circumstances the second part of the integral in equation (3) may be substantial and different in sign from the first part, implying an effect reversal. This reversal is known for contingency tables as the Yule–Simpson paradox (Yule, 1903) and can arise in analysis of variance from unbalanced design (Snedecor and Cochran (1967), pages 472–477). In multiple regression reversal of effect is particularly prone to occur when there is near collinearity in the explanatory variables.

4. Generalization

We have so far treated the random variables as univariate. Multivariate responses would often be treated one component at a time. Multivariate X are most simply studied one contrast at a time, in principle holding other contrasts fixed, although there can in observational studies be difficulties with this which we do not address here. Thus X may consist of several components on an equal footing interrelated in such a way that intervention on any one component has implications for the other components. Blood constituents provide an example.

Very frequently the variable W would be multivariate, especially, for example, if it represents unobserved confounders. In one sense the argument extends immediately, the conditions (1) applying directly to multivariate W. This raises no particular issues for randomized experiments but for observational studies the conditions $Y \perp\!\!\!\perp W|X$ or $X \perp\!\!\!\perp W$ would become strong.

They can, however, be weakened. For this we introduce also the multivariate background variables B that have implicitly been conditioned throughout the previous argument and we assume that Y, X, B and W form a directed acyclic graph (Lauritzen, 1996; Cox and Wermuth, 1996) with Y depending on X, W and B and with X depending on W and B. Any acyclic direction of dependence for nodes in W, B is possible. See Fig. 2(a) for a general set of dependences. Then for the results of the previous section to apply it is enough that $Y \perp\!\!\!\perp W|X, B$ or $X \perp\!\!\!\perp W|B$. Figs 2(b) and 2(c) show one appropriate combination of the two types of conditions in expression (1) which assure also that in a linear system

$$\beta_{YX.W_1W_2B} = \beta_{YX.B}.$$

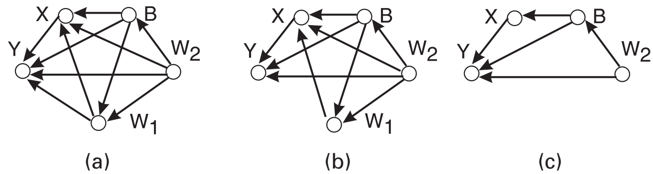

Fig. 2. Relations between response Y, explanatory variable X, in the presence of observed background variables B, and possibly unobserved variables W_1 and W_2: (a) graph without missing edges; jointly sufficient conditions for the preservation of monotonicity of dependence of Y on X after marginalizing over W_1 and W_2 are both of (b) and (c) with (b) independence of Y and W_1 given B, X and W_2 and (c) independence of X and W_2 given B alone

More generally, if both of the two independences

$$Y \perp\!\!\!\perp W_1 | X, B, W_2 \quad \text{and} \quad X \perp\!\!\!\perp W_2 | B$$

are satisfied then monotonicity of dependence of Y on X is preserved after marginalizing over W_1 and W_2.

5. Discussion

The results of this paper have links with other themes in statistics. Indeed the connection with the Yule–Simpson paradox has already been noted. The effects studied here are essentially concerned with the influence of a variable W on the relation between two subsequent other variables Y and X. When such an influence occurs, W may be called a moderating variable. Such a moderating effect can occur even within totally linear systems, such as the multivariate Gaussian distribution, by failure of conditions (1). It is conceptually different from an interactive effect of W on the relation between Y and X, which could not arise in a multivariate Gaussian system.

The absence of a moderating effect implies the possibility of collapsing the data, i.e. essentially ignoring the moderating variable by marginalizing over it without the danger of coming to qualitatively different conclusions about the direction of dependence of Y on X.

Condition (1) coincides for contingency tables with a sufficient condition for simple collapsibility over W of relative risks for Y with respect to X (see Geng (1992), theorem 2, and Wermuth (1987), propositions 1 and 4).

A further general implication is that results about the dependence of Y on X which are qualitatively replicated under a variety of conditions can be viewed formally as examples of monotone dependence. Trust in the generality of such results is enhanced if there is evidence that X is at most weakly related to other important predictors of Y. Estimation of the magnitude of effects and of the relevant precision in general needs inclusion of strata parameters even in balanced data (Gail, 1988). This is relevant to many of the syntheses that are common especially in the clinical trial and epidemiological literature.

References

Cochran, W. G. (1938) The omission or addition of an independent variate in multiple linear regression. *J. R. Statist. Soc.*, suppl., **5**, 171–176.

Cox, D. R. and Wermuth, N. (1996) *Multivariate Dependencies*. London: Chapman and Hall.

Gail, M. H. (1988) The effect of pooling across strata in perfectly balanced studies. *Biometrics*, **44**, 151–162.

Geng, Z. (1992) Collapsibility of relative risk in contingency tables with a response variable. *J. R. Statist. Soc.* B, **54**, 585–593.

Lauritzen, S. L. (1996) *Graphical Models*. Oxford: Oxford University Press.

Rosenbaum, P. R. (2002) *Observational Studies*, 2nd edn. New York: Springer.
Snedecor, G. W. and Cochran, W. G. (1967) *Statistical Methods*. Ames: University of Iowa Press.
Wermuth, N. (1987) Parametric collapsibility and the lack of moderating effects in contingency tables with a dichotomous response variable. *J. R. Statist. Soc.* B, **49**, 353–364.
Yule, G. U. (1903) Notes on the theory of association of attributes in statistics. *Biometrika*, **2**, 121–134.