

3.8.4 Mediation in Linear Systems

When we can assume linear relationships between variables, mediation analysis becomes much simpler than the analysis conducted in nonlinear or nonparametric systems (Section 3.7). Estimating the direct effect of X on Y , for instance, amounts to estimating the path coefficient between the two variables, and this reduces to estimating correlation coefficients, using the techniques introduced in Section 3.8.3. The indirect effect, similarly, is computed via the difference $IE = \tau - DE$, where τ , the total effect, can be estimated by regression in the manner shown in Figure 3.14. In nonlinear systems, on the other hand, the direct effect is defined through expressions such as (3.18), or

$$DE = E[Y|do(x, z)] - E[Y|do(x', z)]$$

where $Z = z$ represents a specific stratum of all other parents of Y (besides X). Even when the identification conditions are satisfied, and we are able to reduce the $do()$ operators (by adjustments) to ordinary conditional expectations, the result will still depend on the specific values of x, x' , and z . Moreover, the indirect effect cannot be given a definition in terms as do -expressions, since we cannot disable the capacity of Y to respond to X by holding variables constant. Nor can the indirect effect be defined as the difference between the total and direct effects, since differences do not faithfully reflect operations in non-linear systems. to X .

Such an operation will be introduced in Chapter 4 (Sections 4.4.5 and 4.5.2) using the language of counterfactuals.

Bibliographical Notes for Chapter 3

Study question 3.3.2 is a version of Lord's paradox (Lord 1967), and is described in Glymour (2006); Hernández-Díaz et al. (2006); Senn (2006); Wainer (1991). A unifying treatment is given in Pearl (2014b). The definition of the do -operator and "ACE" in terms of a modified model, has its conceptual origin with the economist Trygve Haavelmo (1943), who was the first to simulate interventions by modifying equations in the model (see Pearl (2014b) for historical account). Strotz and Wold (1960) later advocated "wiping out" the equation determining X , and Spirtes et al. (1993) gave it a graphical representation in a form of a "manipulated graph." The "adjustment formula" of Eq. (3.5) as well as the "truncated product formula" first appeared in Spirtes et al. (1993), though these are implicit in the G -computation formula of Robins (1986), which was derived using counterfactual assumptions (see Chapter 4). The backdoor criterion of Definition 3.3.1 and its implications for adjustments were introduced in Pearl (1993). The front-door criterion and a general calculus for identifying causal effects (named do -calculus) were introduced in Pearl (1995) and further improved in Tian and Pearl (2002) and Shpitser and Pearl (2007). Section 3.7, and the identification of conditional interventions and c -specific effects is based on Pearl (2009, pp. 113–114). Its extension to dynamic, time-varying policies is described in Pearl and Robins (1995) and Pearl (2009, pp. 119–126). The role of covariate-specific effects in assessing interaction, moderation or effect modification is described in Morgan and Winship (2014) and VanderWeele (2015), whereas applications of Rule 2 to the detection of latent heterogeneity are described in Pearl (2015b). Additional discussions on the use of inverse probability weighting (Section 3.6) can be found in Hernán and Robins (2006). Our

discussion of mediation (Section 3.7), and the identification of CDEs are based on Pearl (2009, pp. 126–130), whereas the fallibility of “conditioning” on a mediator to assess direct effects is demonstrated in Pearl (1998) as well as Cole and Hernán (2002).

The analysis of mediation has become extremely active in the past 15 years, primarily due to the advent of counterfactual logic (see Section 4.4.5); a comprehensive account of this progress is given in VanderWeele (2015). A tutorial survey of causal inference in linear systems (Section 3.8), focusing on parameter identification, is provided by Chen and Pearl (2014). Additional discussion on the confusion of regression versus structural equations can be found in Bollen and Pearl (2013).

A classic, and still the best textbook on the relationships between structural and regression coefficients is Heise (1975) (available online: http://www.indiana.edu/~socpsy/public_files/CausalAnalysis.zip). Other classics are Duncan (1975); Kenny (1979) and Bollen (1989). Classical texts, however, fall short of providing graphical tools of identification, such as those invoking backdoor and G_α (see Study question 3.8.1). A recent exception is Kline (2016).

Introductions to instrumental variables can be found in Greenland (2000) and in many textbooks of econometrics (e.g., Bowden and Turkington (1984); Wooldridge (2013)). Generalized instrumental variables, extending the classical definition of Section 3.8.3 were introduced in (Brito and Pearl 2002).

The program DAGitty (which is available online: <http://www.dagitty.net/dags.html>), permits users to search the graph for generalized instrumental variables, and reports the resulting IV estimators (Textor et al. 2011).