

Confounding of Three Binary-Variable Counterfactual Model with DAG

Jingwei Liu, Shuang Hu

School of Mathematics and System Sciences, Beihang University, Beijing, China

Email: jwliu@buaa.edu.cn

Received May 30, 2013; revised June 30, 2013; accepted July 7, 2013

Copyright © 2013 Jingwei Liu, Shuang Hu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT

Confounding of three binary-variable counterfactual model with directed acyclic graph (DAG) is discussed in this paper. According to the effect between the control variable and the covariate variable, we investigate three causal counterfactual models: the control variable is independent of the covariate variable, the control variable has the effect on the covariate variable and the covariate variable affects the control variable. Using the ancillary information based on conditional independence hypotheses and ignorability, the sufficient conditions to determine whether the covariate variable is an irrelevant factor or whether there is no confounding in each counterfactual model are obtained.

Keywords: Causal Effect; Independence Hypothesis; Counterfactual Model; Confounding Bias; Irrelevant; Ancillary Information; Directed Acyclic Graph

1. Introduction

Causal inference has become an important research field in statistics, data mining, epidemiology and machine learning etc. in recent decades [1-7], and directed acyclic graph (DAG) is involved in describing the relationship between causal connections [4]. Confounding and confounder are two basic concepts for epidemiology causal inference [1,3]. Several models have been presented for causal inference, two of which are the causal diagram model and counterfactual model [6,8,9].

To assess confounding and confounder, two main approaches, “collapsibility-based” and “comparability-based”, are discussed in [10], which regard confounding bias as arising from differences between stratified measures of association and the corresponding original measure or from the exposed and unexposed populations which are not comparable. The comparability-based approach determines a factor to be a confounder if adjusting for it can reduce confounding bias [3,10]. Geng *et al.* (2002) [11] point out that the effect of exposure on the rate of a disease cannot be assessed correctly in the presence of confounding bias. They propose probability criteria for confounding and discuss confounding with multi-value covariate variables. However, their work does not clearly analyze general causal DAG even with three binary-variables, since the simple case of their definition about covariate can only be expressed by **Figure 1** (see Figure

6.2 in p. 61, [12]).

As to three binary-variable DAGs, [5,13] discussed identifiability of the causal effect of the other two kinds of counterfactual models (**Figures 2** and **3**) using the independence hypotheses respectively. Yet, the confounding and confounder in these two simple causal DAGs are not discussed explicitly. For **Figure 2** (see Figure 6.5 in p. 64, [12]), the covariate C is an intermediate variable in the causal chain. [6,12] (p. 30) discuss the intermediate variable causal chain, however more variables are involved

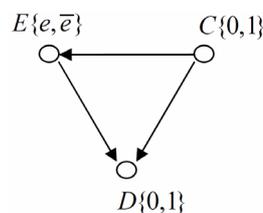


Figure 1. The first model.

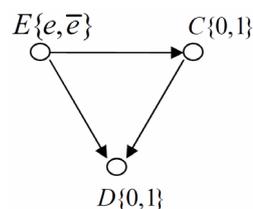


Figure 2. The second model.

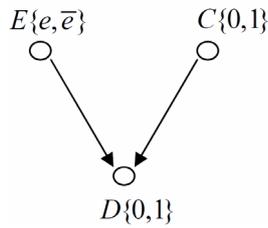


Figure 3. The third model.

in fitting “Back-door” formula and “Front-door” formula. Traditionally, a confounding variable (the precise definition of a confounder) is a variable which is a common cause of both the control variable and the response variable [14] (see Figure 1). Whether the covariate variable, which is not a common cause of both the control variable and the response variable in three binary-variable counterfactual models, is a confounder? [1,10] develop a qualitative definition of confounder: controlling a variable can reduce confounding, then the variable is called a confounder. Hence, the covariate variable, which is not a common cause of both the control variable and the response variable, but affects the response variable, may be a confounder. Recently, the confounder and confounding detection attracts more attention in gene network discovery [15], the question arises from how to investigate the causal DAG in a pure gene network and how to analyze the role of covariate from statistical data if we think that there is causation structure in gene network? It is necessary to discuss the confounder and confounding in the general causal DAG diagram. These motivations drive us to investigate the confounding and confounder in general causal DAG with the definitions in [11].

In this paper, according to the precise definition, one model as shown in Figure 1 is discussed: the covariate variable affects the control variable and the response variable at the same time, and the control variable affects the response variable. By the qualitative definition, we investigate other two models: one, as shown in Figure 2, is that the control variable has the effect on the covariate variable and the covariate variable affects the response variable; the other, as shown in Figure 3, is that the control variable is independent to the covariate variable and the covariate variable affects the response variable. Obviously, the third model is the special case of the other two models with independence of the control variable and the covariate variable. Then we use the formal definitions of a confounder and an irrelevant factor in [11] and the ancillary information based on conditional independence hypotheses [5,13] to discuss the confounding of above-mentioned counterfactual models.

The rest of the paper is organized as follows: In Section 2, we introduce the main notation and definitions, and discuss the relationship between confounder and ir-

relevant factor. In Section 3, confounding and irrelevant factor of three kinds of three binary-variable counterfactual models with DAGs are discussed respectively. The conclusion is given in Section 4.

2. Notation and Definitions

Let E, D, C be binary variables. Let the control variable E be an exposure with the values e and \bar{e} representing “exposed” and “unexposed” respectively. Let the response variable D be an outcome with the values 0 and 1 denoting the presence or absence of a disease, where D_e is the corresponding response when $E = e$ and $D_{\bar{e}}$ is the corresponding response when $E = \bar{e}$, both of which take values 1 or 0 denoting the presence or absence of a disease. Let C be a covariate variable with possible values 0 or 1.

Many kinds of studies focus on the effects of exposure on the rate of a disease in the exposed population. Let $P(D_{\bar{e}} = 1 | E = \bar{e})$ and $P(D_e = 1 | E = e)$ be the proportions of diseased individuals in the unexposed population and the exposed population. Let $P(D_{\bar{e}} = 1 | E = e)$ be the hypothetical proportion of individuals in the exposed population who would have attacked by the disease even if they had not been exposed. Since $P(D_{\bar{e}} = 1 | E = e)$ is a hypothetical proportion, the model is a counterfactual model [8,9].

In order to identify the casual effect of exposure on response, confounding bias B is defined as the difference between the hypothetical proportion of diseased individuals in the exposed population [16,17], that is

$$B = P(D_{\bar{e}} = 1 | E = e) - P(D_{\bar{e}} = 1 | E = \bar{e}). \tag{2.1}$$

If $B = 0$, then there is no confounding.

By the common standardization in epidemiology [1,2, 11,18-20], the standardized proportion $P_{\Delta}(D_{\bar{e}} = 1 | E = \bar{e})$, which is obtained by adjusting the distribution of C in the unexposed population to that in the exposed population, is

$$P_{\Delta}(D_{\bar{e}} = 1 | E = \bar{e}) = \sum_{k=0}^1 P(D_{\bar{e}} = 1 | E = \bar{e}, C = k) P(C = k | E = e) \tag{2.2}$$

Definition 1 [11]. A covariate C is a confounder if

$$\left| P(D_{\bar{e}} = 1 | E = e) - P_{\Delta}(D_{\bar{e}} = 1 | E = \bar{e}) \right| < |B|. \tag{2.3}$$

From the definition, we find that the standardized proportion $P_{\Delta}(D_{\bar{e}} = 1 | E = \bar{e})$ obtained by adjusting for the irrelevant factor is closer to the hypothetical proportion $P(D_{\bar{e}} = 1 | E = e)$ than the observed proportion $P(D_{\bar{e}} = 1 | E = \bar{e})$.

Definition 2 [11]. A covariate C is an irrelevant factor if

$$P_{\Delta}(D_{\bar{e}} = 1|E = \bar{e}) = P(D_{\bar{e}} = 1|E = \bar{e}). \tag{2.4}$$

Since the estimation of the hypothetical proportion is still unchanged after being adjusted for an irrelevant factor, we do not need to adjust it to reduce confounding bias. And, the relationship between irrelevant factor and confounder is obtained in Lemma 1:

Lemma 1. If a covariate C is an irrelevant factor, it is not a confounder. Inversely, if C is a confounder, it is not an irrelevant factor.

Proof.

According to the condition that C is an irrelevant factor, we can obtain that

$$P_{\Delta}(D_{\bar{e}} = 1|E = \bar{e}) = P(D_{\bar{e}} = 1|E = \bar{e}).$$

Then,

$$\begin{aligned} &|P(D_{\bar{e}} = 1|E = e) - P_{\Delta}(D_{\bar{e}} = 1|E = \bar{e})| \\ &= |P(D_{\bar{e}} = 1|E = e) - P(D_{\bar{e}} = 1|E = \bar{e})| = |B|. \end{aligned}$$

which means C is not a confounder.

From the condition that C is a confounder, we can obtain

$$\begin{aligned} &|P(D_{\bar{e}} = 1|E = e) - P_{\Delta}(D_{\bar{e}} = 1|E = \bar{e})| \\ &< |B| = |P(D_{\bar{e}} = 1|E = e) - P(D_{\bar{e}} = 1|E = \bar{e})| \end{aligned}$$

Then,

$$P_{\Delta}(D_{\bar{e}} = 1|E = \bar{e}) \neq P(D_{\bar{e}} = 1|E = \bar{e}).$$

In fact, if

$$P_{\Delta}(D_{\bar{e}} = 1|E = \bar{e}) = P(D_{\bar{e}} = 1|E = \bar{e}),$$

Then,

$$\begin{aligned} &|P(D_{\bar{e}} = 1|E = e) - P_{\Delta}(D_{\bar{e}} = 1|E = \bar{e})| \\ &= |P(D_{\bar{e}} = 1|E = e) - P(D_{\bar{e}} = 1|E = \bar{e})| = |B|. \end{aligned}$$

This is a contradiction!

Hence, C is not an irrelevant factor. \square

[11] (pp. 7-8) gives an example, and illustrates two cases of irrelevant factor and confounder respectively. To illuminate conceptions of confounding and irrelevant factor and Lemma 1, we continue to discuss the relationship based on their original example and give two examples as follows.

Example 1. For example in [11]. Let a factor C express groups categorized by every 10 years of age, and its values 1, 2, 3 and 4 denote the original age groups 20 - 29, 30 - 39, 40 - 49 and 50 - 59 years respectively, we denote it as $\Delta = \{\{1\}, \{2\}, \{3\}, \{4\}\}$. Suppose that there is no exposure effect, *i.e.* there are only individuals of type 1 (individual 'doomed') and type 4 (individual immune to disease), and that the joint distribution of disease, ex-

posure and a factor C is given in **Table 1** of [11] (p. 7), where

$$\begin{aligned} P(D_{\bar{e}} = 1|E = e) &= 0.52, \\ P(D_{\bar{e}} = 1|E = \bar{e}) &= 0.58, \\ |B| &= 0.06. \end{aligned}$$

When the individuals are regrouped by "younger than 50", we denote it as $p = \{\{1, 2, 3\}, \{4\}\}$, which means we adjust the distribution of C , a coarse subpopulation is given in **Table 1**.

Then,

$$\begin{aligned} &P_p(D_{\bar{e}} = 1|E = \bar{e}) \\ &= \frac{122}{200} \times \frac{250}{300} + \frac{52}{100} \times \frac{50}{300} \\ &= 0.595 \neq 0.58 = P(D_{\bar{e}} = 1|E = \bar{e}) \end{aligned}$$

And,

$$\begin{aligned} &|P(D_{\bar{e}} = 1|E = e) - P_p(D_{\bar{e}} = 1|E = \bar{e})| \\ &= |0.595 - 0.52| = 0.075 > |B|. \end{aligned}$$

That is, C is not a confounder, and it is not an irrelevant factor.

Example 2. To continue the discussion of above example in [11], when the individuals are regrouped by "younger than 40 but older than 30", we denote it as $p = \{\{1, 3, 4\}, \{2\}\}$, we can obtain a coarse subpopulation given in **Table 2**.

Then,

$$\begin{aligned} P_p(D_{\bar{e}} = 1|E = \bar{e}) &= \frac{26}{50} \times \frac{100}{300} + \frac{148}{250} \times \frac{200}{300} \\ &= 0.568 \neq 0.58 = P(D_{\bar{e}} = 1|E = \bar{e}) \end{aligned}$$

Table 1. Example of a factor C which is neither a confounder, nor an irrelevant factor.

Type	Distribution for the values of C			
	$C \in \{1, 2, 3\}$		$C \in \{4\}$	
	$E = e$	$E = \bar{e}$	$E = e$	$E = \bar{e}$
1 ("doomed")	133	122	23	52
4 ("immune")	117	78	27	48
Total	250	200	50	100

Table 2. Example of a factor C which is not an irrelevant factor, but is a confounder.

Type	Distribution for the values of C			
	$C \in \{1, 3, 4\}$		$C \in \{2\}$	
	$E = e$	$E = \bar{e}$	$E = e$	$E = \bar{e}$
1 ("doomed")	46	26	110	148
4 ("immune")	54	24	90	102
Total	100	50	200	250

Hence,

$$\begin{aligned} & \left| P(D_{\bar{e}} = 1 | E = e) - P_p(D_{\bar{e}} = 1 | E = \bar{e}) \right| \\ &= |0.568 - 0.52| = 0.048 < |B|. \end{aligned}$$

To sum up, C is not an irrelevant factor, but is a confounder.

As announced in [11], regrouping $p = \{\{1, 2\}, \{3, 4\}\}$, C is a confounder, but not an irrelevant factor. **Example 2** shows that confounder is not unique, $p = \{\{1, 3, 4\}, \{2\}\}$ is another case, and they have different values of $|B|$.

Conclusion: According to Lemma 1, Example 1, Example 2 and [11], whether a factor C is a confounder or an irrelevant factor depends on the adjusting distribution of C . That is, even for a fixed factor C in a specific experiment, for example, age, how to judge it as a confounder or an irrelevant factor relies on the “right” adjustment of its distribution. And, more important, the non-uniqueness of confounder makes the causal analysis be more complex.

If we transform the adjustment of covariate variable in **Figure 1** to the intervention distribution in counterfactual models in **Figures 2** and **3**, the definition 1 and definition 2 would be easily employed in the discussion of confounding and irrelevant factor in the other general causal DAGs.

3. Confounding of Counterfactual Model

Considering the confounding of three binary-variable counterfactual models, there are three counterfactual models of causal DAGs as follows (**Figures 1-3**):

To discuss whether there be confounding in our considering models, we use the conditional independence hypotheses as follows as the ancillary information (H):

- 1) $E \perp D_{\bar{e}}$
- 2) $E \perp D_{\bar{e}} | C = 0$
- 3) $E \perp D_{\bar{e}} | C = 1$.
- 4) $E \perp C$.
- 5) $D_{\bar{e}} \perp C$.
- 6) $D_{\bar{e}} \perp C | E = \bar{e}$.
- 7) $D_{\bar{e}} \perp C | E = e$.

3.1. The First Model

As shown in **Figure 1**, C has effect on E and D at the same time, and C affects E . In order to calculate simply, suppose that

$$\begin{aligned} P(C = 1) &= t, \\ P(E = e | C = 1) &= a_1, P(E = e | C = 0) = a_0, \\ P(D_{\bar{e}} = 1 | E = \bar{e}, C = j) &= b_j, \quad j = 0, 1, \\ P(D_{\bar{e}} = 1 | E = e, C = 0) &= u_0, \\ P(D_{\bar{e}} = 1 | E = e, C = 1) &= u_1, \end{aligned}$$

$$\bar{t} = 1 - t, \bar{a}_0 = 1 - a_0, \bar{a}_1 = 1 - a_1, \bar{b}_j = 1 - b_j.$$

where t, a_0, a_1, b_j can be observed from original data, but u_0, u_1 can not be observed because they are hypothetical proportions.

Then, we obtain the following formulae,

$$\begin{aligned} & P_{\bar{e}}(D_{\bar{e}} = 1 | E = \bar{e}) \\ &= \sum_{k=0}^1 P(D_{\bar{e}} = 1 | E = \bar{e}, C = k) P(C = k | E = \bar{e}) \\ &= \frac{b_0 a_0 \bar{t} + b_1 a_1 t}{a_0 \bar{t} + a_1 t} \\ & P(D_{\bar{e}} = 1 | E = e) \\ &= \sum_{k=0}^1 P(D_{\bar{e}} = 1 | E = e, C = k) P(C = k | E = e) \\ &= \frac{u_0 a_0 \bar{t} + u_1 a_1 t}{a_0 \bar{t} + a_1 t} \\ & P(D_{\bar{e}} = 1 | E = \bar{e}) \\ &= \sum_{k=0}^1 P(D_{\bar{e}} = 1 | E = \bar{e}, C = k) P(C = k | E = \bar{e}) \\ &= \frac{b_0 \bar{a}_0 \bar{t} + b_1 \bar{a}_1 t}{\bar{a}_0 \bar{t} + \bar{a}_1 t} \end{aligned}$$

And,

$$\begin{aligned} B &= P(D_{\bar{e}} = 1 | E = e) - P(D_{\bar{e}} = 1 | E = \bar{e}) \\ &= \frac{b_0 a_0 \bar{t} + b_1 a_1 t}{a_0 \bar{t} + a_1 t} - \frac{b_0 \bar{a}_0 \bar{t} + b_1 \bar{a}_1 t}{\bar{a}_0 \bar{t} + \bar{a}_1 t} \end{aligned}$$

Using the above formulae, we translate each condition of (H) into parameter form:

- 1) $E \perp D_{\bar{e}}$ i.e. $\frac{b_0 a_0 \bar{t} + b_1 a_1 t}{a_0 \bar{t} + a_1 t} = \frac{b_0 \bar{a}_0 \bar{t} + b_1 \bar{a}_1 t}{\bar{a}_0 \bar{t} + \bar{a}_1 t}$
- 2) $E \perp D_{\bar{e}} | C = 0$ i.e. $u_0 = b_0$
- 3) $E \perp D_{\bar{e}} | C = 1$ i.e. $u_1 = b_1$
- 4) $C \perp D_{\bar{e}}$ i.e. $u_0 a_0 + b_0 \bar{a}_0 = u_1 a_1 + b_1 \bar{a}_1$
- 5) $C \perp D_{\bar{e}} | E = \bar{e}$ i.e. $b_0 = b_1$
- 6) $C \perp D_{\bar{e}} | E = e$ i.e. $u_0 = u_1$
- 7) $C \perp E$ i.e. $a_0 = a_1$

Theorem 1. If one of the following conditions holds,

- a) $E \perp C$
- b) $D_{\bar{e}} \perp C | E = \bar{e}$
- c) $D_{\bar{e}} \perp C | E = e, E \perp D_{\bar{e}} | C$

The covariate C is an irrelevant factor.

Proof.

In order to prove C is an irrelevant factor, we only need to prove

$$\frac{b_0 a_0 \bar{t} + b_1 a_1 t}{a_0 \bar{t} + a_1 t} = \frac{b_0 \bar{a}_0 \bar{t} + b_1 \bar{a}_1 t}{\bar{a}_0 \bar{t} + \bar{a}_1 t}.$$

That is,

$$(b_0 - b_1)(a_0 - a_1) = 0.$$

a) From the condition $E \perp C$, we can obtain

$$a_0 = a_1.$$

Then,

$$P_{\Delta}(D_{\bar{e}} = 1 | E = \bar{e}) = P(D_{\bar{e}} = 1 | E = \bar{e}).$$

b) From the condition $D_{\bar{e}} \perp C | E = \bar{e}$, we can obtain

$$b_0 = b_1.$$

Then,

$$P_{\Delta}(D_{\bar{e}} = 1 | E = \bar{e}) = P(D_{\bar{e}} = 1 | E = \bar{e}).$$

c) From the condition $E \perp D_{\bar{e}} | C$, we can obtain

$$E \perp D_{\bar{e}} | C = 0 \quad i.e. \quad u_0 = b_0,$$

$$E \perp D_{\bar{e}} | C = 1 \quad i.e. \quad u_1 = b_1.$$

Since,

$$C \perp D_{\bar{e}} | E = e \quad i.e. \quad u_0 = u_1.$$

We obtain,

$$b_0 = b_1.$$

Hence,

$$P_{\Delta}(D_{\bar{e}} = 1 | E = \bar{e}) = P(D_{\bar{e}} = 1 | E = \bar{e}). \quad \square$$

Theorem 2. If one of the following conditions holds,

- a) $E \perp D_{\bar{e}}$
- b) $C \perp D_{\bar{e}} | E = \bar{e}, E \perp D_{\bar{e}} | C$
- c) $E \perp D_{\bar{e}} | C = 0, C \perp D_{\bar{e}} | E$
- d) $E \perp D_{\bar{e}} | C = 1, C \perp D_{\bar{e}} | E$
- e) $E \perp D_{\bar{e}} | C, E \perp C$

There is no confounding.

Proof.

a) From the condition $E \perp D_{\bar{e}}$, we can obtain

$$\frac{b_0 a_0 \bar{t} + b_1 a_1 t}{a_0 \bar{t} + a_1 t} = \frac{b_0 \bar{a}_0 \bar{t} + b_1 \bar{a}_1 t}{\bar{a}_0 \bar{t} + \bar{a}_1 t}.$$

Then,

$$B = \frac{b_0 a_0 \bar{t} + b_1 a_1 t}{a_0 \bar{t} + a_1 t} - \frac{b_0 \bar{a}_0 \bar{t} + b_1 \bar{a}_1 t}{\bar{a}_0 \bar{t} + \bar{a}_1 t} = 0.$$

b) From the condition $E \perp D_{\bar{e}} | C$, we can obtain

$$E \perp D_{\bar{e}} | C = 0 \quad i.e. \quad u_0 = b_0.$$

$$E \perp D_{\bar{e}} | C = 1 \quad i.e. \quad u_1 = b_1.$$

Furthermore,

$$C \perp D_{\bar{e}} | E = \bar{e} \quad i.e. \quad b_0 = b_1.$$

We obtain,

$$B = \frac{b_0 a_0 \bar{t} + b_1 a_1 t}{a_0 \bar{t} + a_1 t} - \frac{b_0 \bar{a}_0 \bar{t} + b_1 \bar{a}_1 t}{\bar{a}_0 \bar{t} + \bar{a}_1 t} = 0.$$

c) From the condition $C \perp D_{\bar{e}} | E$, we can obtain

$$C \perp D_{\bar{e}} | E = \bar{e} \quad i.e. \quad b_0 = b_1;$$

$$C \perp D_{\bar{e}} | E = e \quad i.e. \quad u_0 = u_1.$$

From the other condition, we obtain

$$E \perp D_{\bar{e}} | C = 0 \quad i.e. \quad u_0 = b_0.$$

Then,

$$B = \frac{b_0 a_0 \bar{t} + b_1 a_1 t}{a_0 \bar{t} + a_1 t} - \frac{b_0 \bar{a}_0 \bar{t} + b_1 \bar{a}_1 t}{\bar{a}_0 \bar{t} + \bar{a}_1 t} = u_0 - b_0 = 0.$$

d) From the condition $C \perp D_{\bar{e}} | E$, we can obtain

$$C \perp D_{\bar{e}} | E = \bar{e} \quad i.e. \quad b_0 = b_1,$$

$$C \perp D_{\bar{e}} | E = e \quad i.e. \quad u_0 = u_1.$$

Furthermore, according to the next condition, we have

$$E \perp D_{\bar{e}} | C = 1 \quad i.e. \quad u_1 = b_1.$$

Then,

$$B = \frac{b_0 a_0 \bar{t} + b_1 a_1 t}{a_0 \bar{t} + a_1 t} - \frac{b_0 \bar{a}_0 \bar{t} + b_1 \bar{a}_1 t}{\bar{a}_0 \bar{t} + \bar{a}_1 t} = u_1 - b_1 = 0.$$

e) From the condition $E \perp D_{\bar{e}} | C$, we can obtain

$$E \perp D_{\bar{e}} | C = 0 \quad i.e. \quad u_0 = b_0;$$

$$E \perp D_{\bar{e}} | C = 1 \quad i.e. \quad u_1 = b_1.$$

Furthermore,

$$C \perp E \quad i.e. \quad a_0 = a_1.$$

Then,

$$B = \frac{b_0 a_0 \bar{t} + b_1 a_1 t}{a_0 \bar{t} + a_1 t} - \frac{b_0 \bar{a}_0 \bar{t} + b_1 \bar{a}_1 t}{\bar{a}_0 \bar{t} + \bar{a}_1 t} = 0. \quad \square$$

3.2. The Second Model

As shown in **Figure 2**, E and C have effect on D at the same time, and E affects C . In order to calculate simply, suppose:

$$P(E = e) = a,$$

$$P(C = 1 | E = e) = c_1, P(C = 1 | E = \bar{e}) = c_0,$$

$$P(D_{\bar{e}} = 1 | E = \bar{e}, C = j) = b_j,$$

$$P(D_{\bar{e}} = 1 | E = e, C = 0) = u_0,$$

$$P(D_{\bar{e}} = 1 | E = e, C = 1) = u_1,$$

$$\bar{a} = 1 - a, \bar{c}_0 = 1 - c_0, \bar{c}_1 = 1 - c_1, \bar{b}_j = 1 - b_j.$$

where a, b_j can be observed from original data. But c_0, c_1, u_0, u_1 can not be observed because they are hypothetical proportions, also can be treated as counterfactual model by intervention [13].

Then, we obtain the following formulae,

$$\begin{aligned} & P_{\Delta}(D_{\bar{e}} = 1|E = \bar{e}) \\ &= \sum_{k=0}^1 P(D_{\bar{e}} = 1|E = \bar{e}, C = k)P(C = k|E = e) \\ &= b_0\bar{c}_1 + b_1c_1 \\ & P(D_{\bar{e}} = 1|E = e) \\ &= \sum_{k=0}^1 P(D_{\bar{e}} = 1|E = e, C = k)P(C = k|E = e) \\ &= u_0\bar{c}_1 + u_1c_1 \\ & P(D_{\bar{e}} = 1|E = \bar{e}) \\ &= \sum_{k=0}^1 P(D_{\bar{e}} = 1|E = \bar{e}, C = k)P(C = k|E = \bar{e}) \\ &= b_0\bar{c}_0 + b_1c_0 \end{aligned}$$

Then,

$$\begin{aligned} B &= P(D_{\bar{e}} = 1|E = e) - P(D_{\bar{e}} = 1|E = \bar{e}) \\ &= u_0\bar{c}_1 + u_1c_1 - b_0\bar{c}_0 - b_1c_0 \end{aligned}$$

Using the above formulae, we translate each condition of (H) into parameter form:

- 1) $E \perp D_{\bar{e}}$ i.e. $u_0\bar{c}_1 + u_1c_1 = b_0\bar{c}_0 + b_1c_0$
- 2) $E \perp D_{\bar{e}}|C = 0$ i.e. $u_0 = b_0$
- 3) $E \perp D_{\bar{e}}|C = 1$ i.e. $u_1 = b_1$
- 4) $C \perp D_{\bar{e}}$ i.e. $\frac{u_0\bar{c}_0\bar{a} + b_0c_1a}{\bar{c}_1a + c_0\bar{a}} = \frac{u_1\bar{c}_1\bar{a} + b_1c_0\bar{a}}{c_0\bar{a} + c_1a}$
- 5) $C \perp D_{\bar{e}}|E = \bar{e}$ i.e. $b_0 = b_1$
- 6) $C \perp D_{\bar{e}}|E = e$ i.e. $u_0 = u_1$
- 7) $C \perp E$ i.e. $c_0 = c_1$

Theorem 3. If one of the following conditions holds,

- a) $E \perp C$
- b) $D_{\bar{e}} \perp C|E = \bar{e}$
- c) $D_{\bar{e}} \perp C|E = e, E \perp D_{\bar{e}}|C$

The covariate C is an irrelevant factor.

Proof.

a) From the condition $E \perp C$, we can obtain

$$c_0 = c_1.$$

Then,

$$\begin{aligned} P_{\Delta}(D_{\bar{e}} = 1|E = \bar{e}) &= b_0\bar{c}_1 + b_1c_1 = b_0\bar{c}_0 + b_1c_0 \\ &= P(D_{\bar{e}} = 1|E = \bar{e}) \end{aligned}$$

b) From the condition $D_{\bar{e}} \perp C|E = \bar{e}$, we can obtain

$$b_0 = b_1.$$

Then,

$$\begin{aligned} P_{\Delta}(D_{\bar{e}} = 1|E = \bar{e}) &= b_0\bar{c}_1 + b_1c_1 = b_0(\bar{c}_1 + c_1) \\ &= b_0(\bar{c}_0 + c_0) = b_0\bar{c}_0 + b_1c_0 = P(D_{\bar{e}} = 1|E = \bar{e}). \end{aligned}$$

c) From the condition $D_{\bar{e}} \perp C|E = e$, we can obtain

$$E \perp D_{\bar{e}}|C = 0 \text{ i.e. } b_0 = b_1,$$

$$E \perp D_{\bar{e}}|C = 1 \text{ i.e. } u_0 = u_1.$$

Furthermore,

$$D_{\bar{e}} \perp C|E = e \text{ i.e. } b_0 = b_1.$$

Then,

$$P_{\Delta}(D_{\bar{e}} = 1|E = \bar{e}) = P(D_{\bar{e}} = 1|E = \bar{e}). \quad \square$$

Theorem 4. If one of the following condition holds,

- a) $E \perp D_{\bar{e}}$
- b) $E \perp D_{\bar{e}}|C = 0, C \perp D_{\bar{e}}|E$
- c) $E \perp D_{\bar{e}}|C = 1, C \perp D_{\bar{e}}|E$
- d) $E \perp D_{\bar{e}}|C, E \perp C$

There is no confounding.

Proof.

a) From the condition that $E \perp D_{\bar{e}}$, we can obtain

$$u_0\bar{c}_1 + u_1c_1 = b_0\bar{c}_0 + b_1c_0.$$

Then,

$$B = u_0\bar{t} + u_1t - b_0\bar{t} - b_1t = 0.$$

b) From the condition $C \perp D_{\bar{e}}|E$, we can obtain

$$C \perp D_{\bar{e}}|E = \bar{e} \text{ i.e. } b_0 = b_1,$$

$$C \perp D_{\bar{e}}|E = e \text{ i.e. } u_0 = u_1.$$

Furthermore,

$$E \perp D_{\bar{e}}|C = 0 \text{ i.e. } u_0 = b_0.$$

Then,

$$B = u_0\bar{c}_1 + u_1c_1 - b_0\bar{c}_0 - b_1c_0 = u_0 - b_0 = 0.$$

c) From the condition $C \perp D_{\bar{e}}|E$, we can obtain

$$C \perp D_{\bar{e}}|E = \bar{e} \text{ i.e. } b_0 = b_1,$$

$$C \perp D_{\bar{e}}|E = e \text{ i.e. } u_0 = u_1.$$

Furthermore,

$$E \perp D_{\bar{e}}|C = 1 \text{ i.e. } u_1 = b_1.$$

Then,

$$B = u_0\bar{c}_1 + u_1c_1 - b_0\bar{c}_0 - b_1c_0 = u_1 - b_1 = 0.$$

d) From the condition $E \perp D_{\bar{e}}|C$, we can obtain

$$E \perp D_{\bar{e}}|C = 0 \text{ i.e. } u_0 = b_0,$$

$$E \perp D_{\bar{e}}|C = 1 \text{ i.e. } u_1 = b_1.$$

Furthermore,

$$C \perp E \text{ i.e. } c_0 = c_1.$$

Then,

$$\begin{aligned} B &= u_0\bar{t} + u_1t - b_0\bar{t} - b_1t \\ &= (u_0 - b_0)\bar{c}_0 + (u_1t - b_1)c_0 = 0. \quad \square \end{aligned}$$

3.3. The Third Model

As shown in **Figure 3**, both E and C have effects on D , and $E \perp C$.

Denote,

$$\begin{aligned} P(E = e) &= a, P(C = 1) = t, \\ P(D_{\bar{e}} = 1 | E = \bar{e}, C = j) &= b_j, \quad j = 0, 1, \\ P(D_{\bar{e}} = 1 | E = e, C = 0) &= u_0, \\ P(D_{\bar{e}} = 1 | E = e, C = 1) &= u_1, \\ \bar{a} &= 1 - a, \bar{t} = 1 - t, \bar{b}_j = 1 - b_j. \end{aligned}$$

where a, t, b_j can be observed from original data, but u_0, u_1 can not be observed because they are hypothetical proportions, or the probability with intervention [5].

Then,

$$\begin{aligned} &P_{\Delta}(D_{\bar{e}} = 1 | E = \bar{e}) \\ &= \sum_{k=0}^1 P(D_{\bar{e}} = 1 | E = \bar{e}, C = k) P(C = k | E = e) \\ &= \sum_{k=0}^1 P(D_{\bar{e}} = 1 | E = \bar{e}, C = k) P(C = k) \\ &= b_0\bar{t} + b_1t \\ &P(D_{\bar{e}} = 1 | E = e) \\ &= \sum_{k=0}^1 P(D_{\bar{e}} = 1 | E = e, C = k) P(C = k | E = e) \\ &= \sum_{k=0}^1 P(D_{\bar{e}} = 1 | E = e, C = k) P(C = k) \\ &= u_0\bar{t} + u_1t \\ &P(D_{\bar{e}} = 1 | E = \bar{e}) \\ &= \sum_{k=0}^1 P(D_{\bar{e}} = 1 | E = \bar{e}, C = k) P(C = k | E = \bar{e}) \\ &= \sum_{k=0}^1 P(D_{\bar{e}} = 1 | E = \bar{e}, C = k) P(C = k) \\ &= b_0\bar{t} + b_1t \end{aligned}$$

Then,

$$\begin{aligned} B &= P(D_{\bar{e}} = 1 | E = e) - P(D_{\bar{e}} = 1 | E = \bar{e}) \\ &= u_0\bar{t} + u_1t - b_0\bar{t} - b_1t. \end{aligned}$$

and,

$$P_{\Delta}(D_{\bar{e}} = 1 | E = \bar{e}) = P(D_{\bar{e}} = 1 | E = \bar{e}).$$

Hence, the covariate C is an irrelevant factor, but not a confounder, and it can not reduce confounding.

Using the above formulae, we translate each condition of (H) into parameter form, where $E \perp C$ is naturally true in **Figure 3**.

- 1) $E \perp D_{\bar{e}}$ i.e. $u_0\bar{t} + u_1t = b_0\bar{t} + b_1t$
- 2) $E \perp D_{\bar{e}} | C = 0$ i.e. $u_0 = b_0$
- 3) $E \perp D_{\bar{e}} | C = 1$ i.e. $u_1 = b_1$
- 4) $C \perp D_{\bar{e}}$ i.e. $b_0\bar{a} + u_0a = b_1\bar{a} + u_1a$
- 5) $C \perp D_{\bar{e}} | E = \bar{e}$ i.e. $b_0 = b_1$
- 6) $C \perp D_{\bar{e}} | E = e$ i.e. $u_0 = u_1$

As discussed above, in the causal DAG **Figure 3** with $E \perp C$, the covariate C is naturally an irrelevant factor. To keep the same expression as other DAGs, we have the following theorem.

Theorem 5. In the causal DAG of **Figure 3** with $E \perp C$, the covariate C is an irrelevant factor.

Theorem 5 shows that in the causal DAG **Figure 3**, covariate C is always an irrelevant factor regardless of any adjustment or interventions on it.

Theorem 6. If one of the following conditions holds,

- a) $E \perp D_{\bar{e}}$
- b) $E \perp D_{\bar{e}} | C = 0, E \perp D_{\bar{e}} | C = 1$
- c) $D_{\bar{e}} \perp C | E, E \perp D_{\bar{e}} | C = 0$
- b) $D_{\bar{e}} \perp C | E, E \perp D_{\bar{e}} | C = 1$

There is no confounding.

Proof.

a) From the condition $E \perp D_{\bar{e}}$, we can obtain

$$u_0\bar{t} + u_1t = b_0\bar{t} + b_1t.$$

Then,

$$B = u_0\bar{t} + u_1t - b_0\bar{t} - b_1t = 0.$$

b) From the conditions

$$E \perp D_{\bar{e}} | C = 0 \text{ and } E \perp D_{\bar{e}} | C = 1,$$

we can obtain

$$u_0 = b_0 \text{ and } u_1 = b_1.$$

Then,

$$B = u_0\bar{t} + u_1t - b_0\bar{t} - b_1t = 0.$$

c) From the condition $D_{\bar{e}} \perp C | E$, we can obtain

$$D_{\bar{e}} \perp C | E = \bar{e} \text{ i.e. } b_0 = b_1,$$

$$C \perp D_{\bar{e}} | E = e \text{ i.e. } u_0 = u_1.$$

Furthermore,

$$E \perp D_{\bar{e}} | C = 0 \text{ i.e. } u_0 = b_0$$

Then,

$$B = u_0\bar{t} + u_1t - b_0\bar{t} - b_1t = u_0 - b_0 = 0.$$

d) The proof is similar to c). \square

4. Conclusions

Using the formal definitions of confounder, non-confounding and irrelevant factor, we discuss the confounding of three kinds of three binary-variable counterfactual models in epidemiology studies and statistics, where the general causal DAG is involved in the discussion. The sufficient conditions of determining non-confounding and irrelevant factor in a three binary-variable causal DAG are discussed. Our work focuses on the general three binary-variable causal DAG, and no other variables are involved in discussion.

Furthermore, confounder and confounding are two different conceptions based on probability criteria, our discussions are definitely different from relative literatures, for example, [5,11,13]. In addition, the non-uniqueness of irrelevant factor and confounder in theory makes it more difficult to detect them and discuss the sufficient and necessary condition. The ancillary information (H) involved in our discussion is only a part of [5,13], hence we only obtain some sufficient conditions, another cause of this design lies in the thought that we want to discuss the causation along causal path. The sufficient and necessary condition discussion would be more complex as shown in our results. The future work will extend the three-variable counterfactual model to multi-variable counterfactual model. And, we will apply the theoretical results to the confounder detection in gene network.

5. Acknowledgements

This work is partially supported by National Natural Science Foundation of China for Youths (NSFC: 10801019). The author would like to thank the anonymous reviewer for the valuable comments and suggestions to make great improvement of this paper.

REFERENCES

- [1] D. G. Kleinbaum, L. L. Kupper and H. Morgenstern, "Epidemiologic Research: Principle and Quantitative Methods," Van Nostrand Reinhold, New York, 1982.
- [2] K. J. Rothman, "Modern Epidemiology," Little Brown, Boston, 1986.
- [3] S. Greenland, J. M. Robins and J. Pearl, "Confounding and Collapsibility in Causal Inference," *Statistical Science*, Vol. 14, No. 1, 1999, pp. 29-46. <http://dx.doi.org/10.1214/ss/1009211805>
- [4] J. Pearl, "Causality: Models, Reasoning and Inference," Cambridge University Press, Cambridge, 2000.
- [5] Z. G. Zheng, Y. Y. Zhang and X. W. Tong, "Identifiability of Causal Effect for a Simple Causal Model," *Science in China*, Vol. 45, No. 3, 2002, pp. 335-341.
- [6] Z. Geng, Y. B. He and X. L. Wang, "Relationship of Causal Effects in a Causal Chain and Related Inference," *Science in China Series A: Mathematics*, Vol. 47, No. 5, 2004, pp. 730-740. <http://dx.doi.org/10.1360/02ys0374>
- [7] X. H. Xie and Z. Geng, "A Recursive Method for Structural Learning of Directed Acyclic Graphs," *Journal of Machine Learning Research*, Vol. 9, 2008, pp. 459-483.
- [8] D. B. Rubin, "Estimating Causal Effects of Treatments in Randomized and Non Randomized Studies," *Journal of Educational Psychology*, Vol. 66, No. 5, 1974, pp. 688-701. <http://dx.doi.org/10.1037/h0037350>
- [9] P. W. Holland, "Statistics and Causal Inference," *Journal of the American Statistical Association*, Vol. 81, No. 396, 1986, pp. 945-970. <http://dx.doi.org/10.1080/01621459.1986.10478354>
- [10] S. Greenland and J. M. Robins, "Identifiability Exchangeability and Epidemiologic Confounding," *International Journal of Epidemiology*, Vol. 15, No. 3, 1986, pp. 413-419. <http://dx.doi.org/10.1093/ije/15.3.413>
- [11] Z. Geng, J. H. Guo and W. K. Fung, "Criteria for Confounders in Epidemiological Studies," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, Vol. 64, No. 1, 2002, pp. 3-15. <http://dx.doi.org/10.1111/1467-9868.00321>
- [12] C. Berzuini, P. Dawid and L. Bernardinelli, "Causality: Statistical Perspectives and Applications," John Wiley & Sons, Hoboken, 2012. <http://dx.doi.org/10.1002/9781119945710>
- [13] Y. Liang and Z. G. Zheng, "The Identifiability Condition of Causal Effect for a Simple Causal Model," *Acta Mathematica Scientia*, Vol. 23A, No. 4, 2003, pp. 456-463.
- [14] G. Wunsch, "Confounding and Control," *Demographic Research*, Vol. 16, No. 4, 2007, pp. 97-120. <http://dx.doi.org/10.4054/DemRes.2007.16.4>
- [15] J. E. Aten, "Causal Not Confounded Gene Networks: Inferring Acyclic and Non-acyclic Gene Bayesian Networks in mRNA Expression Studies using Recursive V-Structures, Genetic Variation, and Orthogonal Causal Anchor Structural Equation Models," Ph.D. Dissertation, University of California, Oakland, 2008.
- [16] P. J. Wickramaratne and T. R. Holford, "Confounding in Epidemiologic Studies: The Adequacy of the Control Groups as a Measure of Confounding," *Biometrics*, Vol. 43, No. 4, 1987, pp. 751-765. <http://dx.doi.org/10.2307/2531530>
- [17] P. W. Holland, "Reader Reactions: Confounding in Epidemiologic Studies," *Biometrics*, Vol. 45, No. 4, 1989, pp. 1310-1316.
- [18] O. S. Miettinen, "Standardization of Risk Ratios," *American Journal of Epidemiology*, Vol. 96, No. 6, 1972, pp. 383-388.
- [19] Z. Geng, X. L. Wang and Y. B. He, "Conditions for Confounding of Causal Diagrams," *Chinese Journal of Epidemiology*, Vol. 23, 2002, pp. 77-78.
- [20] Y. H. Hu and Z. Geng, "Discussion about the Concept of Confounding," *Chinese Journal of Epidemiology*, Vol. 22, No. 6, 2001, pp. 459-461.