因果路径的推断问题

耿直

学生：丁鹏，何平，吴镇国，严巍

北京大学

Methodology

- Reductionism（还原论）: a complex system can be understood by reducing it to a combination of separate parts.
- Whether can a complicated system be studied by reducing it into the basic components?
- For a causal path $T \to S \to Y$,
  to understand the causal mechanism of T and Y, whether can we reduce it into two small mechanisms, one is for T and S, the other is S and Y?

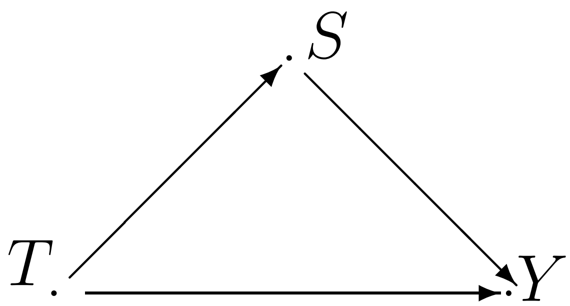
Reducible example

For a causal network (causal DAG),
the full set V is separated into A, B, C.
Suppose $A \perp\!\!\!\perp B | C$. Then for the structural learning on (A, B, C) can be separated into those on (A, C) and (B, C):

1.
$$a \perp\!\!\!\perp c | S, \text{ for } a \in A,\ c \in A \cup C,\ S \subset V$$
$$\iff a \perp\!\!\!\perp c | S', \text{ for } S' \subset A \cup C.$$

2.
$$c_1 \perp\!\!\!\perp c_2 | S, \text{ for } c_1, c_2 \in C,\ S \subset V$$
$$\iff c_1 \perp\!\!\!\perp c_2 | S', \text{ for } S' \subset A \cup C \text{ or } B \cup C.$$

直接作用和间接作用



- 直接和间接作用广泛用于生命科学, 经济学和社会科学. 生物标记物和替代指标等是因果路径上的中间变量.
- 吸烟导致高血脂、胆固醇等问题，引起心血管疾病。高血脂在其中占了多大比例的作用?
- 某种治疗疾病Y的药物T，其副作用是头痛，头痛导致服阿司匹林S, 而阿司匹林S对该病也有疗效。问该药物T对疾病Y的作用有多大?
- 如何从数据中识别直接作用和间接作用等交互作用?

传统的方法

线性模型

$$
\begin{aligned}
Y &= \alpha S + \beta T + \epsilon_y, \\
S &= \gamma T + \epsilon_s.
\end{aligned}
$$

$$
\begin{tikzcd}
 & S \arrow[dr, "\alpha"] & \\
T \arrow[ur, "\gamma"] \arrow[rr, "\beta"] & & Y
\end{tikzcd}
$$

$$Y = (\alpha\gamma + \beta)T + \epsilon_y'.$$

β: T对Y的直接作用;
$\alpha\gamma$: T对Y的间接接作用;
$(\beta + \alpha\gamma)$: T对Y的总作用。

传统的方法

- 如果条件独立$T \perp\!\!\!\perp Y | S$，
  那么，$\beta = 0$，
  因此，T对Y没有直接作用。
- 对于同是高血脂水平人群，不吸烟的人会有吸烟以外的原因导致高血脂，而这些原因可能是心血管疾病的危险因素。
  如果他们与吸烟的人有相同患心血管疾病可能性，那么，吸烟 该有高血脂以外的危害。

直接作用和间接作用

为了进一步说明传统方法的问题，引入因果的概 ：

- w表示个体；
- 观测结果：
  $T(w)$： 个体w的处理；
  $S(w)$： 个体w的中间变量；
  $Y(w)$： 个体w的结果变量；
- 潜在结果：
  $Y_t(w)$，
  $S_t(w)$，
  $Y_{ts}(w)$。

Causal effects

- Individual causal effect of T on S:

$$S_{T=1} - S_{T=0}.$$

- Average causal effect (ACE) of T on S:

$$ACE(T \to S) = E(S_{T=1} - S_{T=0}).$$

- Distributional causal effect (DCE) of T on S:

$$DCE[T \to (S > s)] = P(S_{T=1} > s) - P(S_{T=0} > s).$$

- Similarly

$$ACE(T \to Y) = E(Y_{T=1} - Y_{T=0}),$$

$$ACE(S \to Y|s, s') = E(Y_s - Y_{s'}).$$

主分层的直接作用和间接作用

Rubin的定义采用Y_t，而不采用Y_{ts}:

1. 主分层： $Q(w) = \{S_t(w),\ t \in \mathcal{T}\} = (S_1(w), S_0(w))$
   主分层$Q = q$中的作用，Y_t与Y_{t^*}的比较。
2. 主分层直接作用：当$S_t = S_{t^*}$的主分层，有$Y_t \neq Y_{t^*}$，那么称T对Y有主分层直接作用。
3. 主分层间接作用：如果$P(Y_t - Y_{t^*}|Q = q)$与q有关，那么称T对Y有主分层间接作用。
   或者，如果$P(Y_t - Y_{t^*}|S_t = s, S_{t^*} = s')$与$s$或$s'$有关，那么称$T$对$Y$有主分层间接作用。
   即，改变中间状态$(S_t, S_{t'})$，结果Y_t和$Y_{t'}$)的差别将变化。
   无主分层间接作用表现为，不管怎么改变中间因素的数值(s, s')都不影响结果Y。

传统作用与主分层作用的差别

疫苗W：低剂量1, 高剂量2;
蛋白免疫$S(w)$：高H, 低L;
生存Y：生1，死0;
主分层LH：低疫苗$W=1$，有低免疫$S(1)=L$； 高疫苗$W=2$，有高免疫$S(2)=H$.
假定HL不符合科学意义。 Table 3中有3个主分层。

Display 3. *Two examples of the presence and absence of direct effects*

Principal stratum (equal sized)	Potential outcomes: Surrogates $S(1)$	Surrogates $S(2)$	Survival % $Y(1)$	Survival % $Y(2)$	Observed data $(S_{obs}, \bar{Y}_{obs})$ given treatment assignment: $W_{obs}=1$	$W_{obs}=2$
a. Case where there is a direct causal effect of W on Y given $S(1)$, $S(2)$, but W_{obs} and Y_{obs} are conditionally independent given S_{obs}						
1.	L	L	0	20	L, 20	L, 20
2.	L	H	40	60		H, 80
3.	H	H	80	100	H, 80	
b. Case where there is no direct causal effect of W on Y given $S(1)$, $S(2)$, but W_{obs} and Y_{obs} are conditionally dependent given S_{obs}						
1.	L	L	0	0	L, 20	L, 0
2.	L	H	40	60		H, 70
3.	H	H	80	80	H, 80	

主分层作用的问题

- 因为$S(1)$和$S(2)$只能观察到一个，
不能确定每个个体属于哪个主分层，
因此，主分层的直接作用和间接作用 的识别需要更多的假定。
- 总因果作用不能用主分层的直接作用和间接作用表示。
  - 平均主分层直接因果作用：对于$S_t = S_{t^*}$的主分层$Q = q$，
$$E(Y_t - Y_{t^*}|Q = q).$$
  - 平均主分层间接作用：
$$E(Y_t - Y_{t^*}|Q = q_1) - E(Y_t - Y_{t^*}|Q = q_2).$$
因为$Q = q_1$和$Q = q_2$可能是不同的总体，这个定义是不同总体之间的比较.

主分层作用的识别

- 定义主分层因果参数:

$$\theta_{y|qt} = P(Y(t) = y|q) = P(y|q, t).$$

注意$q = (s(1), s(0))$不可观测。

- 主分层的直接作用:

$$E(Y(1) - Y(0)|q) = \theta_{y|q1} - \theta_{y|q0}.$$

- 引入二值工具变量B，帮助预测主分层。
- 假定单调性$S(T = 1) \geq S(T = 0)$。
- 假定$B \perp\!\!\!\perp (Y, T)|Q$.

主分层作用的识别

- 识别主分层的分布$\pi_{q|b}$: 因为$S(1)=0 \Rightarrow S(0)=0$, $S(0)=1 \Rightarrow S(1)=1$, 所以,

$$P(Q=(0,0)|b) = P(S(1)=0|b,T=1) = P(S=0|b,T=1),$$
$$P(Q=(1,1)|b) = P(S(0)=1|b,T=0) = P(S=1|b,T=0),$$
$$P(Q=(1,0)|b) = 1 - P(Q=(0,0)|b) - P(Q=(1,1)|b).$$

主分层作用的识别

- 识别因果参数$\theta_{y|qt}$: 由$B \perp\!\!\!\perp (T,Y)|Q$,

$$
\begin{aligned}
P(S=0,y|T=1,b) &= \pi_{(0,0)|b}\theta_{y|(0,0)1},\\
P(S=1,y|T=0,b) &= \pi_{(1,1)|b}\theta_{y|(1,1)0},\\
P(S=1,y|T=1,B=1) &= \pi_{(1,1)|1}\theta_{y|(1,1)1} + \pi_{(1,0)|1}\theta_{y|(1,0)1},\\
P(S=1,y|T=1,B=2) &= \pi_{(1,1)|2}\theta_{y|(1,1)1} + \pi_{(1,0)|2}\theta_{y|(1,0)1},\\
P(S=0,y|T=0,B=1) &= \pi_{(0,0)|1}\theta_{y|(0,0)0} + \pi_{(1,0)|1}\theta_{y|(1,0)0},\\
P(S=0,y|T=0,B=2) &= \pi_{(0,0)|2}\theta_{y|(0,0)0} + \pi_{(1,0)|2}\theta_{y|(1,0)0}.
\end{aligned}
$$

控制的直接作用和间接作用

Pearl的定义采用Y_{ts}:

1. 控制直接作用 (Control Direct Effect):

$$CDE_{t,t^*}(s) = Y_{ts} - Y_{t^*s}.$$

2. 控制间接作用 (Control Indirect Effect):

$$CIE_{s,s^*}(t) = Y_{ts} - Y_{ts^*}.$$

3. 总作用TE:

$$\begin{aligned} Y_t - Y_{t^*} &= Y_{ts_t} - Y_{t^*s_{t^*}} = Y_{ts_t} - Y_{ts_{t^*}} + Y_{ts_{t^*}} - Y_{t^*s_{t^*}} \\ &= CIE_{s_t,s_{t^*}}(t) + CDE_{t,t^*}(s_{t^*}). \end{aligned}$$

然的直接作用和间接作用

1. 自然直接作用:

$$NDE_{t,t^*}(t') = Y_{tS_{t'}} - Y_{t^*S_{t'}}.$$

$$NDE_{t,0}(0) = Y_{tS_0} - Y_{0S_0}.$$

2. 自然间接作用:

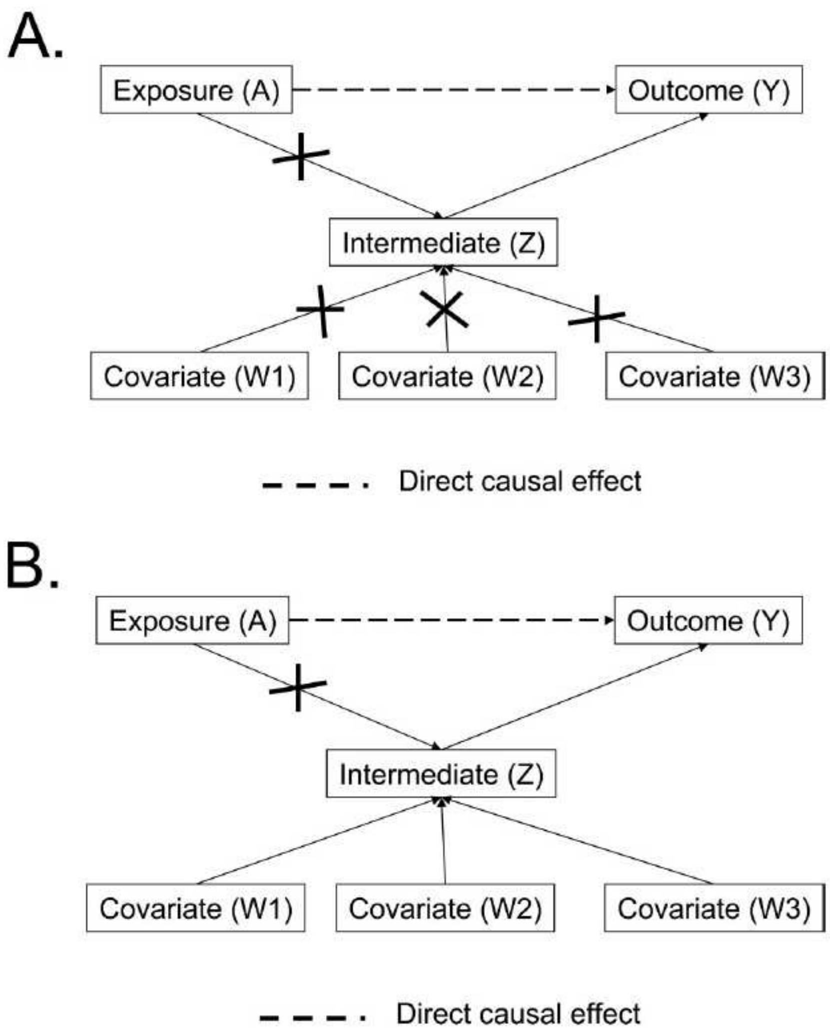
$$NIE_{t,t^*}(t') = Y_{t'S_t} - Y_{t'S_{t^*}}.$$

3. 总作用TE:

$$\begin{aligned} Y_t - Y_{t^*} &= Y_{tS_t} - Y_{t^*S_{t^*}} = Y_{tS_t} - Y_{tS_{t^*}} + Y_{tS_{t^*}} - Y_{t^*S_{t^*}} \\ &= NIE_{t,t^*}(t) + NDE_{t,t^*}(t^*). \end{aligned}$$

然作用和控制作用的差别

Petersen et al. (2006)



(A) 控制直接作用；(B) 自然直接作用

然作用和控制作用的差别

- 考虑一位病人当且仅当接受T治疗时，还服用阿司匹林S； 并且对这位病人治疗T有疗效当且仅当服用阿司匹林S。
- 对于这个病人，有控制的直接作用，因为保持服用阿司匹林的情况下，治疗是有效的：

$$Y_{T=1,S=1} - Y_{T=0,S=1} > 0.$$

- 但是，没有自然的直接作用， 因为保持阿司匹林在当前处理前的零水平时，治疗是无效的

$$Y_{T=1,S_0=0} - Y_{T=0,S_0=0} = 0.$$

识别控制直接作用的条件

令C为混杂变量的集合。

控制直接作用的识别条件:

- $$Y_{ts} \perp\!\!\!\perp T | C.$$
- $$Y_{ts} \perp\!\!\!\perp S | (T, C).$$

识别　然直接作用的条件

自然直接作用的识别还需要条件:

- $$S_T \perp\!\!\!\perp T | C,$$

  和直接作用假定

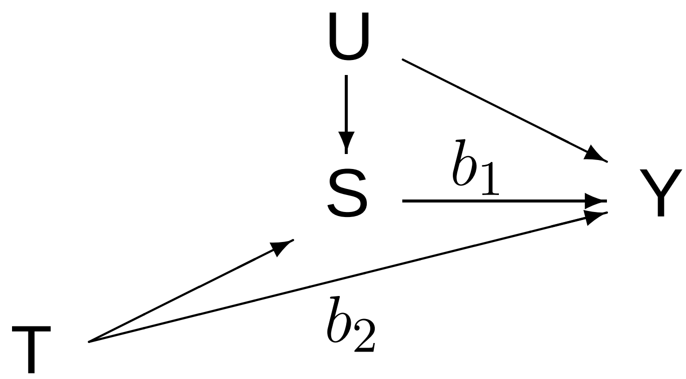
  $$E(Y_{ts} - Y_{0s} | S_0 = s, C) = E(Y_{ts} - Y_{0s} | C).$$

之 关系

VanderWeele (2008)

- 自 用 $NDE_{t,t^*}(t^*)(w) = 0, \forall w,$ 么，T Y的 用（比 t t^*）。
- 用 $CDE_{t,t^*}(s)(w) = 0, \forall s, w,$ 么，T Y的 用（比 t t^*）。
- 无 用，可 自 用.
  无 间 用，可 自 间 用。
  间 用，可 无自 间 用。

因果 模型



不可 测混 因 U

- 型

$$\begin{aligned} Y_{st} &= b_1 s + b_2 t + \phi(U, \varepsilon_Y), \\ S_t &= \psi(t, U, \varepsilon_S). \end{aligned}$$

中$\phi(\cdot)$ $\psi(\cdot)$ 未知函 .

因果　模型

- 因　用:

$$
\begin{aligned}
TE_{t,t'} &= E(Y_t - Y_{t'}) = b_1[E(S|t) - E(S|t')] + b_2(t - t') \\
CDE_{t,t'}(s) &= E(Y_{ts} - Y_{t's}) = b_2(t - t'), \\
NDE_{t,t'}(t') &= E(Y_{t,S_{t'}} - Y_{t'}) = b_2(t - t').
\end{aligned}
$$

因此，$CDE_{t,t'}(s) = NDE_{t,t'}(t') = DE_{t,t'}$.

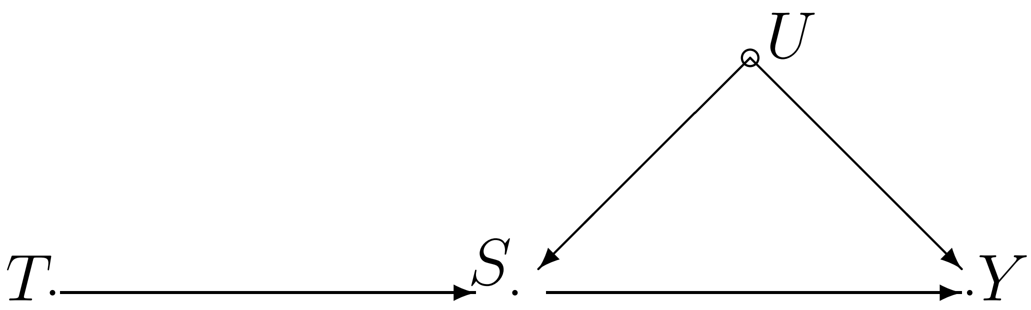
$$
IE(t, t') = TE_{t,t'} - DE_{t,t'} = b_1[E(S|t) - E(S|t')],
$$

- 因　用的识别就　b_1　b_2的识别。

$$
Y_{st} = b_1 s + b_2 t + \phi(U, \varepsilon_Y).
$$

因为S与未　侧的U　，使得　不可识别。

因果推断　　解



- T 　Y 　　　　因　　用的　况下，
  否T　Y的因　推断可以　解为
  T　S的因　推断　S　T的因　推断?

No statistical direct effect

- Prentice's Conditional independence criterion:

$$Y \perp\!\!\!\perp T | S.$$

- Then

$$p(y|t) = \int p(y|s,t)p(s|t)ds = \int p(y|s)p(s|t)ds.$$

$$CE(T \to S) = 0 \Longrightarrow CE(T \to Y) = 0.$$

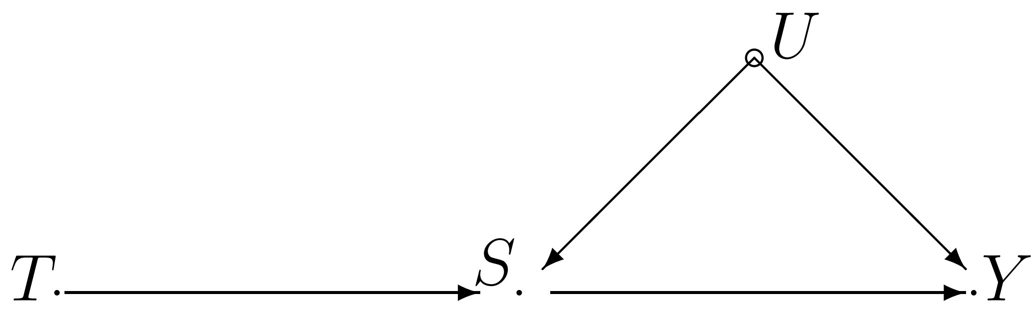
No principal strata direct effect

- Rubin (2004): No principal strata direct effect means an effect of T on Y can occur
only if an effect of T on S has occurred.
- That is, for each s, comparison between two sets

$$\{Y_i(1) : S_i(1) = S_i(0) = s\} \text{ \& } \{Y_i(0) : S_i(1) = S_i(0) = s\}$$

results in equality.

No controlled or natural direct effect



where U is an unobserved variable.

- No causal effect of T on S
  $\Longrightarrow$ no causal effect of T on Y.
- No controlled direct effect implies no principal strata direct effect.
- But these conditions cannot qualitatively predict $CE(T \rightarrow Y)$ with $CE(T \rightarrow S)$ and $CE(S \rightarrow Y)$.

Numerical example

- T: treatment ($T = 1$ treated, $T = 0$ untreated),
- S: Correction of irregular heartbeat ($S = 1$ corrected, $S = 0$ not),
- Y: the survival years.

Assume

1. all effects of the treatment T on the survival Y are through the intermediator S, that is, $Y(t, s) = Y(t', s) = Y(s)$,
2. correction of the heartbeat can increase the survival time for every patient, that is, for all i

$$y_i(S = 1) > y_i(S = 0).$$

Numerical example (continued)

Group	No.	$S(T=0)$	$S(T=1)$	$Y(S=0)$	$Y(S=1)$	$Y(T=0)$	$Y(T=1)$
1	20	0	0	9	10	9	9
2	40	0	1	6	7	6	7
3	20	1	0	5	8	8	5
4	20	1	1	3	5	5	5

$$ACE(T \to S) = \frac{40+20}{100} - \frac{20+20}{100} = \frac{20}{100} > 0,$$

but

$$ACE(T \to Y) = \frac{9 \times 20 + 7 \times 40 \cdots}{100} - \frac{\cdots + 5 \times 20}{100} = 6.6 - 6.8 < 0.$$

Even if $y_i(S=1) > y_i(S=0)$ for every individual,
we cannot use the effect of T on S
to predict the sign of T on Y.

因果 预测 不可 解性

- 即使 统计的 用，
  的 用，
  的 用，
  用，
  根 $DCE(T \to S)$ $DCE(S \to Y)$
  不可 测 $DCE(T \to Y)$的正负。

Conditions based on Causation

- Suppose that the causal diagram holds:

$$\begin{tikzcd} & & U \arrow[dl] \arrow[dr] & \\ T \arrow[r] & S \arrow[rr] & & Y \end{tikzcd}$$

  By the results in Ju and Geng (2010, JRSSB), we have

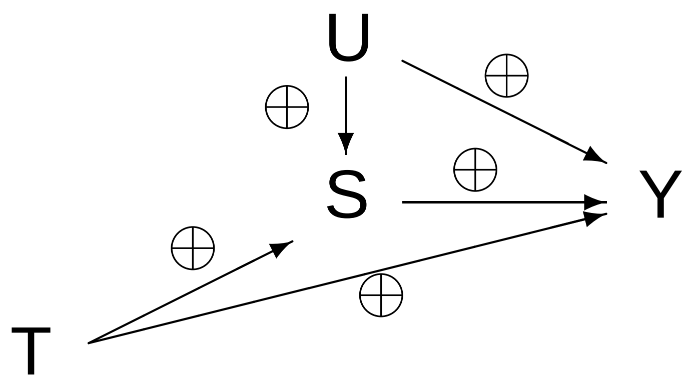
- **Theorem 1.** Suppose that U^* is a subset of U that blocks all back-door paths from S to Y. Then the sign of $DCE(T \to Y)$ is predictable by the signs of $DCE(T \to S)$ and $DCE(S \to Y)$ if
  (1) the DCE of S on Y conditional on $U^* = u^*$ has the same sign for all u^*, and
  (2) the DCE of T on S conditional on $U^* = u^*$ has the same sign for all u^*.

Comments on conditions

- Condition 1 means that S is a risk factor to Y conditionally on U for the treatment group or the control group.
- Condition 2 means a distribution monotonicity of S with respect to the treatment T conditional on U.
- Since U is not observed, conditions of Theorem 1 are untestable.
- There is no direct effect of T on Y.

Signed directed acyclic graphs

- A signed DAG (VanderWeele and Robins, 2010):



- $T \oplus \to S$: a positive effect of T on S: for all $pa_S^* = U$,

$$P(S > s | pa_S^*, t_1) \geq P(S > s | pa_S^*, t_2), \quad \forall t_1 \geq t_2,$$

  where pa_S^* are the parents of S other than T.

Condition based on the signed DAG

- By VanderWeele and Robins (2010), we have
  If the sign of every directed path from T to Y is positive, then $DCE[T \to (Y > y)] \geq 0$.
- Comments
  - The condition requires a complete causal diagram,
  - Since U is not observed, the condition

$$P(Y > y|u, t, s) \geq P(Y > y|u, t, s')$$

    is untestable.

Conditions based on Association

- **Theorem 2.** Suppose that

  1. $P(Y > y|s, T = 1)$ or $P(Y > y|s, T = 0)$ monotonically increases as s increases, and
  2. $P(Y > y|s, T = 1) \geq P(Y > y|s, T = 0)$ for all s.

  Then T has a non-negative DCE on Y if T has a non-negative DCE on S.

- The conditions required in the result are testable since $P(Y > y|s, t)$ is the distribution of observed variables.

Necessary and sufficient conditions

- The one-parameter exponential family has its density

$$p(x;\theta) = C(\theta)\exp\{Q(\theta)\cdot x\}h(x),$$

  where Q is strictly monotonic.

- **Theorem 3.** Suppose that
  1. Prentice's criterion $Y \perp\!\!\!\perp T | S$ holds,
  2. $P(Y > y|s) > P(Y > y|s')$ for $s > s'$, and
  3. S is from the one-parameter exponential family conditional on T.

  Then $ACE(T \to S)$, $DCE(T \to S)$, $ACE(T \to Y)$ and $DCE(T \to Y)$ have the same sign (null, positive, or negative).

For generalized models

- Consider the generalized model for a binary T

$$h[E(Y|S = s, T = t)] \quad = \quad a(s) + bt + c.$$

- **Corollary 1.** Suppose that
  1. $a(s)$ strictly monotonically increases as s increases,
  2. $b \geq 0$.

  - If T has a positive DCE on S, then T has a positive ACE on Y.
  - Further if Y given S and T is from the one-parameter exponential family, then T has a positive DCE on Y.

Necessary and sufficient conditions

- **Corollary 2.** Suppose that
  1. Prentice's criterion $Y \perp\!\!\!\perp T | S$ holds,
  2. $a(s)$ strictly monotonically increases as s increases, and
  3. S given T and Y given S are from the one-parameter exponential family.

  Then $DCE(T \rightarrow S)$ and $DCE(T \rightarrow Y)$ have the same sign (null, positive, or negative).

Comments on conditions

- No direct effects ensure only

$$DCE(T \to S) = 0 \Rightarrow DCE(T \to Y) = 0.$$

- Some causation-based and association-based conditions ensure the implication relations:

$$\begin{aligned} DCE(T \to S) = 0 &\Rightarrow DCE(T \to Y) = 0, \\ DCE(T \to S) > 0 &\Rightarrow DCE(T \to Y) > 0, \\ DCE(T \to S) < 0 &\Rightarrow DCE(T \to Y) < 0. \end{aligned}$$

Comments on conditions

- Additional conditions of the generalized model and the one-parameter exponential family ensure the equivalence relations:

$$
\begin{aligned}
DCE(T \rightarrow S) = 0 \quad &\Leftrightarrow \quad DCE(T \rightarrow Y) = 0, \\
DCE(T \rightarrow S) > 0 \quad &\Leftrightarrow \quad DCE(T \rightarrow Y) > 0, \\
DCE(T \rightarrow S) < 0 \quad &\Leftrightarrow \quad DCE(T \rightarrow Y) < 0.
\end{aligned}
$$

结局

各　　研究中，　　　　法被　　　用。
　理地　　一　复　问　为一系列简　问　？
　法　　面，　待　　一　探索。

- 乌纳穆诺（Miguel de Unamuno, 1864-1936）：

  “理　的最高成　　　引　了人　　　　　　　的怀疑。”

  《人　的悲剧感　》

- 波普(K. R. Popper): “　　始　问　,终　问　”。《猜想与反驳》
- 拉卡托斯(I. Lkatos):　　　　始　问　，终　问　。《　研究纲领　法　》

Thank you!

This is a joint work with my students:
Peng Ding, Ping He, Zhenguo Wu, Wei Yan.