综述：观察性研究中因果效应界的识别与估计

史长浩

2023 年 1 月 9 日

摘要

在观察性研究中，因果参数不可识别，很多研究者考虑为其寻找一个可识别的界，以实现因果参数的“部分识别性”。本文回顾了 5 篇相关的工作：[Balke and Pearl, 1994] 首次提出用线性规划方法求解处理效应的界；[Balke and Pearl, 1997] 利用该技术求解部分依从问题中处理效应的界；[Kuroki et al., 2010] 进一步发展了该方法，并用其求解病例对照研究中因果效应的界；[Gabriel et al., 2022] 将这类问题归入结果依赖型采样的框架下，计算了 8 种因果图对应的因果效应界；[Jun and Lee, 2021] 考虑了仅基于单调性假设时，结果依赖型采样的因果界，但关心的是用协变量分层后的因果参数。其中涉及的思想和技术也许能对未来的工作产生一些启发。

关键字： 观察性研究，因果效应的界，线性规划，结果依赖型采样

1 问题描述

随机化试验是因果推断的金标准，然而在很多问题中，出于对伦理或成本等的考虑，无法控制对受试者的处理分配，这时研究者只能进行观察性研究。如果没有不可检验的假定，根据观察性研究进行因果推断是不可能的，这是因为在观察性研究中，病例组和对照组中的混杂无法得到有效的平衡和控制。因此，根据观察性研究进行因果推断是一个非常具有挑战的课题。

表 1: 2×2 列联表：病例对照研究（总体）。

	$Y=1$	$Y=0$
$X=1$	$P(X=1 \mid Y=1)$	$P(X=1 \mid Y=0)$
$X=0$	$P(X=0 \mid Y=1)$	$P(X=0 \mid Y=1)$

在病例对照研究中，列联表中的参数 $P(X=1 \mid Y=1)$ 与 $P(X=1 \mid Y=0)$ 皆是可识别的，如表1所示。如果知道疾病在人群中的发病率 $P(Y=1)$，则联合分布 $P(X=x, Y=y)$ 可识别。这似乎表明：问题的关键仅仅在于 $P(Y=1)$ 的识别与估计，但事实上，我们关心的因果参数为 CRR(Causal Relative Risk) 和 CRD(Causal Risk Difference)[Kuroki et al., 2010]，其定义为：

$$\mathrm{CRR} = \frac{P(Y(1)=1)}{P(Y(0)=1)}, \tag{1}$$

$$\mathrm{CRD} = P(Y(1)=1) - P(Y(0)=1). \tag{2}$$

利用贝叶斯公式分解包含潜在结果的概率之后，又会出现两个不可识别的参数：$P(Y(1)=1|X=0)$ 与 $P(Y(0)=1|X=1)$，这两个参数连同 $P(Y=1)$ 一同导致了问题的复杂性。

近年来，很多学者对这个问题发起了挑战，主要的思路是为不可识别的因果参数寻找一个可识别的界，或者寻找一个随机区间，以一定的概率覆盖因果参数，实现因果参数的“部分识别性”。

2 文献综述

本节回顾了 5 篇相关的工作：[Balke and Pearl, 1994] 首次提出用线性规划方法求解处理效应的界；[Balke and Pearl, 1997] 利用该技术求解部分依从问题中处理效应的界；[Kuroki et al., 2010] 进一步发展了该技术，并用其求解病例对照研究中因果效应的界；[Gabriel et al., 2022] 将这类问题归入结果依赖型采样的框架下，计算了 8 种因果图对应的因果效应界；[Jun and Lee, 2021] 考虑了仅基于单调性假设时，结果依赖型采样的因果界，但关心的是用协变量分层后的因果参数。

2.1 不依从问题中处理效应的界 [Balke and Pearl, 1994, 1997]

作者针对试验性研究中的不依从问题，即处理分配是随机的，但受试者却不一定完美服从分配的问题，建立了平均处理效应的非参数界。图1为随机临床试验中部分依从问题的因果示意图。

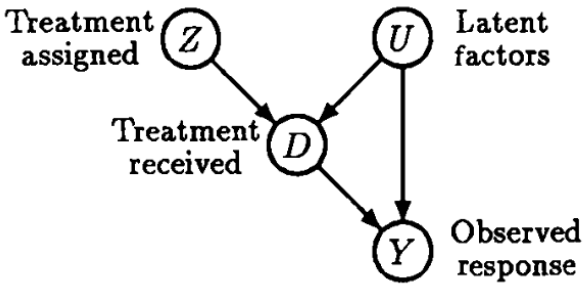


图 1: 随机临床试验中部分依从问题的因果图表示。

这篇文章本来与我们的主题并不相关，然而，这篇文章却是最早研究“因果界”的工作之一。作者提出的用线性规划求解因果界的方法在之后的研究中被广泛使用。

本文总体思想如下，通过对受试者群体进行分类，定义变量 r_d 和 r_y，分别表示受试者的依从类型和治愈（关于用药的）类型。具体地，取值于 0、1、2、3 的 r_d 分别对应：无论是否给药都不吃药、给药则吃不给药则不吃、给药则不吃不给药则吃、无论是否给药都吃药；取值于 0、1、2、3 的 r_y 分别对应：无论是否吃药都不治愈、不吃药不治愈吃药则治愈、不吃药治愈吃药则不治愈、无论是否吃药都治愈。定义 $q_{jk}=P(r_d=j,r_y=k)$，则可以推出 $\mathrm{ACE}(D\to Y)=q_{01}+q_{11}+q_{21}+q_{31}-q_{02}-q_{12}-q_{22}-q_{32}$，且 q 满足线性约束，得线性规划问题如下：

$$
\begin{aligned}
\min \quad & q_{01}+q_{11}+q_{21}+q_{31}-q_{02}-q_{12}-q_{22}-q_{32}. \\
\text{s.t.} \quad & \sum_{j=0}^{3}\sum_{k=0}^{3} q_{jk}=1, \\
& \bar{P}\vec{q}=\vec{p}, \\
& q_{jk}\geq 0 \text{ for } j,k\in\{0,1,2,3\}.
\end{aligned}
\tag{3}
$$

其中

$$
\begin{aligned}
p_{00.0} &= P\left(y_0, d_0 \mid z_0\right), \quad p_{00.1} = P\left(y_0, d_0 \mid z_1\right), \\
p_{01.0} &= P\left(y_0, d_1 \mid z_0\right), \quad p_{01.1} = P\left(y_0, d_1 \mid z_1\right), \\
p_{10.0} &= P\left(y_1, d_0 \mid z_0\right), \quad p_{10.1} = P\left(y_1, d_0 \mid z_1\right), \\
p_{11.0} &= P\left(y_1, d_1 \mid z_0\right), \quad p_{11.1} = P\left(y_1, d_1 \mid z_1\right).
\end{aligned}
$$

及

$$
\begin{aligned}
p_{00.0} &= q_{00} + q_{01} + q_{10} + q_{11}, \\
p_{01.0} &= q_{20} + q_{22} + q_{30} + q_{32}, \\
p_{10.0} &= q_{02} + q_{03} + q_{12} + q_{13}, \\
p_{11.0} &= q_{21} + q_{23} + q_{31} + q_{33}, \\
p_{00.1} &= q_{00} + q_{01} + q_{20} + q_{21}, \\
p_{01.1} &= q_{10} + q_{12} + q_{30} + q_{32}, \\
p_{10.1} &= q_{02} + q_{03} + q_{22} + q_{23}, \\
p_{11.1} &= q_{11} + q_{13} + q_{31} + q_{33}.
\end{aligned}
$$

求解从理论上说非常简单，因为优化空间的维度并不算太高，且线性规划的最优解在可行域顶点处取得，故只需要将各顶点的结果计算出来，取最大和最小即可。算得的结果为

$$
\mathrm{ACE}(D \to Y) \geq \max \left\{ \begin{array}{c}
p_{11.1} + p_{00.0} - 1 \\
p_{11.0} + p_{00.1} - 1 \\
p_{11.0} - p_{11.1} - p_{10.1} - p_{01.0} - p_{10.0} \\
p_{11.1} - p_{11.0} - p_{10.0} - p_{01.1} - p_{10.1} \\
-p_{01.1} - p_{10.1} \\
-p_{01.0} - p_{10.0} \\
p_{00.1} - p_{01.1} - p_{10.1} - p_{01.0} - p_{00.0} \\
p_{00.0} - p_{01.0} - p_{10.0} - p_{01.1} - p_{00.1}
\end{array} \right\},
$$

及

$$
\mathrm{ACE}(D \to Y) \leq \min \left\{ \begin{array}{c}
1 - p_{01.1} - p_{10.0} \\
1 - p_{01.0} - p_{10.1} \\
-p_{01.0} + p_{01.1} + p_{00.1} + p_{11.0} + p_{00.0} \\
-p_{01.1} + p_{11.1} + p_{00.1} + p_{01.0} + p_{00.0} \\
p_{11.1} + p_{00.1} \\
p_{11.0} + p_{00.0} \\
-p_{10.1} + p_{11.1} + p_{00.1} + p_{11.0} + p_{10.0} \\
-p_{10.0} + p_{11.0} + p_{00.0} + p_{11.1} + p_{10.1}
\end{array} \right\}.
$$

类似的技术和思想在之后的文章中会多次出现。

2.2 病例对照研究中的因果界 [Kuroki et al., 2010]

在本文中，作者发展了 [Balke and Pearl, 1994, 1997] 的工作，在不作任何额外假设的情况下，导出了病例对照研究中因果参数 CRD 和 CRR 的非参数界：

$$
\min\left\{-\operatorname{pr}\left(x_0 \mid y_1\right), -\operatorname{pr}\left(x_1 \mid y_0\right)\right\} \leqslant \mathrm{CRD}(x; y), \tag{4}
$$

$$\mathrm{CRD}(x;y) \leqslant \max\left\{\mathrm{pr}\left(x_1 \mid y_1\right), \mathrm{pr}\left(x_0 \mid y_0\right)\right\}. \tag{5}$$

$$0 \leqslant \mathrm{CRR}(x;y) \leqslant \infty. \tag{6}$$

其中 CRR 的非参数界不比其自然的值域更窄，这表明若不做额外假设，数据无法提供对 CRR 有效的信息。

在单调性假设下，CRD 和 CRR 的下界分别取 0 和 1。

在计算 CRR 的非参数界时，作者援引了一个分式线性规划的定理，将一个分式非线性约束转化为线性约束，使之前的方法得以适用。

进一步地，若假设 $0 < a \leqslant \mathrm{pr}\left(y_1\right)$，即假定了发病率的一个严格正下界，并作单调性假设，可得 CRR 有意义的因果界：

$$1 \leqslant \mathrm{CRR}(x;y) \leqslant \frac{\mathrm{pr}\left(x_0 \mid y_0\right)}{\mathrm{pr}\left(x_0 \mid y_1\right) a} + \frac{\mathrm{pr}\left(x_1 \mid y_0\right)}{\mathrm{pr}\left(x_0 \mid y_1\right)}. \tag{7}$$

作者指出，上述结果可以使用简单的代数推理得到，并不依赖繁琐的规划解法。

最后，作者还研究了 3 种缺失机制下因果界的估计问题，因与本文主旨不契合，暂不在此陈述。

2.3 结果依赖型采样的因果界 [Gabriel et al., 2022]

在本文中，作者从一个更普适的观点看待此前的问题，作者认为：病例对照研究、队列研究等观察性研究本质上都是一种所谓的“结果依赖型抽样（outcome-dependent sampling）”的特例。作者强调“选择（Selection）”这一过程，用 $S = 1$ 表示个体入选研究，只有给定 $S = 1$ 条件下的分布是可从观测数据中识别的。

以之前讨论的病例对照研究为例，我们认为 $P(X = x \mid Y = y)$ 是可从数据中识别的分布，但在本文的观点下，$P(X = x \mid Y = y, S = 1)$ 才是可识别的。只有当“选择”仅依赖于 Y 时，如图2，此时 $X \perp\!\!\!\perp S \mid Y$，两个概率才是相等的。

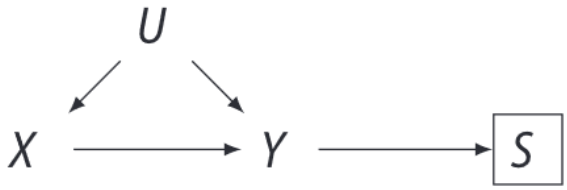
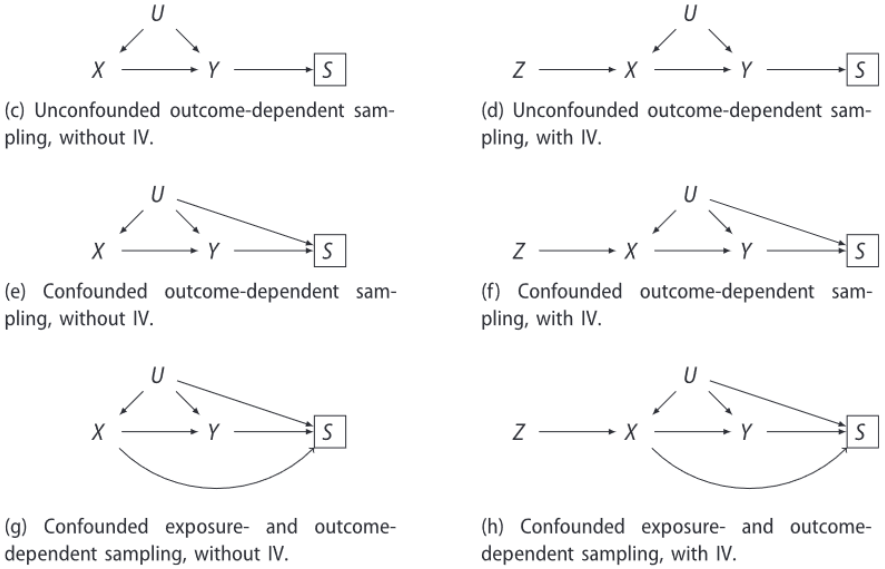


图 2: 因果图：仅依赖 Y 的采样。

类似地，作者又讨论了 $S = 1$ 依赖于混杂 U 和处理 X 的情况，并考虑了存在可用的工具变量时，相应的因果效应界，如图3。作者关心的因果参数为 $\mathrm{CRD} = P(Y(1) = 1) - P(Y(0) = 1)$。

因为 $S = 1$ 的存在，约束不再是线性的，因此此前的线性规划方法不能直接使用，作者的办法是分两步走：首先将目标函数合理地拆成可观测部分和不可观测部分，可观测部分直接使用前人得到的界；下一步集中精力处理不可观测部分（不是很容易，附录中包含 29 页证明），最后把两部分整合起来就得到最后的结果。



(c) Unconfounded outcome-dependent sampling, without IV.

(d) Unconfounded outcome-dependent sampling, with IV.

(e) Confounded outcome-dependent sampling, without IV.

(f) Confounded outcome-dependent sampling, with IV.

(g) Confounded exposure- and outcome-dependent sampling, without IV.

(h) Confounded exposure- and outcome-dependent sampling, with IV.

图 3: 因果图：结果依赖型采样。

但这种方法得到的结果也是非常不简洁的：

$$\theta \geq \max \left\{ \begin{array}{l} p_{11.11} r_1 + p_{00.01} r_0 - 1 \\ p_{11.01} r_0 + p_{00.11} r_1 - 1 \\ \left(p_{11.01} - p_{10.01} - p_{01.01}\right) r_0 - \left(p_{11.11} + p_{01.11}\right) r_1 - \frac{B(0,0,1)}{1+B(0,0,1)} (1 - r_0) - (1 - r_1) \\ \left(p_{11.11} - p_{10.11} - p_{01.11}\right) r_1 - \left(p_{11.01} + p_{01.01}\right) r_0 - \frac{B(0,1,1)}{1+B(0,1,1)} (1 - r_1) - (1 - r_0) \\ -\left(p_{10.11} + p_{01.11}\right) r_1 - \max\left\{ \frac{1}{1+B(1,1,1)}, \frac{B(0,1,1)}{1+B(0,1,1)} \right\} (1 - r_1) \\ -\left(p_{10.01} + p_{01.01}\right) r_0 - \max\left\{ \frac{1}{1+B(1,0,1)}, \frac{B(0,0,1)}{1+B(0,0,1)} \right\} (1 - r_0) \\ \left(p_{00.11} - p_{10.11} - p_{01.11}\right) r_1 - \left(p_{10.01} + p_{00.01}\right) r_0 - \max\left\{ \frac{1}{1+B(1,1,1)}, -\frac{1-B(0,1,1)}{1+B(0,1,1)} \right\} (1 - r_1) - (1 - r_0) \\ \left(p_{00.01} - p_{10.01} - p_{01.01}\right) r_0 - \left(p_{10.11} + p_{00.11}\right) r_1 - \max\left\{ \frac{1}{1+B(1,0,1)}, -\frac{1-B(0,0,1)}{1+B(0,0,1)} \right\} (1 - r_0) - (1 - r_1) \end{array} \right\},$$

$$\theta \leq \min \left\{ \begin{array}{l} 1 - p_{10.11} r_1 - p_{01.01} r_0 \\ 1 - p_{10.01} r_0 - p_{01.11} r_1 \\ \left(p_{10.11} + p_{00.11}\right) r_1 + \left(p_{11.01} + p_{00.01} - p_{10.01}\right) r_0 + (1 - r_1) + \max\left\{ \frac{1-B(0,0,1)}{1+B(0,0,1)}, \frac{B(1,0,1)}{1+B(1,0,1)} \right\} r_0 \\ \left(p_{11.11} + p_{00.11} - p_{10.11}\right) r_1 + \left(p_{10.01} + p_{00.01}\right) r_0 + \max\left\{ \frac{1-B(0,1,1)}{1+B(0,1,1)}, \frac{B(1,1,1)}{B(1,1,1)} \right\} (1 - r_1) + (1 - r_0) \\ \left(p_{11.11} + p_{00.11}\right) r_1 + \max\left\{ \frac{1}{1+B(0,1,1)}, \frac{B(1,1,1)}{1+B(1,1,1)} \right\} (1 - r_1) \\ \left(p_{11.01} + p_{00.01}\right) r_0 + \max\left\{ \frac{1}{1+B(0,0,1)}, \frac{B(1,0,1)}{1+B(1,0,1)} \right\} (1 - r_0) \\ \left(p_{11.11} + p_{00.11} - p_{01.11}\right) r_1 + \left(p_{11.01} + p_{0101}\right) r_0 + \max\left\{ \frac{1}{B(0,1,1)}, -\frac{1-B(1,1,1)}{1+B(1,1,1)} \right\} (1 - r_1) + (1 - r_0) \\ \left(p_{11.01} + p_{00.01} - p_{01.01}\right) r_0 + \left(p_{11.11} + p_{01.11}\right) r_1 + \max\left\{ \frac{1}{B(0,0,1)}, -\frac{1-B(1,0,1)}{1+B(1,0,1)} \right\} (1 - r_0) + (1 - r_1) \end{array} \right\}.$$

值得指出的是，作者认为个体入选研究的概率 $P(S = 1)$ 是已知的或者可估计的，具体来说，被纳入研究的 n 个个体通常来自一个更大的容量为 N 有限样本，而这个有限样本可以视作超总体的一个随机样本，因此可以将 $\frac{n}{N}$ 作为 $P(S = 1)$ 的估计值。

对此我认为拆成下面两句话去理解更容易些：(1) 人类的认识是有边界的，我们所做的任何统计推断都被天然地限制在某个有限总体中，所有因果分析的结论仅在这个有限总体范围内是有效的。(2) 如果该有限总体是来自超总体的一个随机样本，那么，关于该有限总体的统计结论便可外推至整个超总体中。

最后，作者求解 CRD 界的思路和技术也许可以用到 CRR 界的求解中，需要克服的困难是由于做比值而出现的更多非线性项。

2.4 结果依赖型采样的因果界：基于单调性假设 [Jun and Lee, 2021]

本文研究的重点是在单调性假设下做因果推断，关心的因果参数是用协变量分层后的 CRR 和 CRD：

$$\theta(x) := \frac{\mathbb{P}\{Y^*(1)=1 \mid X^*=x\}}{\mathbb{P}\{Y^*(0)=1 \mid X^*=x\}},$$

$$\theta_{\mathrm{AR}}(x) := \mathbb{P}\{Y^*(1)=1 \mid X^*=x\} - \mathbb{P}\{Y^*(0)=1 \mid X^*=x\}.$$

作者证明了在单调性假设下有 $1 \le \theta(x) \le \mathrm{OR}(x)$。

此外，作者定义了一些所谓的聚合参数如：

$$\beta(y) := \int \log \mathrm{OR}(x) dF_{X|Y}(x \mid y),$$

并构造了 $\beta(y)$ 的估计，之后的篇幅是在讨论该类估计的有效性等。

3 数值模拟

在 [Balke and Pearl, 1994, 1997] 线性规划技术的基础上，[Gabriel et al., 2022] 较为全面地讨论了所有情况下 CRD 的因果界（相较本文提到的其他工作而言），因此本节尝试复现其结果，比较不同因果图情形下因果界的数值结果。

为了确保不同方法之间有可比性，且最大程度还原现实问题的复杂性，基于最复杂的情况（图3(h)）生成数据，其生成机制为：

$$\left.\begin{aligned}
p\{U=1\} &\sim \mathrm{Unif}(0,1) \\
p\{Z=1\} &\sim \mathrm{Unif}(0,1) \\
p\{X=1 \mid U,Z\} &= \mathrm{expit}\left(\alpha_1+\alpha_2 U+\alpha_3 Z+\alpha_4 UZ\right) \\
p\{Y=1 \mid U,X\} &= \mathrm{expit}\left(\beta_1+\beta_2 U+\beta_3 X+\beta_4 UX\right) \\
p\{S=1 \mid U,Y,X\} &= \mathrm{expit}\left(\gamma_1+\gamma_2 Y+\gamma_3 U+\gamma_4 X\right) \\
(\alpha_1,\alpha_2,\alpha_3,\alpha_4,\beta_1,\beta_2,\beta_3,\beta_4,\gamma_1,\gamma_2) &\sim N\left(0,5^2\right) \\
\gamma_3 &\sim N\left(0,\sigma_U^2\right) \\
\gamma_4 &\sim N\left(0,\sigma_X^2\right)
\end{aligned}\right\}, \tag{8}$$

其中 $\mathrm{expit}(x) = e^x/(1+e^x)$，感兴趣的因果参数 $\theta = p\{Y(x=1)=1\} - p\{Y(x=0)=1\} = \sum_u \left[\mathrm{expit}\left(\beta_1+\beta_2 u+\beta_3+\beta_4 u\right) - \mathrm{expit}\left(\beta_1+\beta_2 u\right)\right] p\{U=u\}$，即真值依赖于 $p\{U=u\}$ 和 β。在这个模型中，σ_U 和 σ_X 分别决定了采样的暴露依赖程度和混杂程度，即当 $\sigma_U=0$ 时，“选择” 或采样不依赖于混杂 U，σ_U 越大，采样依赖混杂 U 的程度越大；$\sigma_X=0$ 时，采样不依赖于暴露 X，σ_X 越大，采样依赖暴露 X 的程度越大。

在 $\sigma_U=0$ 且 $\sigma_X=0$ 的情况下（图3(d)），做 $N=100000$ 次模拟实验，即从模型 (8) 中产生 100000 个 $p\{U,Z,X,Y,S\}$ 的分布，考察 (c)(d)(e)(f)(g)(h) 对应方法的因果界宽度，如图4左；以及在 θ 不为 0 时，各种方法拒绝 θ 为 0 的能力（频率），如图4右。还可以再增加两种情形，即理想的随机试验情形和有工具变量的随机试验情形。结果表明，适用于更复杂情形下的方法在简单情形（图3(d)）下的表现更差，有更宽的因果界；工具变量的加入会使得因果界变窄（对比 (c) 与 (d)，(e) 与 (f)，(g) 与 (h)）；带有工具变量的方法在零假设 $\theta=0$ 不成立的情况下拒绝零假设的能力随着 $|\theta|$ 增加而增大，而不带有工具变量的方法则不具有这种能力，表现为其因果界总是将 0 包含在内。

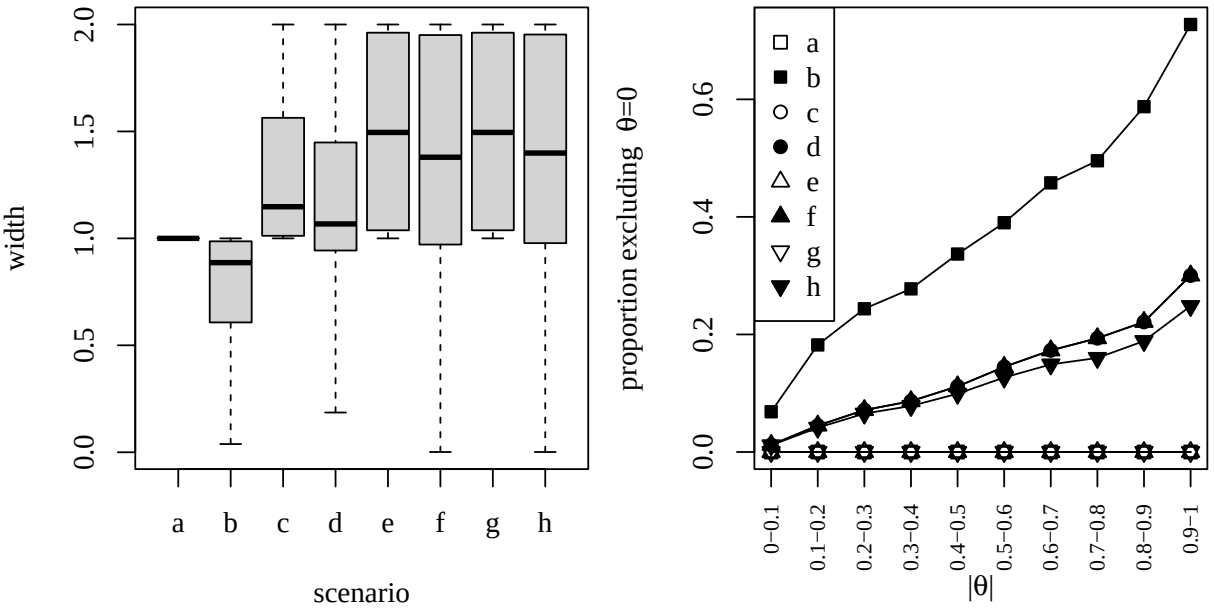


图 4: 左：因果界宽度关于 8 种方法的箱线图；右：8 种方法随 $|\theta|$ 增加而拒绝 $\theta = 0$ 的频率。

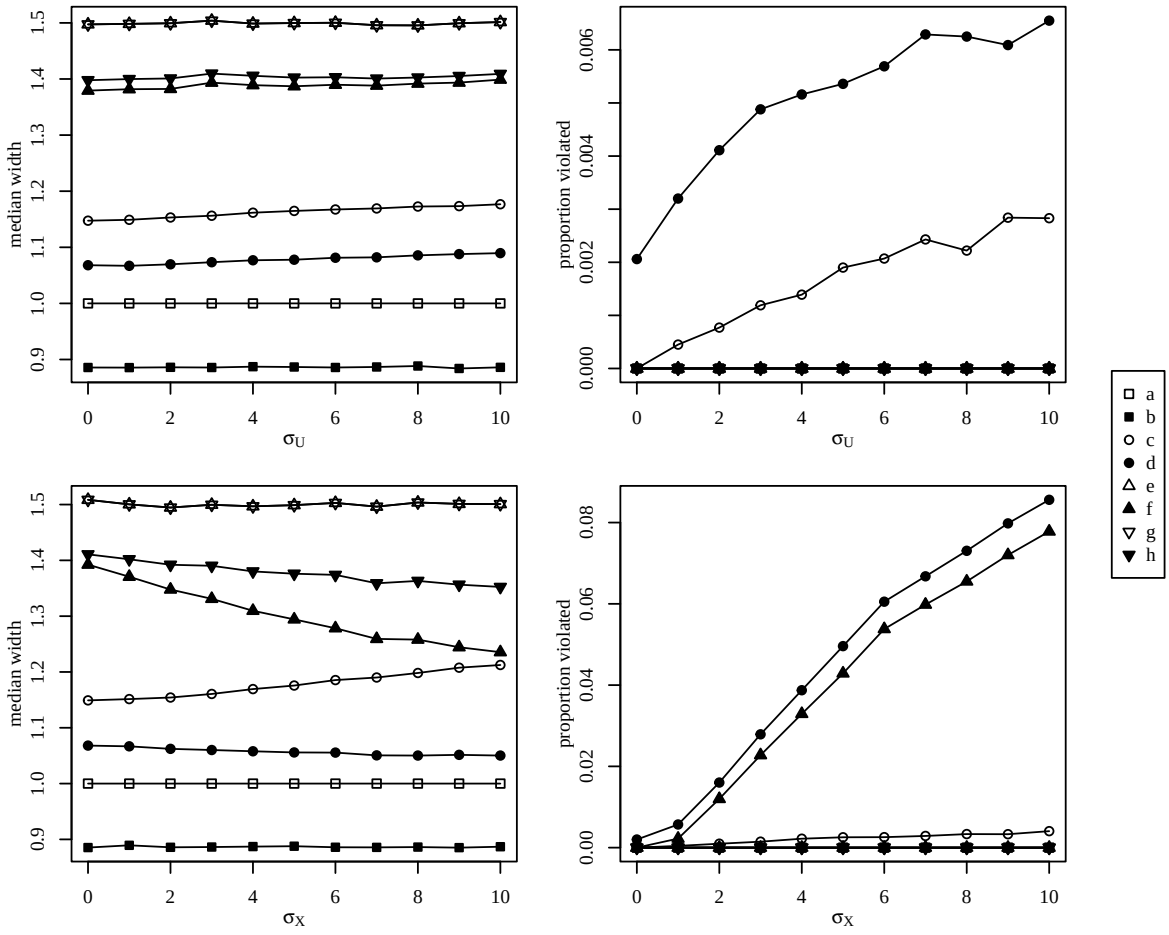


图 5: 因果界的中值宽度 (左列) 和违反比例 (右列)，作为 σ_U(上行) 和 σ_X(下行) 的函数。

更进一步实验，令 σ_U 与 σ_X 从 0 到 10 以间隔 1 逐渐增加，其中一个为正时控制另一个为 0，对应因果图为图3(d) 或 (f)。在每种组合下都进行 100000 次实验，考察各种方法所得因果界宽度的中位数以及未包含真值的频率随 σ_U 与 σ_X 的变化规律，结果如图5所示。结果表明，当 σ_U 增大，即采样越来越依赖于混杂时，各种方法所得因果界的中值宽度几乎不变，而 (c)(d) 因果界不覆盖真值的比例逐渐增加，这是可以期待的，因为 (c)(d) 两种方法恰好是针对无混杂采样的，因此当采样越依赖于混杂，其精度理应下降，有趣的是，这种下降其实是非常有限的，即使 $\sigma_U = 10$，其覆盖真值的频率依然达到了 99%；当 σ_X 增大，即采样越来越依赖于暴露时，(f)(h) 的宽度变窄，而 (c) 的宽度增加，(c)(d)(f) 的违反比例（不覆盖真值的比例）增加，(h) 的违反比例不变，这表明只有 (h) 的效果是在变好的，而 (c) 的效果则在变差（很微弱地变差，比较稳健），在 $\sigma_X = 10$ 时，最差的方法覆盖真值的概率也有 90%。两个结果对比，似乎我们宁可接受有混杂的采样，也不愿意采样受到暴露影响，且在可以容忍一定犯错率（不覆盖真值的比例）的条件下，我们更愿意采用具有更窄的界的方法，如 (c)。

4 讨论

在观察性研究中，因果参数不可识别，研究者通过寻找其可识别的边界来实现部分识别性。但一种更受欢迎的办法是找一个随机区间以一定的概率覆盖感兴趣的因果参数，这是更加困难的一个问题。

最后，关于因果效应界的应用问题，可以参考 [Gabriel and Sachs, 2021, Gabriel et al., 2022] 和 [Jun and Lee, 2021] 等，这些文献中提供了基于因果效应界的有意义的案例。在实际中，这方面的工作可以帮助研究者明确正效应和负效应的边界，以做出更明智的判断。

参考文献

A. Balke and J. Pearl. Counterfactual probabilities: Computational methods, bounds and applications. In *Uncertainty Proceedings 1994*, pages 46–54. Elsevier, 1994.

A. Balke and J. Pearl. Bounds on Treatment Effects from Studies with Imperfect Compliance. *Journal of the American Statistical Association*, 92(439):1171–1176, Sept. 1997.

E. E. Gabriel and M. C. Sachs. On the Use of Nonparametric Bounds for Causal Effects in Null Randomized Trials. *American Journal of Epidemiology*, 190(10):2231–2231, Oct. 2021.

E. E. Gabriel, M. C. Sachs, and A. Sjölander. Causal Bounds for Outcome-Dependent Sampling in Observational Studies. *Journal of the American Statistical Association*, 117(538):939–950, Apr. 2022.

S. J. Jun and S. Lee. Causal inference under outcome-based sampling with monotonicity assumptions. *arXiv preprint arXiv:2004.08318*, 2021.

M. Kuroki, Z. Cai, and Z. Geng. Sharp bounds on causal effects in case-control and cohort studies. *Biometrika*, 97(1):123–132, Mar. 2010.