# 主分层

罗姗姗

March, 2023

# 目录

# 目录

# 后处理混杂变量

到目前为止，我们讨论的大多数问题都是调整前处理混杂变量，即协变量。

在治疗后（但在最终结果之前）发生混杂会对因果推断产生不同的挑战。

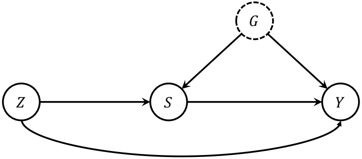
后处理混杂：一个后处理中间变量 $S$ 位于 $Z$ 和 $Y$ 之间的因果路径上：

$$Z \longrightarrow S \longrightarrow Y$$

在经济学中被称为“内生选择问题”。

Rosenbaum（1984）表明：以与前处理协变量 $X$ 相同的方式调整后处理变量 $D$ 会导致偏倚的因果作用。包括一系列（看似不同的）问题。

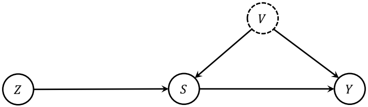
# 主分层框架




| $S_1$ | $S_0$ | $G$ | Description |
|---|---|---|---|
| 1 | 1 | $ss$ | Always-taker |
| 1 | 0 | $s\bar{s}$ | Complier |
| 0 | 1 | $\bar{s}s$ | Defier |
| 0 | 0 | $\bar{s}\bar{s}$ | Never-taker |

- $Z$: 二值处理变量,$Z=1$ 表示接受处理, $Z=0$ 表示接受对照
- $S$: 二值中间变量
- $Y$: 表示感兴趣的结果变量
- $S_z$: 表示在接受处理分配 $Z=z$ 后的中间变量的潜在结果
- $G$: 主分层变量, 定义为 $G=(S_1,S_0)$.

# 不依从问题




| $S_1$ | $S_0$ | $G$ | Description |
|---|---|---|---|
| 1 | 1 | $ss$ | Always-taker |
| 1 | 0 | $s\bar{s}$ | Complier |
| 0 | 1 | $\bar{s}s$ | Defier |
| 0 | 0 | $\bar{s}\bar{s}$ | Never-taker |

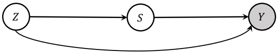
令 $Z=1$ 表示成长于大学附近，$Z=0$ 表示没有成长于大学附近；$S=1$ 表示完成了高中学业，$S=0$ 表示没有. 令 $Y$ 表示对数收入. Angrist et al. (1996) 讨论了依从组上的因果作用的识别性：

$$\mathbb{E}(Y_{S=1}-Y_{S=0} \mid S_1=S_0=1).$$

关于上述参数的识别和估计参考工具变量在单调性假设下的识别性。

# 目录

# 死亡删失




| $S_1$ | $S_0$ | $G$ | Description | $Y_1$ | $Y_0$ |
|---|---|---|---|---|---|
| 1 | 1 | $ss$ | Always-survivors | ✓ | ✓ |
| 1 | 0 | $s\bar{s}$ | Protected | ✓ | × |
| 0 | 1 | $\bar{s}s$ | Harmed | × | ✓ |
| 0 | 0 | $\bar{s}\bar{s}$ | Doomed | × | × |

处理 $Z=1$ 表示接受安慰剂, $Z=0$ 表示接受毒性处理; $S=1$ 表示存活, $S=0$ 表示死亡; $Y$ 是感兴趣的结果变量 (e.g., 体重, 注意 $S=0$ 时所对应的 $Y$ 没有定义).

# 定义及记号

- $Z$：表示二值的处理变量:$Z \in \{1, \ldots, m\}$
- $X$：表示处理前协变量
- $S$：表示生存状态，其中 $S=1$ 表示生存，$S=0$ 表示死亡
- $Y$：表示感兴趣的结果变量
- $S_z$：表示在接受处理分配 $Z=z$ 后的生存状况的潜在结果
- $Y_z$：表示在接受处理分配 $Z=z$ 后的潜在结果
- $G$：主分层变量，定义为 $G=(S_0\ S_1)$

**Assumption 1 (可忽略性)**

$(S_0, S_1, Y_0, Y_1) \perp\!\!\!\perp Z \mid X.$

**Assumption 2 (单调性)**

$S_1 \geq S_0$ .

# 单调性假设下的主分层变量

| $S_1$ | $S_0$ | $G$ | Description | $Y_1$ | $Y_0$ |
|---|---|---|---|---|---|
| 1 | 1 | $ss$ | Always-survivors | $\checkmark$ | $\checkmark$ |
| 1 | 0 | $s\bar{s}$ | Protected | $\checkmark$ | $\times$ |
| 0 | 0 | $\bar{s}\bar{s}$ | Doomed | $\times$ | $\times$ |

我们将感兴趣的参数定义为:

$$\Delta = E(Y_{Z=1} - Y_{Z=0} \mid G = ss).$$

- 在可忽略性假设及单调性假设下: 主分层的权重 $\pi_g(X) = \mathrm{pr}(G = g \mid X)$ 可识别.

$$
\begin{aligned}
\mathrm{pr}(S=1 \mid Z=1, X) &= \mathrm{pr}(S_{Z=1}=1 \mid X) \\
&= \mathrm{pr}(S_{Z=1}=1, S_{Z=0}=1 \mid X) + \mathrm{pr}(S_{Z=1}=1, S_{Z=0}=0 \mid X) \\
&= \mathrm{pr}(G = ss \mid X) + \mathrm{pr}(G = s\bar{s} \mid X), \\
\mathrm{pr}(S=1 \mid Z=0, X) &= \mathrm{pr}(S_{Z=0}=1 \mid X) \\
&= \mathrm{pr}(S_{Z=0}=1, S_{Z=1}=1 \mid X) + \mathrm{pr}(S_{Z=0}=1, S_{Z=1}=0 \mid X) \\
&= \mathrm{pr}(G = ss \mid X) + 0, \\
\mathrm{pr}(S=0 \mid Z=1, X) &= \mathrm{pr}(S_{Z=1}=0 \mid X) \\
&= \mathrm{pr}(S_{Z=0}=1, S_{Z=1}=0 \mid X) + \mathrm{pr}(S_{Z=0}=0, S_{Z=1}=0 \mid X) \\
&= 0 + \mathrm{pr}(G = \bar{s}\bar{s} \mid X).
\end{aligned}
$$

在 $X$ 的每一层上, 等式左边均可观测, 等式右边有三个未知数, 可识别。

- 在可忽略性假设下:

$$\begin{aligned}
\mu_z =& E(Y_z \mid G = ss) \\
&= E\{\underbrace{E(Y_z \mid G = ss, X)}_{\mu_z(X)=E(Y|Z=z,G=ss)} \mid G = ss\} \\
&= E\{\mu_z(X) \mid G = ss\} \\
&= \frac{E\{\mu_z(X)\mathbb{I}(G = ss)\}}{\mathrm{pr}(G = ss)} \\
&= \frac{E[E\{\mu_z(X)\mathbb{I}(G = ss) \mid X\}]}{\mathrm{pr}(G = ss)} \\
&= \frac{E[\mu_z(X)E\{\mathbb{I}(G = ss) \mid X\}]}{\mathrm{pr}(G = ss)} \\
&= \frac{E\big\{\mu_z(X)\pi_{ss}(X)\big\}}{E\{\pi_{ss}(X)\}},
\end{aligned}$$

其中 $\pi_g(X) = \mathrm{pr}(G = g \mid X)$ 是可识别的，我们只需要证明 $\mu_z(X)$ 的可识别性.

我们将感兴趣的参数定义为:

$$
\begin{aligned}
\Delta &= E(Y_{Z=1} - Y_{Z=0} \mid G = ss) \\
&= \underbrace{E(Y_{Z=1} \mid G = ss)}_{\mu_1} - \underbrace{E(Y_{Z=0} \mid G = ss)}_{\mu_0}.
\end{aligned}
$$

我们首先说明 $\mu_0$ 是可识别的，只需要 $E(Y \mid Z = 0, G = ss, X)$ 的识别性即可

$$
\begin{aligned}
\mu_0 &= \frac{E\{\mu_0(X)\pi_{ss}(X)\}}{E\{\pi_{ss}(X)\}} \\
&= \frac{E\{E(Y \mid Z = 0, G = ss, X)\pi_{ss}(X)\}}{E\{\pi_{ss}(X)\}} \\
&= \frac{E\{E(Y \mid Z = 0, S = 1, X)\pi_{ss}(X)\}}{E\{\pi_{ss}(X)\}}
\end{aligned}
$$

类似的，关于参数 $\mu_1$, 我们也只需要证明 $\mu_1(X) = E(Y \mid Z = 1, G = ss, X)$ 的识别性.

# Outline

# 关键识别假设

**Assumption 3 (辅助变量)**

我们假设辅助变量 $X$ 满足: $Y \perp\!\!\!\perp X \mid (Z, G)$.

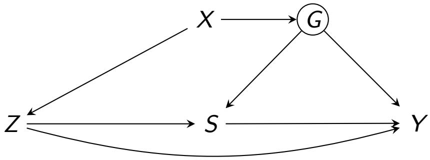



图: 辅助变量

$$E(Y \mid Z = 1, G = ss, X) = E(Y \mid Z = 1, G = ss)$$

# 识别性 I

## 定理 1 (非参识别性)

给定可忽略性、单调性及替代变量假设, 因果参数 $\mu_1(X) = E(Y \mid Z = 1, G = ss, X)$ 在一些正则条件下是可识别的.

## 证明

$$
\begin{aligned}
E(Y \mid Z = 1, S = 1, X) &= E(Y \mid Z = 1, S = 1, X, G = ss)\mathrm{pr}(G = ss \mid X) \\
&\quad + E(Y \mid Z = 1, S = 1, X, G = s\bar{s})\mathrm{pr}(G = s\bar{s} \mid X) \\
&= E(Y \mid Z = 1, G = ss)\mathrm{pr}(G = ss \mid X) \\
&\quad + E(Y \mid Z = 1, G = s\bar{s})\mathrm{pr}(G = s\bar{s} \mid X)
\end{aligned}
$$

# 识别性 II

**证明**

我们取 $X=1,0$,

$$
\begin{aligned}
E(Y \mid Z=1,S=1,X=1) &= E(Y \mid Z=1, G=ss)\mathrm{pr}(G=ss \mid X=1) \\
&\quad + E(Y \mid Z=1, G=s\bar{s})\mathrm{pr}(G=s\bar{s} \mid X=1) \\
E(Y \mid Z=1,S=1,X=0) &= E(Y \mid Z=1, G=ss)\mathrm{pr}(G=ss \mid X=0) \\
&\quad + E(Y \mid Z=1, G=s\bar{s})\mathrm{pr}(G=s\bar{s} \mid X=0)
\end{aligned}
$$

由于 $\mathrm{pr}(G=s\bar{s} \mid X)$ 是可识别的，上述两个方程，两个未知数，在正则性条件得到保证的基础上，参数 $\mu_1(X)=E(Y \mid Z=1, G=ss, X)=E(Y \mid Z=1, G=ss)$和 $E(Y \mid Z=1, G=s\bar{s})$ 是可识别的.

# 非参数估计 I

对于 $X$ 的每个可能取值 $x$，简便起见，我们假设 $x = 0, 1$

1. 我们如下估计 $\hat{\text{pr}}(G = ss \mid X = x) = \frac{\sum_{i=1}^{N} \mathbb{I}(S_i=1, Z_i=0, X_i=x)}{\sum_{i=1}^{N} \mathbb{I}(Z_i=0, X_i=x)}$.
2. 我们如下估计 $\hat{\text{pr}}(G = \overline{ss} \mid X = x) = \frac{\sum_{i=1}^{N} \mathbb{I}(S_i=0, Z_i=1, X_i=x)}{\sum_{i=1}^{N} \mathbb{I}(Z_i=1, X_i=x)}$.
3. 我们如下估计 $\hat{\text{pr}}(G = s\bar{s} \mid X = x) = 1 - \hat{\text{pr}}(G_i = ss \mid X_i = x) - \hat{\text{pr}}(G_i = \overline{ss} \mid X_i = x)$.
4. 我们如下估计 $\hat{E}(Y \mid Z = 0, S = 1, X = x) = \frac{\sum_{i=1}^{N} \mathbb{I}(S_i=1, Z_i=0, X_i=x, Y_i)}{\sum_{i=1}^{N} \mathbb{I}(Z_i=0, S_i=1, X_i=x)}$.
5. 我们如下估计 $\mu_0 = E(Y_0 \mid G = ss)$,
$$\hat{\mu}_0 = \frac{\sum_{i=1}^{N} \hat{E}(Y \mid Z = 0, S = 1, X_i) \hat{\text{pr}}(G_i = ss \mid X_i)}{\sum_{i=1}^{N} \hat{\text{pr}}(G_i = ss \mid X_i)}$$
6. 我们如下估计 $\hat{E}(Y \mid Z = 1, S = 1, X = x) = \frac{\sum_{i=1}^{N} \mathbb{I}(S_i=1, Z_i=1, X_i=x, Y_i)}{\sum_{i=1}^{N} \mathbb{I}(Z_i=1, S_i=1, X_i=x)}$.

# 非参数估计 II

7. 我们如下估计 $\hat{E}(Y \mid Z=1, G=ss)$:

$$\hat{E}(Y \mid Z=1, S=1, X=1) = \gamma_{ss}\hat{\text{pr}}(G=ss \mid X=1) + \gamma_{s\bar{s}}\hat{\text{pr}}(G=s\bar{s} \mid X=1),$$
$$\hat{E}(Y \mid Z=1, S=1, X=0) = \gamma_{ss}\hat{\text{pr}}(G=ss \mid X=0) + \gamma_{s\bar{s}}\hat{\text{pr}}(G=s\bar{s} \mid X=0),$$

   其中 $\gamma_{ss} = E(Y \mid Z=1, G=ss)$, $\gamma_{s\bar{s}} = E(Y \mid Z=1, G=s\bar{s})$.

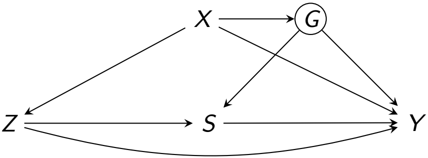
8. 我们如下估计 $\mu_1 = E(Y_1 \mid G=ss)$,

$$\hat{\mu}_1 = \frac{\sum_{i=1}^{N} \gamma_{ss}\hat{\text{pr}}(G_i=ss \mid X_i)}{\sum_{i=1}^{N} \hat{\text{pr}}(G_i=ss \mid X_i)} = \gamma_{ss}.$$

# Outline

# 识别性 I

在现实中，当难以找到辅助变量满足图 1 时，考虑参数模型是更为合适的。



# 识别性 II

不失一般性，我们考虑以下线性模型:

$$E(Y \mid Z=1, G=g, X=x)=\beta_0+\beta_{ss}\mathbb{I}(G=ss)+\beta_x X.$$

$$\Rightarrow E(Y \mid Z=1, G=ss, X=x)=\beta_0+\beta_{ss}+\beta_x X.$$

$$\Rightarrow E(Y \mid Z=1, G=s\bar{s}, X=x)=\beta_0+\beta_x X.$$

**定理 2 (线性识别)**

给定可忽略性、单调性及线性模型假设，因果参数 $\mu_1(X)=E(Y \mid Z=1, G=ss, X)$ 在一些正则条件下是可识别的.

## 识别性 III

**证明**

$$\begin{aligned} E(Y \mid Z=1, S=1, X) &= E(Y \mid Z=1, S=1, X, G=ss)\mathrm{pr}(G=ss \mid X) \\ &\quad + E(Y \mid Z=1, S=1, X, G=s\bar{s})\mathrm{pr}(G=s\bar{s} \mid X) \\ &= (\beta_0+\beta_{ss}+\beta_x X)\mathrm{pr}(G=ss \mid X) + (\beta_0+\beta_x X)\mathrm{pr}(G=s\bar{s} \mid X) \\ &= \beta_0 + \beta_{ss}\mathrm{pr}(G=ss \mid X) + \beta_x X \end{aligned}$$

$$\begin{aligned} E(Y \mid Z=1, S=1, X=0) &= \beta_0 + \beta_{ss}\mathrm{pr}(G=ss \mid X=0) + 0\beta_x \\ E(Y \mid Z=1, S=1, X=1) &= \beta_0 + \beta_{ss}\mathrm{pr}(G=ss \mid X=1) + \beta_x, \\ E(Y \mid Z=1, S=1, X=2) &= \beta_0 + \beta_{ss}\mathrm{pr}(G=ss \mid X=2) + 2\beta_x. \end{aligned}$$

三个方程，三个未知数，上述参数 $(\beta_0, \beta_{ss}, \beta_x)$ 是可识别的.

# 非参数估计 I

对于 $X$ 的每个可能取值 $x$，简便起见，我们假设 $x=0,1,2$

1. 我们如下估计 $\hat{\text{pr}}(G=ss \mid X=x)=\frac{\sum_{i=1}^{N}\mathbb{I}(S_i=1,Z_i=0,X_i=x)}{\sum_{i=1}^{N}\mathbb{I}(Z_i=0,X_i=x)}$.
2. 我们如下估计 $\hat{\text{pr}}(G=\bar{s}\bar{s} \mid X=x)=\frac{\sum_{i=1}^{N}\mathbb{I}(S_i=0,Z_i=1,X_i=x)}{\sum_{i=1}^{N}\mathbb{I}(Z_i=1,X_i=x)}$.
3. 我们如下估计 $\hat{\text{pr}}(G=s\bar{s} \mid X=x)=1-\hat{\text{pr}}(G_i=ss \mid X_i=x)-\hat{\text{pr}}(G_i=\bar{s}\bar{s} \mid X_i=x)$.
4. 我们如下估计 $\hat{E}(Y \mid Z=0,S=1,X=x)=\frac{\sum_{i=1}^{N}\mathbb{I}(S_i=1,Z_i=0,X_i=x,Y_i)}{\sum_{i=1}^{N}\mathbb{I}(Z_i=0,S_i=1,X_i=x)}$.
5. 我们如下估计 $\mu_0=E(Y_0 \mid G=ss)$,

$$\hat{\mu}_0=\frac{\sum_{i=1}^{N}\hat{E}(Y \mid Z=0,S=1,X_i)\hat{\text{pr}}(G_i=ss \mid X_i)}{\sum_{i=1}^{N}\hat{\text{pr}}(G_i=ss \mid X_i)}.$$

6. 我们如下估计 $\hat{E}(Y \mid Z=1,S=1,X=x)=\frac{\sum_{i=1}^{N}\mathbb{I}(S_i=1,Z_i=1,X_i=x,Y_i)}{\sum_{i=1}^{N}\mathbb{I}(Z_i=1,S_i=1,X_i=x)}$.

# 非参数估计 II

7. 我们通过求解下述方程获得参数 $\hat{\beta}_0, \hat{\beta}_{ss}, \hat{\beta}_x$:

$$\hat{E}(Y \mid Z=1, S=1, X=0) = \beta_0 + \beta_{ss}\hat{\text{pr}}(G=ss \mid X=0) + 0\beta_x$$

$$\hat{E}(Y \mid Z=1, S=1, X=1) = \beta_0 + \beta_{ss}\hat{\text{pr}}(G=ss \mid X=1) + \beta_x,$$

$$\hat{E}(Y \mid Z=1, S=1, X=2) = \beta_0 + \beta_{ss}\hat{\text{pr}}(G=ss \mid X=2) + 2\beta_x.$$

8. 我们如下估计 $\hat{\mu}_1(X)$: $\hat{\mu}_1(X) = \beta_0 + \beta_{ss}\mathbb{I}(G=ss) + \beta_x X$.
9. 我们如下估计 $\mu_1 = E(Y_1 \mid G=ss)$,

$$\hat{\mu}_1 = \frac{\sum_{i=1}^N \hat{\mu}_1(X)\hat{\text{pr}}(G_i=ss \mid X_i)}{\sum_{i=1}^N \hat{\text{pr}}(G_i=ss \mid X_i)}.$$

# 目录

# Outline

1. 不依从问题
2. 死亡删失
3. 替代指标评估
    - 替代指标悖论
    - 替代指标评估

# 悖论

- 令 $Z$ 为二值的随机处理，1 表示接受一种新的处理，0 表示接受对照处理。
- $S$ 表示心律失常是否纠正，1 表示纠正了，0 表示未纠正。
- $S_z$ 表示接受处理 $Z=z$ 情形下是否纠正心律失常的潜在结果。
- $Y_{sz}$ 表示在处理 $Z=z$ 且心律失常纠正与否 $S=s$ 情形下的潜在生存时间。
- 假定处理 $Z$ 对生存时间 $Y$ 的作用完全通过中间变量 $S$ 起作用，对任意的 $s\in\{0,1\}$，即

$$Y_{1s}=Y_{0s}=Y_{S=s}.$$

# 悖论

| 主分层 | 人数 | $S_{Z=0}$ | $S_{Z=1}$ | $Y_{S=0}$ | $Y_{S=1}$ | $Y_{Z=0}$ | $Y_{Z=1}$ |
|---|---|---|---|---|---|---|---|
| 1 | 20 | 0 | 0 | 3 | 5 | 3 | 3 |
| 2 | 40 | 0 | 1 | 6 | 7 | 6 | 7 |
| 3 | 20 | 1 | 0 | 5 | 8 | 8 | 5 |
| 4 | 20 | 1 | 1 | 9 | 10 | 10 | 10 |

**表**: 100 位心律失常患者的总体

# 悖论 I

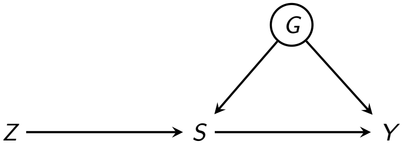



图: 工具变量图释

因此，当处理 $Z$ 对心律 $S$ 没有个体因果作用 $S_{Z=1}=S_{Z=0}=s$ 时，处理 $Z$ 对生存时间 $Y$ 也没有因果作用 $Y_{1s}=Y_{0s}$，此时排除约束假设成立，处理 $Z$ 对 $Y$ 没有直接作用，可以由图2所表示。令 $\mathrm{ACE}_{Z\to S}=\mathbb{E}\left(S_{Z=1}-S_{Z=0}\right)$ 以及 $\mathrm{ACE}_{Z\to Y}=E\left(Y_{Z=1}-Y_{Z=0}\right)$。在随机化假设下，

$$\mathrm{ACE}_{Z\to S}=\mathrm{pr}(G=s\bar{s})-\mathrm{pr}(G=\bar{s}s)=\frac{40-20}{100}=\frac{1}{5}>0.$$

# 悖论 II

我们关于主分层变量 $G$ 进行调整，可以算出 $Z$ 对 $Y$ 的因果作用

$$
\begin{aligned}
\mathrm{ACE}_{Z\to Y} &= \sum_g \mathbb{E}(Y_{Z=1} - Y_{Z=0} \mid G = g)\mathrm{pr}(G = g) \\
&= 0 \cdot \mathrm{pr}(G = \bar{s}\bar{s}) + 1 \cdot \mathrm{pr}(G = s\bar{s}) - 3 \cdot \mathrm{pr}(G = \bar{s}s) + 0 \cdot \mathrm{pr}(G = ss) \\
&= -\frac{1}{5}.
\end{aligned}
$$

我们关于主分层变量 $G$ 进行调整，可以算出 $S$ 对 $Y$ 的因果作用

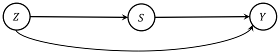
$$
\begin{aligned}
\mathrm{ACE}_{S\to Y} &= \sum_g \mathbb{E}(Y_{S=1} - Y_{S=0} \mid G = g)\mathrm{pr}(G = g) \\
&= 2 \cdot \mathrm{pr}(G = \bar{s}\bar{s}) + 1 \cdot \mathrm{pr}(G = s\bar{s}) + 3 \cdot \mathrm{pr}(G = \bar{s}s) + 1 \cdot \mathrm{pr}(G = ss) \\
&= \frac{8}{5}.
\end{aligned}
$$

# 悖论 III

- 上面的例子表明了处理 $Z$ 对纠正心律失常 $S$ 有正的因果作用
- 纠正心律失常 $S$ 对寿命 $Y$ 的因果作用也为正
- 但是处理 $Z$ 对寿命 $Y$ 的因果作用却为负。
- 这表明因果作用的统计结论不具有传递性，即变量 $Z$ 能提高变量 $S$, 变量 $S$ 能提高变量 $Y$, 但是根据这两个因果结论不能推出变量 $Z$ 能提高变量 $Y$。

# Outline

# 替代指标评估 I




| $S_1$ | $S_0$ | $G$ | 描述 | $E(Y_1 - Y_0 \mid G)$ |
|---|---|---|---|---|
| 1 | 1 | $ss$ | causal necessity | 0 |
| 1 | 0 | $s\bar{s}$ | causal sufficiency | $\neq 0$ |
| 0 | 1 | $\bar{s}s$ | causal sufficiency | $\neq 0$ |
| 0 | 0 | $\bar{s}\bar{s}$ | causal necessity | 0 |

- 处理 $Z = 1$ 表示接受处理，$Z = 0$ 表示接受对照；$S = 1$ 表示三年内未复发癌症，$S = 0$ 表示复发；$Y = 1$ 表示存活，$Y = 0$ 表示死亡。
- 文献 Frangakis & Rubin (2022) 指出一个好的替代指标应该满足“因果必要性”(causal necessity)
- “因果必要性”是指只要处理变量 $Z$ 对替代指标 $S$ 没有影响，那么 $Z$ 就应该对结果变量 $Y$ 也没有影响，也就是 $\text{ACE}_{ss} = 0$ 和 $\text{ACE}_{\bar{s}\bar{s}} = 0$，满足因果必要性的替代指标也常被称为主替代指标（principal surrogate）。

# 替代指标评估 II

- 但尽管如此，在更弱的因果必要性及单调性假设下，我们可以建立类似非依从框架下的识别结果（工具变量识别性），即

$$\begin{aligned}\mathsf{ACE}_{Z\to Y} &= \mathsf{ACE}_{Z\to S} \times \mathbb{E}(Y_{S=1} - Y_{S=0} \mid G = ss) \\ &= \mathsf{ACE}_{Z\to S} \times \mathbb{E}(Y_{Z=1} - Y_{Z=0} \mid G = ss),\end{aligned}$$

  其中第二个等式成立是因为在依从组上 $Z$ 和 $S$ 有相同的取值。

- 上式有助于我们利用已有的信息在新数据集下进行替代指标评估。

# 替代指标评估 III

**定理 1**

在 $(S_{Z=1}, S_{Z=0}, Y_{Z=0}, Y_{Z=1}) \perp\!\!\!\perp Z$，替代指标 $S$ 满足因果必要性，及 $\mathsf{ACE}_{s\bar{s}} > 0$ 时，我们假设替代指标 $S$ 满足因果必要性，即

(i) 单调性成立时，$\mathsf{ACE}_{Z\to Y} = \mathsf{ACE}_{Z\to S} \times \mathsf{ACE}_{s\bar{s}}$；

(ii) 单调性不成立时，假设 $\mathsf{ACE}_{Z\to S} > 0$，我们可以得到 $\mathsf{ACE}_{Z\to Y}$ 的上下界：如果 $\mathsf{ACE}_{s\bar{s}} + \mathsf{ACE}_{\bar{s}s} \geq 0$，则

$$\mathsf{ACE}_{Z\to S} \times \mathsf{ACE}_{s\bar{s}} \leq \mathsf{ACE}_{Z\to Y} \leq \frac{\mathsf{ACE}_{s\bar{s}} + \mathsf{ACE}_{\bar{s}} + \mathsf{ACE}_{Z\to S} \times (\mathsf{ACE}_{s\bar{s}} - \mathsf{ACE}_{\bar{s}})}{2}.$$

否则

$$\frac{\mathsf{ACE}_{s\bar{s}} + \mathsf{ACE}_{\bar{s}s} + \mathsf{ACE}_{Z\to S} \times (\mathsf{ACE}_{s\bar{s}} - \mathsf{ACE}_{\bar{s}s})}{2} \leq \mathsf{ACE}_{Z\to Y} \leq \mathsf{ACE}_{Z\to S} \times \mathsf{ACE}_{s\bar{s}}.$$

定理1给出了处理对替代指标的作用与处理对结果变量作用之间的关系，从而让我们可以直接用于另一个 $Y$ 未观测时的数据。

# 替代指标评估 IV

**定理 2**

在 $(S_{Z=1}, S_{Z=0}, Y_{Z=0}, Y_{Z=1}) \perp\!\!\!\perp Z$，替代指标 $S$ 满足因果必要性，及 $\text{ACE}_{s\bar{s}} > 0$ 时，

(i) 单调性2成立时，$\text{ACE}_{Z\to S} > 0$ (或者 $= 0$ ) 蕴含 $\text{ACE}_{Z\to Y} > 0$ (或者 $= 0$ );

(ii) 单调性2不成立时，$\text{ACE}_{s\bar{s}} + \text{ACE}_{\bar{s}s} \geq 0$ 和 $\text{ACE}_{Z\to S} > 0$ 蕴含 $\text{ACE}_{Z\to Y} > 0$.

- 在定理2(i) 中，单调性成立时，一个满足因果必要性和充分性的替代指标可以避免替代指标悖论。
- 没有单调性时，因为对 $Y$ 的作用在 $g = s\bar{s}$ 和 $\bar{s}s$ 层可能不一样，上面的结论 (i) 不再成立。

# 替代指标评估 V

- 所以，定理2(ii) 需要 $\mathsf{ACE}_{s\bar{s}} + \mathsf{ACE}_{\bar{s}s} \geq 0$，也就是说，在 $s\bar{s}$ 层中的正因果作用可以抵消 $\bar{s}s$ 层中负的因果作用。
- 因为不能观测到主分层变量 $G$，即使观测到终点指标 $Y$，这两个条件也不能用数据进行检验。
- 文献 Jiang et al. (2016) 利用辅助变量假设3建立了主分层因果作用 $\mathsf{ACE}_g$ 的非参数识别性结果，并利用定理2进行替代指标的评估。