



MIE: An approach to obtaining causal effect estimation in R

Bo Liu^{*}
Peking University

Zhixuan Shao^{*}
Peking University

Xiaohua Zhou[†]
Peking University

Abstract

Keywords: causal effect, survival, treatment, substitution variable, R.

1. Introduction

Researches regarding causal effect allure much attention of statisticians due to its various applications in scholarship as well as industry. In the field of biostatistics, one essential application is to determine the causal effect on outcomes a specific treatment may exert. However, the distinguishing characteristic of ineluctable deaths among the examinees may possibly invalidate the identifiability of the group we focus on, where patients will survive both with and without the treatment. A number of new methods have been proposed to eliminate the obstacle. One pertinent paper Wang, Zhou, and Richardson (2017) suggests the identifiability can be vindicated by introducing a substitution variable, and both the estimated causal effect and its asymptotic variance are provided. The R package (?) is designed in light of this method.

The paper is organized as follows. Section 2 retraces the model Wang *et al.* (2017) used to define SACE, and Section 3 demonstrates the theoretical deduction of how to estimate SACE. Section 4 introduces the elements in the package **mieSACE** in a detailed manner. In Section 5, a simulating dataset is generated and tested by the method, and an example on real data is presented in Section 6.

^{*}Equally contributed

[†]Corresponding Author

2. The Causal Effect Model

2.1. Survivor Average Causal Effect(SACE)

Consider a trial for a treatment with a single follow-up visit. Let Z be the exposure indicator to the treatment, i.e., $Z = 1$ means the patient receives the treatment and $Z = 0$ means not. Each patient may have two potential survival statuses $S(1)$ and $S(0)$, defined as whether the patient would be alive at the time of follow-up visit had the subject been exposed or not respectively. [Hernán \(2004\)](#) has defined the causal effect to be $\mathbb{E}(S(1)) - \mathbb{E}(S(0))$, which indicates the difference in the probability of survival between subjects with and without treatment. [Rubin \(2006\)](#) represents with two letters L and D for 4 potential principle strata of the subjects, where L represents live and D stands for die. For example, if a person will live if treated but will die if not, he belongs to the stratum of LD. [Rubin \(2006\)](#) extends the causal effect from survival status to the outcomes observed. Similarly, we define $Y(1)$ and $Y(0)$ as the potential outcomes had the subject been exposed or not, respectively. However, these outcomes can only be observed if the subjects are still alive. Therefore, the causal effect of treatment on outcome is defined as

$$\Delta_{LL} = \mathbb{E}\{Y(1) - Y(0)|G = LL\}. \quad (1)$$

This is also known as the survivor average causal effect (SACE).

2.2. Assumptions of the Model

In real experiments, it is impossible to divide the subjects in the four strata *a priori*. Those alive and treated belong to strata LL or LD, while those alive but not treated belong to LL or DL. The specific stratum LL is unidentifiable.

[Egleston, Scharfstein, Freeman, and West \(2007\)](#) assumes that

$$S(1) \geq S(0) \text{ a.s.} \quad (2)$$

Let W be the covariates of subjects, and [Rosenbaum and Rubin \(1983\)](#) referred as "strong ignorability of treatment assignment" to the assumption

$$Z \perp\!\!\!\perp S(z)|W \text{ for } z = 0, 1. \quad (3)$$

In light of the "missing at random" assumption ([Rubin \(1976\)](#)), we have

$$Z \perp\!\!\!\perp Y(z)|W, G = LL \text{ for } z = 0, 1. \quad (4)$$

Let baseline covariates $W = (X, A)$ where X is similar to a confounder while A is a substitution variable, which satisfies the substitution and exclusion assumption ([Stephen Burgess \(2016\)](#))

$$\begin{aligned} & A \perp\!\!\!\perp Y(1)|Z = 1, G, X, \\ & {}^n A \not\perp\!\!\!\perp G|Z = 1, S = 1, X, \end{aligned} \quad (5)$$

and the SACE (Eqn.1) is identifiable with assumptions Eqn.2-5 on the variables ([Wang et al. \(2017\)](#)).

2.3. Identification of SACE

We illustrate hereafter the main points in the proof by Wang *et al.* (2017).

$$\mathbb{E}(Y(z)|G = \text{LL}) = \frac{\mathbb{E}_w(\mu_{z,\text{LL},w} \cdot \pi_{\text{LL}|w})}{\mathbb{E}_w(\pi_{\text{LL}|w})}, \quad (6)$$

where $\mu_{z,g,w} = \mathbb{E}(Y(z)|G = g, W = w)$ and $\pi_{g|w} = \mathbb{P}(G = g|W = w)$.

Under Eqn.4,

$$\mu_{z,g,w} = \mathbb{E}(Y(z)|Z = z, G = g, W = w) = \mathbb{E}(Y|Z = z, G = g, W = w). \quad (7)$$

Under Eqn.2 and Eqn.3,

$$\pi_{\text{LL}|w} = \mathbb{P}(S(0) = 1, S(1) = 1|W = w) = \mathbb{P}(S(0) = 1|W = w) = \mathbb{P}(S = 1|Z = 0, W = w). \quad (8)$$

$$\pi_{\text{LD}|w} = \mathbb{P}(S(0) = 0, S(1) = 1|W = w) = \mathbb{P}(S = 1|Z = 1, W = w) - \mathbb{P}(S = 1|Z = 0, W = w). \quad (9)$$

From Eqn.6,

$$\mu_{0,\text{LL},w} = \mathbb{E}(Y|Z = 0, S = 1, W = w). \quad (10)$$

Since Y, Z, S and W are all observable from the data, $\pi_{\text{LL}|w}$, $\pi_{\text{LD}|w}$ and $\mu_{0,\text{LL},w}$ are identifiable. Denote W in the form of (X, A) , and define

$$p_{g|z,x,a,s} = \frac{\mathbb{P}(G = g|X = x, A = a)}{\mathbb{P}(S = s|Z = z, X = x, A = a)}. \quad (11)$$

It can be proven that under Eqn.5,

$$\mathbb{E}(Y|Z = 1, S = 1, X = x, A = a) = p_{\text{LL}|1,x,a,1}\mu_{1,\text{LL},x} + (1 - p_{\text{LL}|1,x,a,1})\mu_{1,\text{LD},x}. \quad (12)$$

For $\forall x$, there exist different a_1 and a_2 satisfying Eqn.12 (under Eqn.5), and thus $\mu_{1,\text{LL},x}$ can be identified.

2.4. Model Parameterization

For simplicity, we use (generalized) linear model as our assumption.

$$\mu_{0,\text{LL},W} = \mathbb{E}(Y|Z = 0, S = 1, X, A) = \alpha_{10} + X\alpha_{11} + A\alpha_{12}. \quad (13)$$

$$\mu_{1,\text{L},W} = \mathbb{E}(Y|Z = 1, S(0) = L, S(1) = 1, X, A) = \alpha_{20} + X\alpha_{21} + L\alpha_{22}. \quad (14)$$

In Eqn.14, A is dropped because under Eqn.5, $A \perp\!\!\!\perp Y(1)|Z = 1, S(0) = L, S(1) = 1, X$. L is not observable when $S(1) = 1$, so we need to figure out later how to estimate L .

$$\mathbb{P}(S = 1|Z = 1, W) = \text{expit}(\beta_0 + X\beta_1 + A\beta_2). \quad (15)$$

$$\frac{\mathbb{P}(S = 1|Z = 0, W)}{\mathbb{P}(S = 1|Z = 1, W)} = \text{expit}(\gamma_0 + X\gamma_1 + A\gamma_2). \quad (16)$$

One can verify that these assumptions on the parameterization satisfy the requirements of the model.

3. SACE Estimation

3.1. Data

The data input should include multiple observations on random variables Z, S, Y, X and A . Suppose the i -th observation can be denoted by $(Z^i, S^i, Y^i, X^i, A^i)$, and they denote respectively:

- Z^i : an integer valued 0 or 1, exposure indicator of the i -th subject;
- S^i : an integer valued 0 or 1, survival indicator of the i -th subject;
- Y^i : a numeric vector, indicating the survival outcome of the i -th subject;
 - Y^i is discarded where $S^i = 0$.
 - A warning will be raised if Y^i is NA where $S^i = 1$.
- X^i : a numeric row vector, the confounder of covariate W of the i -th subject.
- A^i : a numeric row vector, the substitution variable of covariate W of the i -th subject.

Accordingly, the data type of observed data matrix $\mathbf{Z}, \mathbf{S}, \mathbf{Y}, \mathbf{X}$ and \mathbf{A} are defined as Table 3.2.

| Variable | Data Type | Size |
|--------------|-----------------------|--------------|
| \mathbf{Z} | Logical Column Vector | n |
| \mathbf{S} | Logical Column Vector | n |
| \mathbf{Y} | Numeric Matrix | $n \times k$ |
| \mathbf{X} | Numeric Matrix | $n \times p$ |
| \mathbf{A} | Numeric Matrix | $n \times q$ |

Table 1: Data Types of Observed Variables

3.2. Parameters

In Sec. 2.4, we put forward four models with parameters $\alpha_0, \alpha_1, \beta$ and γ . Here we put it in the matrix form.

Let

$$\alpha_1 = \begin{pmatrix} \alpha_{10} \\ \alpha_{11} \\ \alpha_{12} \end{pmatrix}, \alpha_2 = \begin{pmatrix} \alpha_{20} \\ \alpha_{21} \\ \alpha_{22} \end{pmatrix}, \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix}, \gamma = \begin{pmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \end{pmatrix}, \quad (17)$$

be the parameters, and let

$$\begin{aligned}\mathbf{W} &= (\mathbf{X} \ \mathbf{A}), \\ \widetilde{\mathbf{W}} &= (\mathbf{1}_n \ \mathbf{W}), \\ \widetilde{\mathbf{X}} &= (\mathbf{1}_n \ \mathbf{X}).\end{aligned}\tag{18}$$

be the observed data matrix of random variables W , \widetilde{W} and \widetilde{X} respectively, then our parametric model can be rewritten as

$$\begin{aligned}\mathbb{E}(Y|Z = 0, S = 1, X, A) &= \widetilde{W}\alpha_1, \\ \mathbb{E}(Y|Z = 1, S(0) = L, S(1) = 1, X, A) &= (\widetilde{X} \ L)\alpha_2, \\ \mathbb{P}(S = 1|Z = 1, W) &= \text{expit}(\widetilde{W}\beta), \\ \frac{\mathbb{P}(S = 1|Z = 0, W)}{\mathbb{P}(S = 1|Z = 1, W)} &= \text{expit}(\widetilde{W}\gamma).\end{aligned}\tag{19}$$

| Parameter | Data Type | Size |
|------------|----------------|------------------------|
| α_1 | Numeric Matrix | $(p + q + 1) \times k$ |
| α_2 | Numeric Matrix | $(p + 2) \times k$ |
| β | Numeric Vector | $p + q + 1$ |
| γ | Numeric Vector | $p + q + 1$ |

Table 2: Data Types of Parameters

3.3. Estimating the Parameters

Estimating α_1

Since

$$\mathbb{E}(Y|Z = 0, S = 1, X, A) = \widetilde{W}\alpha_1,$$

we apply the ordinary least square method (OLS) to estimate α_1 . Let \mathbf{Y}_{01} and $\widetilde{\mathbf{W}}_{01}$ be the observed data of Y and \widetilde{W} in the subset where $Z = 0$ and $S = 1$, and the estimator of α_1 can be

$$\hat{\alpha}_1 = (\widetilde{\mathbf{W}}_{01}^T \widetilde{\mathbf{W}}_{01})^{-1} \widetilde{\mathbf{W}}_{01}^T \mathbf{Y}_{01}.\tag{20}$$

Estimating α_2

It may be natural to apply here the same method how we estimate α_1 , but it can be rather tricky because L is not observable. By doing simple algebra, one can avoid the problem by

taking expectation over L .

$$\begin{aligned}
\mathbb{E}(Y|Z=1, S(1)=1, X, A) &= \mathbb{E}(Y|Z=1, S(0)=0, S(1)=1, X, A)\mathbb{P}(S(0)=0|Z=1, S(1)=1, W) \\
&\quad + \mathbb{E}(Y|Z=1, S(0)=1, S(1)=1, X, A)\mathbb{P}(S(0)=1|Z=1, S(1)=1, W) \\
&= (\tilde{X}-0)\alpha_2\mathbb{P}(S(0)=0|Z=1, S(1)=1, W) \\
&\quad + (\tilde{X}-1)\alpha_2\mathbb{P}(S(0)=1|Z=1, S(1)=1, W) \\
&= (\tilde{X}-\mathbb{P}(S(0)=1|Z=1, S(1)=1, W)) \\
&= \left(\tilde{X} - \frac{\mathbb{P}(S=1|Z=0, W)}{\mathbb{P}(S=1|Z=1, W)}\right) \\
&= \left(\tilde{X} - \text{expit}(\tilde{W}\gamma)\right).
\end{aligned} \tag{21}$$

Let \mathbf{Y}_{11} , $\tilde{\mathbf{X}}_{11}$ and $\tilde{\mathbf{W}}_{11}$ be the observed data of Y , \tilde{X} and \tilde{W} in the subset where $Z=1$ and $S=1$, $\hat{\gamma}$ be the estimator of γ . Denote $\tilde{\mathbf{X}}\tilde{\mathbf{L}}_{11} = \left(\tilde{\mathbf{X}}_{11} - \text{expit}(\tilde{\mathbf{W}}_{11}\hat{\gamma})\right)$, and the estimator of α_2 can be

$$\hat{\alpha}_2 = (\tilde{\mathbf{X}}\tilde{\mathbf{L}}_{11}^T \tilde{\mathbf{X}}\tilde{\mathbf{L}}_{11})^{-1} \tilde{\mathbf{X}}\tilde{\mathbf{L}}_{11}^T \mathbf{Y}_{11}. \tag{22}$$

Estimating β and γ

Note that the model involving β and γ is equivalent to

$$\begin{aligned}
S|Z=1, W &\sim \mathcal{B}(1, \text{expit}(\tilde{W}\beta)), \\
S|Z=0, W &\sim \mathcal{B}(1, \text{expit}(\tilde{W}\beta)\text{expit}(\tilde{W}\gamma)).
\end{aligned} \tag{23}$$

Define

$$\begin{aligned}
L(\beta, \gamma|z, s, w) &= \mathbb{P}(S=s|\beta, \gamma, Z=z, W=w) \\
&= \begin{cases} \text{expit}(\tilde{w}\beta), & z=1, s=1 \\ 1 - \text{expit}(\tilde{w}\beta), & z=1, s=0 \\ \text{expit}(\tilde{w}\beta)\text{expit}(\tilde{w}\gamma), & z=0, s=1 \\ 1 - \text{expit}(\tilde{w}\beta)\text{expit}(\tilde{w}\gamma), & z=0, s=0 \end{cases}
\end{aligned} \tag{24}$$

and

$$\begin{aligned}
D_\beta L(\beta, \gamma|z, s, w) &= \frac{\partial}{\partial \beta} \ln L(\beta, \gamma|z, s, w) \\
&= \begin{cases} (1 - \text{expit}(\tilde{w}\beta))\tilde{w}, & z=1, s=1 \\ -\text{expit}(\tilde{w}\beta)\tilde{w}, & z=1, s=0 \\ (1 - \text{expit}(\tilde{w}\beta))\tilde{w}, & z=0, s=1 \\ -\frac{(1 - \text{expit}(\tilde{w}\beta))}{\text{expit}^{-1}(\tilde{w}\beta)\text{expit}^{-1}(\tilde{w}\gamma) - 1}\tilde{w}, & z=0, s=0 \end{cases}
\end{aligned} \tag{25}$$

$$\begin{aligned}
D_\gamma L(\beta, \gamma|z, s, w) &= \frac{\partial}{\partial \gamma} \ln L(\beta, \gamma|z, s, w) \\
&= \begin{cases} 0, & z = 1 \\ (1 - \text{expit}(\tilde{w}\gamma))\tilde{w}, & z = 0, s = 1 \\ -\frac{(1 - \text{expit}(\tilde{w}\gamma))}{\text{expit}^{-1}(\tilde{w}\beta)\text{expit}^{-1}(\tilde{w}\gamma) - 1}\tilde{w}. & z = 0, s = 0 \end{cases} \quad (26)
\end{aligned}$$

Using maximum likelihood estimation,

$$(\hat{\beta}, \hat{\gamma}) = \arg \max \sum_{i=1}^n \ln L(\beta, \gamma|z^i, s^i, w^i). \quad (27)$$

3.4. Estimating SACE

With regards to Eqn.6, SACE is defined as

$$\begin{aligned}
\Delta_{LL} &= \frac{\mathbb{E}_w(\mu_{1,LL,w} \cdot \pi_{LL|w})}{\mathbb{E}(\pi_{LL|w})} - \frac{\mathbb{E}_w(\mu_{0,LL,w} \cdot \pi_{LL|w})}{\mathbb{E}(\pi_{LL|w})} \\
&\triangleq \mu_{1,LL} - \mu_{0,LL}. \quad (28)
\end{aligned}$$

To estimate Δ_{LL} , we use the empirical distribution of W , and

$$\hat{\mu}_{\cdot,LL} = \frac{\sum_{i=1}^n (\hat{\mu}_{\cdot,LL,w_i} \cdot \hat{\pi}_{LL|w_i})}{\sum_{i=1}^n \hat{\pi}_{LL|w_i}}. \quad (29)$$

For simplicity, we define $\text{unit}(\alpha)$ to be the vector parallel to α which has all its elements summed up to 1, and the estimation of $\mu_{\cdot,LL}$ has a rather simple expression

$$\hat{\mu}_{\cdot,LL} = \sum_{i=1}^n \hat{\mu}_{\cdot,LL,w_i} \cdot \text{unit}(\hat{\pi}_{LL|w_i}). \quad (30)$$

Let $\hat{\mu}_{\cdot,LL,\mathbf{W}}$ be a matrix whose i -th row equals $\hat{\mu}_{\cdot,LL,w_i}$, and $\hat{\pi}_{LL|\mathbf{W}}$ be a column vector composed of $\hat{\pi}_{LL|w_i}$.

$$\hat{\mu}_{\cdot,LL} = \text{unit}(\hat{\pi}_{LL|\mathbf{W}}^T) \cdot \hat{\mu}_{\cdot,LL,\mathbf{W}}. \quad (31)$$

Using the notations forementioned,

$$\hat{\pi}_{LL|\mathbf{W}} = \text{expit}(\tilde{\mathbf{W}}\hat{\beta})\text{expit}(\tilde{\mathbf{W}}\hat{\gamma}), \quad (32)$$

where the multiplication should be operated element-wise.

$$\hat{\mu}_{0,LL|\mathbf{W}} = \tilde{\mathbf{W}}\hat{\alpha}_1. \quad (33)$$

$$\hat{\mu}_{1,LL|\mathbf{W}} = (\tilde{\mathbf{X}} \quad \mathbf{1}_n)\hat{\alpha}_2. \quad (34)$$

Thus we are able to estimate the SACE

$$\hat{\Delta}_{LL} = \hat{\mu}_{1,LL} - \hat{\mu}_{0,LL}. \quad (35)$$

4. SACE Analysis

The R package of **mieSACE** is written in an pseudo-object-oriented form in order that it should be easy to use and at the same time, efficient. This is realized by the function **SACE** which returns a list of public functions.

To establish a SACE object, one need to provide some or all of the 6 parameters shown in Table 4.

| Parameter | Data Type | Nullable | Description |
|-----------|----------------|----------|--|
| Z | Logical Vector | No | Was the subject exposed to treatment? |
| S | Logical Vector | No | Did the subject survive to observation? |
| Y | Numeric Matrix | No | The outcomes on observation |
| X | Numeric Matrix | Yes | The confounders of the subject at baseline |
| A | Numeric Matrix | Yes | The substitution variable satisfying Eqn.5 |
| subset | Logical Vector | Yes | Indicating the subjects to be included |

Table 3: Parameters of **SACE**

The function **SACE** returns a list of 20 functions for users to manipulate the data and obtain results in a simple and safe way (Table 4).

| Type | Function Name | Description |
|------------------|--------------------------|--|
| Get | Get* ¹ | Return private.* as read-only |
| | GetSubset | Return private.subset as read-only |
| | Get*Subset | Return a subset of private.* as read-only |
| Set ² | Set* | Set private.* |
| | SetSubset | Set private.subset |
| | UnsetSubset | Set private.subset to NULL |
| MIE | | Return a list indicating the result of SACE estimation |
| ConfInt | | Return the confident interval of estimated SACE |

¹ * can be Z, S, Y, X or A.

² All the Set functions are implemented with certain data type checks.

Table 4: Returned functions of **SACE**

It is simple to establish a SACE object using **SACE**. To illustrate it, we first generate random data as our input parameters, namely Z, S, Y, X, A, **subset**.

```
R> NUM_SUBJECTS <- 60
R> Z <- rbinom(NUM_SUBJECTS, 1, 0.7)
R> S <- rbinom(NUM_SUBJECTS, 1, 0.6)
R> Y <- matrix(rnorm(NUM_SUBJECTS * 3, 0, 1), nrow = NUM_SUBJECTS)
R> X <- matrix(rnorm(NUM_SUBJECTS * 5, 0, 1), nrow = NUM_SUBJECTS)
R> A <- rnorm(NUM_SUBJECTS, 0, 1)
```

And we establish a corresponding SACE object by


```
R> sace.obj <- SACE(Z, S, Y, X, A)
```

4.1. How is the data stored?

In the function `SACE`, some variables are defined with names like `private.Z`, and we actually store the parameter `Z` there on call of `SACE`. In order to be support read and write operations, these variables are indeed defined in the outer environment of the function (which may be directly manipulate upon outside the function). The modifying word *private* should always recommend the users visit these variables through provided functions.

Assume that we have to increase `Y` by 1 due to some scaling concerns, we can update `sace.obj` instead of creating a new object.

```
R> sace.obj$SetY(Y + 1)
```

To limit the calculation on only the first 5 subjects, we can set the subset by

```
R> sace.obj$SetSubset(1:NUM_SUBJECTS <= 5)
```

and view their outcomes.

```
R> sace.obj$GetYSubset()
```

| | [,1] | [,2] | [,3] |
|------|-----------|-----------|-----------|
| [1,] | 0.7733092 | 2.2316216 | 0.8320438 |
| [2,] | 0.9861842 | 0.2492520 | 0.6805767 |
| [3,] | 0.8744999 | 1.2124410 | 1.0647277 |
| [4,] | 2.0399051 | 0.7039607 | 0.3285627 |
| [5,] | 1.5003603 | 1.0102448 | 2.8764985 |

4.2. Data Type and Restrictions Check

Our model includes numbers of derived variables from given input, such as $\tilde{W} = (\mathbf{1}_n \ X \ A)$ as mentioned before. If we design as a function of X and A the process of calculating \tilde{W} , we may have to do the computation again each time we need \tilde{W} , which could be rather time-consuming. If we otherwise, define \tilde{W} as variable, there may be problems when the value of \tilde{W} does not coincide with that of $(\mathbf{1}_n \ X \ A)$, which probably happens when X is changed without updating \tilde{W} .

We managed to solve the problem with inspiration from the concept of cache, a commonly used architecture in computer science. A boolean variable `private.checked` is designed to indicate whether some "caches" (all with "check" in their names) are still valid. Once the original data has changed, `private.checked` is set `FALSE` indicating the caches may be implausible. When a function requires the the value of a particular cached variable, it first checks `private.checked`, and will only redo the calculation when `private.checked` is `FALSE`. If the calculation succeeds, the caches should be updated, and `private.checked` should be reset `TRUE`.

The caches are updated in function `private.check`, with certain data type and restriction checks previously. These checks includes size checks(whether the matrices have proper size for addition or multiplication), NA checks(whether a matrix can have NA values inside) and some restriction checks.

4.3. Estimating SACE and its confident interval

One can obtain the estimated SACE through function `MIE`, which is quite simple to use.

```
R> mie <- sace.obj$MIE()
```

The return value is a list of 6 elements among which the most important two are `sace` and `estimate`.

```
R> mie$sace
```

```
[1] 0.3774153 -0.4699400 0.0484997
```

```
R> mie$estimate
```

```
$alpha1.estimate
```

| | [,1] | [,2] | [,3] |
|------|------------|------------|-------------|
| [1,] | 0.6859891 | 1.0895737 | 0.82491572 |
| [2,] | 0.3610506 | -0.2503188 | 0.18028269 |
| [3,] | 0.1270728 | -0.1378440 | -0.41900916 |
| [4,] | -0.3806892 | 0.2237894 | 0.36467945 |
| [5,] | -0.1758697 | 0.5373655 | -0.02100811 |
| [6,] | -0.6007420 | -0.1115998 | 0.34522913 |
| [7,] | -0.5808966 | 0.3598921 | 0.27266321 |

```
$alpha2.estimate
```

| | [,1] | [,2] | [,3] |
|------|-------------|-------------|------------|
| [1,] | -1.57489345 | 1.26644065 | 1.3354401 |
| [2,] | 0.07001789 | 0.05115594 | -0.2451194 |
| [3,] | -0.01509677 | 0.12985987 | -0.0248238 |
| [4,] | -0.26250149 | 0.12282619 | 0.0194298 |
| [5,] | 0.23341392 | -0.14980854 | -0.2675334 |
| [6,] | 0.35501926 | 0.20347317 | 0.1334751 |
| [7,] | 2.77466962 | -0.54183713 | -0.3390038 |

```
$beta.estimate
```

```
[1] 0.9021380 0.2096719 -0.2572589 -0.1038979 0.2279089 -0.5438395 -0.5594817
```

```
$gamma.estimate
```

```
[1] 181.30895 37.94398 -48.07846 125.71451 -54.83176 -44.57635 41.90218
```

```
$beta_gamma.convergence
```

```
[1] 0

$mu0.estimate
[1] 0.8081135 1.1201891 0.9097668

$mu1.estimate
[1] 1.1855288 0.6502491 0.9582665
```

We added the return list class attribute "mie", and implemented a function `print.mie`.

```
R> print(mie)
```

Call:

```
private.model1.mie(thres = thres, optim.method = optim.method,
  max.step = max.step, singular.ok = singular.ok, need.variance = need.variance)
```

```
sample size: 60
```

```
average potential outcomes among control group: 0.8081135 1.120189 0.9097668
```

```
average potential outcomes among treatment group: 1.185529 0.6502491 0.9582665
```

```
SACE (survivor average causal effect): 0.3774153 -0.46994 0.0484997
```

Currently, the bootstrap method is implemented to compute confident intervals of the estimators.

```
R> confint <- sace.obj$ConfInt(alpha = 0.05, print.progress = F)
R> confint$ci
```

```
      [,1]      [,2]      [,3]
2.5% -1.874308 -2.525785 -3.173551
97.5%  2.430036  1.902488  2.672909
```

5. Simulation

We use **mieSACE** to repeat the simulation Wang *et al.* (2017) did. Suppose $X = (X_1, X_2, X_3)$ where X_1 takes value 1 and -1 with equal possibility and (X_2, X_3) follows a multivariate normal distribution $\mathcal{N}(\mu, \Sigma)$ in which

$$\mu = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \Sigma = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}.$$

A follows a Bernoulli distribution with $\mathbb{P}(A = 1|X) = \text{expit}(uX)$, where $u = (1, 1, 1)/2$. Z is generated following a logistic model $\mathbb{P}(Z = 1|X, A) = \text{expit}(uX + A)$. The survival type G satisfies that $\mathbb{P}(G = \text{LL}|X, A) = \text{expit}(2 + \beta X + A)\text{expit}(\gamma X + A)$, $\mathbb{P}(G = \text{LD}|X, A) = \text{expit}(2 + \beta X + A)(1 - \text{expit}(\gamma X + A))$, and $\mathbb{P}(G = \text{DD}|X, A) = 1 - \text{expit}(2 + \beta X + A)$, where $\beta =$

$(0.5, 0.5, 0.5)$ and $\gamma = (-1.5, 0.5, 0.5)$. The potential outcomes Y is generated with conditional distribution $Y(1)|G = LL, X, A \sim \mathcal{N}(uX, 0.5^2)$, $Y(0)|G = LL, X, A \sim \mathcal{N}(-1 + uX, 0.5^2)$, $Y(1)|G = LD, X, A \sim \mathcal{N}(1 + uX, 0.5^2)$. The true value for SACE is 1.

5.1. Generating Data

```
R> expit <- function(x) { exp(x) / (1 + exp(x)) }
R> size <- 5000
R> X1 <- sample(c(1,-1),size,prob=c(0.5,0.5),replace = T)
R> mu <- c(1,-1)
R> sigma <- matrix(c(1,0.5,0.5,1),2,2)
R> X23 <- mvrnorm(size, mu, sigma)
R> X <- cbind(X1,X23)
R> A <- rbinom(size, 1, expit(X %*% c(0.5,0.5,0.5)))
R> Z <- rbinom(size, 1, expit(X %*% c(0.5,0.5,0.5) + 1))
R> S1 <- rbinom(size, 1, expit(X %*% c(0.5,0.5,0.5) + A + 2))
R> S0 <- rbinom(size, 1, expit(X %*% c(-1.5,0.5,0.5) + A))
R> S0[S1 == 0] <- 0
R> S <- S1 * Z + S0 * (1 - Z)
R> Y1LL <- rnorm(size, X %*% c(0.5,0.5,0.5), 0.5^2)
R> Y0LL <- rnorm(size, X %*% c(0.5,0.5,0.5) - 1, 0.5^2)
R> Y1LD <- rnorm(size, X %*% c(0.5,0.5,0.5) + 1, 0.5^2)
R> Y1 <- S0 * Y1LL + (1 - S0) * Y1LD
R> Y0 <- Y0LL
R> Y <- Y1 * Z + Y0 * (1 - Z)
R> Y[S == 0] <- NA
```

5.2. Calculating SACE and its 95% confident interval

```
R> sace <- SACE(Z,S,Y,X,A)
R> print(sace$MIE())
```

Call:

```
private.model1.mie(thres = thres, optim.method = optim.method,
  max.step = max.step, singular.ok = singular.ok, need.variance = need.variance)
```

```
sample size: 5000
average potential outcomes among control group: -0.8351665
average potential outcomes among treatment group: 0.1825276
SACE (survivor average causal effect): 1.017694
```

```
R> print(sace$ConfInt(nboot=500,print.progress = F)$ci)
```

```
      [,1]
2.5%  0.913946
97.5% 1.108069
```

6. Example: SACE of ISIS 396443 on SMA patients

We use **mieSACE** on data collected from a study of ISIS 396443 in patients with infantile-onset SMA. Spinal muscular atrophy (SMA) is a rare neuromuscular disorder characterised by loss of lower motor neurons and progressive muscle wasting, often leading to early death ([Wikipedia contributors \(2018\)](#)). The experiment involved 121 subjects with 80 treated with ISIS 396443, and the number of serious respiratory events were measured upon these subjects both at screening and 394 days afterward. The change of the number is the outcome we concern, which indicates in what extent the patients deteriorated during the period. We choose some other variables at screen time as confounders, including the age at screening, the age at first dose, sex, age at symptom onset, indicator that disease duration is longer than 12 weeks, symptoms and baseline hours of ventilator support. The number of serious respiratory events at screening is selected as the substitution variable, because it is believed to affect survival status directly but impact the outcome only indirectly. The estimated SACE is 7.406508 with a 95% confident interval of $(-0.8260246, 16.8671560)$.

7. Summary and discussion

■ As usual ...

Computational details

■ If necessary or useful, information about certain computational details such as version numbers, operating systems, or compilers could be included in an unnumbered section. Also, auxiliary packages (say, for visualizations, maps, tables, ...) that are not cited in the main text can be credited here.

The results in this paper were obtained using R 3.4.1 with the **MASS** 7.3.47 package. R itself and all packages used are available from the Comprehensive R Archive Network (CRAN) at <https://CRAN.R-project.org/>.

Acknowledgments

■ All acknowledgments (note the AE spelling) should be collected in this unnumbered section before the references. It may contain the usual information about funding and feedback from colleagues/reviewers/etc. Furthermore, information such as relative contributions of the authors may be added here (if any).

References

Egleston BL, Scharfstein DO, Freeman EE, West SK (2007). “Causal inference for non-mortality outcomes in the presence of death.” *Biostatistics*, 8(3), 526–545. doi:10.1093/

- biostatistics/kxl027. /oup/backfile/content_public/journal/biostatistics/8/3/10.1093_biostatistics_kxl027/1/kxl027.pdf, URL <http://dx.doi.org/10.1093/biostatistics/kxl027>.
- Hernán MA (2004). “A definition of causal effect for epidemiological research.” *Journal of Epidemiology & Community Health*, **58**(4), 265–271. ISSN 0143-005X. doi:10.1136/jech.2002.006361. <https://jech.bmj.com/content/58/4/265.full.pdf>, URL <https://jech.bmj.com/content/58/4/265>.
- Rosenbaum PR, Rubin DB (1983). “The central role of the propensity score in observational studies for causal effects.” *Biometrika*, **70**(1), 41–55. doi:10.1093/biomet/70.1.41. /oup/backfile/content_public/journal/biomet/70/1/10.1093/biomet/70.1.41/2/70-1-41.pdf, URL <http://dx.doi.org/10.1093/biomet/70.1.41>.
- Rubin DB (1976). “Inference and missing data.” *Biometrika*, **63**(3), 581–592. doi:10.1093/biomet/63.3.581. /oup/backfile/content_public/journal/biomet/63/3/10.1093/biomet/63.3.581/2/63-3-581.pdf, URL <http://dx.doi.org/10.1093/biomet/63.3.581>.
- Rubin DB (2006). “Causal Inference Through Potential Outcomes and Principal Stratification: Application to Studies with Censoring Due to Death.” *Statist. Sci.*, **21**(3), 299–309. doi:10.1214/088342306000000114. URL <https://doi.org/10.1214/088342306000000114>.
- Stephen Burgess DSS (2016). “Predicting the Direction of Causal Effect Based on an Instrumental Variable Analysis: A Cautionary Tale.” *Journal of Causal Inference*, **4**(1), 49–59. doi:10.1515/jci-2015-0024. URL <https://doi.org/10.1515/jci-2015-0024>.
- Wang L, Zhou XH, Richardson TS (2017). “Identification and estimation of causal effects with outcomes truncated by death.” *Biometrika*, **104**(3), 597–612.
- Wikipedia contributors (2018). “Spinal muscular atrophy — Wikipedia, The Free Encyclopedia.” https://en.wikipedia.org/w/index.php?title=Spinal_muscular_atrophy&oldid=868791733. [Online; accessed 23-November-2018].

A. More technical details

Appendices can be included after the bibliography (with a page break). Each section within the appendix should have a proper section title (rather than just *Appendix*).

For more technical style details, please check out JSS's style FAQ at <https://www.jstatsoft.org/pages/view/style#frequently-asked-questions> which includes the following topics:

- Title vs. sentence case.
- Graphics formatting.
- Naming conventions.
- Turning JSS manuscripts into R package vignettes.
- Trouble shooting.
- Many other potentially helpful details...

B. Using BibT_EX

References need to be provided in a BibT_EX file (`.bib`). All references should be made with `\cite`, `\citet`, `\citep`, `\citealp` etc. (and never hard-coded). These commands yield different formats of author-year citations and allow to include additional details (e.g., pages, chapters, ...) in brackets. In case you are not familiar with these commands see the JSS style FAQ for details.

Cleaning up BibT_EX files is a somewhat tedious task – especially when acquiring the entries automatically from mixed online sources. However, it is important that informations are complete and presented in a consistent style to avoid confusions. JSS requires the following format.

- JSS-specific markup (`\proglang`, `\pkg`, `\code`) should be used in the references.
- Titles should be in title case.
- Journal titles should not be abbreviated and in title case.
- DOIs should be included where available.
- Software should be properly cited as well. For R packages `citation("pkgname")` typically provides a good starting point.

Affiliation:

Xiaohua Zhou

Professor

Beijing International Center for Mathematical Research

Peking University

5 Yiheyuanlu Road

100871 Beijing, China

E-mail: azhou@bicmr.pku.edu.cn

URL: <http://faculty.washington.edu/azhou>