Contents

1 Brief description of the causal mediation

The literature on the causal mediation is vast [VanderWeele, 2015] and is evolving, thus, only the pieces relevant for the current software are reviewed here.

1.1 Decomposition of total effect

Let Y be the outcome variable of interest, A be the treatment variable of interest, M be the mediator variable of interest, and $\mathbf{C}$ be the potentially vector-valued pre-treatment baseline covariates necessary for exchangeability. The treatment contrast of interest is a_1 vs a_0, the second being the reference level. The counterfactual $Y_{a,m}$ is the value of Y for an individual when, possibly contrary to the fact, the treatment level a and mediator level m are assigned.

Given these notations, the effects are defined as follows at the covariate level $\mathbf{C} = \mathbf{c}$ and the mediator level $M = m$ (only for CDE).

$$CDE(m) = E[Y_{a_1,m}|\mathbf{C} = \mathbf{c}] - E[Y_{a_0,m}|\mathbf{C} = \mathbf{c}]$$

$$\begin{aligned} PNDE &= E[Y_{a_1,M_{a_0}}|\mathbf{C} = \mathbf{c}] - E[Y_{a_0,M_{a_0}}|\mathbf{C} = \mathbf{c}] \\ TNIE &= E[Y_{a_1,M_{a_1}}|\mathbf{C} = \mathbf{c}] - E[Y_{a_1,M_{a_0}}|\mathbf{C} = \mathbf{c}] \end{aligned}$$

$$\begin{aligned} TNDE &= E[Y_{a_1,M_{a_1}}|\mathbf{C} = \mathbf{c}] - E[Y_{a_0,M_{a_1}}|\mathbf{C} = \mathbf{c}] \\ PNIE &= E[Y_{a_0,M_{a_1}}|\mathbf{C} = \mathbf{c}] - E[Y_{a_0,M_{a_0}}|\mathbf{C} = \mathbf{c}] \end{aligned}$$

$$TE = E[Y_{a_1}|\mathbf{C} = \mathbf{c}] - E[Y_{a_0}|\mathbf{C} = \mathbf{c}]$$

The total effect (TE) can be decomposed into the direct (non-mediated) effect and indirect (mediated) effect in two different ways [Robins and Greenland, 1992, VanderWeele, 2013].

The decomposition of TE into the pure (natural) direct effect (PNDE) and the total (natural) indirect effect (TNIE) is the usual decomposition [Pearl, 2001]. Note that the treatment value indexing the mediator M is fixed at a_0 in the PNDE, whereas the treatment value indexing the outcome Y is fixed at a_1 in the TNIE. These are emphasized with $\utilde{\quad}$ throughout the document.

$$\begin{aligned} PNDE &= E[Y_{a_1,\utilde{M_{a_0}}}|\mathbf{C} = \mathbf{c}] - E[Y_{a_0,\utilde{M_{a_0}}}|\mathbf{C} = \mathbf{c}] \\ TNIE &= E[Y_{\utilde{a_1},M_{a_1}}|\mathbf{C} = \mathbf{c}] - E[Y_{\utilde{a_1},M_{a_0}}|\mathbf{C} = \mathbf{c}] \end{aligned}$$

The other decomposition of TE is into the pure (natural) indirect effect (PNIE) and the total (natural) direct effect (TNDE) [Robins and Greenland, 1992]. Note that the treatment value indexing the mediator M is fixed at a_1 in the TNDE, whereas the treatment value indexing the outcome Y is fixed at a_0 in the PNIE. That is, these flipped in this decomposition. These are emphasized with $\utilde{\quad}$ throughout the document.

$$\begin{aligned} TNDE &= E[Y_{a_1,\utilde{M_{a_1}}}|\mathbf{C} = \mathbf{c}] - E[Y_{a_0,\utilde{M_{a_1}}}|\mathbf{C} = \mathbf{c}] \\ PNIE &= E[Y_{\utilde{a_0},M_{a_1}}|\mathbf{C} = \mathbf{c}] - E[Y_{\utilde{a_0},M_{a_0}}|\mathbf{C} = \mathbf{c}] \end{aligned}$$

More intuitively, these decomposition differs in the cross-world counterfactual state that is used as the partition. The usual PNDE+TNIE decomposition uses $Y_{a_1,M_{a_0}}$ (treatment indexing on the outcome goes up to a_1 from a_0 first), whereas the PNIE+TNDE decomposition uses $Y_{a_0,M_{a_1}}$ (treatment indexing the mediator goes up to a_1 from

a_0 first).

In either case, the effect that has the reference counterfactual outcome $Y_{a_0} = Y_{a_0,M_{a_0}}$ is the "pure" natural direct/indirect effect and the effect that has $Y_{a_1} = Y_{a_1,M_{a_1}}$ is the total natural direct/indirect effect. See [VanderWeele, 2013] for the meaning of these two decompositions in terms of causal interaction.

More in general, we can consider the effects on the link function scale as follows [Starkopf et al., 2017].

$$CDE(m) = g(E[Y_{a_1,m}|\mathbf{C} = \mathbf{c}]) - g(E[Y_{a_0,m}|\mathbf{C} = \mathbf{c}])$$

$$\begin{aligned} PNDE &= g(E[Y_{a_1,M_{a_0}}|\mathbf{C} = \mathbf{c}]) - g(E[Y_{a_0,M_{a_0}}|\mathbf{C} = \mathbf{c}]) \\ TNIE &= g(E[Y_{a_1,M_{a_1}}|\mathbf{C} = \mathbf{c}]) - g(E[Y_{a_1,M_{a_0}}|\mathbf{C} = \mathbf{c}]) \end{aligned}$$

$$\begin{aligned} TNDE &= g(E[Y_{a_1,M_{a_1}}|\mathbf{C} = \mathbf{c}]) - g(E[Y_{a_0,M_{a_1}}|\mathbf{C} = \mathbf{c}]) \\ PNIE &= g(E[Y_{a_0,M_{a_1}}|\mathbf{C} = \mathbf{c}]) - g(E[Y_{a_0,M_{a_0}}|\mathbf{C} = \mathbf{c}]) \end{aligned}$$

$$TE = g(E[Y_{a_1}|\mathbf{C} = \mathbf{c}]) - g(E[Y_{a_0}|\mathbf{C} = \mathbf{c}])$$

1.2 Identification of natural effects

Several conditional exchangeabilities must be assumed for identification of effects in the causal mediation framework. See [VanderWeele, 2015] (p463) for details.

$$\begin{aligned} &\text{A1} \\ Y_{a,m} &\perp\!\!\!\perp A|\mathbf{C} \\ &\text{A2} \\ Y_{a,m} &\perp\!\!\!\perp M|\left\{A, \mathbf{C}\right\} \\ &\text{A3} \\ M_a &\perp\!\!\!\perp A|\mathbf{C} \\ &\text{A4} \\ Y_{a,m} &\perp\!\!\!\perp M_{a^*}|\mathbf{C} \end{aligned}$$

The controlled direct effect (CDE) is identified with only Assumptions A1 and A2. The natural effects require all four assumptions. Intuitively, the identification of CDE involves handling the treatment and the mediator as a sequence of exposures, whose causal effects on the outcome must be identified. Thus, the conditional exchangeabilities for the treatment as an exposure (A1) and the mediator as an exposure (A2) are required.

Additionally, the identification of the natural effects require identifying the causal effect of the treatment on the mediator acting as an "outcome" (thus, Assumption A3). The identification of the partitioning counterfactual state mentioned above, in which the treatment value indexing the outcome Y and the treatment value indexing the mediator M differ requires Assumption A4.

Given these four assumptions, the mean counterfactual with different treatment values indexing the outcome Y and the mediator M can be identified as follows. See [VanderWeele, 2015] (p465) for the proof. For a continuous M, the summation is replaced with an integration.

$$E[Y_{a,M_{a^*}}|\mathbf{C} = \mathbf{c}] = \sum_m E[Y|A = a, M = m, \mathbf{C} = \mathbf{c}]p(M = m|A = a^*, \mathbf{C} = \mathbf{c})$$

We can observe that the first treatment value a in the counterfactual $Y_{a,M_{a^*}}$ indexes the outcome model $E[Y|A = \utilde{a}, M = m, \mathbf{C} = \mathbf{c}]$, whereas the second treatment value a^* in the counterfactual $Y_{a,\utilde{M_{a^*}}}$ indexes the mediator model $p(M = m|A = \utilde{a^*}, \mathbf{C} = \mathbf{c})$.

As a result, the identification formulas for the two natural *direct* effects are the following. Note the change in the treatment value indexing the *mediator model* from a_0 in the PNDE (usual NDE) to a_1 in the TNDE (annotated with $\utilde{\quad}$). Within each effect, only the treatment values indexing the *outcome model* vary (ones not annotated) because each NDE represent the direct effect of the treatment when the mediator is fixed at the natural value it would take under one treatment value (annotated with $\utilde{\quad}$).

$$\begin{aligned}
PNDE &= E[Y_{a_1,\utilde{M_{a_0}}}|\mathbf{C} = \mathbf{c}] - E[Y_{a_0,\utilde{M_{a_0}}}|\mathbf{C} = \mathbf{c}] \\
&= \sum_m E[Y|A = a_1, M = m, \mathbf{C} = \mathbf{c}]p(M = m|A = \utilde{a_0}, \mathbf{C} = \mathbf{c}) \\
&\quad - \sum_m E[Y|A = a_0, M = m, \mathbf{C} = \mathbf{c}]p(M = m|A = \utilde{a_0}, \mathbf{C} = \mathbf{c}) \\
&= \sum_m \{E[Y|A = a_1, M = m, \mathbf{C} = \mathbf{c}] - E[Y|A = a_0, M = m, \mathbf{C} = \mathbf{c}]\} \\
&\quad \times p(M = m|A = \utilde{a_0}, \mathbf{C} = \mathbf{c})
\end{aligned}$$

$$\begin{aligned}
TNDE &= E[Y_{a_1,\utilde{M_{a_1}}}|\mathbf{C} = \mathbf{c}] - E[Y_{a_0,\utilde{M_{a_1}}}|\mathbf{C} = \mathbf{c}] \\
&= \sum_m E[Y|A = a_1, M = m, \mathbf{C} = \mathbf{c}]p(M = m|A = \utilde{a_1}, \mathbf{C} = \mathbf{c}) \\
&\quad - \sum_m E[Y|A = a_0, M = m, \mathbf{C} = \mathbf{c}]p(M = m|A = \utilde{a_1}, \mathbf{C} = \mathbf{c}) \\
&= \sum_m \{E[Y|A = a_1, M = m, \mathbf{C} = \mathbf{c}] - E[Y|A = a_0, M = m, \mathbf{C} = \mathbf{c}]\} \\
&\quad \times p(M = m|A = \utilde{a_1}, \mathbf{C} = \mathbf{c})
\end{aligned}$$

Also, the identification formulas for the two natural *indirect* effects are the following. Note the change in the treatment value indexing the *outcome model* from a_1 in the TNIE (usual NIE) to a_0 in the PNIE (annotated with). Within each effect, only the treatment values indexing the *mediator model* vary (ones not annotated) because each NIE represent the indirect effect of the treatment when its effect on the mediator is "turned on", while the treatment value representing the direct path is fixed at the natural value it would take under one treatment value (annotated with $\utilde{\quad}$).

$$\begin{aligned}
TNIE &= E[Y_{\utilde{a_1},M_{a_1}}|\mathbf{C} = \mathbf{c}] - E[Y_{\utilde{a_1},M_{a_0}}|\mathbf{C} = \mathbf{c}] \\
&= \sum_m E[Y|A = \utilde{a_1}, M = m, \mathbf{C} = \mathbf{c}]p(M = m|A = a_1, \mathbf{C} = \mathbf{c}) \\
&\quad - \sum_m E[Y|A = \utilde{a_1}, M = m, \mathbf{C} = \mathbf{c}]p(M = m|A = a_0, \mathbf{C} = \mathbf{c}) \\
&= \sum_m E[Y|A = \utilde{a_1}, M = m, \mathbf{C} = \mathbf{c}] \\
&\quad \times \{p(M = m|A = a_1, \mathbf{C} = \mathbf{c}) - p(M = m|A = a_0, \mathbf{C} = \mathbf{c})\} \\
PNIE &= E[Y_{\utilde{a_0},M_{a_1}}|\mathbf{C} = \mathbf{c}] - E[Y_{\utilde{a_0},M_{a_0}}|\mathbf{C} = \mathbf{c}] \\
&= \sum_m E[Y|A = \utilde{a_0}, M = m, \mathbf{C} = \mathbf{c}]p(M = m|A = a_1, \mathbf{C} = \mathbf{c})
\end{aligned}$$

$$
\begin{aligned}
&- \sum_m E[Y|A = \underset{\sim}{a_0}, M = m, \mathbf{C} = \mathbf{c}]p(M = m|A = a_0, \mathbf{C} = \mathbf{c}) \\
&= \sum_m E[Y|A = \underset{\sim}{a_0}, M = m, \mathbf{C} = \mathbf{c}] \\
&\quad \times \{p(M = m|A = a_1, \mathbf{C} = \mathbf{c}) - p(M = m|A = a_0, \mathbf{C} = \mathbf{c})\}
\end{aligned}
$$

2 Understanding the approach

Here we describe the formulas implemented in `regmendint`, using the notational convention in [VanderWeele, 2015, Valeri and Vanderweele, 2013, Valeri and VanderWeele, 2015].

2.1 Parametrizing the mediation effect formulas

A seen above, there are two models involved in identification of natural effects: the outcome model ($E[Y|A = a, M = m, \mathbf{C} = \mathbf{c}]$) and the mediator model ($p(M|A = a, \mathbf{C} = \mathbf{c})$). The identification formulas do not specify any particular model structure (non-parametric). In the method described in [Valeri and Vanderweele, 2013, Valeri and VanderWeele, 2015], a simple parametric model is proposed for each.

Under this parametric modeling assumption, each effect of interest can be written as a function of the parameters (coefficients) of the the mediator model (β) and the outcome model (θ). Because of the product configuration (outcome model $\times$ mediator model), each natural effect is a non-linear function of the parameters (model coefficients). The maximum likelihood estimates (MLE) of these effects are the ones with these parameters replaced with their respective MLEs from the two models.

2.2 Obtaining standard errors via multivariate delta method

Each effect of interest is estimated as a scalar-valued, non-linear function of estiamted coefficients for the mediator model and the outcome model. Thus, we can obtain the standard error of each effect estimate using the variance covariance matrix for the coefficients and multivariate delta method [Hoef, 2012].

Let the scalar quantity of interest be Q, a function of parameter vector $(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T)^T$. Then, its gradient (vector of partial derivatives) with respect to the parameter vector $(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T)^T$ is the following.

$$
\begin{aligned}
\nabla Q &= \frac{\partial Q}{\partial(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T)^T} \\
&= \begin{bmatrix} \frac{\partial Q}{\partial \beta_0} \\ \frac{\partial Q}{\partial \beta_1} \\ \frac{\partial Q}{\partial \boldsymbol{\beta}_2} \\ \frac{\partial Q}{\partial \theta_0} \\ \frac{\partial Q}{\partial \theta_1} \\ \frac{\partial Q}{\partial \theta_2} \\ \frac{\partial Q}{\partial \theta_3} \\ \frac{\partial Q}{\partial \boldsymbol{\theta}_4} \end{bmatrix}
\end{aligned}
$$

In the case of a linear mediator model and a non-linear outcome model, there is an additional element $\frac{\partial Q}{\partial \sigma^2}$ at the bottom of the gradient vector.

By the large sample approximation using the multivariate delta method, the variance of the quantity of interest evaluated at the MLEs $(\widehat{\boldsymbol{\beta}}^T, \widehat{\boldsymbol{\theta}}^T)^T$ is the following.

$$
\underbrace{Var\left[Q\left\{(\widehat{\boldsymbol{\beta}}^T, \widehat{\boldsymbol{\theta}}^T)^T\right\}\right]}_{\text{scalar}} \approx \underbrace{\left[\nabla Q\left((\widehat{\boldsymbol{\beta}}^T, \widehat{\boldsymbol{\theta}}^T)^T\right)\right]^T}_{\text{row vector}} \underbrace{Var((\widehat{\boldsymbol{\beta}}^T, \widehat{\boldsymbol{\theta}}^T)^T)}_{\text{matrix}} \underbrace{\left[\nabla Q\left((\widehat{\boldsymbol{\beta}}^T, \widehat{\boldsymbol{\theta}}^T)^T\right)\right]}_{\text{column vector}}
$$

This expression is abbreviated as $\Gamma\Sigma\Gamma'$ in [VanderWeele, 2015, Valeri and Vanderweele, 2013, Valeri and VanderWeele, 2015]. In these references, the treatment contrast $(a_1 - a_0)$ is factored out from $\nabla Q\left((\widehat{\boldsymbol{\beta}}^T, \widehat{\boldsymbol{\theta}}^T)^T\right)$ when possible.

2.3 Linear mediator model, linear outcome model

2.3.1 Effect formulas

The function `calc_myreg_mreg_linear_yreg_linear_est()` implements the effect formulas in [VanderWeele, 2015] (p466).

$$\text{Models}$$

$$\begin{aligned}
E[Y|A=a, M=m, \mathbf{C}=\mathbf{c}] &= \theta_0 + \theta_1 a + \theta_2 m + \theta_3 am + \boldsymbol{\theta}_4^T \mathbf{c} \\
E[M|A=a, \mathbf{C}=\mathbf{c}] &= \beta_0 + \beta_1 a + \boldsymbol{\beta}_2^T \mathbf{c}
\end{aligned}$$

$$\text{Effects}$$

$$\begin{aligned}
CDE(m) &= E[Y_{a_1,m}|\mathbf{C}=\mathbf{c}] - E[Y_{a_0,m}|\mathbf{C}=\mathbf{c}] \\
&= (\theta_1 + \theta_3 m)(a_1 - a_0)
\end{aligned}$$

$$\begin{aligned}
PNDE &= E[Y_{a_1,\utilde{M_{a_0}}}|\mathbf{C}=\mathbf{c}] - E[Y_{a_0,\utilde{M_{a_0}}}|\mathbf{C}=\mathbf{c}] \\
&= \left\{\theta_1 + \theta_3(\beta_0 + \beta_1 \utilde{a_0} + \boldsymbol{\beta}_2^T \mathbf{c})\right\}(a_1 - a_0) \\
TNIE &= E[Y_{\utilde{a_1},M_{a_1}}|\mathbf{C}=\mathbf{c}] - E[Y_{\utilde{a_1},M_{a_0}}|\mathbf{C}=\mathbf{c}] \\
&= \beta_1(\theta_2 + \theta_3 \utilde{a_1})(a_1 - a_0)
\end{aligned}$$

$$\begin{aligned}
TNDE &= E[Y_{a_1,\utilde{M_{a_1}}}|\mathbf{C}=\mathbf{c}] - E[Y_{a_0,\utilde{M_{a_1}}}|\mathbf{C}=\mathbf{c}] \\
&= \left\{\theta_1 + \theta_3(\beta_0 + \beta_1 \utilde{a_1} + \boldsymbol{\beta}_2^T \mathbf{c})\right\}(a_1 - a_0) \\
PNIE &= E[Y_{\utilde{a_0},M_{a_1}}|\mathbf{C}=\mathbf{c}] - E[Y_{\utilde{a_0},M_{a_0}}|\mathbf{C}=\mathbf{c}] \\
&= \beta_1(\theta_2 + \theta_3 \utilde{a_0})(a_1 - a_0)
\end{aligned}$$

$$\begin{aligned}
TE &= PNDE + TNIE \\
PM &= \frac{TNIE}{PNDE + TNIE}
\end{aligned}$$

2.3.2 Variance formulas

The function `calc_myreg_mreg_linear_yreg_linear_se()` implements the standard error formulas in [VanderWeele, 2015] (p466).

$$(a_1 - a_0)\Gamma_{CDE(m)} = \frac{\partial CDE(m)}{\partial(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T)^T}$$

$$= (a_1 - a_0) \begin{bmatrix} 0 \\ 0 \\ \underline{0} \\ 0 \\ 1 \\ 0 \\ m \\ \underline{0} \end{bmatrix}$$

$$\begin{aligned}
(a_1 - a_0)\Gamma_{PNDE} &= \frac{\partial PNDE}{\partial(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T)^T} \\
&= (a_1 - a_0) \begin{bmatrix} \theta_3 \\ \theta_3 \underset{\sim}{a_0} \\ \theta_3 \mathbf{c} \\ 0 \\ 1 \\ 0 \\ \beta_0 + \beta_1 \underset{\sim}{a_0} + \boldsymbol{\beta}_2^T \mathbf{c} \\ \underline{0} \end{bmatrix}
\end{aligned}$$

$$\begin{aligned}
(a_1 - a_0)\Gamma_{TNIE} &= \frac{\partial TNIE}{\partial(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T)^T} \\
&= (a_1 - a_0) \begin{bmatrix} 0 \\ \theta_2 + \theta_3 \underset{\sim}{a_1} \\ \underline{0} \\ 0 \\ 0 \\ \beta_1 \\ \beta_1 \underset{\sim}{a_1} \\ \underline{0} \end{bmatrix}
\end{aligned}$$

$$\begin{aligned}
(a_1 - a_0)\Gamma_{TNDE} &= \frac{\partial TNDE}{\partial(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T)^T} \\
&= (a_1 - a_0) \begin{bmatrix} \theta_3 \\ \theta_3 \underset{\sim}{a_1} \\ \theta_3 \mathbf{c} \\ 0 \\ 1 \\ 0 \\ \beta_0 + \beta_1 \underset{\sim}{a_1} + \boldsymbol{\beta}_2^T \mathbf{c} \\ \underline{0} \end{bmatrix}
\end{aligned}$$

$$\begin{aligned}
(a_1 - a_0)\Gamma_{PNIE} &= \frac{\partial PNIE}{\partial(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T)^T} \\
&= (a_1 - a_0) \begin{bmatrix} 0 \\ \theta_2 + \theta_3 \underset{\sim}{a_0} \\ \underline{0} \\ 0 \\ 0 \\ \beta_1 \\ \beta_1 \underset{\sim}{a_0} \\ \underline{0} \end{bmatrix}
\end{aligned}$$

$$
\begin{aligned}
(a_1 - a_0)\Gamma_{TE} &= \frac{\partial TE}{\partial(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T)^T} \\
&= \frac{\partial(PNDE + TNIE)}{\partial(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T)^T} \\
&= (a_1 - a_0)(\Gamma_{PNDE} + \Gamma_{TNIE})
\end{aligned}
$$

$$
\begin{aligned}
(a_1 - a_0)\Gamma_{PM} &= \frac{\partial PM}{\partial(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T)^T} \\
&\quad \text{By multivariate chain rule} \\
&= \frac{\partial PM}{\partial PNDE}\frac{\partial PNDE}{\partial(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T)^T} + \frac{\partial PM}{\partial TNIE}\frac{\partial TNIE}{\partial(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T)^T} \\
&= \frac{\partial PM}{\partial PNDE}(a_1 - a_0)\Gamma_{PNDE} + \frac{\partial PM}{\partial TNIE}(a_1 - a_0)\Gamma_{TNIE} \\
&= \frac{-TNIE}{(PNDE + TNIE)^2}(a_1 - a_0)\Gamma_{PNDE} + \frac{PNDE}{(PNDE + TNIE)^2}(a_1 - a_0)\Gamma_{TNIE} \\
&= (a_1 - a_0)\frac{-TNIE\ \Gamma_{PNDE} + PNDE\ \Gamma_{TNIE}}{(PNDE + TNIE)^2}
\end{aligned}
$$

$$
\begin{aligned}
&\text{Variance-covariance matrix from two models} \\
\boldsymbol{\Sigma} &= \begin{bmatrix} \boldsymbol{\Sigma}_{\boldsymbol{\beta}} & 0 \\ 0 & \boldsymbol{\Sigma}_{\boldsymbol{\theta}} \end{bmatrix} \\
SE(\widehat{CDE}(m)) &= \sqrt{\Gamma_{CDE(m)}^T\ \boldsymbol{\Sigma}\ \Gamma_{CDE(m)}}\ |a_1 - a_0| \\
SE(\widehat{PNDE}) &= \sqrt{\Gamma_{PNDE}^T\ \boldsymbol{\Sigma}\ \Gamma_{PNDE}}\ |a_1 - a_0| \\
SE(\widehat{TNIE}) &= \sqrt{\Gamma_{TNIE}^T\ \boldsymbol{\Sigma}\ \Gamma_{TNIE}}\ |a_1 - a_0| \\
SE(\widehat{TNDE}) &= \sqrt{\Gamma_{TNDE}^T\ \boldsymbol{\Sigma}\ \Gamma_{TNDE}}\ |a_1 - a_0| \\
SE(\widehat{PNIE}) &= \sqrt{\Gamma_{PNIE}^T\ \boldsymbol{\Sigma}\ \Gamma_{PNIE}}\ |a_1 - a_0| \\
SE(\widehat{TE}) &= \sqrt{\Gamma_{TE}^T\ \boldsymbol{\Sigma}\ \Gamma_{TE}}\ |a_1 - a_0| \\
SE(\widehat{PM}) &= \sqrt{\Gamma_{PM}^T\ \boldsymbol{\Sigma}\ \Gamma_{PM}}\ |a_1 - a_0|
\end{aligned}
$$

2.4 Linear mediator model, non-linear outcome model

2.4.1 Effect formulas

The function `calc_myreg_mreg_linear_yreg_logistic_est()` implements the effect formulas in [VanderWeele, 2015] (p468).

$$
\begin{aligned}
&\text{Models} \\
\text{logit}(E[Y|A = a, M = m, \mathbf{C} = \mathbf{c}]) &= \theta_0 + \theta_1 a + \theta_2 m + \theta_3 am + \boldsymbol{\theta}_4^T \mathbf{c} \\
E[M|A = a, \mathbf{C} = \mathbf{c}] &= \beta_0 + \beta_1 a + \boldsymbol{\beta}_2^T \mathbf{c}
\end{aligned}
$$

$$
\begin{aligned}
&\text{Effects on link function scale} \\
CDE(m) &= \text{logit}(E[Y_{a_1,m}|\mathbf{C} = \mathbf{c}]) - \text{logit}(E[Y_{a_0,m}|\mathbf{C} = \mathbf{c}]) \\
&= (\theta_1 + \theta_3 m)(a_1 - a_0)
\end{aligned}
$$

$$
\begin{aligned}
PNDE &= \text{logit}(E[Y_{a_1,\underset{\sim}{M_{a_0}}}|\mathbf{C}=\mathbf{c}]) - \text{logit}(E[Y_{a_0,\underset{\sim}{M_{a_0}}}|\mathbf{C}=\mathbf{c}]) \\
&\approx \left\{\theta_1 + \theta_3(\beta_0 + \beta_1 \underset{\sim}{a_0} + \boldsymbol{\beta}_2^T\mathbf{c} + \theta_2\sigma^2)\right\}(a_1 - a_0) + \frac{1}{2}\theta_3^2\sigma^2(a_1^2 - a_0^2) \\
TNIE &= \text{logit}(E[Y_{\underset{\sim}{a_1},M_{a_1}}|\mathbf{C}=\mathbf{c}]) - \text{logit}(E[Y_{\underset{\sim}{a_1},M_{a_0}}|\mathbf{C}=\mathbf{c}]) \\
&\approx \beta_1(\theta_2 + \theta_3 \underset{\sim}{a_1})(a_1 - a_0)
\end{aligned}
$$

$$
\begin{aligned}
TNDE &= \text{logit}(E[Y_{a_1,\underset{\sim}{M_{a_1}}}|\mathbf{C}=\mathbf{c}]) - \text{logit}(E[Y_{a_0,\underset{\sim}{M_{a_1}}}|\mathbf{C}=\mathbf{c}]) \\
&\approx \left\{\theta_1 + \theta_3(\beta_0 + \beta_1 \underset{\sim}{a_1} + \boldsymbol{\beta}_2^T\mathbf{c} + \theta_2\sigma^2)\right\}(a_1 - a_0) + \frac{1}{2}\theta_3^2\sigma^2(a_1^2 - a_0^2) \\
PNIE &= \text{logit}(E[Y_{\underset{\sim}{a_0},M_{a_1}}|\mathbf{C}=\mathbf{c}]) - \text{logit}(E[Y_{\underset{\sim}{a_0},M_{a_0}}|\mathbf{C}=\mathbf{c}]) \\
&\approx \beta_1(\theta_2 + \theta_3 \underset{\sim}{a_0})(a_1 - a_0)
\end{aligned}
$$

$$
\begin{aligned}
TE &= PNDE + TNIE \\
PM &= \frac{\exp(PNDE)(\exp(TNIE) - 1)}{\exp(PNDE)\exp(TNIE) - 1}
\end{aligned}
$$

2.4.2 Variance formulas

The function `calc_myreg_mreg_linear_yreg_logistic_se()` implements the standard error formulas in [VanderWeele, 2015] (p468).

$$
\begin{aligned}
(a_1 - a_0)\Gamma_{CDE(m)} &= \frac{\partial CDE(m)}{\partial(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T, \sigma^2)^T} \\
&= (a_1 - a_0)\begin{bmatrix} 0 \\ 0 \\ \underline{0} \\ 0 \\ 1 \\ 0 \\ m \\ \underline{0} \\ 0 \end{bmatrix}
\end{aligned}
$$

$$
\begin{aligned}
(a_1 - a_0)\Gamma_{PNDE} &= \frac{\partial PNDE}{\partial(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T, \sigma^2)^T} \\
&= (a_1 - a_0)\begin{bmatrix} \theta_3 \\ \theta_3 \underset{\sim}{a_0} \\ \theta_3\mathbf{c} \\ 0 \\ 1 \\ \theta_3\sigma^2 \\ \beta_0 + \beta_1 \underset{\sim}{a_0} + \boldsymbol{\beta}_2^T\mathbf{c} + \theta_2\sigma^2 + \theta_3\sigma^2(a_0 + a_1) \\ \underline{0} \\ \theta_3\theta_2 + \frac{1}{2}\theta_3^2(a_1 + a_0) \end{bmatrix} \\
(a_1 - a_0)\Gamma_{TNIE} &= \frac{\partial TNIE}{\partial(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T, \sigma^2)^T}
\end{aligned}
$$

$$= (a_1 - a_0) \begin{bmatrix} 0 \\ \theta_2 + \theta_3 \underset{\sim}{a_1} \\ \underline{0} \\ 0 \\ 0 \\ \beta_1 \\ \beta_1 \underset{\sim}{a_1} \\ \underline{0} \\ 0 \end{bmatrix}$$

$$\begin{aligned}
(a_1 - a_0)\Gamma_{TNDE} &= \frac{\partial TNDE}{\partial(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T, \sigma^2)^T} \\
&= (a_1 - a_0) \begin{bmatrix} \theta_3 \\ \theta_3 \underset{\sim}{a_1} \\ \theta_3 \mathbf{c} \\ 0 \\ 1 \\ \theta_3 \sigma^2 \\ \beta_0 + \beta_1 \underset{\sim}{a_1} + \boldsymbol{\beta}_2^T \mathbf{c} + \theta_2 \sigma^2 + \theta_3 \sigma^2 (a_0 + a_1) \\ \underline{0} \\ \theta_3 \theta_2 + \frac{1}{2} \theta_3^2 (a_1 + a_0) \end{bmatrix}
\end{aligned}$$

$$\begin{aligned}
(a_1 - a_0)\Gamma_{PNIE} &= \frac{\partial PNIE}{\partial(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T, \sigma^2)^T} \\
&= (a_1 - a_0) \begin{bmatrix} 0 \\ \theta_2 + \theta_3 \underset{\sim}{a_0} \\ \underline{0} \\ 0 \\ 0 \\ \beta_1 \\ \beta_1 \underset{\sim}{a_0} \\ \underline{0} \\ 0 \end{bmatrix}
\end{aligned}$$

$$\begin{aligned}
(a_1 - a_0)\Gamma_{TE} &= \frac{\partial TE}{\partial(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T, \sigma^2)^T} \\
&= \frac{\partial (PNDE + TNIE)}{\partial(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T, \sigma^2)^T} \\
&= (a_1 - a_0)(\Gamma_{PNDE} + \Gamma_{TNIE})
\end{aligned}$$

$$\begin{aligned}
(a_1 - a_0)\Gamma_{PM} &= \frac{\partial PM}{\partial(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T, \sigma^2)^T} \\
&\quad \text{By multivariate chain rule} \\
&= \frac{\partial PM}{\partial PNDE} \frac{\partial PNDE}{\partial(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T, \sigma^2)^T} + \frac{\partial PM}{\partial TNIE} \frac{\partial TNIE}{\partial(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T, \sigma^2)^T} \\
&= \frac{\partial PM}{\partial PNDE} (a_1 - a_0)\Gamma_{PNDE} + \frac{\partial PM}{\partial TNIE} (a_1 - a_0)\Gamma_{TNIE} \\
&= -\frac{\exp(PNDE)\left\{\exp(TNIE) - 1\right\}}{\left\{\exp(PNDE)\exp(TNIE) - 1\right\}^2} (a_1 - a_0)\Gamma_{PNDE}
\end{aligned}$$

$$+ \frac{\exp(PNDE)\exp(TNIE)\{\exp(PNDE)-1\}}{\{\exp(PNDE)\exp(TNIE)-1\}^2}(a_1-a_0)\Gamma_{TNIE}$$

$$\text{Variance-covariance matrix from two models}$$
$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma_\beta} & 0 & 0 \\ 0 & \boldsymbol{\Sigma_\theta} & 0 \\ 0 & 0 & \Sigma_{\sigma^2} \end{bmatrix}$$
$$\Sigma_{\sigma^2} = \frac{2(\sigma^2)^2}{n-p} \text{ where } p = \text{length}(\boldsymbol{\beta})$$
$$SE(\widehat{CDE}(m)) = \sqrt{\Gamma^T_{CDE(m)}\ \boldsymbol{\Sigma}\ \Gamma_{CDE(m)}}\ |a_1 - a_0|$$
$$SE(\widehat{PNDE}) = \sqrt{\Gamma^T_{PNDE}\ \boldsymbol{\Sigma}\ \Gamma_{PNDE}}\ |a_1 - a_0|$$
$$SE(\widehat{TNIE}) = \sqrt{\Gamma^T_{TNIE}\ \boldsymbol{\Sigma}\ \Gamma_{TNIE}}\ |a_1 - a_0|$$
$$SE(\widehat{TNDE}) = \sqrt{\Gamma^T_{TNDE}\ \boldsymbol{\Sigma}\ \Gamma_{TNDE}}\ |a_1 - a_0|$$
$$SE(\widehat{PNIE}) = \sqrt{\Gamma^T_{PNIE}\ \boldsymbol{\Sigma}\ \Gamma_{PNIE}}\ |a_1 - a_0|$$
$$SE(\widehat{TE}) = \sqrt{\Gamma^T_{TE}\ \boldsymbol{\Sigma}\ \Gamma_{TE}}\ |a_1 - a_0|$$
$$SE(\widehat{PM}) = \sqrt{\Gamma^T_{PM}\ \boldsymbol{\Sigma}\ \Gamma_{PM}}\ |a_1 - a_0|$$

2.5 Logistic mediator model, linear outcome model

2.5.1 Effect formulas

The function `calc_myreg_mreg_logistic_yreg_linear_est()` implements the effect formulas in [VanderWeele, 2015] (p471).

$$\text{Models}$$
$$E[Y|A=a, M=m, \mathbf{C}=\mathbf{c}] = \theta_0 + \theta_1 a + \theta_2 m + \theta_3 am + \boldsymbol{\theta}_4^T \mathbf{c}$$
$$\text{logit}(E[M|A=a, \mathbf{C}=\mathbf{c}]) = \beta_0 + \beta_1 a + \boldsymbol{\beta}_2^T \mathbf{c}$$

$$\text{Effects}$$
$$\begin{aligned} CDE(m) &= E[Y_{a_1,m}|\mathbf{C}=\mathbf{c}] - E[Y_{a_0,m}|\mathbf{C}=\mathbf{c}] \\ &= (\theta_1 + \theta_3 m)(a_1 - a_0) \end{aligned}$$

$$\begin{aligned} PNDE &= E[Y_{a_1,\utilde{M_{a_0}}}|\mathbf{C}=\mathbf{c}] - E[Y_{a_0,\utilde{M_{a_0}}}|\mathbf{C}=\mathbf{c}] \\ &= \{\theta_1(a_1-a_0)\} + \{\theta_3(a_1-a_0)\}\frac{\exp(\beta_0 + \beta_1 \utilde{a_0} + \boldsymbol{\beta}_2^T \mathbf{c})}{1+\exp(\beta_0 + \beta_1 \utilde{a_0} + \boldsymbol{\beta}_2^T \mathbf{c})} \\ TNIE &= E[Y_{\utilde{a_1},M_{a_1}}|\mathbf{C}=\mathbf{c}] - E[Y_{\utilde{a_1},M_{a_0}}|\mathbf{C}=\mathbf{c}] \\ &= (\theta_2 + \theta_3 \utilde{a_1})\left\{\frac{\exp(\beta_0 + \beta_1 a_1 + \boldsymbol{\beta}_2^T \mathbf{c})}{1+\exp(\beta_0 + \beta_1 a_1 + \boldsymbol{\beta}_2^T \mathbf{c})} - \frac{\exp(\beta_0 + \beta_1 a_0 + \boldsymbol{\beta}_2^T \mathbf{c})}{1+\exp(\beta_0 + \beta_1 a_0 + \boldsymbol{\beta}_2^T \mathbf{c})}\right\} \\ TNDE &= E[Y_{a_1,\utilde{M_{a_1}}}|\mathbf{C}=\mathbf{c}] - E[Y_{a_0,\utilde{M_{a_1}}}|\mathbf{C}=\mathbf{c}] \\ &= \{\theta_1(a_1-a_0)\} + \{\theta_3(a_1-a_0)\}\frac{\exp(\beta_0 + \beta_1 \utilde{a_1} + \boldsymbol{\beta}_2^T \mathbf{c})}{1+\exp(\beta_0 + \beta_1 \utilde{a_1} + \boldsymbol{\beta}_2^T \mathbf{c})} \end{aligned}$$

$$
\begin{aligned}
PNIE &= E[Y_{\underset{\sim}{a_0},M_{a_1}}|\mathbf{C}=\mathbf{c}] - E[Y_{\underset{\sim}{a_0},M_{a_0}}|\mathbf{C}=\mathbf{c}] \\
&= (\theta_2 + \theta_3 \underset{\sim}{a_0}) \left\{ \frac{\exp(\beta_0 + \beta_1 a_1 + \boldsymbol{\beta}_2^T \mathbf{c})}{1+\exp(\beta_0 + \beta_1 a_1 + \boldsymbol{\beta}_2^T \mathbf{c})} - \frac{\exp(\beta_0 + \beta_1 a_0 + \boldsymbol{\beta}_2^T \mathbf{c})}{1+\exp(\beta_0 + \beta_1 a_0 + \boldsymbol{\beta}_2^T \mathbf{c})} \right\} \\
TE &= PNDE + TNIE \\
PM &= \frac{TNIE}{PNDE + TNIE}
\end{aligned}
$$

2.5.2 Variance formulas

The function `calc_myreg_mreg_logistic_yreg_linear_se()` implements the standard error formulas in [VanderWeele, 2015] (p471).

$$
\begin{aligned}
(a_1 - a_0)\Gamma_{CDE(m)} &= \frac{\partial CDE(m)}{\partial(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T)^T} \\
&= (a_1 - a_0) \begin{bmatrix} 0 \\ 0 \\ \underset{\sim}{0} \\ 0 \\ 1 \\ 0 \\ m \\ \underset{\sim}{0} \end{bmatrix}
\end{aligned}
$$

$$
\begin{aligned}
d_{1,PNDE} &= \theta_3 \frac{\exp(\beta_0 + \beta_1 \underset{\sim}{a_0} + \boldsymbol{\beta}_2^T \mathbf{c})}{1+\exp(\beta_0 + \beta_1 \underset{\sim}{a_0} + \boldsymbol{\beta}_2^T \mathbf{c})} \\
d_{2,PNDE} &= \underset{\sim}{a_0} \theta_3 \frac{\exp(\beta_0 + \beta_1 \underset{\sim}{a_0} + \boldsymbol{\beta}_2^T \mathbf{c})}{1+\exp(\beta_0 + \beta_1 \underset{\sim}{a_0} + \boldsymbol{\beta}_2^T \mathbf{c})} \\
d_{3,PNDE} &= \mathbf{c}\theta_3 \frac{\exp(\beta_0 + \beta_1 \underset{\sim}{a_0} + \boldsymbol{\beta}_2^T \mathbf{c})}{1+\exp(\beta_0 + \beta_1 \underset{\sim}{a_0} + \boldsymbol{\beta}_2^T \mathbf{c})} \\
d_{4,PNDE} &= 0 \\
d_{5,PNDE} &= 1 \\
d_{6,PNDE} &= 0 \\
d_{7,PNDE} &= \frac{\exp(\beta_0 + \beta_1 \underset{\sim}{a_0} + \boldsymbol{\beta}_2^T \mathbf{c})}{1+\exp(\beta_0 + \beta_1 \underset{\sim}{a_0} + \boldsymbol{\beta}_2^T \mathbf{c})} \\
d_{8,PNDE} &= \underset{\sim}{0} \\
(a_1 - a_0)\Gamma_{PNDE} &= \frac{\partial PNDE}{\partial(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T)^T} \\
&= (a_1 - a_0) \begin{bmatrix} d_{1,PNDE} \\ d_{2,PNDE} \\ d_{3,PNDE} \\ d_{4,PNDE} \\ d_{5,PNDE} \\ d_{6,PNDE} \\ d_{7,PNDE} \\ d_{8,PNDE} \end{bmatrix}
\end{aligned}
$$

$$
\begin{aligned}
Q_{TNIE} &= \frac{\exp(\beta_0 + \beta_1 a_1 + \boldsymbol{\beta}_2^T \mathbf{c})}{\left\{1 + \exp(\beta_0 + \beta_1 a_1 + \boldsymbol{\beta}_2^T \mathbf{c})\right\}^2} \\
B_{TNIE} &= \frac{\exp(\beta_0 + \beta_1 a_0 + \boldsymbol{\beta}_2^T \mathbf{c})}{\left\{1 + \exp(\beta_0 + \beta_1 a_0 + \boldsymbol{\beta}_2^T \mathbf{c})\right\}^2} \\
K_{TNIE} &= \frac{\exp(\beta_0 + \beta_1 a_1 + \boldsymbol{\beta}_2^T \mathbf{c})}{1 + \exp(\beta_0 + \beta_1 a_1 + \boldsymbol{\beta}_2^T \mathbf{c})} \\
D_{TNIE} &= \frac{\exp(\beta_0 + \beta_1 a_0 + \boldsymbol{\beta}_2^T \mathbf{c})}{1 + \exp(\beta_0 + \beta_1 a_0 + \boldsymbol{\beta}_2^T \mathbf{c})} \\
d_{1,TNIE} &= (\theta_2 + \theta_3 \underset{\sim}{a_1})(Q_{TNIE} - B_{TNIE}) \\
d_{2,TNIE} &= (\theta_2 + \theta_3 \underset{\sim}{a_1})(a_1 Q_{TNIE} - a_0 B_{TNIE}) \\
d_{3,TNIE} &= (\theta_2 + \theta_3 \underset{\sim}{a_1})\mathbf{c}(Q_{TNIE} - B_{TNIE}) \\
d_{4,TNIE} &= 0 \\
d_{5,TNIE} &= 0 \\
d_{6,TNIE} &= K_{TNIE} - D_{TNIE} \\
d_{7,TNIE} &= \underset{\sim}{a_1}(K_{TNIE} - D_{TNIE}) \\
d_{8,TNIE} &= \underline{0} \\
&\quad \text{Note the lack of the common factor } (a_1 - a_0) \\
\Gamma_{TNIE} &= \frac{\partial TNIE}{\partial(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T)^T} \\
&= \begin{bmatrix} d_{1,TNIE} \\ d_{2,TNIE} \\ d_{3,TNIE} \\ d_{4,TNIE} \\ d_{5,TNIE} \\ d_{6,TNIE} \\ d_{7,TNIE} \\ d_{8,TNIE} \end{bmatrix}
\end{aligned}
$$

$$
\begin{aligned}
d_{1,TNDE} &= \theta_3 \frac{\exp(\beta_1 + \beta_1 \underset{\sim}{a_1} + \boldsymbol{\beta}_2^T \mathbf{c})}{1 + \exp(\beta_1 + \beta_1 \underset{\sim}{a_1} + \boldsymbol{\beta}_2^T \mathbf{c})} \\
d_{2,TNDE} &= \underset{\sim}{a_1} \theta_3 \frac{\exp(\beta_1 + \beta_1 \underset{\sim}{a_1} + \boldsymbol{\beta}_2^T \mathbf{c})}{1 + \exp(\beta_1 + \beta_1 \underset{\sim}{a_1} + \boldsymbol{\beta}_2^T \mathbf{c})} \\
d_{3,TNDE} &= \mathbf{c} \theta_3 \frac{\exp(\beta_1 + \beta_1 \underset{\sim}{a_1} + \boldsymbol{\beta}_2^T \mathbf{c})}{1 + \exp(\beta_1 + \beta_1 \underset{\sim}{a_1} + \boldsymbol{\beta}_2^T \mathbf{c})} \\
d_{4,TNDE} &= 0 \\
d_{5,TNDE} &= 1 \\
d_{6,TNDE} &= 0 \\
d_{7,TNDE} &= \frac{\exp(\beta_1 + \beta_1 \underset{\sim}{a_1} + \boldsymbol{\beta}_2^T \mathbf{c})}{1 + \exp(\beta_1 + \beta_1 \underset{\sim}{a_1} + \boldsymbol{\beta}_2^T \mathbf{c})}
\end{aligned}
$$

$$
\begin{aligned}
d_{8,TNDE} &= \underset{\sim}{0} \\
(a_1 - a_0)\Gamma_{TNDE} &= \frac{\partial TNDE}{\partial(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T)^T} \\
&= (a_1 - a_0)\begin{bmatrix} d_{1,TNDE} \\ d_{2,TNDE} \\ d_{3,TNDE} \\ d_{4,TNDE} \\ d_{5,TNDE} \\ d_{6,TNDE} \\ d_{7,TNDE} \\ d_{8,TNDE} \end{bmatrix}
\end{aligned}
$$

$$
\begin{aligned}
Q_{PNIE} &= \frac{\exp(\beta_0 + \beta_1 a_1 + \boldsymbol{\beta}_2^T \mathbf{c})}{\left\{1 + \exp(\beta_0 + \beta_1 a_1 + \boldsymbol{\beta}_2^T \mathbf{c})\right\}^2} \\
B_{PNIE} &= \frac{\exp(\beta_0 + \beta_1 a_0 + \boldsymbol{\beta}_2^T \mathbf{c})}{\left\{1 + \exp(\beta_0 + \beta_1 a_0 + \boldsymbol{\beta}_2^T \mathbf{c})\right\}^2} \\
K_{PNIE} &= \frac{\exp(\beta_0 + \beta_1 a_1 + \boldsymbol{\beta}_2^T \mathbf{c})}{1 + \exp(\beta_0 + \beta_1 a_1 + \boldsymbol{\beta}_2^T \mathbf{c})} \\
D_{PNIE} &= \theta_3 \frac{\exp(\beta_0 + \beta_1 a_0 + \boldsymbol{\beta}_2^T \mathbf{c})}{1 + \exp(\beta_0 + \beta_1 a_0 + \boldsymbol{\beta}_2^T \mathbf{c})} \\
d_{1,PNIE} &= (\theta_2 + \theta_3 \underset{\sim}{a_0})(Q_{PNIE} - B_{PNIE}) \\
d_{2,PNIE} &= (\theta_2 + \theta_3 \underset{\sim}{a_0})(a_1 Q_{PNIE} - a_0 B_{PNIE}) \\
d_{3,PNIE} &= (\theta_2 + \theta_3 \underset{\sim}{a_0})\mathbf{c}(Q_{PNIE} - B_{PNIE}) \\
d_{4,PNIE} &= 0 \\
d_{5,PNIE} &= 0 \\
d_{6,PNIE} &= K_{PNIE} - D_{PNIE} \\
d_{7,PNIE} &= \underset{\sim}{a_0}(K_{PNIE} - D_{PNIE}) \\
d_{8,PNIE} &= \underset{\sim}{0} \\
&\quad \text{Note the lack of the common factor } (a_1 - a_0) \\
\Gamma_{PNIE} &= \frac{\partial PNIE}{\partial(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T)^T} \\
&= \begin{bmatrix} d_{1,PNIE} \\ d_{2,PNIE} \\ d_{3,PNIE} \\ d_{4,PNIE} \\ d_{5,PNIE} \\ d_{6,PNIE} \\ d_{7,PNIE} \\ d_{8,PNIE} \end{bmatrix}
\end{aligned}
$$

$$
\begin{aligned}
&\quad \text{Note the lack of the common factor } (a_1 - a_0) \\
\Gamma_{TE} &= \frac{\partial TE}{\partial(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T)^T} \\
&= \frac{\partial(PNDE + TNIE)}{\partial(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T)^T}
\end{aligned}
$$

$$= (a_1 - a_0)\Gamma_{PNDE} + \Gamma_{TNIE}$$

Note the lack of the common factor $(a_1 - a_0)$

$$\Gamma_{PM} = \frac{\partial PM}{\partial(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T)^T}$$

By multivariate chain rule

$$= \frac{\partial PM}{\partial PNDE}\frac{\partial PNDE}{\partial(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T)^T} + \frac{\partial PM}{\partial TNIE}\frac{\partial TNIE}{\partial(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T)^T}$$
$$= \frac{\partial PM}{\partial PNDE}(a_1 - a_0)\Gamma_{PNDE} + \frac{\partial PM}{\partial TNIE}\Gamma_{TNIE}$$
$$= \frac{-TNIE}{(PNDE + TNIE)^2}(a_1 - a_0)\Gamma_{PNDE} + \frac{PNDE}{(PNDE + TNIE)^2}\Gamma_{TNIE}$$
$$= \frac{-TNIE\ (a_1 - a_0)\Gamma_{PNDE} + PNDE\ \Gamma_{TNIE}}{(PNDE + TNIE)^2}$$

Variance-covariance matrix from two models

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma_\beta} & 0 \\ 0 & \boldsymbol{\Sigma_\theta} \end{bmatrix}$$
$$SE(\widehat{CDE}(m)) = \sqrt{\Gamma^T_{CDE(m)}\ \boldsymbol{\Sigma}\ \Gamma_{CDE(m)}}\ |a_1 - a_0|$$
$$SE(\widehat{PNDE}) = \sqrt{\Gamma^T_{PNDE}\ \boldsymbol{\Sigma}\ \Gamma_{PNDE}}\ |a_1 - a_0|$$
$$SE(\widehat{TNIE}) = \sqrt{\Gamma^T_{TNIE}\ \boldsymbol{\Sigma}\ \Gamma_{TNIE}}$$
$$SE(\widehat{TNDE}) = \sqrt{\Gamma^T_{TNDE}\ \boldsymbol{\Sigma}\ \Gamma_{TNDE}}\ |a_1 - a_0|$$
$$SE(\widehat{PNIE}) = \sqrt{\Gamma^T_{PNIE}\ \boldsymbol{\Sigma}\ \Gamma_{PNIE}}$$
$$SE(\widehat{TE}) = \sqrt{\Gamma^T_{TE}\ \boldsymbol{\Sigma}\ \Gamma_{TE}}$$
$$SE(\widehat{PM}) = \sqrt{\Gamma^T_{PM}\ \boldsymbol{\Sigma}\ \Gamma_{PM}}$$

2.6 Logistic mediator model, non-linear outcome model

2.6.1 Effect formulas

The function `calc_myreg_mreg_logistic_yreg_logistic_est()` implements the effect formulas in [VanderWeele, 2015] (p473).

Models

$$\text{logit}(E[Y|A = a, M = m, \mathbf{C} = \mathbf{c}]) = \theta_0 + \theta_1 a + \theta_2 m + \theta_3 am + \boldsymbol{\theta}_4^T \mathbf{c}$$
$$\text{logit}(E[M|A = a, \mathbf{C} = \mathbf{c}]) = \beta_0 + \beta_1 a + \boldsymbol{\beta}_2^T \mathbf{c}$$

Effects on link function scale

$$CDE(m) = \text{logit}(E[Y_{a_1,m}|\mathbf{C} = \mathbf{c}]) - \text{logit}(E[Y_{a_0,m}|\mathbf{C} = \mathbf{c}])$$
$$= (\theta_1 + \theta_3 m)(a_1 - a_0)$$

$$PNDE = \text{logit}(E[Y_{a_1,\utilde{M_{a_0}}}|\mathbf{C} = \mathbf{c}]) - \text{logit}(E[Y_{a_0,\utilde{M_{a_0}}}|\mathbf{C} = \mathbf{c}])$$
$$\approx \theta_1(a_1 - a_0)$$

$$
\begin{aligned}
&+ \log(1 + \exp(\theta_2 + \theta_3 a_1 + \beta_0 + \beta_1 \underset{\sim}{a_0} + \boldsymbol{\beta}_2^T \mathbf{c})) \\
&- \log(1 + \exp(\theta_2 + \theta_3 a_0 + \beta_0 + \beta_1 \underset{\sim}{a_0} + \boldsymbol{\beta}_2^T \mathbf{c}))
\end{aligned}
$$

$$
\begin{aligned}
TNIE &= \text{logit}(E[Y_{\underset{\sim}{a_1}, M_{a_1}} | \mathbf{C} = \mathbf{c}]) - \text{logit}(E[Y_{\underset{\sim}{a_1}, M_{a_0}} | \mathbf{C} = \mathbf{c}]) \\
&\approx \log(1 + \exp(\beta_0 + \beta_1 a_0 + \boldsymbol{\beta}_2^T \mathbf{c})) \\
&\quad - \log(1 + \exp(\beta_0 + \beta_1 a_1 + \boldsymbol{\beta}_2^T \mathbf{c})) \\
&\quad + \log(1 + \exp(\theta_2 + \theta_3 \underset{\sim}{a_1} + \beta_0 + \beta_1 a_1 + \boldsymbol{\beta}_2^T \mathbf{c})) \\
&\quad - \log(1 + \exp(\theta_2 + \theta_3 \underset{\sim}{a_1} + \beta_0 + \beta_1 a_0 + \boldsymbol{\beta}_2^T \mathbf{c}))
\end{aligned}
$$

Note the $a_0 \to a_1$ changes associated with β_1.

$$
\begin{aligned}
TNDE &= \text{logit}(E[Y_{a_1, \underset{\sim}{M_{a_1}}} | \mathbf{C} = \mathbf{c}]) - \text{logit}(E[Y_{a_0, \underset{\sim}{M_{a_1}}} | \mathbf{C} = \mathbf{c}]) \\
&\approx \theta_1 (a_1 - a_0) \\
&\quad + \log(1 + \exp(\theta_2 + \theta_3 a_1 + \beta_0 + \beta_1 \underset{\sim}{a_1} + \boldsymbol{\beta}_2^T \mathbf{c})) \\
&\quad - \log(1 + \exp(\theta_2 + \theta_3 a_0 + \beta_0 + \beta_1 \underset{\sim}{a_1} + \boldsymbol{\beta}_2^T \mathbf{c}))
\end{aligned}
$$

Note the $a_1 \to a_0$ changes associated with θ_3.

$$
\begin{aligned}
PNIE &= \text{logit}(E[Y_{\underset{\sim}{a_0}, M_{a_1}} | \mathbf{C} = \mathbf{c}]) - \text{logit}(E[Y_{\underset{\sim}{a_0}, M_{a_0}} | \mathbf{C} = \mathbf{c}]) \\
&\approx \log(1 + \exp(\beta_0 + \beta_1 a_0 + \boldsymbol{\beta}_2^T \mathbf{c})) \\
&\quad - \log(1 + \exp(\beta_0 + \beta_1 a_1 + \boldsymbol{\beta}_2^T \mathbf{c})) \\
&\quad + \log(1 + \exp(\theta_2 + \theta_3 \underset{\sim}{a_0} + \beta_0 + \beta_1 a_1 + \boldsymbol{\beta}_2^T \mathbf{c})) \\
&\quad - \log(1 + \exp(\theta_2 + \theta_3 \underset{\sim}{a_0} + \beta_0 + \beta_1 a_0 + \boldsymbol{\beta}_2^T \mathbf{c}))
\end{aligned}
$$

$$
\begin{aligned}
TE &= PNDE + TNIE \\
PM &= \frac{\exp(PNDE)(\exp(TNIE) - 1)}{\exp(PNDE)\exp(TNIE) - 1}
\end{aligned}
$$

2.6.2 Variance formulas

The function `calc_myreg_mreg_logistic_yreg_logistic_se()` implements the standard error formulas in [VanderWeele, 2015] (p473).

Note the lack of the common factor $(a_1 - a_0)$ throughout.

$$
\begin{aligned}
\Gamma_{CDE(m)} &= \frac{\partial CDE(m)}{\partial (\boldsymbol{\beta}^T, \boldsymbol{\theta}^T)^T} \\
&= (a_1 - a_0) \begin{bmatrix} 0 \\ 0 \\ \underline{0} \\ 0 \\ (a_1 - a_0) \\ 0 \\ m(a_1 - a_0) \\ \underline{0} \end{bmatrix}
\end{aligned}
$$

$$\begin{aligned}
Q_{PNDE} &= \frac{\exp(\theta_2 + \theta_3 a_1 + \beta_0 + \beta_1 \underset{\sim}{a_0} + \boldsymbol{\beta}_2^T \mathbf{c})}{1 + \exp(\theta_2 + \theta_3 a_1 + \beta_0 + \beta_1 \underset{\sim}{a_0} + \boldsymbol{\beta}_2^T \mathbf{c})} \\
B_{PNDE} &= \frac{\exp(\theta_2 + \theta_3 a_0 + \beta_0 + \beta_1 \underset{\sim}{a_0} + \boldsymbol{\beta}_2^T \mathbf{c})}{1 + \exp(\theta_2 + \theta_3 a_0 + \beta_0 + \beta_1 \underset{\sim}{a_0} + \boldsymbol{\beta}_2^T \mathbf{c})} \\
d_{1,PNDE} &= Q_{PNDE} - B_{PNDE} \\
d_{2,PNDE} &= \underset{\sim}{a_0}(Q_{PNDE} - B_{PNDE}) \\
d_{3,PNDE} &= \mathbf{c}(Q_{PNDE} - B_{PNDE}) \\
d_{4,PNDE} &= 0 \\
d_{5,PNDE} &= a_1 - a_0 \\
d_{6,PNDE} &= Q_{PNDE} - B_{PNDE} \\
d_{7,PNDE} &= a_1 Q_{PNDE} - a_0 B_{PNDE} \\
d_{8,PNDE} &= \underline{0} \\
\Gamma_{PNDE} &= \frac{\partial PNDE}{\partial(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T)^T} \\
&= \begin{bmatrix} d_{1,PNDE} \\ d_{2,PNDE} \\ d_{3,PNDE} \\ d_{4,PNDE} \\ d_{5,PNDE} \\ d_{6,PNDE} \\ d_{7,PNDE} \\ d_{8,PNDE} \end{bmatrix}
\end{aligned}$$

$$\begin{aligned}
Q_{TNIE} &= \frac{\exp(\theta_2 + \theta_3 \underset{\sim}{a_1} + \beta_0 + \beta_1 a_1 + \boldsymbol{\beta}_2^T \mathbf{c})}{1 + \exp(\theta_2 + \theta_3 \underset{\sim}{a_1} + \beta_0 + \beta_1 a_1 + \boldsymbol{\beta}_2^T \mathbf{c})} \\
B_{TNIE} &= \frac{\exp(\theta_2 + \theta_3 \underset{\sim}{a_1} + \beta_0 + \beta_1 a_0 + \boldsymbol{\beta}_2^T \mathbf{c})}{1 + \exp(\theta_2 + \theta_3 \underset{\sim}{a_1} + \beta_0 + \beta_1 a_0 + \boldsymbol{\beta}_2^T \mathbf{c})} \\
K_{TNIE} &= \frac{\exp(\beta_0 + \beta_1 a_1 + \boldsymbol{\beta}_2^T \mathbf{c})}{1 + \exp(\beta_0 + \beta_1 a_1 + \boldsymbol{\beta}_2^T \mathbf{c})} \\
D_{TNIE} &= \theta_3 \frac{\exp(\beta_0 + \beta_1 a_0 + \boldsymbol{\beta}_2^T \mathbf{c})}{1 + \exp(\beta_0 + \beta_1 a_0 + \boldsymbol{\beta}_2^T \mathbf{c})} \\
d_{1,TNIE} &= (D_{TNIE} + Q_{TNIE}) - (K_{TNIE} + B_{TNIE}) \\
d_{2,TNIE} &= a_0(D_{TNIE} - B_{TNIE}) + a_1(Q_{TNIE} - K_{TNIE}) \\
d_{3,TNIE} &= \mathbf{c}\left\{(D_{TNIE} + Q_{TNIE}) - (K_{TNIE} + B_{TNIE})\right\} \\
d_{4,TNIE} &= 0 \\
d_{5,TNIE} &= 0 \\
d_{6,TNIE} &= Q_{TNIE} - B_{TNIE} \\
d_{7,TNIE} &= \underset{\sim}{a_1}(Q_{TNIE} - B_{TNIE}) \\
d_{8,TNIE} &= \underline{0} \\
\Gamma_{TNIE} &= \frac{\partial TNIE}{\partial(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T)^T}
\end{aligned}$$

$$
= \begin{bmatrix} d_{1,TNIE} \\ d_{2,TNIE} \\ d_{3,TNIE} \\ d_{4,TNIE} \\ d_{5,TNIE} \\ d_{6,TNIE} \\ d_{7,TNIE} \\ d_{8,TNIE} \end{bmatrix}
$$

$$
\begin{aligned}
Q_{TNDE} &= \frac{\exp(\theta_2 + \theta_3 a_1 + \beta_0 + \beta_1 \underset{\sim}{a_1} + \boldsymbol{\beta}_2^T \mathbf{c})}{1 + \exp(\theta_2 + \theta_3 a_1 + \beta_0 + \beta_1 \underset{\sim}{a_1} + \boldsymbol{\beta}_2^T \mathbf{c})} \\
B_{TNDE} &= \frac{\exp(\theta_2 + \theta_3 a_0 + \beta_0 + \beta_1 \underset{\sim}{a_1} + \boldsymbol{\beta}_2^T \mathbf{c})}{1 + \exp(\theta_2 + \theta_3 a_0 + \beta_0 + \beta_1 \underset{\sim}{a_1} + \boldsymbol{\beta}_2^T \mathbf{c})} \\
d_{1,TNDE} &= Q_{TNDE} - B_{TNDE} \\
d_{2,TNDE} &= \underset{\sim}{a_1}(Q_{TNDE} - B_{TNDE}) \\
d_{3,TNDE} &= \mathbf{c}(Q_{TNDE} - B_{TNDE}) \\
d_{4,TNDE} &= 0 \\
d_{5,TNDE} &= a_1 - a_0 \\
d_{6,TNDE} &= Q_{TNDE} - B_{TNDE} \\
d_{7,TNDE} &= a_1 Q_{TNDE} - a_0 B_{TNDE} \\
d_{8,TNDE} &= \underline{0} \\
\Gamma_{TNDE} &= \frac{\partial TNDE}{\partial(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T)^T} \\
&= \begin{bmatrix} d_{1,TNDE} \\ d_{2,TNDE} \\ d_{3,TNDE} \\ d_{4,TNDE} \\ d_{5,TNDE} \\ d_{6,TNDE} \\ d_{7,TNDE} \\ d_{8,TNDE} \end{bmatrix}
\end{aligned}
$$

$$
\begin{aligned}
Q_{PNIE} &= \frac{\exp(\theta_2 + \theta_3 \underset{\sim}{a_0} + \beta_0 + \beta_1 a_1 + \boldsymbol{\beta}_2^T \mathbf{c})}{1 + \exp(\theta_2 + \theta_3 \underset{\sim}{a_0} + \beta_0 + \beta_1 a_1 + \boldsymbol{\beta}_2^T \mathbf{c})} \\
B_{PNIE} &= \frac{\exp(\theta_2 + \theta_3 \underset{\sim}{a_0} + \beta_0 + \beta_1 a_0 + \boldsymbol{\beta}_2^T \mathbf{c})}{1 + \exp(\theta_2 + \theta_3 \underset{\sim}{a_0} + \beta_0 + \beta_1 a_0 + \boldsymbol{\beta}_2^T \mathbf{c})} \\
K_{PNIE} &= \frac{\exp(\beta_0 + \beta_1 a_1 + \boldsymbol{\beta}_2^T \mathbf{c})}{1 + \exp(\beta_0 + \beta_1 a_1 + \boldsymbol{\beta}_2^T \mathbf{c})} \\
D_{PNIE} &= \theta_3 \frac{\exp(\beta_0 + \beta_1 a_0 + \boldsymbol{\beta}_2^T \mathbf{c})}{1 + \exp(\beta_0 + \beta_1 a_0 + \boldsymbol{\beta}_2^T \mathbf{c})} \\
d_{1,PNIE} &= (D_{PNIE} + Q_{PNIE}) - (K_{PNIE} + B_{PNIE}) \\
d_{2,PNIE} &= a_0(D_{PNIE} - B_{PNIE}) + a_1(Q_{PNIE} - K_{PNIE}) \\
d_{3,PNIE} &= \mathbf{c}\left\{(D_{PNIE} + Q_{PNIE}) - (K_{PNIE} + B_{PNIE})\right\} \\
d_{4,PNIE} &= 0
\end{aligned}
$$

$$
\begin{aligned}
d_{5,PNIE} &= 0 \\
d_{6,PNIE} &= Q_{PNIE} - B_{PNIE} \\
d_{7,PNIE} &= \underset{\sim}{a_0}(Q_{PNIE} - B_{PNIE}) \\
d_{8,PNIE} &= \underset{\sim}{0} \\
\Gamma_{PNIE} &= \frac{\partial PNIE}{\partial(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T)^T} \\
&= \begin{bmatrix} d_{1,PNIE} \\ d_{2,PNIE} \\ d_{3,PNIE} \\ d_{4,PNIE} \\ d_{5,PNIE} \\ d_{6,PNIE} \\ d_{7,PNIE} \\ d_{8,PNIE} \end{bmatrix}
\end{aligned}
$$

$$
\begin{aligned}
\Gamma_{TE} &= \frac{\partial TE}{\partial(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T)^T} \\
&= \frac{\partial(PNDE + TNIE)}{\partial(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T)^T} \\
&= \Gamma_{PNDE} + \Gamma_{TNIE}
\end{aligned}
$$

$$
\begin{aligned}
\Gamma_{PM} &= \frac{\partial PM}{\partial(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T)^T} \\
&\quad \text{By multivariate chain rule} \\
&= \frac{\partial PM}{\partial PNDE} \frac{\partial PNDE}{\partial(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T)^T} + \frac{\partial PM}{\partial TNIE} \frac{\partial TNIE}{\partial(\boldsymbol{\beta}^T, \boldsymbol{\theta}^T)^T} \\
&= \frac{\partial PM}{\partial PNDE} \Gamma_{PNDE} + \frac{\partial PM}{\partial TNIE} \Gamma_{TNIE} \\
&= -\frac{\exp(PNDE)\left\{\exp(TNIE) - 1\right\}}{\left\{\exp(PNDE)\exp(TNIE) - 1\right\}^2} \Gamma_{PNDE} \\
&\quad + \frac{\exp(PNDE)\exp(TNIE)\left\{\exp(PNDE) - 1\right\}}{\left\{\exp(PNDE)\exp(TNIE) - 1\right\}^2} \Gamma_{TNIE}
\end{aligned}
$$

Variance-covariance matrix from two models

$$
\begin{aligned}
\boldsymbol{\Sigma} &= \begin{bmatrix} \boldsymbol{\Sigma}_{\boldsymbol{\beta}} & 0 \\ 0 & \boldsymbol{\Sigma}_{\boldsymbol{\theta}} \end{bmatrix} \\
SE(\widehat{CDE}(m)) &= \sqrt{\Gamma_{CDE(m)}^T \ \boldsymbol{\Sigma} \ \Gamma_{CDE(m)}} \\
SE(\widehat{PNDE}) &= \sqrt{\Gamma_{PNDE}^T \ \boldsymbol{\Sigma} \ \Gamma_{PNDE}} \\
SE(\widehat{TNIE}) &= \sqrt{\Gamma_{TNIE}^T \ \boldsymbol{\Sigma} \ \Gamma_{TNIE}} \\
SE(\widehat{TNDE}) &= \sqrt{\Gamma_{TNDE}^T \ \boldsymbol{\Sigma} \ \Gamma_{TNDE}} \\
SE(\widehat{PNIE}) &= \sqrt{\Gamma_{PNIE}^T \ \boldsymbol{\Sigma} \ \Gamma_{PNIE}} \\
SE(\widehat{TE}) &= \sqrt{\Gamma_{TE}^T \ \boldsymbol{\Sigma} \ \Gamma_{TE}} \\
SE(\widehat{PM}) &= \sqrt{\Gamma_{PM}^T \ \boldsymbol{\Sigma} \ \Gamma_{PM}}
\end{aligned}
$$

3 Software

The detailed explanation of the software implementation and its use are given on its website (`https://kaz-yos.github.io/regmedint/index.html`).

4 Bibliography

[Hoef, 2012] Hoef, J. M. V. (2012). Who Invented the Delta Method? *The American Statistician*, 66(2):124–127. _eprint: https://doi.org/10.1080/00031305.2012.687494.

[Pearl, 2001] Pearl, J. (2001). Direct and Indirect Effects. In *Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence*, UAI'01, pages 411–420, San Francisco, CA, USA. Morgan Kaufmann Publishers Inc.

[Robins and Greenland, 1992] Robins, J. M. and Greenland, S. (1992). Identifiability and exchangeability for direct and indirect effects. *Epidemiology*, 3(2):143–155.

[Starkopf et al., 2017] Starkopf, L., Andersen, M. P., Gerds, T. A., Torp-Pedersen, C., and Lange, T. (2017). Comparison of five software solutions to mediation analysis. page 77.

[Valeri and Vanderweele, 2013] Valeri, L. and Vanderweele, T. J. (2013). Mediation analysis allowing for exposure-mediator interactions and causal interpretation: Theoretical assumptions and implementation with SAS and SPSS macros. *Psychol Methods*, 18(2):137–150.

[Valeri and VanderWeele, 2015] Valeri, L. and VanderWeele, T. J. (2015). SAS macro for causal mediation analysis with survival data. *Epidemiology*, 26(2):e23–24.

[VanderWeele, 2015] VanderWeele, T. (2015). *Explanation in Causal Inference: Methods for Mediation and Interaction*. Oxford University Press, New York, 1 edition edition.

[VanderWeele, 2013] VanderWeele, T. J. (2013). A three-way decomposition of a total effect into direct, indirect, and interactive effects. *Epidemiology*, 24(2):224–232.