Mediation analysis

February 9, 2023

Denote by Y an oucome, M a mediator, E an exposure, and C possible confounders (between Y and E, or Y and M, or E and M).

- $Y(e,m) = Y(e,m,c)$ will be the counterfactual outcome, i.e. the outcome value had the exposure be set to e, the mediator to m, and the confounder set to c.
- $M(e) = M(e,c)$ will be the counterfactual mediator, i.e. the mediator value had the exposure be set to e and the confounder set to c.

We define (all at a fixed confounder value c):

- the total effect as $Y(e, M(e)) - Y(e^*, M(e^*))$
- the natural direct effect as $Y(e, M(e)) - Y(e^*, M(e))$
- the indirect direct effect as $Y(e^*, M(e)) - Y(e^*, M(e^*))$

1 Binary outcome, continuous exposure

1.1 Theory

We consider the following models:

$$\begin{aligned}
\operatorname{logit}\left(\mathbb{P}\left[Y=1|E,M,C\right]\right) &= \beta_0 + \beta_1 E + \beta_2 M + \beta_3 C \\
\mathbb{E}\left[M|E,C\right] &= \alpha_0 + \alpha_1 E + \alpha_3 C
\end{aligned}$$

Therefore the counterfactual probability can be expressed as:

$$\begin{aligned}
Y(e, M(e^*), c) &= \frac{1}{1+\exp^{-\beta_0-\beta_1 e-\beta_2 M(e^*)+\beta_3 c}} \\
&= \frac{1}{1+\exp^{-(\beta_0+\beta_2\alpha_0)-(\beta_1 e+\beta_2\alpha_1 e^*)+(\beta_3+\beta_2\alpha_3)c}} \\
\operatorname{logit}\left(\mathbb{P}\left[Y(e, M(e^*), c)\right]\right) &= (\beta_0+\beta_2\alpha_0) + (\beta_1 e + \beta_2\alpha_1 e^*) + (\beta_3+\beta_2\alpha_3)c
\end{aligned}$$

So the odd for the total effect is:

$$OR^{TE} = \frac{\mathbb{P}\left[Y(e, M(e))\right] / \mathbb{P}\left[Y(e, M(e))\right]}{\mathbb{P}\left[Y(e^*, M(e^*))\right] / \mathbb{P}\left[Y(e^*, M(e^*))\right]} = \exp((\beta_1 + \beta_2\alpha_1)(e - e^*))$$

for the natural direct effect:

$$OR^{NDE} = \frac{\mathbb{P}\left[Y(e, M(e))\right] / \mathbb{P}\left[Y(e, M(e))\right]}{\mathbb{P}\left[Y(e^*, M(e))\right] / \mathbb{P}\left[Y(e^*, M(e))\right]} = \exp(\beta_1(e - e^*))$$

for the natural indirect effect:

$$OR^{NIE} = \frac{\mathbb{P}\left[Y(e^*, M(e))\right] / \mathbb{P}\left[Y(e^*, M(e))\right]}{\mathbb{P}\left[Y(e^*, M(e^*))\right] / \mathbb{P}\left[Y(e^*, M(e^*))\right]} = \exp(\beta_2\alpha_1(e - e^*))$$

This matches formula in VanderWeele and Vansteelandt (2010)

Note that we can also express the total effect, natural direct effect, and natural indirect effect on the probability scale, e.g.:

$$\begin{aligned}
&\mathbb{P}\left[Y(e, M(e), c) = 1\right] - \mathbb{P}\left[Y(e^*, M(e), c) = 1\right] \\
&= \frac{1}{1 + \exp^{-(\beta_0+\beta_2\alpha_0)-(\beta_1 e+\beta_2\alpha_1 e)+(\beta_3+\beta_2\alpha_3)c}} - \frac{1}{1 + \exp^{-(\beta_0+\beta_2\alpha_0)-(\beta_1 e^*+\beta_2\alpha_1 e)+(\beta_3+\beta_2\alpha_3)c}}
\end{aligned}$$

which will be a function of the confounder value. To get a single estimate we could average it out over the confounder distribution in our population.

1.2 Practice

```
library(medflex)
data(UPBdata)
head(UPBdata)
sum(is.na(UPBdata))
```

```
         att attbin attcat     negaff  initiator gender educ age UPB
1  1.0005617      1      M  0.8404610     myself      F    M  41   1
2 -0.7085889      0      L -1.2574650       both      M    M  42   0
3 -0.7085889      0      L -1.2022564       both      F    H  43   0
4  0.6061423      1      M -0.3741277 ex-partner      M    H  52   1
5  0.2117230      1      M  1.9446325 ex-partner      M    M  32   1
6  2.0523467      1      H -0.8157964 ex-partner      M    H  47   0
[1] 0
```

Manually

```
e.lm <- glm(negaff ~ factor(attbin) + gender + educ + age, data = UPBdata)
e.logit <- glm(UPB ~ attbin + negaff + gender + educ + age, data = UPBdata
   )
```

```
expData <- neWeight(negaff ~ factor(attbin) + gender + educ + age, family
    = gaussian, data = UPBdata)
neMod1 <- neModel(UPB ~ attbin0 + attbin1 + gender + educ + age, family =
    binomial("logit"), expData = expData)
```

```
exp(cbind(estimate = coef(neMod1),confint(neMod1))[c("attbin01", "attbin11
    "), ])
```

```
         estimate  95% LCL  95% UCL
attbin01 1.485757 0.946311 2.294548
attbin11 1.421865 1.188970 1.678737
```

2 References

VanderWeele, T. J. and Vansteelandt, S. (2010). Odds ratios for mediation analysis for a dichotomous outcome. *American journal of epidemiology*, 172(12):1339–1348.