Online Resource 2

Electronic supporting materials for "Peer victimization and suicidal thoughts and behaviors in depressed adolescents" submitted to *Journal of Abnormal Child Psychology*.

Caption: Mathematical derivation of a multiple mediation model in which one mediator is a count variable and the other is linear with a count outcome under the counterfactual framework.

Definition of causal effects under the counterfactual framework for a single mediator

We let Y_a and M_a denote respectively the values of the outcome and mediator that would have been observed had the exposure A been set to level a. We let Y_{am} denote the value of the outcome that would have been observed had the exposure, A, and mediator, M, been set to levels a and m, respectively.

The average controlled direct effect comparing exposure level a to a^* and fixing the mediator to level m is defined by $CDE_{a,a^*}(m) = E[Y_{am} - Y_{a^*m}]$. The average natural direct effect is then defined by $NDE_{a,a^*}(a^*) = E[Y_{aM_{a^*}} - Y_{a^*M_{a^*}}]$. The average natural indirect effect can be defined as $NIE_{a,a^*}(a) = E[Y_{aM_a} - Y_{aM_{a^*}}]$, which compares the effect of the mediator at levels M_a and M_{a^*} on the outcome when exposure A is set to a. Controlled direct effects and natural direct and indirect effects within strata of $C = c$ are then defined by: $CDE_{a,a^*|c}(m) = E[Y_{am} - Y_{a^*m}|c]$, $NDE_{a,a^*|c}(a^*) = E[Y_{aM_{a^*}} - Y_{a^*M_{a^*}}|c]$ and $NIE_{a,a^*|c}(a) = E[Y_{aM_a} - Y_{aM_{a^*}}|c]$ respectively.

If we let $X \perp Y|Z$ denote that X is independent of Y conditional on Z then the identification assumptions for the causal effects previously defined can be expressed formally in terms of counterfactual independence as (i) $Y_{am} \perp A|C$, (ii) $Y_{am} \perp M|\{A, C\}$, (iii) $M_a \perp A|C$, and (iv) $Y_{am} \perp M_{a^*}|C$. Assumptions (i) and (ii) suffice to identify controlled direct effects; assumptions (i)-(iv) suffice to identify natural direct and indirect effects (Pearl, 2001; VanderWeele and Vansteelandt, 2009). The intuitive interpretation of these assuptions follows from the theory of causal diagrams (Pearl, 2001). Alternative identification assumptions have also been proposed (Imai 2010a; Hafeman and VanderWeele, 2011). However, it has been shown that the intuitive graphical interpretation of these alternative assumptions are in fact equivalent (Shpitser and VanderWeele, 2011). Technical examples can be constructed where one set of identifiation assumptions holds and another does not, but on a causal diagram corresponding to a set of non-parametric structural equations, whenever one set of the assumptions among those in VanderWeele and Vansteelandt (2009), Imai (2010a), and Hafeman and VanderWeele (2011) holds, the others will also.

Definition of causal effects under the counterfactual framework for multiple mediators

Suppose now that there are multiple mediators of interest, $\mathbf{M} = (M^{(1)}, ..., M^{(K)})$ and that we are interested in the effects mediated through $(M^{(1)}, ..., M^{(K)})$ jointly and the effects independent of $(M^{(1)}, ..., M^{(K)})$. We can define controlled direct effects and natural direct and indirect effects in a similar way as before simply replacing our single mediator $\mathbf{M}$ with the entire vector of mediators $\mathbf{M} = (M^{(1)}, ..., M^{(K)})$. Thus, let $\mathbf{M}_a$ be the counterfactual value of $\mathbf{M}$ if exposure A were set to the value a and let $Y_{a\mathbf{m}}$ denote the counterfactual value for Y if A were set to a and $\mathbf{M}$ were set to $\mathbf{m}$. The controlled direct effect is defined by $Y_{a\mathbf{m}} - Y_{a^*\mathbf{m}}$; the natural direct effect is defined as $Y_{a\mathbf{M}a} - Y_{a^*\mathbf{M}_a}$; the natural indirect effect is defined as $Y_{a\mathbf{M}_a} - Y_{a\mathbf{M}_{a^*}}$; and once again the total effect can be decomposed into a natural

direct and indirect effect. Suppose again that the four assumptions about confounding hold but now with respect to the whole set of mediators $\mathbf{M} = (M^{(1)}, ..., M^{(K)})$. We once again need to control for all exposure-outcome, mediator- outcome, and exposure-mediator confounders, but note that now for the second and third assumptions control must be made for the mediator-outcome confounders for all of the mediators, not just one and likewise control must be made for the exposure-outcome confounders for all of the mediators, not just one. The fourth assumption again requires that there be no effect of the exposure that confounds the mediator-outcome relationship for any of the mediators. If there were such a variable then to proceed it would have to be included in the mediator vector $\mathbf{M}$ if the assumption were not to be violated.

Count Mediators and Outcome with mean modeled using log link

Suppose that two mediators the outcome are count variables and that allowing for exposure-mediator interaction the following models fit the observed data:

$$log[E(M^{(1)}|a,c)] = \beta_0^{(1)} + \beta_1^{(1)}a + \beta_2^{(1)\prime}c \tag{1}$$

$$log[E(M^{(2)}|a,c)] = \beta_0^{(2)} + \beta_1^{(2)}a + \beta_2^{(2)\prime}c \tag{2}$$

$$log[E(Y|a,m^{(1)},m^{(2)},c)] = \theta_0 + \theta_1 a + \theta_2^{(1)}m^{(1)} + \theta_2^{(2)}m^{(2)} + \theta_3^{(1)}a * m^{(1)} + \theta_3^{(2)}a * m^{(2)} + \theta_4^{\prime}c \tag{3}$$

If the covariates C satisfied the no-unmeasured confounding assumptions (i)-(iv) above and the models are correctly specified, then the average natural direct and indirect effects can be derived. Extending the work of Valeri and VanderWeele (2013) on direct and indirect effect estimators when generalized linear models are employed in the presence of exposure-mediator interaction we obtain:

$$\begin{aligned}
&NDE = E[Y_{a\mathbf{M}a} - Y_{a^*\mathbf{M}_a}|C=c] \\
&= exp(\theta_0 + \theta_1 a + \theta_4^{\prime}c)exp\{exp(\beta_0^{(1)} + \beta_1^{(1)}a^* + \beta_2^{(1)\prime}c)[exp(\theta_2^{(1)} + \theta_3^{(1)}a) - 1]\} \\
&\times exp\{exp(\beta_0^{(2)} + \beta_1^{(2)}a^* + \beta_2^{(2)\prime}c)[exp(\theta_2^{(2)} + \theta_3^{(2)}a) - 1]\} + \\
&-exp(\theta_0 + \theta_1 a^* + \theta_4^{\prime}c)exp\{exp(\beta_0^{(1)} + \beta_1^{(1)}a^* + \beta_2^{(1)\prime}c)[exp(\theta_2^{(1)} + \theta_3^{(1)}a^*) - 1]\} \\
&\times exp\{exp(\beta_0^{(2)} + \beta_1^{(2)}a^* + \beta_2^{(2)\prime}c)[exp(\theta_2^{(2)} + \theta_3^{(2)}a^*) - 1]\}.
\end{aligned}$$

Moreover under the same assumptions we can compute the natural indirect effects by:

$$\begin{aligned}
&NIE = E[Y_{a\mathbf{M}a} - Y_{a\mathbf{M}_a^*|C=c}] \\
&= exp(\theta_0 + \theta_1 a + \theta_4^{\prime}c)exp\{exp(\beta_0^{(1)} + \beta_1^{(1)}a + \beta_2^{(1)\prime}c)[exp(\theta_2^{(1)} + \theta_3^{(1)}a) - 1]\} \\
&\times exp\{exp(\beta_0^{(2)} + \beta_1^{(2)}a + \beta_2^{(2)\prime}c)[exp(\theta_2^{(2)} + \theta_3^{(2)}a) - 1]\} + \\
&-exp(\theta_0 + \theta_1 a + \theta_4^{\prime}c)exp\{exp(\beta_0^{(1)} + \beta_1^{(1)}a^* + \beta_2^{(1)\prime}c)[exp(\theta_2^{(1)} + \theta_3^{(1)}a) - 1]\} \\
&\times exp\{exp(\beta_0^{(2)} + \beta_1^{(2)}a^* + \beta_2^{(2)\prime}c)[exp(\theta_2^{(2)} + \theta_3^{(2)}a) - 1]\}.
\end{aligned}$$

Standard errors for these estimators can ben obtained either via bootstrap procedure or by the delta method (Valeri and VanderWeele, 2013).

Count and Continuous Mediators and Count Outcome

Suppose that one mediator and the outcome are count variables modeled using log link and one mediator is continuous modeled with linear link. Allowing for exposure-mediator interaction the following models fit the observed data:

$$log[E(M^{(1)}|a,c)] = \beta_0^{(1)} + \beta_1^{(1)}a + \beta_2^{(1)\prime}c \tag{4}$$

$$E(M^{(2)}|a,c) = \beta_0^{(2)} + \beta_1^{(2)}a + \beta_2^{(2)\prime}c \tag{5}$$

$$log[E(Y|a,m^{(1)},m^{(2)},c)] = \theta_0 + \theta_1 a + \theta_2^{(1)}m^{(1)} + \theta_2^{(2)}m^{(2)} + \theta_3^{(1)}a*m^{(1)} + \theta_3^{(2)}a*m^{(2)} + \theta_4^{\prime}c \tag{6}$$

If the covariates C satisfied the no-unmeasured confounding assumptions (i)-(iv) above and the models are correctly specified, then the average natural direct and indirect effects can be derived.

$$\begin{aligned}
&NDE = E[Y_{a\mathbf{M}a} - Y_{a^*\mathbf{M}_a}|C=c] \\
&= exp(\theta_0 + \theta_1 a + \theta_4^{\prime}c)exp\{exp(\beta_0^{(1)} + \beta_1^{(1)}a^* + \beta_2^{(1)\prime}c)[exp(\theta_2^{(1)} + \theta_3^{(1)}a) - 1]\} \\
&\times exp\{(\beta_0^{(2)} + \beta_1^{(2)}a^* + \beta_2^{(2)\prime}c)(\theta_2^{(2)} + \theta_3^{(2)}a) + \tfrac{1}{2}\sigma^2_{m^{(2)}}(\theta_2^{(2)} + \theta_3^{(2)}a)^2\} + \\
&-exp(\theta_0 + \theta_1 a^* + \theta_4^{\prime}c)exp\{exp(\beta_0^{(1)} + \beta_1^{(1)}a^* + \beta_2^{(1)\prime}c)[exp(\theta_2^{(1)} + \theta_3^{(1)}a^*) - 1]\} \\
&\times exp\{(\beta_0^{(2)} + \beta_1^{(2)}a^* + \beta_2^{(2)\prime}c)(\theta_2^{(2)} + \theta_3^{(2)}a^*) + \tfrac{1}{2}\sigma^2_{m^{(2)}}(\theta_2^{(2)} + \theta_3^{(2)}a^*)^2\}.
\end{aligned}$$

Moreover under the same assumptions we can compute the natural indirect effects by:

$$\begin{aligned}
&NIE = E[Y_{aM_a} - Y_{a^*M_{a^*}}|C=c] \\
&= exp(\theta_0 + \theta_1 a + \theta_4^{\prime}c)exp\{exp(\beta_0^{(1)} + \beta_1^{(1)}a + \beta_2^{(1)\prime}c)[exp(\theta_2^{(1)} + \theta_3^{(1)}a) - 1]\} \\
&exp\{(\beta_0^{(2)} + \beta_1^{(2)}a + \beta_2^{(2)\prime}c)(\theta_2^{(2)} + \theta_3^{(2)}a) + \tfrac{1}{2}\sigma^2_{m^{(2)}}(\theta_2^{(2)} + \theta_3^{(2)}a)^2\} + \\
&-exp(\theta_0 + \theta_1 a + \theta_4^{\prime}c)exp\{exp(\beta_0^{(1)} + \beta_1^{(1)}a^* + \beta_2^{(1)\prime}c)[exp(\theta_2^{(1)} + \theta_3^{(1)}a) - 1]\} \\
&\times exp\{(\beta_0^{(2)} + \beta_1^{(2)}a^* + \beta_2^{(2)\prime}c)(\theta_2^{(2)} + \theta_3^{(2)}a) + \tfrac{1}{2}\sigma^2_{m^{(2)}}(\theta_2^{(2)} + \theta_3^{(2)}a)^2\}.
\end{aligned}$$

Standard errors for these estimators can ben obtained either via bootstrap procedure or by the delta method (Valeri and VanderWeele, 2013).