

Discussion of Methods for Mediation Analysis

Michael R. Elliott

Assistant Professor of Biostatistics

University of Michigan School of Public Health

Assistant Research Scientist

Institute for Social Research

Mid-Atlantic Causal Inference Conference

Baltimore, MD May 19, 2008.

Egleston, Scharfstein, Munoz, West:
Mediation with non-metaphysical counterfactuals

- Goal is to ascertain to what extent UVB exposure mediates relationship between eyeglasses use and development of cataracts.
- Have "observable" counterfactual – estimate of UVB exposure in the absence of glasses use
 - Assumes behavior would have been unchanged if hadn't worn glasses.

ESMW: Mediation with non-metaphysical counterfactuals

Mediation

Barron and Kenny (1986) consider linear regression on treatment Z

1. of outcome Y unadjusted for mediator M
2. of mediator M
3. of outcome Y adjusted for mediator M

If M and Z are associated in 2), then can use 1) and 3) to consider mediation

- Unadjusted and adjusted treatment effects equal: no mediation
- Adjusted treatment effect zero: complete mediation
- Adjusted treatment effect attenuated: partial mediation

ESMW: Mediation with non-metaphysical counterfactuals

Direct and Indirect Effects

$Y(z, m)$: outcome under treatment $Z = z$, mediator $M = m$

$M(z)$: mediator under treatment $Z = z$.

- (Prescriptive) Direct Effect: $E\{Y(1, m) - Y(0, m)\}$
- (Prescriptive) Indirect Effect:
 $E\{Y(1) - Y(0)\} - E\{Y(1, m) - Y(0, m)\}$

Direct effect of 0 implies complete mediation (all of treatment effect in indirect, i.e., through the mediator).

ESMW: Mediation with non-metaphysical counterfactuals
Direct and Indirect Effects

If have

1. Z independent $Y(z, m)$, $M(z)$ (randomization)
2. Given Z , $Y(z, m)$ or $Y(1 - z, m)$ independent of $M(z)$
(independence)

then can interpret adjusted regression estimator of treatment effect in linear model as direct effect.

If also have (Robins 2003)

1. $Y(1, m) - Y(0, m)$ constant with respect to m (no interaction)

then can interpret difference between unadjusted and adjusted regression estimator of treatment effect in linear model as indirect effect.

ESMW: Mediation with non-metaphysical counterfactuals

<u>Z = 1</u>	<u>Z = 0</u>	<u>M(1), M(0)</u>
$E(Y(1, 1)) = \alpha_{11}^{(1,1)}$	$E(Y(0, 1)) = \alpha_{11}^{(0,1)}$	$M(1) = 1, M(0) = 1$
$E(Y(1, 1)) = \alpha_{10}^{(1,1)}$	$E(Y(0, 0)) = \alpha_{10}^{(0,0)}$	$M(1) = 1, M(0) = 0$
$E(Y(1, 0)) = \alpha_{01}^{(1,0)}$	$E(Y(0, 1)) = \alpha_{01}^{(0,1)}$	$M(1) = 0, M(0) = 1$
$E(Y(1, 0)) = \alpha_{00}^{(1,0)}$	$E(Y(0, 0)) = \alpha_{00}^{(0,0)}$	$M(1) = 0, M(0) = 0$

Assume $P(M(1) = i, M(0) = j) = 1/4$ for $i, j = 0, 1$. Under randomization

$$E(Y | Z = 1, M = 1) - E(Y | Z = 0, M = 1) = \frac{\alpha_{11}^{(1,1)} + \alpha_{10}^{(1,1)}}{2} - \frac{\alpha_{11}^{(0,1)} + \alpha_{01}^{(0,1)}}{2}$$

but direct effect of interest for $m = 1$ when we assume that the mediator is not under the control of the investigator is

$$E(Y(1, M(1) = 1) - Y(0, M(0) = 1)) = \alpha_{11}^{(1,1)} - \alpha_{11}^{(0,1)}$$

ESMW: Mediation with non-metaphysical counterfactuals

Under independence, $\alpha_{11}^{(1,1)} = \alpha_{10}^{(1,1)} = \alpha^{(1,1)}$, $\alpha_{11}^{(0,1)} = \alpha_{01}^{(0,1)} = \alpha^{(0,1)}$, and the observed treatment difference conditional on $M = 1$ equals the direct effect for $m = 1$: $\alpha^{(1,1)} - \alpha^{(0,1)}$.

Under independence and no interaction, $\alpha^{(1,1)} - \alpha^{(0,1)} = \alpha^{(1,0)} - \alpha^{(0,0)} = \alpha^{(1)} - \alpha^{(0)}$, and the adjusted treatment difference equals the direct effect $\alpha^{(1)} - \alpha^{(0)}$.

ESMW: Mediation with non-metaphysical counterfactuals

Principal Strata

Instead of estimating direct and indirect effects in the fashion of Robins (2003) and Pearl (2001), ESMW use the principal stratification approach of Frangakis and Rubin (2002), which conditions on the joint distribution of the potential intermediate variables. Some of the advantages of this approach:

- Focuses on “potentially observable” counterfactuals, not counterfactuals that assume the intermediate variable has been manipulated to a different value than would have been observed under the assigned treatment.
 - Are these “supercounterfactuals” of interest?
- Focus on estimating $P(Y(z) | M(1), M(0))$, rather than $P(Y(z, m))$: require only $Z \perp (Y(z), M(z))$ instead of $(Y(z, m), M(z))$, and $Y(z) \perp M(z) | Z$ instead of $Y(z, m)$. May be easier to achieve in a non-observational setting.
 - How much weaker are these assumptions?
- No longer require linear link functions to obtain parameters with meaningful causal interpretations.
 - $E(Y(1 | S = s)/Y(0 | S = s))$ for dichotomous Y does not exist outside of degenerate cases. But $E(Y(1 | S = s))/E(Y(0 | S = s))$ does – good enough?

ESMW: Mediation with non-metaphysical counterfactuals

Results

- Use of $v(p, m) = \frac{E(Y(1)|M(1)/M(0)=p, M(0)=m)}{E(Y(0)|M(1)/M(0)=p, M(0)=m)}$ as causal estimand of interest illustrates the flexibility in the PS approach; hard to see how to estimate using alternative method.
 - $\frac{E(Y(1)|M(1)=m_1, M(0)=m_0)}{E(Y(0)|M(1)=m_1, M(0)=m_0)}$ might be of interest as well, if “Maryland sun-years” interpretable.
- Figure illustrating relationship between casual relative risk and potential mediation levels is worthy of a Tufte Hall of Fame for information density and interpretability.

ESMW: Mediation with non-metaphysical counterfactuals

Results

- Results generally sensible: little or no effect of glasses use when $p \approx 1$; also less helpful when $M(0)$ is very large. (Missing data due to surgery? Non-linearity in mediation effect?)
 - Contrast with standard regression approach, which found little evidence for evidence of mediation of eyeglasses use effect via UVB.
- Dropoff in mediation around $p = .2$, which is mediation value for those many of those who only wore corrective glasses, not sunglasses.
 - Different mediation effects $v(p, m)$ when treatment Z includes sunglasses use versus when corrective eyeglasses only.
 - Different “strategic use” among corrective eyeglasses only users? UVB, at least as measured by M , not the whole story?

ESMW: Mediation with non-metaphysical counterfactuals

Questions

- What are the implications of a lack of a relationship between $P(Y(0) = 1)$ and $M(0)$, either marginally or conditional on $M(1)/M(0)$?
 - Would seem to damage case for mediation: lack of path between M and Y .
- How realistic is the assumption that $Y(0) \perp M(1) \mid Z, M(0), X$? (Observed UVB in eyeglass wearers not associated with risk of cataracts in the absence of glasses use?)
 - Does this require that eyeglass wearers not change their behavior in the absence of glasses use? Some way to relax this assumption?
- Use propensity score as covariate in generalized linear model rather than as weight?

ESMW: Mediation with non-metaphysical counterfactuals

Use of Missing Data Paradigm

ESMW use a two-stage approach regression approach, but could treat $M(1)$ as missing data, and use either EM algorithm or data augmentation algorithm to estimate model parameters.

Another classic case where the counterfactual value of the intermediate variable is known: non-compliance where subjects assigned to the control arm cannot access the treatment.

- Z is treatment assignment, M is treatment taken, and Y is outcome.
- $M(0) = 0$ for all subjects; $M(1) = M$ if $z = 1$, but is unobserved if $z = 0$.
- $(M(1) = 1, M(0) = 0)$ are “compliers”; $(M(1) = 0, M(0) = 0)$ are “never takers”.

ESMW: Mediation with non-metaphysical counterfactuals
Use of Missing Data Paradigm

No longer require assumption that $Y(0) \perp M(1) \mid Z, M(0), X$ (although still need predictors X of $M(1)$ if Y is dichotomous).

$$\text{logit}(P(Y(z) = 1)) = \beta_0 + \beta_1 z + \beta_2 M(1) + \beta_3 M(1)z$$

$$v(1) = P(Y(1) \mid M(1) = 1) / P(Y(0) \mid M(1) = 1) = \frac{\text{expit}(\beta_0 + \beta_1 + \beta_2 + \beta_3)}{\text{expit}(\beta_0 + \beta_2)}$$

$$v(0) = P(Y(1) \mid M(1) = 0) / P(Y(0) \mid M(1) = 0) = \frac{\text{expit}(\beta_0 + \beta_1)}{\text{expit}(\beta_0)}$$

Simulation study: $X \sim U(0, 1)$, $\text{logit}(P(M(1) = 1)) = \alpha_0 + \alpha_1 x$. Consider $\alpha_0 = -1$, $\alpha_1 = 2$, $\beta_0 = -2$, $\beta_1 = .5$, $\beta_3 = 1.5$, and either $\beta_2 = 0$ (independence assumption holds) or $\beta_2 = 1$ (compliers have better outcomes on average on the control arm than never takers).

ESMW: Mediation with non-metaphysical counterfactuals

Use of Missing Data Paradigm

$\beta_2 = 0$ (Independence of $Y(0)$ and $M(1)$)

	Truth	EM		Two-Stage	
		Bias	RMSE	Bias	RMSE
$v(1)$	4.19	.186	1.089	-.001	.264
$v(0)$	1.53	.062	.374	-.000	.128

$\beta_2 = 1$ ($Y(0)$ depends on $M(1)$)

	Truth	EM		Two-Stage	
		Bias	RMSE	Bias	RMSE
$v(1)$	2.72	.050	.341	.953	.969
$v(0)$	1.53	.109	.545	-.557	.561

ESMW: Mediation with non-metaphysical counterfactuals
Use of Missing Data Paradigm

- Harder to implement when M is continuous, but numerical methods exist
- Might be easier to use a fully Bayesian approach and a data augmentation algorithm
 - $\eta \mid \text{rest} \propto f(p \mid M(0), X, \eta)p(\eta)$
 - $\beta \mid \text{rest} \propto P(Y(z) \mid p, M(0), X, z, \beta)p(\beta)$
 - $p \mid \text{rest} \propto P(Y(z) \mid p, M(0), X, z, \beta)f(p \mid M(0), X, \eta)$

Joffe, Small, Ten Have, Brunelli, Feldman:
Extended Instrumental Variables

Standard instrumental variable (IV) approaches use variables associated with a treatment of interest but independent of the potential outcomes under the various treatment arms to account for unobserved confounders of treatment and outcome.

- Classic example is non-compliance: the randomized treatment assignment is by definition independent of the potential outcomes but (hopefully) associated with treatment taken.
- In practice, look for factors associated with an exposure of interest but at least quasi-randomized with respect to the potential outcome.

JSTBF: Extended Instrumental Variables

Example

- JSTBF propose method for using IV to deal with unmeasured confounding when interested in direct effect of exposure on outcome in the presence of (potential) mediation.
- Example of vascular access (VA) choice (catheter vs. fistula/grant) in treating end-stage renal disease via hemodialysis. Choice of VA affects amount of dialysis that can be received: also can have direct effect on outcome. Thus “dose” of dialysis mediated the effect of VA on outcome.
- VA not randomized: sicker patients more likely to receive catheter. Use center as an IV?

JSTBF: Extended Instrumental Variables
Standard IV Estimation

$$Y = Y(a) = Y(0) + \Phi_A A + \epsilon$$

$$E(Y | R, X) = E(Y(0) | R, X) + E(A | R, X)\Phi_A = \mu + \phi_0\Phi_A + \phi_r\Phi_A R$$

where $E(Y(0) | R, X) = E(Y(0)) = \mu$ and $E(Z | R, X) = \phi_0 + \phi_r R$.

Thus the total effect of R on Y is given by $\phi_r\Phi_A$, and we can obtain an estimate of ϕ_r by first regressing Z on R to obtain

$\hat{E}(A | R, X) = \hat{\phi}_0 + \hat{\phi}_r R$, and regressing Y on $\hat{E}(A | R, X)$ to obtain $\hat{\Phi}_A$, the effect of A on Y after accounting for unobserved confounders U .

JSTBF: Extended Instrumental Variables
Extended IV Estimation

$$Y = Y(a, s) = Y(0, 0) + \Phi_A A + \Phi_S S + \epsilon$$

$$E(Y | R, X) = E(Y(0, 0)) + E(A | R, X)\Phi_A + E(S | R, X)\Phi_S$$

Obtain $\hat{\Phi}_A$ and $\hat{\Phi}_S$ by regressing Y on $\hat{E}(A | R, X)$ and $\hat{E}(S | R, X)$, where the latter on the predicted values obtained from regressing A and S on R .

But total effect of A on Y is given by direct effect $\hat{\Phi}_A$ plus the indirect effect through S .

- Problem: effect of A on S is confounded by R .
- Solution: Since R is an IV, can estimate direct effect of A on S .

$$S = S(a, r) = S(0, 0) + \phi_a A + \phi_r R + \epsilon$$

$$E(S | R, X) = E(S(0, 0)) + E(A | R, X)\phi_a + R\phi_r$$

Regress S on $\hat{E}(A | R, X)$ and R to obtain ϕ_a and ϕ_r .

Total effect of A on Y is then estimated by $\hat{\Phi}_A + \hat{\phi}_a \hat{\Phi}_S$.

JSTBF: Extended Instrumental Variables

Additional mediators

There may be alternative pathways Q through which R affects Y that do not pass through A .

- Use of other medications to treat side effects of HD might differ by practice

$$Y = Y(a, s) = Y(0, 0) + \Phi_A A + \Phi_S S + \Phi_Q Q + \epsilon$$

$$E(Y | R, X) = E(Y(0, 0)) + E(A | R, X)\Phi_A + E(S | R, X)\Phi_S + E(Q | R, X)\Phi_Q$$

Proceed as before, but now must regress Q on R to obtain $\hat{E}(Q | R, X)$ and now regress Y on $\hat{E}(A | R, X)$, $\hat{E}(S | R, X)$, and $\hat{E}(Q | R, X)$.

Total effect of A on Y remains $\hat{\Phi}_A + \hat{\phi}_a \hat{\Phi}_S$.

JSTBF: Extended Instrumental Variables

Questions

- Does the inclusion of additional mediators help R to perform its IV role?
- No interaction assumption between A and S and A and R : can be accommodated by including interaction terms?
- Linearity assumption (require Y to be approximately continuous).
Any way out of this?

JSTBF: Extended Instrumental Variables

Principal stratification

- Stratify on the potential outcomes of $A(0)$, $A(1)$, $S(0)$, $S(1)$.
- Advantages?
 - No longer require linearity.
 - Allow for direct effects of R on Y (equivalent to dropping the exclusion restriction assumption)?
- Might require good covariates to estimate principal strata.
- Might require at least "partial" monotonicity ($A(1) \geq A(0)$, $S(1) \geq S(0)$ or vice-versa) to obtain well-identified estimators.

Special thanks to Dylan Small, Tom Ten Have, Marshall Joffe, and Brian Egleston.