Causal mediation analysis: a whistle-stop tour and some recent advances

Rhian Daniel
London School of Hygiene and Tropical Medicine

CSM Seminar Series, Causal Inference Theme
1st July, 2016

LONDON SCHOOL of HYGIENE &TROPICAL MEDICINE

Acknowledgements

This talk is based on a paper:

Vansteelandt S, Daniel RM.
Interventional effects for mediation analysis with multiple mediators.
Epidemiology, in press.

Acknowledgements

This talk is based on a paper:

Vansteelandt S, Daniel RM.
Interventional effects for mediation analysis with multiple mediators.
Epidemiology, in press.

Thanks also to Bernard Rachet for access to the NYCRIS data, and to Ruoran Li for assistance with the data analysis.

Outline

Outline

Outline

Motivating example: one mediator

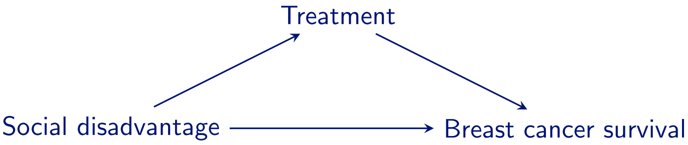
- For nearly a century, statisticians, and researchers in many different substantive disciplines, have been attempting to address questions concerning mediation.

Motivating example: one mediator

- For nearly a century, statisticians, and researchers in many different substantive disciplines, have been attempting to address questions concerning mediation.
  [Wright 1921, 1934; Baron and Kenny 1986; Robins and Greenland 1992; Pearl 2001; Cole and Hernán 2002; VanderWeele and Vansteelandt 2009; VanderWeele 2015.]

Motivating example: one mediator

- For nearly a century, statisticians, and researchers in many different substantive disciplines, have been attempting to address questions concerning mediation.
  [Wright 1921, 1934; Baron and Kenny 1986; Robins and Greenland 1992; Pearl 2001; Cole and Hernán 2002; VanderWeele and Vansteelandt 2009; VanderWeele 2015.]

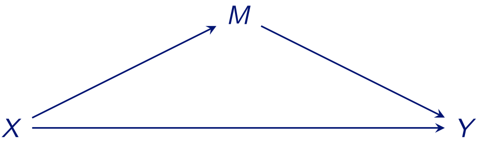


- For example, how much of the effect of social disadvantage on breast cancer survival is explained by treatment choices?

Outline

Traditional approach

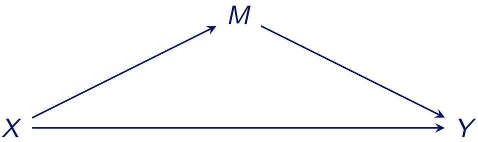
Path tracing rules [Wright 1934]



- Originally, mediation analysis was only attempted using linear models.

Traditional approach

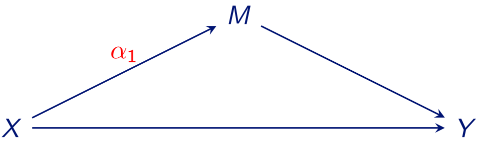
Path tracing rules [Wright 1934]



- Originally, mediation analysis was only attempted using linear models.
- Two models would be fitted:

Traditional approach

Path tracing rules [Wright 1934]

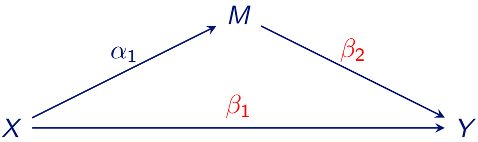


- Originally, mediation analysis was only attempted using linear models.
- Two models would be fitted:

$$E(M|X) = \alpha_0 + \alpha_1 X$$

Traditional approach

Path tracing rules [Wright 1934]

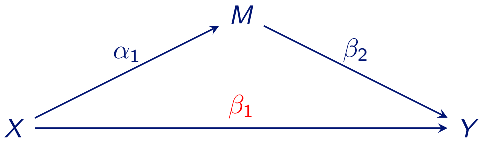


- Originally, mediation analysis was only attempted using linear models.
- Two models would be fitted:

$$E(M|X) = \alpha_0 + \alpha_1 X$$
$$E(Y|X, M) = \beta_0 + \beta_1 X + \beta_2 M$$

Traditional approach

Path tracing rules [Wright 1934]



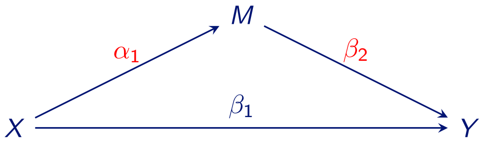
- Originally, mediation analysis was only attempted using linear models.
- Two models would be fitted:

$$E(M|X) = \alpha_0 + \alpha_1 X$$
$$E(Y|X, M) = \beta_0 + \beta_1 X + \beta_2 M$$

- β_1 would then be labelled the direct effect.

Traditional approach

Path tracing rules [Wright 1934]



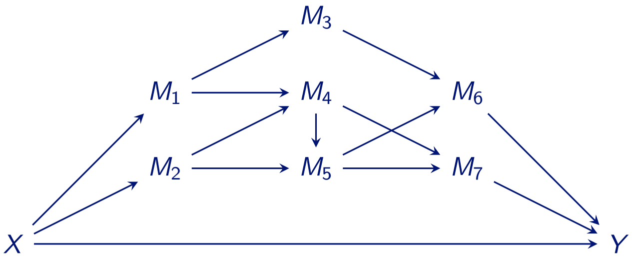
- Originally, mediation analysis was only attempted using linear models.
- Two models would be fitted:

$$E(M|X) = \alpha_0 + \alpha_1 X$$
$$E(Y|X, M) = \beta_0 + \beta_1 X + \beta_2 M$$

- β_1 would then be labelled the direct effect.
- And $\alpha_1\beta_2$ the indirect effect.

More complex diagrams

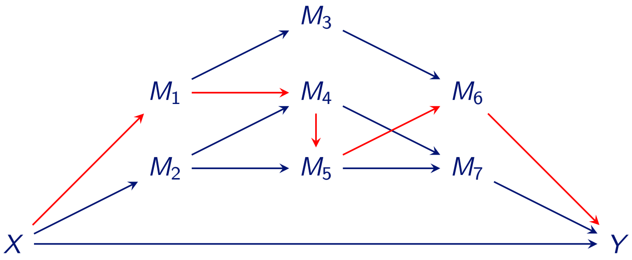
Path tracing rules [Wright 1934]



- This simple method extends to arbitrarily complex diagrams, as long as all models are simple linear regressions (with no interaction terms).

More complex diagrams

Path tracing rules [Wright 1934]



- This simple method extends to arbitrarily complex diagrams, as long as all models are simple linear regressions (with no interaction terms).
- The path-specific effect along a particular pathway is equal to the product of the coefficients along that path.

Outline

Causal inference 'investigates'

- In the early 1990s, the 'causal inference' school became interested in this area [Robins and Greenland 1992].

Causal inference 'investigates'

- In the early 1990s, the 'causal inference' school became interested in this area [Robins and Greenland 1992].
- Mediation is a causal concept: associations are symmetric, but mediation implies an ordered sequence.

Causal inference 'investigates'

- In the early 1990s, the 'causal inference' school became interested in this area [Robins and Greenland 1992].
- Mediation is a causal concept: associations are symmetric, but mediation implies an ordered sequence.
- Core principles of causal inference: (1) what is the estimand? (2) under what assumptions can it be identified? (3) are there more flexible estimation methods than currently used?

Potential outcomes and mediators

- Let $Y(x)$ be the value that Y would take if we intervened on X and set it (possibly counter to fact) to the value x.

Potential outcomes and mediators

- Let $Y(x)$ be the value that Y would take if we intervened on X and set it (possibly counter to fact) to the value x.
- Let $Y(x, m)$ be the value that Y would take if we intervened simultaneously on both X and M and set them to the values x and m.

Potential outcomes and mediators

- Let $Y(x)$ be the value that Y would take if we intervened on X and set it (possibly counter to fact) to the value x.
- Let $Y(x, m)$ be the value that Y would take if we intervened simultaneously on both X and M and set them to the values x and m.
- Let $M(x)$ be the value that M would take if we intervened on X and set it to x.

Potential outcomes and mediators

- Let $Y(x)$ be the value that Y would take if we intervened on X and set it (possibly counter to fact) to the value x.
- Let $Y(x, m)$ be the value that Y would take if we intervened simultaneously on both X and M and set them to the values x and m.
- Let $M(x)$ be the value that M would take if we intervened on X and set it to x.
- Let $Y\{x, M(x^*)\}$ be the value that Y would take if we intervened on X and set it to x whilst simultaneously intervening on M and setting it to $M(x^*)$, the value that M would take under an intervention setting X to x^*, where x and x^* are not necessarily equal.

Potential outcomes and mediators

- Let $Y(x)$ be the value that Y would take if we intervened on X and set it (possibly counter to fact) to the value x.
- Let $Y(x, m)$ be the value that Y would take if we intervened simultaneously on both X and M and set them to the values x and m.
- Let $M(x)$ be the value that M would take if we intervened on X and set it to x.
- Let $Y\{x, M(x^*)\}$ be the value that Y would take if we intervened on X and set it to x whilst simultaneously intervening on M and setting it to $M(x^*)$, the value that M would take under an intervention setting X to x^*, where x and x^* are not necessarily equal.

These hypothetical quantities were used to create model-free definitions of direct/indirect effects that match our intuition.

Outline

Natural direct effect

Pearl 2001; Robins and Greenland 1992

- The natural direct effect of (a binary) X on Y expressed as a marginal mean difference is

$$\text{NDE} = E\left[Y\left\{1, M(0)\right\}\right] - E\left[Y\left\{0, M(0)\right\}\right].$$

Natural direct effect

Pearl 2001; Robins and Greenland 1992

- The natural direct effect of (a binary) X on Y expressed as a marginal mean difference is

$$\text{NDE} = E\left[Y\left\{1, M(0)\right\}\right] - E\left[Y\left\{0, M(0)\right\}\right].$$

- This is a comparison of two hypothetical situations.

Natural direct effect

Pearl 2001; Robins and Greenland 1992

- The natural direct effect of (a binary) X on Y expressed as a marginal mean difference is

$$\text{NDE} = E\left[Y\left\{1, M(0)\right\}\right] - E\left[Y\left\{0, M(0)\right\}\right].$$

- This is a comparison of two hypothetical situations.
- In the first, X is set to 1, and in the second X is set to 0. In both, M is set to $M(0)$, the value it would take if X were set to 0.

Natural direct effect

Pearl 2001; Robins and Greenland 1992

- The natural direct effect of (a binary) X on Y expressed as a marginal mean difference is

$$\text{NDE} = E\left[Y\left\{1, M(0)\right\}\right] - E\left[Y\left\{0, M(0)\right\}\right].$$

- This is a comparison of two hypothetical situations.
- In the first, X is set to 1, and in the second X is set to 0. In both, M is set to $M(0)$, the value it would take if X were set to 0.
- Since M is the same (*within* subject) in both situations, we are intuitively getting at the direct effect of X.

Natural indirect effect

Pearl 2001; Robins and Greenland 1992

- The natural indirect effect of X on Y is

$$\text{NIE} = E\left[Y\left\{1, M(1)\right\}\right] - E\left[Y\left\{1, M(0)\right\}\right].$$

Natural indirect effect

Pearl 2001; Robins and Greenland 1992

- The natural indirect effect of X on Y is

$$\text{NIE} = E\left[Y\left\{1, M(1)\right\}\right] - E\left[Y\left\{1, M(0)\right\}\right].$$

- This is a comparison of two hypothetical situations.

Natural indirect effect

Pearl 2001; Robins and Greenland 1992

- The natural indirect effect of X on Y is

$$\text{NIE} = E\left[Y\left\{1, M(1)\right\}\right] - E\left[Y\left\{1, M(0)\right\}\right].$$

- This is a comparison of two hypothetical situations.
- In the first, M is set to $M(1)$ and in the second M is set to $M(0)$. In both, X is set to 1.

Natural indirect effect

Pearl 2001; Robins and Greenland 1992

- The natural indirect effect of X on Y is

$$\text{NIE} = E\left[Y\left\{1, M(1)\right\}\right] - E\left[Y\left\{1, M(0)\right\}\right].$$

- This is a comparison of two hypothetical situations.
- In the first, M is set to $M(1)$ and in the second M is set to $M(0)$. In both, X is set to 1.
- X is allowed to influence Y only through its influence on M. Thus it intuitively corresponds to an indirect effect through M.

Effect decomposition

The sum of the natural direct and indirect effects is

$$\begin{aligned}\text{NDE} + \text{NIE} = {} & E\left[Y\left\{1, M(0)\right\}\right] - E\left[Y\left\{0, M(0)\right\}\right] \\ & + E\left[Y\left\{1, M(1)\right\}\right] - E\left[Y\left\{1, M(0)\right\}\right]\end{aligned}$$

Effect decomposition

The sum of the natural direct and indirect effects is

$$\begin{aligned}\text{NDE} + \text{NIE} = {} & E\left[Y\left\{1, M(0)\right\}\right] - E\left[Y\left\{0, M(0)\right\}\right] \\ & + E\left[Y\left\{1, M(1)\right\}\right] - E\left[Y\left\{1, M(0)\right\}\right]\end{aligned}$$

Effect decomposition

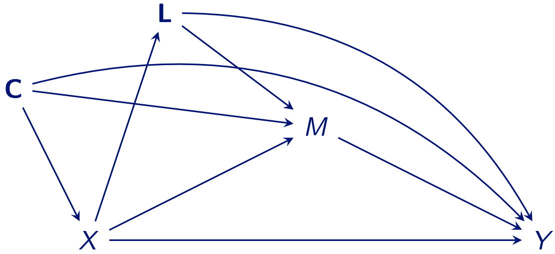
The sum of the natural direct and indirect effects is

$$\begin{aligned} \text{NDE} + \text{NIE} &= E\left[Y\left\{1, M(0)\right\}\right] - E\left[Y\left\{0, M(0)\right\}\right] \\ &\quad + E\left[Y\left\{1, M(1)\right\}\right] - E\left[Y\left\{1, M(0)\right\}\right] \\ &= E\left[Y\left\{1, M(1)\right\}\right] - E\left[Y\left\{0, M(0)\right\}\right] \\ &= E\left\{Y(1)\right\} - E\left\{Y(0)\right\} = \text{TCE}, \end{aligned}$$

the total causal effect of X on Y.

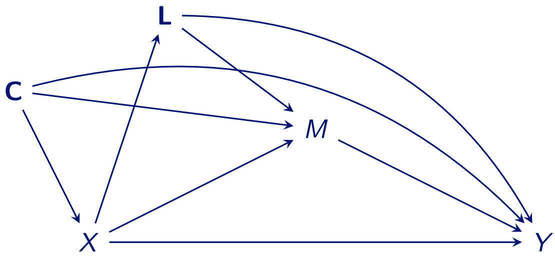
Outline

Assumptions for identification (1)



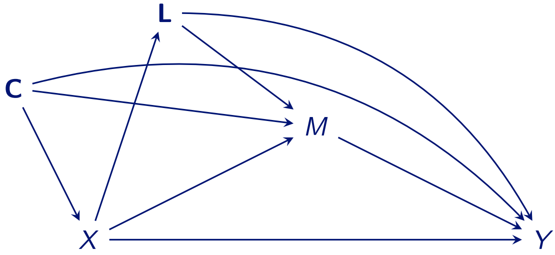
- Consider the setting with baseline confounders **C** and intermediate confounders **L**.

Assumptions for identification (1)



- Consider the setting with baseline confounders **C** and intermediate confounders **L**.
- Sufficient assumptions under which NDE and NIE can be identified: first, technical assumptions of no interference and consistency.

Assumptions for identification (1)

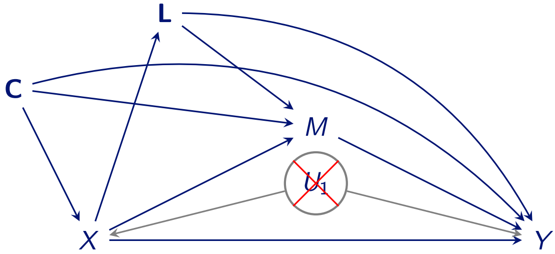


- Consider the setting with baseline confounders **C** and intermediate confounders **L**.
- Sufficient assumptions under which NDE and NIE can be identified: first, technical assumptions of no interference and consistency.
- Then there are sequential conditional exchangeability assumptions:

$$Y(x,m) \perp\!\!\!\perp X \,|\mathbf{C} = \mathbf{c} \ , \forall x, m, \mathbf{c}$$

$$Y(x,m) \perp\!\!\!\perp M \,|\mathbf{C} = \mathbf{c}, X = x, \mathbf{L} = \mathbf{l} \ , \forall x, m, \mathbf{c}, \mathbf{l}$$

Assumptions for identification (1)

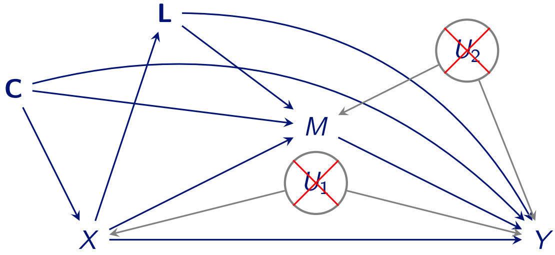


- Consider the setting with baseline confounders **C** and intermediate confounders **L**.
- Sufficient assumptions under which NDE and NIE can be identified: first, technical assumptions of no interference and consistency.
- Then there are sequential conditional exchangeability assumptions:

$$Y(x,m) \perp\!\!\!\perp X \,|\, \mathbf{C} = \mathbf{c} \;,\; \forall x, m, \mathbf{c}$$

$$Y(x,m) \perp\!\!\!\perp M \,|\, \mathbf{C} = \mathbf{c}, X = x, \mathbf{L} = \mathbf{l} \;,\; \forall x, m, \mathbf{c}, \mathbf{l}$$

Assumptions for identification (1)

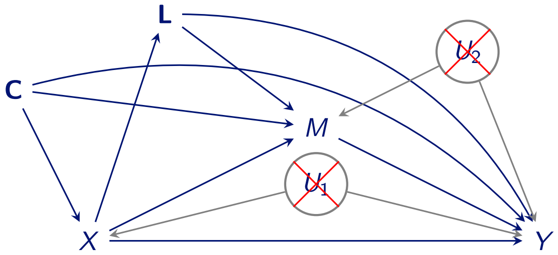


- Consider the setting with baseline confounders **C** and intermediate confounders **L**.
- Sufficient assumptions under which NDE and NIE can be identified: first, technical assumptions of no interference and consistency.
- Then there are sequential conditional exchangeability assumptions:

$$Y(x,m) \perp\!\!\!\perp X \,|\mathbf{C} = \mathbf{c} \;, \forall x, m, \mathbf{c}$$

$$Y(x,m) \perp\!\!\!\perp M \,|\mathbf{C} = \mathbf{c}, X = x, \mathbf{L} = \mathbf{l} \;, \forall x, m, \mathbf{c}, \mathbf{l}$$

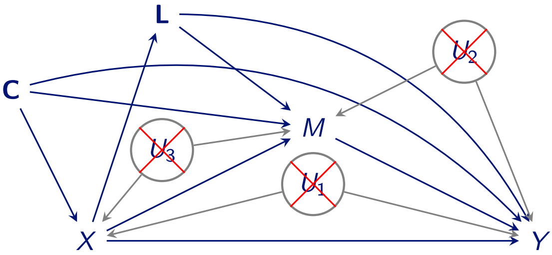
Assumptions for identification (2)



- And:

$$M(x) \perp\!\!\!\perp X \,|\, \mathbf{C} = \mathbf{c} \;, \forall x, \mathbf{c}$$

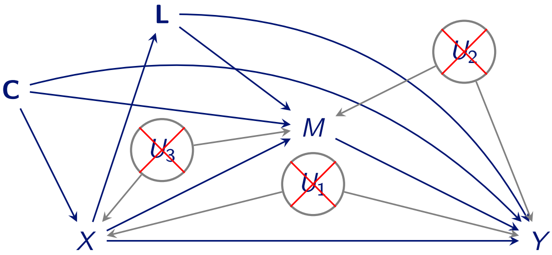
Assumptions for identification (2)



- And:

$$M(x) \perp\!\!\!\perp X \,|\, \mathbf{C} = \mathbf{c} \,,\, \forall x, \mathbf{c}$$

Assumptions for identification (2)

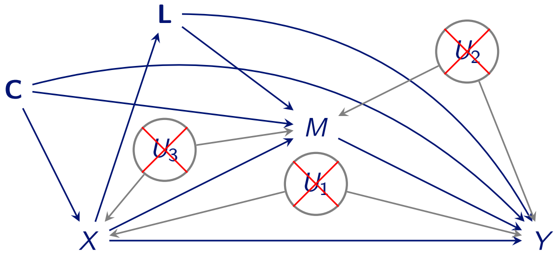


- And:

$$M(x) \perp\!\!\!\perp X \,|\mathbf{C} = \mathbf{c} \,,\, \forall x, \mathbf{c}$$

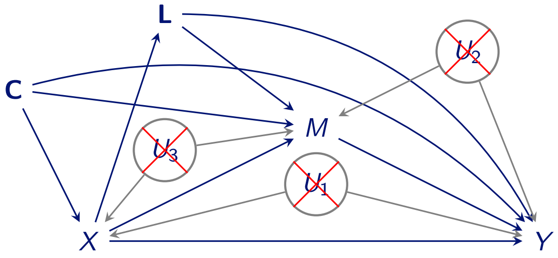
This much, we would probably expect!

Assumptions for identification (3)



- Perhaps surprisingly, these assumptions are not enough.

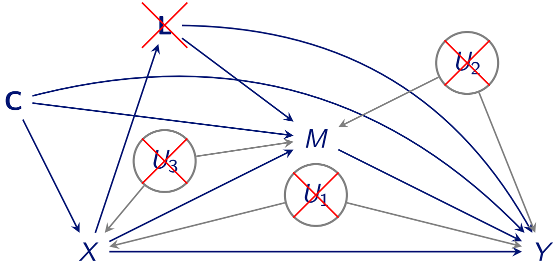
Assumptions for identification (3)



- Perhaps surprisingly, these assumptions are not enough.
- In addition, we need something such as the cross-world independence assumption:

$$M(x^*) \perp\!\!\!\perp Y(x,m) \,|\, \mathbf{C} = \mathbf{c} \;, \forall x, m, x^*, \mathbf{c}$$

Assumptions for identification (3)



- Perhaps surprisingly, these assumptions are not enough.
- In addition, we need something such as the cross-world independence assumption:

$$M(x^*) \perp\!\!\!\perp Y(x,m) \,|\mathbf{C} = \mathbf{c} \,,\, \forall x, m, x^*, \mathbf{c}$$

- This implies (but is not implied by, ie it is stronger than) no **L**.

Relaxing the cross-world independence assumption

- The cross-world independence assumption

$$M(x^*) \perp\!\!\!\perp Y(x,m) \,|\mathbf{C} = \mathbf{c} \,,\, \forall x, m, x^*, \mathbf{c}$$

rules out intermediate confounders **L**.

Relaxing the cross-world independence assumption

- The cross-world independence assumption

$$M(x^*) \perp\!\!\!\perp Y(x,m) \,|\mathbf{C} = \mathbf{c} \;, \forall x, m, x^*, \mathbf{c}$$

  rules out intermediate confounders **L**.

- In fact, a slightly weaker assumption, which does not rule out **L** is sufficient:

  $$E\{Y(1,m) - Y(0,m)\,|\mathbf{C}=\mathbf{c}, M(0)=m\} = E\{Y(1,m) - Y(0,m)\,|\mathbf{C}=\mathbf{c}\}$$

  [Petersen et al 2006]

Relaxing the cross-world independence assumption

- The cross-world independence assumption

$$M(x^*) \perp\!\!\!\perp Y(x,m) \,|\, \mathbf{C}=\mathbf{c} \,,\ \forall x, m, x^*, \mathbf{c}$$

  rules out intermediate confounders **L**.

- In fact, a slightly weaker assumption, which does not rule out **L** is sufficient:

$$E\{Y(1,m)-Y(0,m)\,|\,\mathbf{C}=\mathbf{c}, M(0)=m\} = E\{Y(1,m)-Y(0,m)\,|\,\mathbf{C}=\mathbf{c}\}$$

  [Petersen et al 2006]

- Both assumptions are very strong, and not even a hypothetical experiment exists in which they would hold by design. [Richardson and Robins 2013]

Relaxing the cross-world independence assumption

- The cross-world independence assumption

$$M(x^*) \perp\!\!\!\perp Y(x,m)\,|\mathbf{C} = \mathbf{c}\ ,\ \forall x, m, x^*, \mathbf{c}$$

  rules out intermediate confounders **L**.

- In fact, a slightly weaker assumption, which does not rule out **L** is sufficient:

  $$E\{Y(1,m) - Y(0,m)\,|\mathbf{C} = \mathbf{c}, M(0) = m\} = E\{Y(1,m) - Y(0,m)\,|\mathbf{C} = \mathbf{c}\}$$

  [Petersen et al 2006]

- Both assumptions are very strong, and not even a hypothetical experiment exists in which they would hold by design.
  [Richardson and Robins 2013]

- Even the Petersen assumption places strong parametric restrictions on the relationship between **L** and Y, which can essentially only hold in linear models with no non-linearities involving **L**.
  [De Stavola et al 2015]

Outline

Identification (1)

Pearl 2001

- Identifying $E[Y\{x, M(x^*)\}]$ is sufficient for identifying the NDE and NIE.

Identification (1)

Pearl 2001

- Identifying $E[Y\{x, M(x^*)\}]$ is sufficient for identifying the NDE and NIE.
- First we write:

$$E[Y\{x, M(x^*)\}] =$$

$$\sum_{\mathbf{c},m} E\{Y(x, m)\,|\mathbf{C} = \mathbf{c}, M(x^*) = m\}P\{M(x^*) = m|\mathbf{C} = \mathbf{c}\}P\{\mathbf{C} = \mathbf{c}\}$$

Identification (1)

Pearl 2001

- Identifying $E[Y\{x, M(x^*)\}]$ is sufficient for identifying the NDE and NIE.
- First we write:

$$E[Y\{x, M(x^*)\}] =$$
$$\sum_{\mathbf{c},m} E\{Y(x,m)\,|\mathbf{C}=\mathbf{c}, M(x^*)=m\}P\{M(x^*)=m|\mathbf{C}=\mathbf{c}\}P\{\mathbf{C}=\mathbf{c}\}$$

- By the cross-world independence assumption, this is equal to:

$$\sum_{\mathbf{c},m} E\{Y(x,m)\,|\mathbf{C}=\mathbf{c}\}P\{M(x^*)=m|\mathbf{C}=\mathbf{c}\}P\{\mathbf{C}=\mathbf{c}\}$$

Identification (1)

Pearl 2001

- Identifying $E[Y\{x, M(x^*)\}]$ is sufficient for identifying the NDE and NIE.
- First we write:

$$E[Y\{x, M(x^*)\}] =$$
$$\sum_{\mathbf{c},m} E\{Y(x,m)\,|\mathbf{C}=\mathbf{c}, M(x^*)=m\}P\{M(x^*)=m|\mathbf{C}=\mathbf{c}\}P\{\mathbf{C}=\mathbf{c}\}$$

- By the cross-world independence assumption, this is equal to:

$$\sum_{\mathbf{c},m} E\{Y(x,m)\,|\mathbf{C}=\mathbf{c}\}P\{M(x^*)=m|\mathbf{C}=\mathbf{c}\}P\{\mathbf{C}=\mathbf{c}\}$$

- By conditional exchangeability (now without **L**), this is:

$$\sum_{\mathbf{c},m} E\{Y(x,m)\,|X=x, M=m, \mathbf{C}=\mathbf{c}\}P\{M(x^*)=m|X=x^*, \mathbf{C}=\mathbf{c}\}P\{\mathbf{C}=\mathbf{c}\}$$

Identification (2)

Pearl 2001

$$\sum_{\mathbf{c},m} E\{Y(x,m)\,|X=x, M=m, \mathbf{C}=\mathbf{c}\}P\{M(x^*)=m|X=x^*, \mathbf{C}=\mathbf{c}\}P\{\mathbf{C}=\mathbf{c}\}$$

Identification (2)

Pearl 2001

$$\sum_{\mathbf{c},m} E\{Y(x,m)\,|X=x, M=m, \mathbf{C}=\mathbf{c}\}P\{M(x^*)=m|X=x^*, \mathbf{C}=\mathbf{c}\}P\{\mathbf{C}=\mathbf{c}\}$$

- By consistency, this is:

$$\sum_{\mathbf{c},m} E\{Y\,|X=x, M=m, \mathbf{C}=\mathbf{c}\}P\{M=m|X=x^*, \mathbf{C}=\mathbf{c}\}P\{\mathbf{C}=\mathbf{c}\}$$

Identification (2)

Pearl 2001

$$\sum_{\mathbf{c},m} E\{Y(x,m)\,|X=x, M=m, \mathbf{C}=\mathbf{c}\}P\{M(x^*)=m|X=x^*, \mathbf{C}=\mathbf{c}\}P\{\mathbf{C}=\mathbf{c}\}$$

- By consistency, this is:

$$\sum_{\mathbf{c},m} E\{Y\,|X=x, M=m, \mathbf{C}=\mathbf{c}\}P\{M=m|X=x^*, \mathbf{C}=\mathbf{c}\}P\{\mathbf{C}=\mathbf{c}\}$$

- The hard work is now done.

Identification (2)

Pearl 2001

$$\sum_{\mathbf{c},m} E\{Y(x,m)\,|X=x, M=m, \mathbf{C}=\mathbf{c}\}P\{M(x^*)=m|X=x^*, \mathbf{C}=\mathbf{c}\}P\{\mathbf{C}=\mathbf{c}\}$$

- By consistency, this is:

$$\sum_{\mathbf{c},m} E\{Y\,|X=x, M=m, \mathbf{C}=\mathbf{c}\}P\{M=m|X=x^*, \mathbf{C}=\mathbf{c}\}P\{\mathbf{C}=\mathbf{c}\}$$

- The hard work is now done.
- By substituting different values for x and x^*, we can re-write the NDE and the NIE using only functions of aspects of the distribution of the observed data.

Identification (2)

Pearl 2001

$$\sum_{\mathbf{c},m} E\{Y(x,m)\,|X=x, M=m, \mathbf{C}=\mathbf{c}\}P\{M(x^*)=m|X=x^*, \mathbf{C}=\mathbf{c}\}P\{\mathbf{C}=\mathbf{c}\}$$

- By consistency, this is:

$$\sum_{\mathbf{c},m} E\{Y\,|X=x, M=m, \mathbf{C}=\mathbf{c}\}P\{M=m|X=x^*, \mathbf{C}=\mathbf{c}\}P\{\mathbf{C}=\mathbf{c}\}$$

- The hard work is now done.
- By substituting different values for x and x^*, we can re-write the NDE and the NIE using only functions of aspects of the distribution of the observed data.
- Plug-in or alternative (semiparametric) estimation could then be used. Many many proposals have been made!

Summary so far

- Mediation analysis is not new.

Summary so far

- Mediation analysis is not new.
- When all models are linear (with no interactions) quite complicated structures can be incorporated and path-specific effects estimated.

Summary so far

- Mediation analysis is not new.
- When all models are linear (with no interactions) quite complicated structures can be incorporated and path-specific effects estimated.
- However, in the traditional approach, it was unclear what exactly was being estimated, under what assumptions this was possible, and how things could be extended to non-linear settings.

Summary so far

- Mediation analysis is not new.
- When all models are linear (with no interactions) quite complicated structures can be incorporated and path-specific effects estimated.
- However, in the traditional approach, it was unclear what exactly was being estimated, under what assumptions this was possible, and how things could be extended to non-linear settings.

Summary so far

- Mediation analysis is not new.
- When all models are linear (with no interactions) quite complicated structures can be incorporated and path-specific effects estimated.
- However, in the traditional approach, it was unclear what exactly was being estimated, under what assumptions this was possible, and how things could be extended to non-linear settings.

Summary so far

- Mediation analysis is not new.
- When all models are linear (with no interactions) quite complicated structures can be incorporated and path-specific effects estimated.
- However, in the traditional approach, it was unclear what exactly was being estimated, under what assumptions this was possible, and how things could be extended to non-linear settings.
- The causal inference literature has addressed many of these concerns by giving counterfactual definitions of direct and indirect effects that are independent of any model, and by deriving clear identification assumptions.

Summary so far

- Mediation analysis is not new.
- When all models are linear (with no interactions) quite complicated structures can be incorporated and path-specific effects estimated.
- However, in the traditional approach, it was unclear what exactly was being estimated, under what assumptions this was possible, and how things could be extended to non-linear settings.
- The causal inference literature has addressed many of these concerns by giving counterfactual definitions of direct and indirect effects that are independent of any model, and by deriving clear identification assumptions.

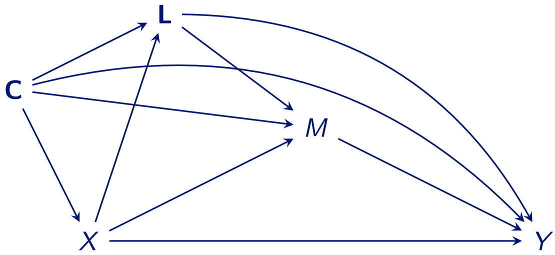
Summary so far

- Mediation analysis is not new.
- When all models are linear (with no interactions) quite complicated structures can be incorporated and path-specific effects estimated.
- However, in the traditional approach, it was unclear what exactly was being estimated, under what assumptions this was possible, and how things could be extended to non-linear settings.
- The causal inference literature has addressed many of these concerns by giving counterfactual definitions of direct and indirect effects that are independent of any model, and by deriving clear identification assumptions.
- The identification expressions can be used to derive estimators of direct and indirect effects in the presence of non-linearities, greatly increasing the flexibility of mediation analysis.

Summary so far

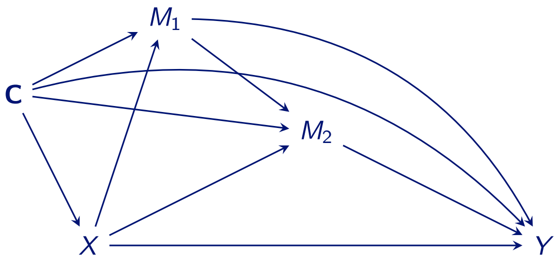
- Mediation analysis is not new.
- When all models are linear (with no interactions) quite complicated structures can be incorporated and path-specific effects estimated.
- However, in the traditional approach, it was unclear what exactly was being estimated, under what assumptions this was possible, and how things could be extended to non-linear settings.
- The causal inference literature has addressed many of these concerns by giving counterfactual definitions of direct and indirect effects that are independent of any model, and by deriving clear identification assumptions.
- The identification expressions can be used to derive estimators of direct and indirect effects in the presence of non-linearities, greatly increasing the flexibility of mediation analysis.
- However, it is plagued by the cross-world assumption; in particular the fact that this almost rules out intermediate confounders.

Consequences for multiple mediators



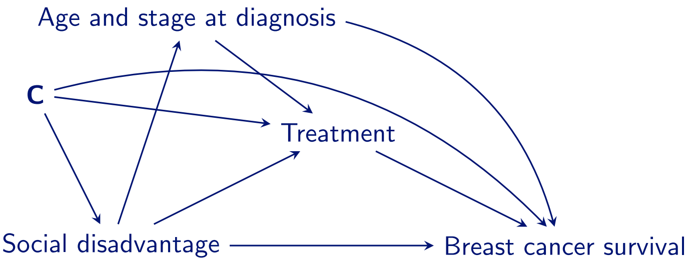
- For the same reason that in general we can't have **L** . . .

Consequences for multiple mediators



- For the same reason that in general we can't have **L** . . .
- . . . settings involving multiple mediators are also problematic.

Consequences for multiple mediators



- For the same reason that in general we can't have **L** ...
- ...settings involving multiple mediators are also problematic.
- eg in our motivating example.

Outline

Outline

Randomised interventional analogues of NDE/NIE

VanderWeele et al 2014

- The randomised interventional analogue of the NDE is

$$\text{RIA-NDE} = E\left\{Y\left(1, M^*_{0|\mathbf{C}}\right)\right\} - E\left\{Y\left(0, M^*_{0|\mathbf{C}}\right)\right\}$$

where $M^*_{x|\mathbf{C}}$ is a random draw from the distribution of M among those with $X = x$ conditional on $\mathbf{C}$.

Randomised interventional analogues of NDE/NIE

VanderWeele et al 2014

- The randomised interventional analogue of the NDE is

$$\text{RIA-NDE} = E\left\{Y\left(1, M^*_{0|\mathbf{C}}\right)\right\} - E\left\{Y\left(0, M^*_{0|\mathbf{C}}\right)\right\}$$

  where $M^*_{x|\mathbf{C}}$ is a random draw from the distribution of M among those with $X = x$ conditional on $\mathbf{C}$.

- The randomised interventional analogue of the NIE of X on Y expressed as a marginal mean difference is

$$\text{RIA-NIE} = E\left\{Y\left(1, M^*_{1|\mathbf{C}}\right)\right\} - E\left\{Y\left(1, M^*_{0|\mathbf{C}}\right)\right\}.$$

Randomised interventional analogues of NDE/NIE

VanderWeele et al 2014

- The randomised interventional analogue of the NDE is

$$\text{RIA-NDE} = E\left\{Y\left(1, M^*_{0|\mathbf{C}}\right)\right\} - E\left\{Y\left(0, M^*_{0|\mathbf{C}}\right)\right\}$$

where $M^*_{x|\mathbf{C}}$ is a random draw from the distribution of M among those with $X = x$ conditional on $\mathbf{C}$.

- The randomised interventional analogue of the NIE of X on Y expressed as a marginal mean difference is

$$\text{RIA-NIE} = E\left\{Y\left(1, M^*_{1|\mathbf{C}}\right)\right\} - E\left\{Y\left(1, M^*_{0|\mathbf{C}}\right)\right\}.$$

- The RIA-NDE, for example, is a direct effect comparing exposure versus no exposure with the mediator in both cases randomly drawn from the distribution of the population when given no exposure (given baseline confounders **C**).

Advantages and disadvantages

- The RIA-NDE and RIA-NIE can be identified under the no interference, consistency and conditional exchangeability assumptions mentioned earlier, but without the additional cross-world (or Petersen) assumption.

Advantages and disadvantages

- The RIA-NDE and RIA-NIE can be identified under the no interference, consistency and conditional exchangeability assumptions mentioned earlier, but without the additional cross-world (or Petersen) assumption.
- Intuitively, the 1st identification step (which is where the cross-world assumption came in) is removed, and the estimand is changed to the quantity in the 2nd line of the identification.

Advantages and disadvantages

- The RIA-NDE and RIA-NIE can be identified under the no interference, consistency and conditional exchangeability assumptions mentioned earlier, but without the additional cross-world (or Petersen) assumption.
- Intuitively, the 1st identification step (which is where the cross-world assumption came in) is removed, and the estimand is changed to the quantity in the 2nd line of the identification.
- If the cross-world assumption does hold, then NDE=RIA-NDE.

Advantages and disadvantages

- The RIA-NDE and RIA-NIE can be identified under the no interference, consistency and conditional exchangeability assumptions mentioned earlier, but without the additional cross-world (or Petersen) assumption.
- Intuitively, the 1st identification step (which is where the cross-world assumption came in) is removed, and the estimand is changed to the quantity in the 2nd line of the identification.
- If the cross-world assumption does hold, then NDE=RIA-NDE.
- If not, then the stronger **C** predicts M, the smaller the difference between NDE and RIA-NDE.

Advantages and disadvantages

- The RIA-NDE and RIA-NIE can be identified under the no interference, consistency and conditional exchangeability assumptions mentioned earlier, but without the additional cross-world (or Petersen) assumption.
- Intuitively, the 1st identification step (which is where the cross-world assumption came in) is removed, and the estimand is changed to the quantity in the 2nd line of the identification.
- If the cross-world assumption does hold, then NDE=RIA-NDE.
- If not, then the stronger $\mathbf{C}$ predicts M, the smaller the difference between NDE and RIA-NDE.
- RIA effects correspond to interventions that could in principle be done.

Advantages and disadvantages

- The RIA-NDE and RIA-NIE can be identified under the no interference, consistency and conditional exchangeability assumptions mentioned earlier, but without the additional cross-world (or Petersen) assumption.
- Intuitively, the 1st identification step (which is where the cross-world assumption came in) is removed, and the estimand is changed to the quantity in the 2nd line of the identification.
- If the cross-world assumption does hold, then NDE=RIA-NDE.
- If not, then the stronger **C** predicts M, the smaller the difference between NDE and RIA-NDE.
- RIA effects correspond to interventions that could in principle be done.
- However, RIA-NDE + RIA-NIE =

$$E\left\{Y\left(1, M^*_{1|\mathbf{C}}\right)\right\} - E\left\{Y\left(0, M^*_{0|\mathbf{C}}\right)\right\}$$

which is **NOT** in general equal to the total causal effect!

Our proposal

- We propose an extension of these interventional effects to multiple mediator settings.

Our proposal

- We propose an extension of these interventional effects to multiple mediator settings.
- The effects we define will sum to the total causal effect.

Our proposal

- We propose an extension of these interventional effects to multiple mediator settings.
- The effects we define will sum to the total causal effect.
- Identification will be possible under no interference, consistency, no unmeasured confounding of X–$\mathbf{M}$, X–Y and $\mathbf{M}$–Y, where the mediators $\mathbf{M}$ are for this purpose considered en bloc.

Our proposal

- We propose an extension of these interventional effects to multiple mediator settings.
- The effects we define will sum to the total causal effect.
- Identification will be possible under no interference, consistency, no unmeasured confounding of X–$\mathbf{M}$, X–Y and $\mathbf{M}$–Y, where the mediators $\mathbf{M}$ are for this purpose considered en bloc.
- We will not need any sort of cross-world assumption, we will not need to assume no unmeasured confounding between different mediators, and we won't require knowledge of the order of the mediators.

Our proposal

- We propose an extension of these interventional effects to multiple mediator settings.
- The effects we define will sum to the total causal effect.
- Identification will be possible under no interference, consistency, no unmeasured confounding of X–$\mathbf{M}$, X–Y and $\mathbf{M}$–Y, where the mediators $\mathbf{M}$ are for this purpose considered en bloc.
- We will not need any sort of cross-world assumption, we will not need to assume no unmeasured confounding between different mediators, and we won't require knowledge of the order of the mediators.
- For simplicity, we describe our approach for two mediators.

Outline

Interventional direct effect through neither M_1 nor M_2

With two mediators we propose the following definition of an interventional direct effect:

$$\sum_{\mathbf{c}} \sum_{m_1} \sum_{m_2} \left[E\left\{ Y(1, m_1, m_2) | \mathbf{C} = \mathbf{c} \right\} - E\left\{ Y(0, m_1, m_2) | \mathbf{C} = \mathbf{c} \right\} \right] \cdot$$
$$P\{M_1(0) = m_1, M_2(0) = m_2 | \mathbf{C} = \mathbf{c}\} P(\mathbf{C} = \mathbf{c})$$

Interventional direct effect through neither M_1 nor M_2

With two mediators we propose the following definition of an interventional direct effect:

$$\sum_{\mathbf{c}} \sum_{m_1} \sum_{m_2} \left[E\left\{ Y(1, m_1, m_2) | \mathbf{C} = \mathbf{c} \right\} - E\left\{ Y(0, m_1, m_2) | \mathbf{C} = \mathbf{c} \right\} \right] \cdot$$
$$P\{M_1(0) = m_1, M_2(0) = m_2 | \mathbf{C} = \mathbf{c}\} P(\mathbf{C} = \mathbf{c})$$

- This expresses the exposure effect when fixing the joint distribution of both mediators (by controlling the mediators for each subject at a random draw from their counterfactual joint distribution with the exposure set at 0, given covariates $\mathbf{C}$).

Interventional indirect effect through M_1

We propose the following definition of an interventional indirect effect throught M_1:

$$\sum_{\mathbf{c}} \sum_{m_1} \sum_{m_2} E\left\{Y(1, m_1, m_2) | \mathbf{C} = \mathbf{c}\right\} \cdot \\ \left[P\{M_1(1) = m_1 | \mathbf{C} = \mathbf{c}\} - P\{M_1(0) = m_1 | \mathbf{C} = \mathbf{c}\}\right] \\ \cdot P\{M_2(0) = m_2 | \mathbf{C} = \mathbf{c}\} P(\mathbf{C} = \mathbf{c})$$

Interventional indirect effect through M_1

We propose the following definition of an interventional indirect effect throught M_1:

$$\sum_{\mathbf{c}} \sum_{m_1} \sum_{m_2} E\left\{Y(1, m_1, m_2) | \mathbf{C} = \mathbf{c}\right\} \cdot [P\{M_1(1) = m_1 | \mathbf{C} = \mathbf{c}\} - P\{M_1(0) = m_1 | \mathbf{C} = \mathbf{c}\}] \cdot P\{M_2(0) = m_2 | \mathbf{C} = \mathbf{c}\} P(\mathbf{C} = \mathbf{c})$$

- This expresses the effect of shifting the distribution of mediator M_1 from the counterfactual distribution (given covariates) at exposure level 0 to that at level 1, while fixing the exposure at 1 and the mediator M_2 to a random subject-specific draw from the counterfactual distribution (given covariates) at level 0 for all subjects.

Interventional indirect effect through M_1

We propose the following definition of an interventional indirect effect throught M_1:

$$\sum_{\mathbf{c}}\sum_{m_1}\sum_{m_2} E\{Y(1,m_1,m_2)|\mathbf{C}=\mathbf{c}\}\cdot [P\{M_1(1)=m_1|\mathbf{C}=\mathbf{c}\}-P\{M_1(0)=m_1|\mathbf{C}=\mathbf{c}\}] \cdot P\{M_2(0)=m_2|\mathbf{C}=\mathbf{c}\}P(\mathbf{C}=\mathbf{c})$$

- This expresses the effect of shifting the distribution of mediator M_1 from the counterfactual distribution (given covariates) at exposure level 0 to that at level 1, while fixing the exposure at 1 and the mediator M_2 to a random subject-specific draw from the counterfactual distribution (given covariates) at level 0 for all subjects.
- This effect captures all of the exposure effect that is mediated by M_1, but not by causal descendants of M_1 in the graph.

Interventional indirect effect through M_2

We propose the following definition of an interventional indirect effect throught M_2:

$$\sum_{\mathbf{c}} \sum_{m_1} \sum_{m_2} E\{Y(1, m_1, m_2) | \mathbf{C} = \mathbf{c}\} \cdot$$
$$[P\{M_2(1) = m_2 | \mathbf{C} = \mathbf{c}\} - P\{M_2(0) = m_2 | \mathbf{C} = \mathbf{c}\}]$$
$$\cdot P\{M_1(0) = m_1 | \mathbf{C} = \mathbf{c}\} P(\mathbf{C} = \mathbf{c})$$

Interventional indirect effect through M_2

We propose the following definition of an interventional indirect effect throught M_2:

$$\sum_{\mathbf{c}} \sum_{m_1} \sum_{m_2} E\left\{Y(1, m_1, m_2) | \mathbf{C} = \mathbf{c}\right\} \cdot [P\{M_2(1) = m_2 | \mathbf{C} = \mathbf{c}\} - P\{M_2(0) = m_2 | \mathbf{C} = \mathbf{c}\}] \cdot P\{M_1(0) = m_1 | \mathbf{C} = \mathbf{c}\} P(\mathbf{C} = \mathbf{c})$$

- This expresses the effect of shifting the distribution of mediator M_2 from the counterfactual distribution (given covariates) at exposure level 0 to that at level 1, while fixing the exposure at 1 and the mediator M_1 to a random subject-specific draw from the counterfactual distribution (given covariates) at level 0 for all subjects.

Interventional indirect effect through M_2

We propose the following definition of an interventional indirect effect throught M_2:

$$\sum_{\mathbf{c}} \sum_{m_1} \sum_{m_2} E\{Y(1, m_1, m_2)|\mathbf{C} = \mathbf{c}\} \cdot [P\{M_2(1) = m_2|\mathbf{C} = \mathbf{c}\} - P\{M_2(0) = m_2|\mathbf{C} = \mathbf{c}\}] \cdot P\{M_1(0) = m_1|\mathbf{C} = \mathbf{c}\} P(\mathbf{C} = \mathbf{c})$$

- This expresses the effect of shifting the distribution of mediator M_2 from the counterfactual distribution (given covariates) at exposure level 0 to that at level 1, while fixing the exposure at 1 and the mediator M_1 to a random subject-specific draw from the counterfactual distribution (given covariates) at level 0 for all subjects.
- This effect captures all of the exposure effect that is mediated by M_2, but not by causal descendants of M_2 in the graph.

Remainder?

Finally, the TCE decomposes into the sum of the three previous effects plus a remainder term:

$$\sum_{\mathbf{c}}\sum_{m_1}\sum_{m_2} E\left\{Y(1,m_1,m_2)|\mathbf{C}=\mathbf{c}\right\}\cdot$$
$$[P\{M_1(1)=m_1, M_2(1)=m_2|\mathbf{C}=\mathbf{c}\}$$
$$-P\{M_1(1)=m_1|\mathbf{C}=\mathbf{c}\}P\{M_2(1)=m_2|\mathbf{C}=\mathbf{c}\}$$
$$-P\{M_1(0)=m_1, M_2(0)=m_2|\mathbf{C}=\mathbf{c}\}$$
$$+P\{M_1(0)=m_1|\mathbf{C}=\mathbf{c}\}P\{M_2(0)=m_2|\mathbf{C}=\mathbf{c}\}]\,P(\mathbf{C}=\mathbf{c})$$

Remainder?

Finally, the TCE decomposes into the sum of the three previous effects plus a remainder term:

$$\sum_{\mathbf{c}} \sum_{m_1} \sum_{m_2} E\left\{Y(1, m_1, m_2) | \mathbf{C} = \mathbf{c}\right\} \cdot$$
$$[P\{M_1(1) = m_1, M_2(1) = m_2 | \mathbf{C} = \mathbf{c}\}$$
$$- P\{M_1(1) = m_1 | \mathbf{C} = \mathbf{c}\} P\{M_2(1) = m_2 | \mathbf{C} = \mathbf{c}\}$$
$$- P\{M_1(0) = m_1, M_2(0) = m_2 | \mathbf{C} = \mathbf{c}\}$$
$$+ P\{M_1(0) = m_1 | \mathbf{C} = \mathbf{c}\} P\{M_2(0) = m_2 | \mathbf{C} = \mathbf{c}\}] \, P(\mathbf{C} = \mathbf{c})$$

- This can be interpreted as the indirect effect of X on Y mediated through the dependence between $M_1(1)$ and $M_2(1)$ (given $\mathbf{C}$).

Suppose the outcome obeys the model:

$$E(Y|X = x, M_1 = m_1, M_2 = m_2, \mathbf{C} = \mathbf{c})$$
$$= \theta_0 + \theta_1 x + \theta_2 m_1 + \theta_3 m_2 + \theta_4 m_1 m_2 + \theta_5 x m_1 + \theta_6 x m_2 + \theta_7^T \mathbf{c}$$

and the mediators (M_1, M_2), conditional on X and $\mathbf{C}$, have means

$$E(M_j|X = x, \mathbf{C} = \mathbf{c}) = \beta_{0j} + \beta_{1j} x + \beta_{2j}^T \mathbf{c},$$

with residual variances σ_j^2, $j = 1, 2$, and covariance σ_{12}.

Then the interventional direct effect is given by

$$E\left\{\theta_1 + \theta_5(\beta_{01} + \beta_{21}^T \mathbf{C}) + \theta_6(\beta_{02} + \beta_{22}^T \mathbf{C})\right\}$$
$$= \theta_1 + \theta_5\{\beta_{01} + \beta_{21}^T E(\mathbf{C})\} + \theta_6\{\beta_{02} + \beta_{22}^T E(\mathbf{C})\}.$$

This is θ_1 in the absence of exposure–mediator interactions.

Suppose the outcome obeys the model:

$$E(Y|X = x, M_1 = m_1, M_2 = m_2, \mathbf{C} = \mathbf{c})$$
$$= \theta_0 + \theta_1 x + \theta_2 m_1 + \theta_3 m_2 + \theta_4 m_1 m_2 + \theta_5 x m_1 + \theta_6 x m_2 + \theta_7^T \mathbf{c}$$

and the mediators (M_1, M_2), conditional on X and $\mathbf{C}$, have means

$$E(M_j|X = x, \mathbf{C} = \mathbf{c}) = \beta_{0j} + \beta_{1j} x + \beta_{2j}^T \mathbf{c},$$

with residual variances σ_j^2, $j = 1, 2$, and covariance σ_{12}.

The interventional indirect effect via M_1 is

$$\left[\theta_2 + \theta_4 \left\{\beta_{02} + \beta_{22}^T E(\mathbf{C})\right\} + \theta_5\right] \beta_{11}$$

which is $\theta_2\beta_{11}$ in the absence of exposure–mediator and mediator–mediator interactions.

Suppose the outcome obeys the model:

$$E(Y|X=x, M_1=m_1, M_2=m_2, \mathbf{C}=\mathbf{c})$$
$$=\theta_0+\theta_1 x+\theta_2 m_1+\theta_3 m_2+\theta_4 m_1 m_2+\theta_5 x m_1+\theta_6 x m_2+\theta_7^T \mathbf{c}$$

and the mediators (M_1, M_2), conditional on X and $\mathbf{C}$, have means

$$E(M_j|X=x, \mathbf{C}=\mathbf{c})=\beta_{0j}+\beta_{1j}x+\beta_{2j}^T\mathbf{c},$$

with residual variances σ_j^2, $j=1,2$, and covariance σ_{12}.

The interventional indirect effect via M_2 is

$$\left[\theta_3+\theta_4\left\{\beta_{01}+\beta_{11}+\beta_{21}^T E(\mathbf{C})\right\}+\theta_6\right]\beta_{12}$$

which is $\theta_3\beta_{12}$ in the absence of exposure–mediator and mediator–mediator interactions.

Suppose the outcome obeys the model:

$$E(Y|X = x, M_1 = m_1, M_2 = m_2, \mathbf{C} = \mathbf{c})$$
$$= \theta_0 + \theta_1 x + \theta_2 m_1 + \theta_3 m_2 + \theta_4 m_1 m_2 + \theta_5 x m_1 + \theta_6 x m_2 + \theta_7^T \mathbf{c}$$

and the mediators (M_1, M_2), conditional on X and $\mathbf{C}$, have means

$$E(M_j|X = x, \mathbf{C} = \mathbf{c}) = \beta_{0j} + \beta_{1j} x + \beta_{2j}^T \mathbf{c},$$

with residual variances σ_j^2, $j = 1, 2$, and covariance σ_{12}.

Finally, the indirect effect resulting from the effect of exposure on the mediators' dependence (the 'remainder' term) is

$$\theta_4 \sigma_{12} - \theta_4 \sigma_{12} = 0$$

Suppose the outcome obeys the model:

$$E(Y|X = x, M_1 = m_1, M_2 = m_2, \mathbf{C} = \mathbf{c})$$
$$= \theta_0 + \theta_1 x + \theta_2 m_1 + \theta_3 m_2 + \theta_4 m_1 m_2 + \theta_5 x m_1 + \theta_6 x m_2 + \theta_7^T \mathbf{c}$$

and the mediators (M_1, M_2), conditional on X and $\mathbf{C}$, have means

$$E(M_1|X = x, \mathbf{C} = \mathbf{c}) = \beta_{01} + \beta_{11} x + \beta_{21}^T \mathbf{c}$$
$$E(M_2|M_1 = m_1, X = x, \mathbf{C} = \mathbf{c}) = \beta_{02} + \beta_{12} x + \beta_{22}^T \mathbf{c} + \beta_{32} m_1 + \beta_{42} x m_1$$

with residual variances σ_j^2, $j = 1, 2$, and covariance σ_{12}.

If instead, X and M_1 interacted in their effect on M_2 in the sense above then the remainder term would be

$$\sigma_1^2 \theta_4 \beta_{42}$$

- This regression approach has the drawback that it requires a new derivation each time a different outcome or mediator model is considered.

- This regression approach has the drawback that it requires a new derivation each time a different outcome or mediator model is considered.
- This can be remedied via a Monte-Carlo approach, which involves sampling counterfactual values of the mediators from their respective distributions.

For instance, to evaluate the first component

$$\sum_{\mathbf{c}} \sum_{m_1} \sum_{m_2} E\left\{Y(1, m_1, m_2) | \mathbf{C} = \mathbf{c}\right\} P\{M_1(1) = m_1 | \mathbf{C} = \mathbf{c}\} P\{M_2(0) = m_2 | \mathbf{C} = \mathbf{c}\} P(\mathbf{C} = \mathbf{c})$$

of the interventional indirect effect through M_1, we can:

For instance, to evaluate the first component

$$\sum_{\mathbf{c}}\sum_{m_1}\sum_{m_2} E\left\{Y(1,m_1,m_2)|\mathbf{C}=\mathbf{c}\right\} P\{M_1(1)=m_1|\mathbf{C}=\mathbf{c}\} P\{M_2(0)=m_2|\mathbf{C}=\mathbf{c}\}P(\mathbf{C}=\mathbf{c})$$

of the interventional indirect effect through M_1, we can:

- take a random draw $M_{2,i}(0)$ for each subject i from the (fitted) distribution $P(M_2|X=0,\mathbf{C}_i)$

For instance, to evaluate the first component

$$\sum_{\mathbf{c}}\sum_{m_1}\sum_{m_2} E\left\{Y(1,m_1,m_2)|\mathbf{C}=\mathbf{c}\right\} P\{M_1(1)=m_1|\mathbf{C}=\mathbf{c}\} P\{M_2(0)=m_2|\mathbf{C}=\mathbf{c}\}P(\mathbf{C}=\mathbf{c})$$

of the interventional indirect effect through M_1, we can:

- take a random draw $M_{2,i}(0)$ for each subject i from the (fitted) distribution $P(M_2|X=0,\mathbf{C}_i)$
- then take a random draw $M_{1,i}(1)$ for each subject i from the (fitted) distribution $P(M_1|X=1,\mathbf{C}_i)$

For instance, to evaluate the first component

$$\sum_{\mathbf{c}} \sum_{m_1} \sum_{m_2} E\{Y(1, m_1, m_2)|\mathbf{C} = \mathbf{c}\} P\{M_1(1) = m_1|\mathbf{C} = \mathbf{c}\} P\{M_2(0) = m_2|\mathbf{C} = \mathbf{c}\} P(\mathbf{C} = \mathbf{c})$$

of the interventional indirect effect through M_1, we can:

- take a random draw $M_{2,i}(0)$ for each subject i from the (fitted) distribution $P(M_2|X = 0, \mathbf{C}_i)$
- then take a random draw $M_{1,i}(1)$ for each subject i from the (fitted) distribution $P(M_1|X = 1, \mathbf{C}_i)$
- Finally, we predict the outcome as the expected outcome under a suitable model with exposure set to 1, M_1 set to $M_{1,i}(1)$, M_2 set to $M_{1,i}(0)$, and covariates $\mathbf{C}_i$.

For instance, to evaluate the first component

$$\sum_{\mathbf{c}} \sum_{m_1} \sum_{m_2} E\{Y(1, m_1, m_2)|\mathbf{C} = \mathbf{c}\} P\{M_1(1) = m_1|\mathbf{C} = \mathbf{c}\} P\{M_2(0) = m_2|\mathbf{C} = \mathbf{c}\} P(\mathbf{C} = \mathbf{c})$$

of the interventional indirect effect through M_1, we can:

- take a random draw $M_{2,i}(0)$ for each subject i from the (fitted) distribution $P(M_2|X = 0, \mathbf{C}_i)$
- then take a random draw $M_{1,i}(1)$ for each subject i from the (fitted) distribution $P(M_1|X = 1, \mathbf{C}_i)$
- Finally, we predict the outcome as the expected outcome under a suitable model with exposure set to 1, M_1 set to $M_{1,i}(1)$, M_2 set to $M_{1,i}(0)$, and covariates $\mathbf{C}_i$.
- The average of these fitted values across subjects then estimates the above component.

- Its performance can be improved by repeating the random sampling many times and averaging the results across the different Monte-Carlo runs.

- Its performance can be improved by repeating the random sampling many times and averaging the results across the different Monte-Carlo runs.
- In practice, we recommend the bootstrap for inference.

Outline

Example: NYCRIS data

- Northern and Yorkshire Cancer Registry Information Service (NYCRIS), a population-based cancer registry covering 12% of the English population

Example: NYCRIS data

- Northern and Yorkshire Cancer Registry Information Service (NYCRIS), a population-based cancer registry covering 12% of the English population
- Survival to 1 year: 95.9% in higher SES women vs. 93.2% in lower SES women

Example: NYCRIS data

- Northern and Yorkshire Cancer Registry Information Service (NYCRIS), a population-based cancer registry covering 12% of the English population
- Survival to 1 year: 95.9% in higher SES women vs. 93.2% in lower SES women
- Survival to 5 years: 64.7% vs. 54.1%

Example: NYCRIS data

- Northern and Yorkshire Cancer Registry Information Service (NYCRIS), a population-based cancer registry covering 12% of the English population
- Survival to 1 year: 95.9% in higher SES women vs. 93.2% in lower SES women
- Survival to 5 years: 64.7% vs. 54.1%
- Question: what explains this? Screening? Treatment?

Example: NYCRIS data

- Data: 29,580 women diagnosed with malignant, invasive breast cancer 2000–2006.

Example: NYCRIS data

- Data: 29,580 women diagnosed with malignant, invasive breast cancer 2000–2006.
- X: SES (dichotomised just for simplicity here, from IMD2001)

Example: NYCRIS data

- Data: 29,580 women diagnosed with malignant, invasive breast cancer 2000–2006.
- X: SES (dichotomised just for simplicity here, from IMD2001)
- $\mathbf{M}_1$: Age and stage (TNM stage 1-2 vs 3-4) at diagnosis

Example: NYCRIS data

- Data: 29,580 women diagnosed with malignant, invasive breast cancer 2000–2006.
- X: SES (dichotomised just for simplicity here, from IMD2001)
- $\mathbf{M}_1$: Age and stage (TNM stage 1-2 vs 3-4) at diagnosis
- M_2: Treatment ('major' vs 'minor or no' surgery)

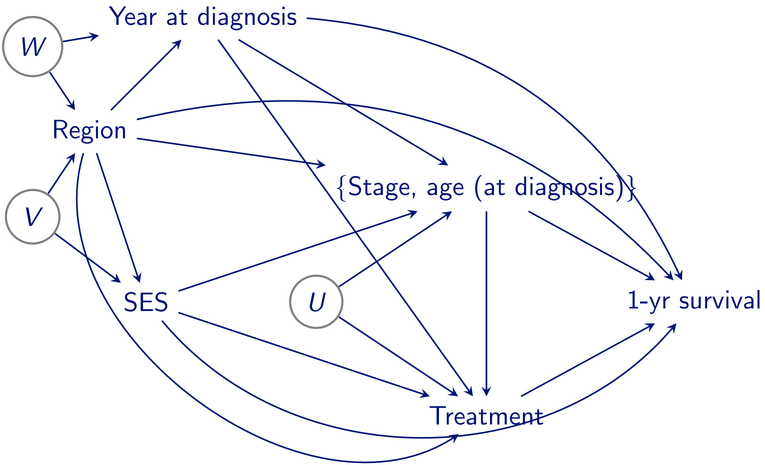
Example: NYCRIS data

- Data: 29,580 women diagnosed with malignant, invasive breast cancer 2000–2006.
- X: SES (dichotomised just for simplicity here, from IMD2001)
- $\mathbf{M}_1$: Age and stage (TNM stage 1-2 vs 3-4) at diagnosis
- M_2: Treatment ('major' vs 'minor or no' surgery)
- Y: Survival to 1-year post diagnosis

Example: NYCRIS data

- Data: 29,580 women diagnosed with malignant, invasive breast cancer 2000–2006.
- X: SES (dichotomised just for simplicity here, from IMD2001)
- $\mathbf{M}_1$: Age and stage (TNM stage 1-2 vs 3-4) at diagnosis
- M_2: Treatment ('major' vs 'minor or no' surgery)
- Y: Survival to 1-year post diagnosis
- $\mathbf{C}$: Year of diagnosis, region

DAG



W
Year at diagnosis
Region
V
{Stage, age (at diagnosis)}
SES
U
1-yr survival
Treatment

Results

- Mediation estimands estimated using Monte Carlo simulation (6,000,000 draws, 1,000 bootstrap samples)

Results

- Mediation estimands estimated using Monte Carlo simulation (6,000,000 draws, 1,000 bootstrap samples)
- All interactions included in all models.

Results

- Mediation estimands estimated using Monte Carlo simulation (6,000,000 draws, 1,000 bootstrap samples)
- All interactions included in all models.

Effect	Estimate	Bootstrap SE	95% CI lower	95% CI upper
Total causal effect	0.028	0.0028	0.023	0.034
Int DE	0.013	0.0027	0.008	0.018
Int IE through $\mathbf{M}_1$	0.007	0.0008	0.005	0.008
Int IE through M_2	0.0002	0.0003	–0.0005	0.0008
Remainder	0.007	0.0009	0.005	0.009

Results

- Mediation estimands estimated using Monte Carlo simulation (6,000,000 draws, 1,000 bootstrap samples)
- All interactions included in all models.

Effect	Estimate	Bootstrap SE	95% CI lower	95% CI upper
Total causal effect	0.028	0.0028	0.023	0.034
Int DE	0.013	0.0027	0.008	0.018
Int IE through $\mathbf{M}_1$	0.007	0.0008	0.005	0.008
Int IE through M_2	0.0002	0.0003	–0.0005	0.0008
Remainder	0.007	0.0009	0.005	0.009

Results: explaining the remainder term

Results of logistic regression of Treatment (M_2) on SES (X), Stage and Age at diagnosis ($\mathbf{M}_1$), and Region and Year of diagnosis ($\mathbf{C}$):

	Estimate	SE	95% CI lower	95% CI upper
Baseline odds*	4.796	0.226	4.373	5.261
Conditional odds ratios				
SES				
higher	0.725	0.026	0.677	0.777
Age at diagnosis (yrs)**	0.937	0.002	0.934	0.941
Stage				
advanced	0.186	0.009	0.169	0.205
SES × Agediag	1.033	0.003	1.027	1.038
SES × Stage	1.799	0.152	1.525	2.123
Agediag × Stage	1.014	0.004	1.007	1.021
SES × Agediag × Stage	0.974	0.006	0.962	0.985
Region				
North-West	1.806	0.155	1.526	2.138
Yorks	0.795	0.025	0.747	0.846
Year of diagnosis				
2001	1.089	0.061	0.976	1.214
2002	1.119	0.062	1.003	1.249
2003	1.248	0.069	1.120	1.390
2004	1.429	0.081	1.280	1.596
2005	1.411	0.079	1.265	1.575
2006	1.442	0.082	1.291	1.611

Treatment is coded 1 for major surgery and 0 for minor or no surgery. * estimated odds of major surgery for women diagnosed in the North East region in 2000, with low SES, age at diagnosis 62 years and early stage. ** centred at the mean age at diagnosis (61.8 years)

Interpretation of results (1)

- Without relying on any cross-world assumptions nor any assumptions about the causal structure of the mediators, our results would suggest that, of the 2.8% (95% CI 2.3%–3.4%) total difference in survival probability, about a quarter of this (0.7%, 95%CI 0.5%–0.9%) is mediated by the dependence of treatment on stage and age at diagnosis.

Interpretation of results (1)

- Without relying on any cross-world assumptions nor any assumptions about the causal structure of the mediators, our results would suggest that, of the 2.8% (95% CI 2.3%–3.4%) total difference in survival probability, about a quarter of this (0.7%, 95%CI 0.5%–0.9%) is mediated by the dependence of treatment on stage and age at diagnosis.
- Recall that we expected this effect to be small, except when there are particular interactions present, as is the case here.

Interpretation of results (1)

- Without relying on any cross-world assumptions nor any assumptions about the causal structure of the mediators, our results would suggest that, of the 2.8% (95% CI 2.3%–3.4%) total difference in survival probability, about a quarter of this (0.7%, 95%CI 0.5%–0.9%) is mediated by the dependence of treatment on stage and age at diagnosis.
- Recall that we expected this effect to be small, except when there are particular interactions present, as is the case here.
- Among women of a lower SES, there is a strong negative association between stage and treatment: those diagnosed at an advanced stage are less likely to receive major surgery.

Interpretation of results (1)

- Without relying on any cross-world assumptions nor any assumptions about the causal structure of the mediators, our results would suggest that, of the 2.8% (95% CI 2.3%–3.4%) total difference in survival probability, about a quarter of this (0.7%, 95%CI 0.5%–0.9%) is mediated by the dependence of treatment on stage and age at diagnosis.
- Recall that we expected this effect to be small, except when there are particular interactions present, as is the case here.
- Among women of a lower SES, there is a strong negative association between stage and treatment: those diagnosed at an advanced stage are less likely to receive major surgery.
- One possible interpretation would be that doctors and/or patients decide that treatment is not likely to be beneficial for patients with advanced disease, or that surgical treatment is substantially delayed for these patients due to tumor-reducing treatments such as chemotherapy being prioritised first.

Interpretation of results (2)

- This negative association is much less pronounced for women of higher SES.

Interpretation of results (2)

- This negative association is much less pronounced for women of higher SES.
- Therefore, we would interpret this estimated 0.7% as the increase in survival that would be expected if the treatment decision, as a function of stage and age at diagnosis (and baseline confounders), would be made for poorer women as it is currently made for higher SES women.

Interpretation of results (2)

- This negative association is much less pronounced for women of higher SES.
- Therefore, we would interpret this estimated 0.7% as the increase in survival that would be expected if the treatment decision, as a function of stage and age at diagnosis (and baseline confounders), would be made for poorer women as it is currently made for higher SES women.
- There is little evidence of further mediation through the treatment variable (estimated effect 0.02%, 95% CI: –0.05, 0.08%), and evidence of an effect through age and stage at diagnosis (estimated effect 0.7%, 95%CI 0.5%–0.8%).

Interpretation of results (2)

- This negative association is much less pronounced for women of higher SES.
- Therefore, we would interpret this estimated 0.7% as the increase in survival that would be expected if the treatment decision, as a function of stage and age at diagnosis (and baseline confounders), would be made for poorer women as it is currently made for higher SES women.
- There is little evidence of further mediation through the treatment variable (estimated effect 0.02%, 95% CI: –0.05, 0.08%), and evidence of an effect through age and stage at diagnosis (estimated effect 0.7%, 95%CI 0.5%–0.8%).
- This would suggest that an additional 0.7% reduction in one-year mortality for lower SES women could be achieved if the distribution of age and stage at diagnosis (given year of diagnosis and region) were changed from that seen in lower SES women to that of higher SES women, a change that could perhaps be affected by encouraging better uptake of screening and other health-seeking behaviour among lower SES women.

Summary (1)

- Mediation analysis, although intuitive and with a long history, is a surprisingly subtle business as soon as there are any non-linearities in the picture.

Summary (1)

- Mediation analysis, although intuitive and with a long history, is a surprisingly subtle business as soon as there are any non-linearities in the picture.
- Advances thanks to the field of causal inference have greatly clarified these subtleties, giving rise to clear estimands that capture the notions of direct and indirect effects, clear assumptions under which these can be identified, and flexible estimation methods.

Summary (1)

- Mediation analysis, although intuitive and with a long history, is a surprisingly subtle business as soon as there are any non-linearities in the picture.
- Advances thanks to the field of causal inference have greatly clarified these subtleties, giving rise to clear estimands that capture the notions of direct and indirect effects, clear assumptions under which these can be identified, and flexible estimation methods.
- However, this endeavour has been limited by the extremely strong and untestable cross-world assumption.

Summary (1)

- Mediation analysis, although intuitive and with a long history, is a surprisingly subtle business as soon as there are any non-linearities in the picture.
- Advances thanks to the field of causal inference have greatly clarified these subtleties, giving rise to clear estimands that capture the notions of direct and indirect effects, clear assumptions under which these can be identified, and flexible estimation methods.
- However, this endeavour has been limited by the extremely strong and untestable cross-world assumption.
- This has effectively prohibited flexible multiple mediation analyses, even though applied problems frequently involve multiple mediators.

Summary (1)

- Mediation analysis, although intuitive and with a long history, is a surprisingly subtle business as soon as there are any non-linearities in the picture.
- Advances thanks to the field of causal inference have greatly clarified these subtleties, giving rise to clear estimands that capture the notions of direct and indirect effects, clear assumptions under which these can be identified, and flexible estimation methods.
- However, this endeavour has been limited by the extremely strong and untestable cross-world assumption.
- This has effectively prohibited flexible multiple mediation analyses, even though applied problems frequently involve multiple mediators.
- Interventional effects are perhaps the way forward, since they don't require this cross-world assumption.

Summary (2)

- We have shown how interventional effects can be used in multiple mediator settings.

Summary (2)

- We have shown how interventional effects can be used in multiple mediator settings.
- A big advantage of our approach is that no assumption need be made regarding the causal structure of the mediators.

Summary (2)

- We have shown how interventional effects can be used in multiple mediator settings.
- A big advantage of our approach is that no assumption need be made regarding the causal structure of the mediators.
- The price we must pay for this is that the decomposition includes a remainder term which can be interpreted as a mediated dependence.

Summary (2)

- We have shown how interventional effects can be used in multiple mediator settings.
- A big advantage of our approach is that no assumption need be made regarding the causal structure of the mediators.
- The price we must pay for this is that the decomposition includes a remainder term which can be interpreted as a mediated dependence.
- We have seen that at least in some settings, this parameter has a real-world interpretation.

Summary (2)

- We have shown how interventional effects can be used in multiple mediator settings.
- A big advantage of our approach is that no assumption need be made regarding the causal structure of the mediators.
- The price we must pay for this is that the decomposition includes a remainder term which can be interpreted as a mediated dependence.
- We have seen that at least in some settings, this parameter has a real-world interpretation.
- The next steps include seeing how well this approach extends to problems with more than 2 mediators.

Outline

References: traditional approach

Wright, S. (1921)
Correlation and causation: Part I - Method of Path Coefficients
Journal of Agriculture Research, 20:557–575.

Wright, S. (1934) The method of path coefficients
Annals of Mathematical Statistics, 5(3)161–215.

Baron, R. M. and Kenny, D. A. (1986)
The moderator-mediator variable distinction in social psychological research: conceptual, strategic and statistical considerations
Journal of Personality and Social Psychology, 51:1173–1182.

References: early investigations by causal inference

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

Pearl, J. (2001)
Direct and indirect effects.
Proceedings of the 17th Conference in Uncertainty in Artificial Intelligence, 411–420.

Cole, S.R. and Hernán, M.A. (2002)
Fallibility in estimating direct effects.
International Journal of Epidemiology, 31(1):163–165.

References: more on identification and estimation

Robins, J. (1999)

Testing and estimation of direct effects by reparameterizing directed acyclic graphs with structural nested models.

In *Computation, Causation, and Discovery*, C. Glymour & G. Cooper, eds. Menlo Park, CA, Cambridge: AAAI Press/The MIT Press, pp. 349–405.

Petersen, M.L., Sinisi, S.E. and van der Laan, M.J. (2006)

Estimation of direct causal effects.

Epidemiology, 17:276–284.

VanderWeele, T.J. and Vansteelandt, S. (2009)

Conceptual issues concerning mediation, interventions and composition.

Statistics and its Interface, 2:457–468.

References: more on identification and estimation

VanderWeele, T.J. (2009)
Marginal structural models for the estimation of direct and indirect effects.
Epidemiology, 20:18–26.

Joffe, M. and Greene, T. (2009)
Related causal frameworks for surrogate outcomes.
Biometrics, 65:530–538.

Tchetgen Tchetgen, E.J. (2013)
Inverse odds ratio-weighted estimation for causal mediation.
Statistics in Medicine, 32:4567–80.

References: key textbook on causal mediation analysis

VanderWeele, T.J. (2015)
Explanation in Causal Inference: Methods for Mediation and Interaction.
Oxford University Press.

References: on interventional effects

Vanderweele, T.J., Vansteelandt, S. and Robins, J.M. (2014)
Effect decomposition in the presence of an exposure-induced mediator-outcome confounder.
Epidemiology, 25:300–306.

Vansteelandt, S. and Daniel, R.M..
Interventional effects for mediation analysis with multiple mediators.
Epidemiology, **in press**.