Causal

~~Casual~~ mediation analysis

Aramayis Dallakyan
StataCorp LLC

April 9, 2024

Stata's `mediate` command

- New in Stata 18: **`mediate`**
- Performs causal mediation analysis for linear and generalized linear models.

Stata's `mediate` command

`mediate` (*ovar* [*omvarlist*, *omodel* <u>nocons</u>tant])
(*mvar* [*mmvarlist*, *mmodel* <u>nocons</u>tant])
(*tvar* [, <u>cont</u>inuous(*numlist*)]) [*if*] [*in*] [*weight*] [, *stat options*]

ovar is a continuous, binary, or count outcome of interest.

omvarlist specifies the covariates in the outcome model.

mvar is the mediator variable and may be continuous, binary, or count.

mmvarlist specifies the covariates in the mediator model.

tvar is the treatment variable and may be binary, multivalued, or continuous.

Stata's `mediate` command

Mediator / *Outcome*	linear	logit	probit	Poisson	exp. mean
linear	X	X	X	X	X
logit		X	X	X	
probit	X	X	X	X	X
Poisson	X	X	X	X	X
exp. mean	X	X	X	X	X

Note: X indicates a supported model combination

Stata's mediate command

```
. mediate (wellbeing) (bonotonin) (exercise)

Final EE criterion =  2.04e-28

Causal mediation analysis                                  Number of obs = 2,000

Outcome model:      Linear
Mediator model:     Linear
Mediator variable: bonotonin
Treatment type:     Binary
```

wellbeing	Coefficient	Robust std. err.	z	P>|z|	[95% conf.	interval]
NIE						
exercise						
(Exercise vs Control)	9.799821	.3943251	24.85	0.000	9.026958	10.57268
NDE						
exercise						
(Exercise vs Control)	2.891453	.2304278	12.55	0.000	2.439823	3.343083
TE						
exercise						
(Exercise vs Control)	12.69127	.4005941	31.68	0.000	11.90612	13.47642

Note: Outcome equation includes treatment–mediator interaction.

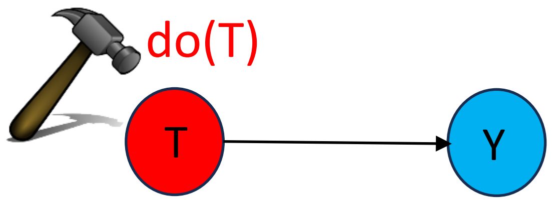
Outline

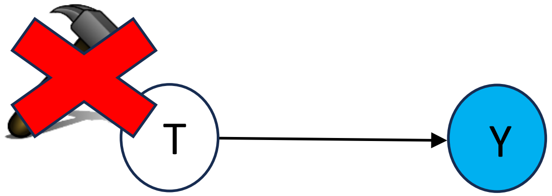
- Basics of causal thinking and inference
    - Introduction and motivation
    - Potential-outcomes framework and DAGs
    - Fundamental steps of causal inference

- Causal mediation analysis
    - Direct and indirect effects
    - Identification
    - Demonstration

Causal thinking and causal inference

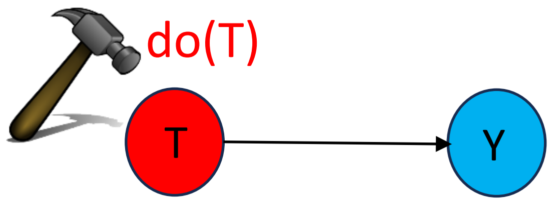
- Causal inference tackles the fundamental questions of cause and effect.
- The causal effect aims to compare the outcome when an action T is taken versus the outcome when the action T is withheld.

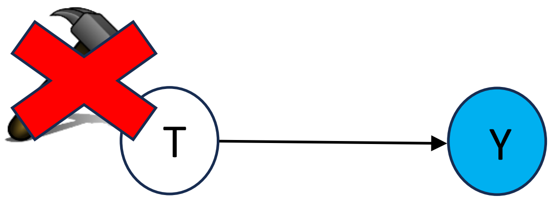




Causal inference

- Causal inference tackles the fundamental questions of cause and effect.
- The causal effect aims to compare the outcome when an action T is taken versus the outcome when the action T is withheld.



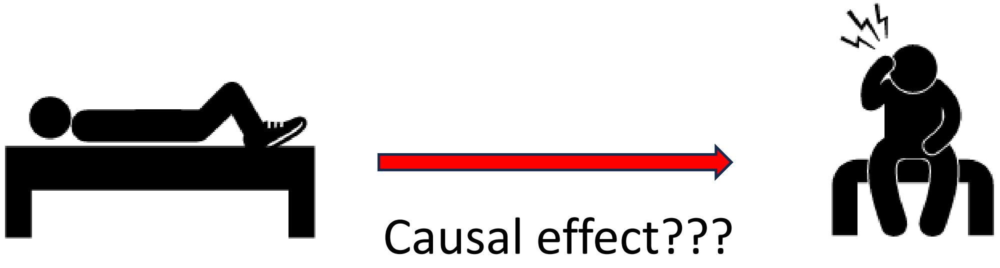


- We refer to action T as an intervention, an exposure, or a treatment.
  - Effect of a treatment/drug/vaccine on a disease;
  - Effect of social media on mental health;
  - Effect of genes on a disease, etc.

Causal inference

- **Why do we need causality?**
- Why association or statistical dependence is not enough?
- Association does not imply causation!
- The amount of association and the amount of causation can be different

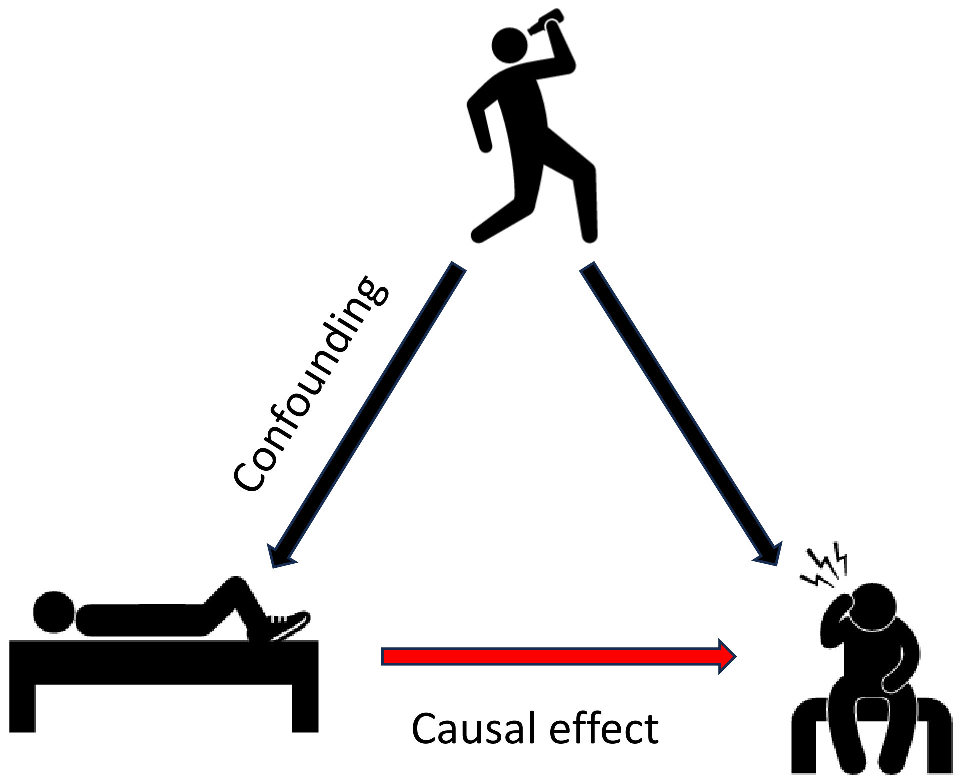
Causal inference



- Suppose we analyze data where the "treatment" is sleeping with shoes on (or not), and the outcome is waking up with a headache (or not) the next day.
- We find that most times when someone wears shoes to bed, that person wakes up with a headache.
- Question: Can we interpret this relationship as causal?

Causal inference

- One possible explanations for association
  - Both treatment and outcome are caused by a common cause: drinking the night before.
  - Such variables are known as confounders and the association as confounding association.
  - Confounding is the main source of differentiating association from causation.



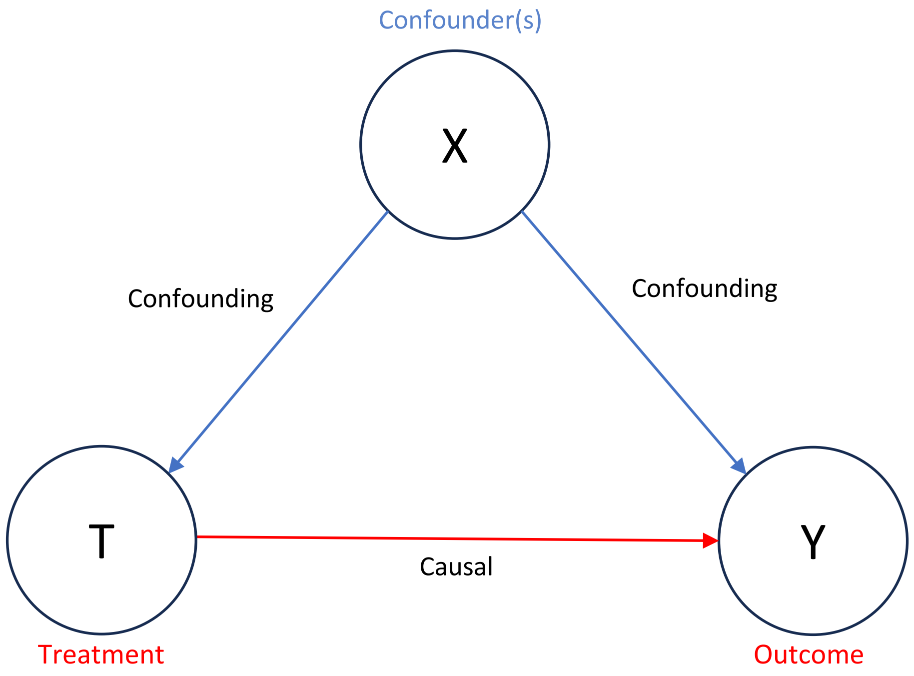
*Borrowed from Neal (2020)

Causal inference

- Our goal: Learn about causal effects
  - Represent the causal structure
  - Characterize the causal effect

- Notation:
  - **T** ϵ {0,1} denotes treatment assignment: Wearing shoes vs not wearing shoes to bed
  - **Y** denotes the outcome: Headache vs no headache
  - **X** denotes potential confounders that affect both **T** and **Y:** Drinking the previous day

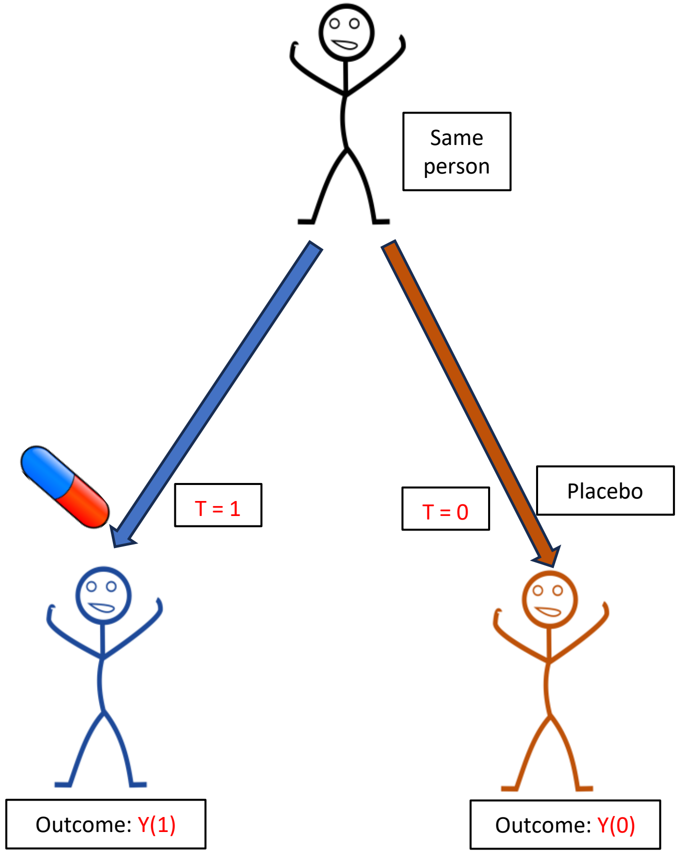
Directed acyclic graphs (DAGs)



- We use DAGs to represent causal relationships and structure.
- Arrows indicate a direct causal effect (not mediated) for at least one subject.
- Informally, **the goal of causal inference** is to estimate the **causal part** of the graph while controlling for the **confounding part**.

Potential-outcomes framework

- To characterize the causal effect we use the potential outcomes framework.



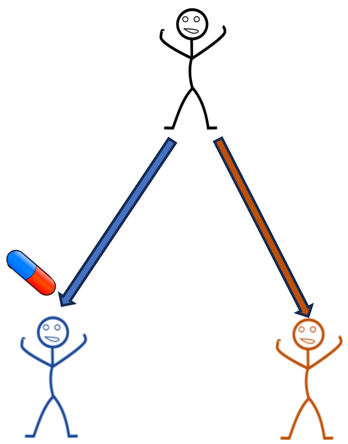
- The potential outcome **Y(T = t) = Y(t)** is the outcome we would have observed had T = t been assigned.
- The causal effect can be measured as Y(1) -Y(0), which is the change due to the treatment keeping everything else the same.

Potential-outcomes framework

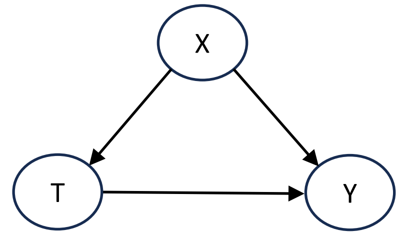
- Fundamental Problem of Causal Inference: Only one of **{Y(1), Y(0)}** is observed.

- The *observed* potential outcome is called **factual.**
- The *unobserved* potential outcome is called **counterfactual**.

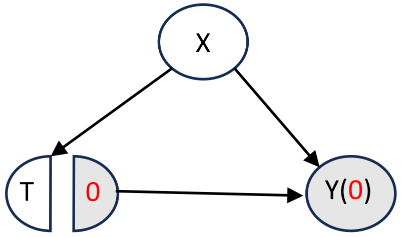
- The causal effect is a **contrast between two parallel worlds**, which we imagine for the same subject.



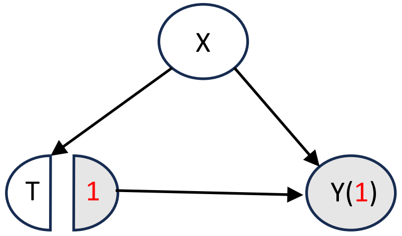
Potential-outcomes framework



- Actual, observed world.



- In this world everything is the same but **T** is set to 0.



- In this world everything is the same but **T** is set to 1.

- **Note** that compared to the observed world, in imaginary worlds the causal link between **X** and treatment **T** is broken.

Potential-outcomes framework

Subject	T	Y	Y(1)	Y(0)	Y(1) – Y(0)
1	0	2.1	?	2.1	?
2	1	3.7	3.7	?	?
3	1	4.2	4.2	?	?
4	0	6.2	?	6.2	?
…	…	…	…	…	…

- The observed outcome:

$$Y = T*Y(1) + (1 - T)*Y(0)$$

- For the subject with treatment T = 1

$$Y = 1*Y(1) + 0*Y(0)$$

- Similarly, for T = 0

$$Y = 0*Y(1) + (1 - 0)*Y(0)$$

- Thus, Y(1) – Y(0) is never observed for subject *i*.

Potential-outcomes framework

Subject	T	Y	Y(1)	Y(0)	Y(1) – Y(0)
1	0	2.1	?	2.1	?
2	1	3.7	3.7	?	?
3	1	4.2	4.2	?	?
4	0	6.2	?	6.2	?
…	…	…	…	…	…

- The observed outcome:

  Y = T*Y(1) + (1 – T)*Y(0)
- For the subject with treatment T = 1

  Y = 1*Y(1) + 0*Y(0)
- Similarly, for T = 0

  Y = 0*Y(1) + (1 – 0)*Y(0)
- Thus, Y(1) – Y(0) is never observed for subject *i*.

- Natural measure of causal effect is the average treatment effect (ATE)

$$\mu = E[Y(1) - Y(0)]$$

Potential-outcomes framework

- **Important question:** Is it possible to estimate the ATE if $Y(1) - Y(0)$ is never observed?
  - Yes, but under certain causal assumptions.

- Causal inference helps in moving observables **(Y, T, X)** to the distribution **{Y(0), Y(1), T , X}**.

- Causal Inference is much more than familiar statistical inference
  - Statistical inference: from sample to population
  - Causal Inference: from sample to counterfactual populations

Causal identification

- Causal identification: the process of learning a causal estimand (ATE) $\mu = \mu_1 - \mu_0$ with $\mu = E[Y(t)]$, t = 0,1 from observed data (Y_i, T_i, X_i).

Causal identification

- Causal identification: the process of learning a causal estimand (ATE) $\mu = \mu_1 - \mu_0$ with $\mu = E[Y(t)]$, t = 0,1 from observed data (Y_i, T_i, X_i).

- Question: Can we naively estimate ATE $\mu = E[Y(1) - Y(0)]$ via the association?

$$\mu_{naive} = E[Y|T = 1] - E[Y|T = 0]$$

Subject	T	Y	Y(1)	Y(0)	Y(1) – Y(0)
1	0	2.1	?	2.1	?
2	1	3.7	3.7	?	?
3	1	4.2	4.2	?	?
4	0	6.2	?	6.2	?
…	…	…	…	…	…

Causal identification

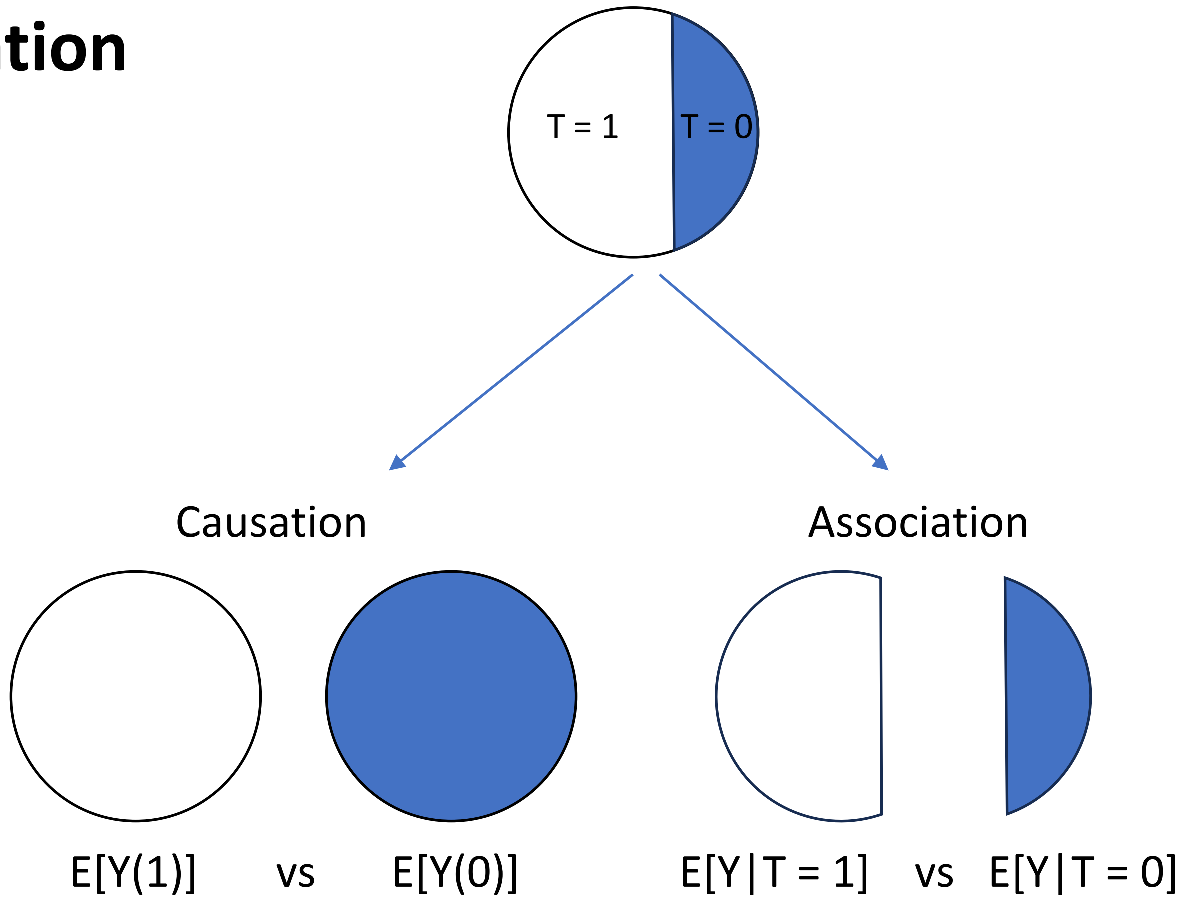
- Causal identification: the process of learning a causal estimand (ATE) $\mu = \mu_1 - \mu_0$ with $\mu = E[Y(t)]$, t = 0,1 from observed data (Y_i, T_i, X_i).

- Question: Can we naively estimate ATE $\mu = E[Y(1) - Y(0)]$ via the association?

$$\mu_{naive} = E[Y|T = 1] - E[Y|T = 0]$$

- Answer: In general NO. Recall the shoe example.

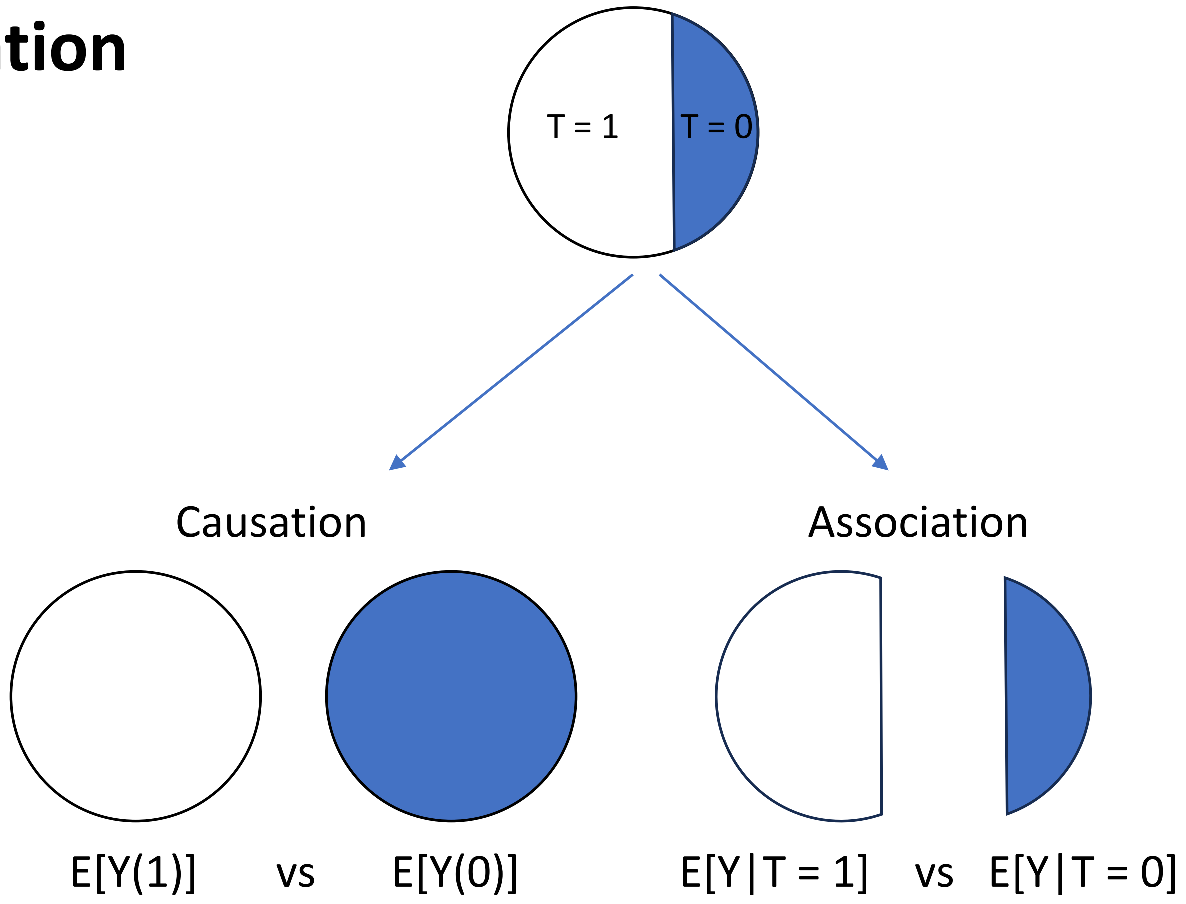
Causal identification



- In general, the causal effect is not the association effect:

$$E[Y(1)] - E[Y(0)] \neq E[Y|T = 1] - E[Y|T = 0]$$

Causal identification

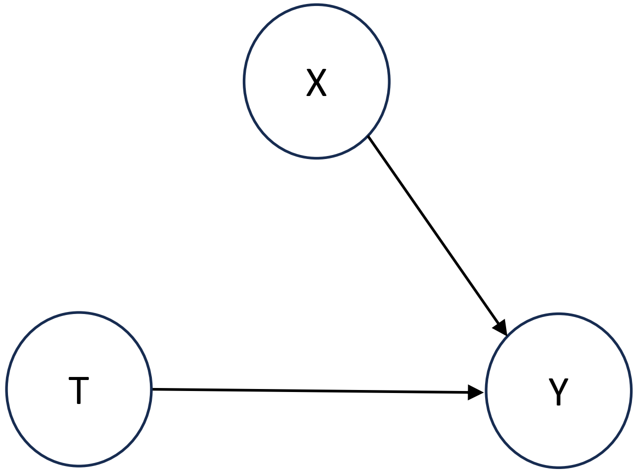


- In general, the causal effect is not the association effect:

$$E[Y(1)] - E[Y(0)] \neq E[Y|T = 1] - E[Y|T = 0]$$

- Question: When are they equal?

Randomized control trials (RCTs)



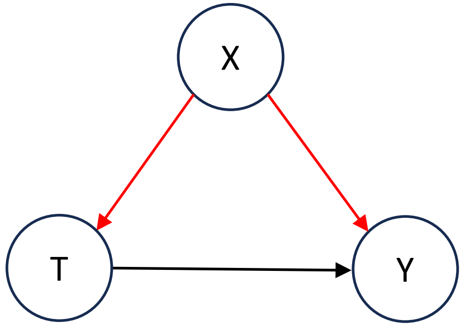
- RCTs randomize **T**, i.e., **T** is independent of **{Y(0),Y(1), X}**.

- Consequently, it removes any confounding effect.

$$E[Y|T = t] = E[Y(t)|T = t] = E[Y(t)]$$

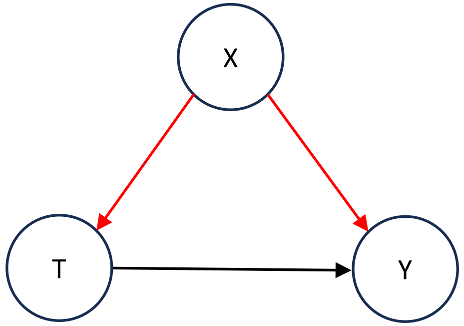
- In other words, in RCTs an observed association between **T** and **Y** is a causal association

Observational data

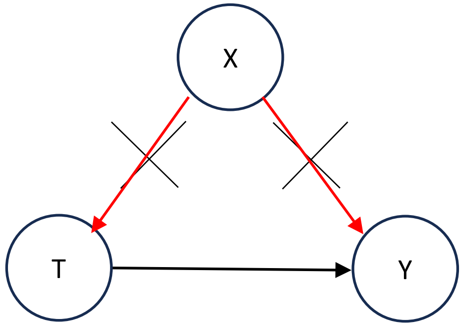


- In observational data, T is not independent of **{Y(0),Y(1), X}**
- Hence, the association between T and Y includes confounding/selection bias.

Observational data



- In observational data, T is not independent of **{Y(0),Y(1), X}**
- Hence, the association between T and Y includes confounding/selection bias.



- We need additional causal assumptions that will block/eliminate the bias.

Causal assumptions

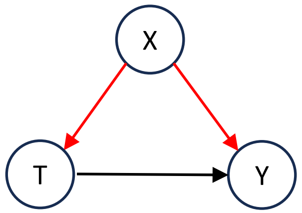
- Conditional ignorability or unconfoudedness assumption:
  - **T** is independent of **Y(1), Y(0)|X**
  - Informally, it says given confounders **X**, the treatment **T** is as good as random.
  - This assumption cannot be tested from the data.

- Other assumptions: Positivity, consistency and SUTVA

- Under the above assumptions, the causal effect is identified:

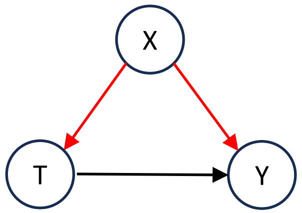
$$E[Y(1)] - E[Y(0)] = E_X\{E[Y|T = 1,X] - E[Y|T = 0,X]\}$$

Summary: Fundamental Steps of Causal Inference

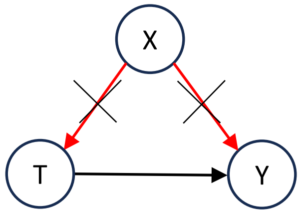


- Hypothetical modeling: Researchers make causal assumptions about relationships among variables based on their understanding and expertise.

Summary: Fundamental Steps of Causal Inference

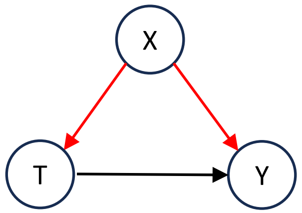


- Hypothetical modeling: Researchers make causal assumptions about relationships among variables based on their understanding and expertise.

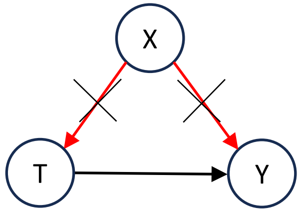


- Causal identification: Based on the previous assumptions, researchers try to determine whether the causal effect is identified, i.e., bias elimination.

Summary: Fundamental Steps of Causal Inference



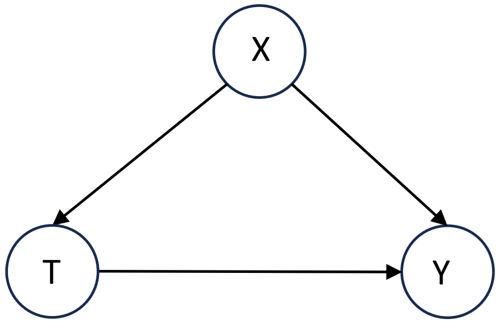
- Hypothetical modeling: Researchers make causal assumptions about relationships among variables based on their understanding and expertise.



- Causal identification: Based on the previous assumptions, researchers try to determine whether the causal effect is identified, i.e., bias elimination.

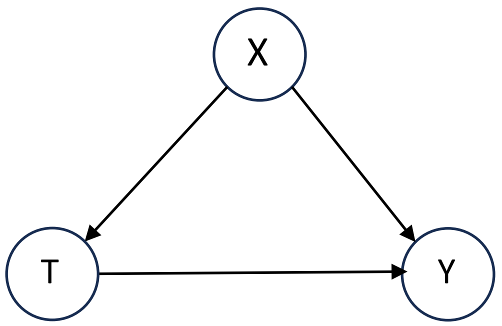
- Parameter estimation: If the answer to the second phase is positive, the researcher can then use various estimation techniques, such as **`teffects`** or **`mediate`** to estimate the causal effect.

Causal mediation

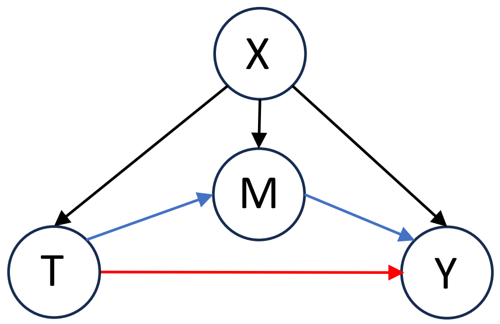


- Suppose using the Fundamental Steps of Causal Analysis (FSCA), a researcher concluded that **exercise**, **T**, has a beneficial causal effect on the perceptions of the **well-being of individuals**, **Y**.

Causal mediation

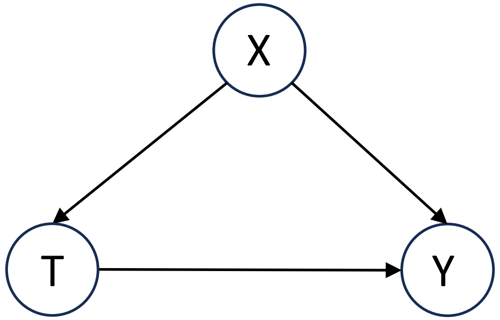


- Suppose using the Fundamental Steps of Causal Analysis (FSCA), a researcher concluded that **exercise**, **T**, has a beneficial causal effect on the perceptions of the **well-being of individuals**, **Y**.

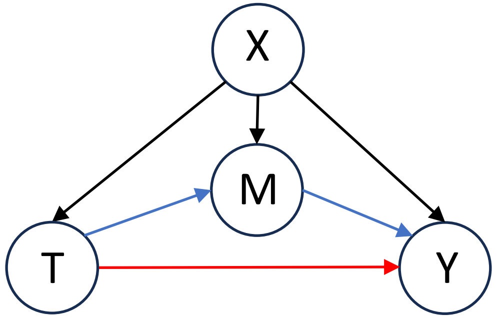


- Now, the researcher wonders whether the benefit is a consequence of the effect of **T** on increasing the level of the **hormone bonotonin**, **M**, which in turn has a positive effect on subjective well-being, **Y**.

Causal mediation



- Suppose using the Fundamental Steps of Causal Analysis (FSCA), a researcher concluded that **exercise**, **T**, has a beneficial causal effect on the perceptions of the **well-being of individuals**, **Y**.



- Now, the researcher wonders whether the benefit is a consequence of the effect of **T** on increasing the level of the **hormone bonotonin**, **M**, which in turn has a positive effect on subjective well-being, **Y**.

- That is, the researcher is interested in decomposing the **total effect of T on Y** into the indirect causal pathway mediated by M and the direct pathway not mediated by M.

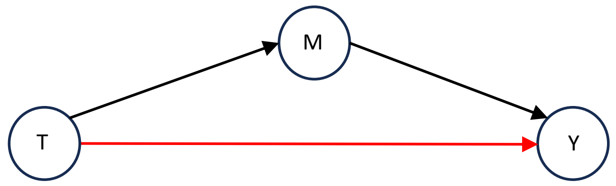
Causal mediation: The fundamental steps of causal analysis

- Suppose we want to estimate the mediation effect of **hormone bonotonin**, **M**, between the effect of **exercise**, **T**, on subjective **wellbeing**, **Y**.

Causal mediation: The fundamental steps of causal analysis

- Suppose we want to estimate the mediation effect of **hormone bonotonin**, **M**, between the effect of **exercise**, **T**, on subjective **wellbeing**, **Y**.

- Step 1: Hypothetical modeling

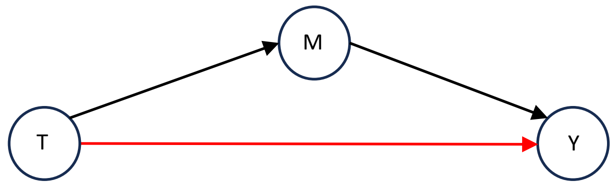


- **T** – exercise
- **M** – bonotonin
- **Y** – well-being

Causal mediation: The fundamental steps of causal analysis

- Suppose we want to estimate the mediation effect of **hormone bonotonin**, **M**, between the effect of **exercise**, **T**, on subjective **wellbeing**, **Y**.

- Step 1: Hypothetical modeling



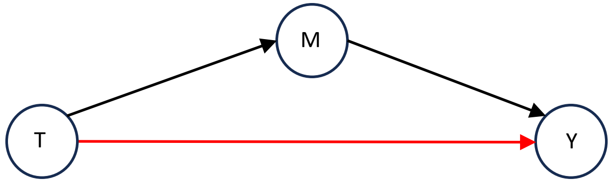
- **T** – exercise
- **M** – bonotonin
- **Y** – well-being

- Step 2: Causal identification – more on this later

Causal mediation: The fundamental steps of causal analysis

- Suppose we want to estimate the mediation effect of **hormone bonotonin**, **M**, between the effect of **exercise**, **T**, on subjective **wellbeing**, **Y**.

- Step 1: Hypothetical modeling



- **T** – exercise
- **M** – bonotonin
- **Y** – well-being

- Step 2: Causal identification – more on this later

- Step 3: Estimation in Stata

Demonstration: The data

```
. webuse wellbeing
(Fictional well-being data)

. list wellbeing bonotonin exercise in 1/5, abbreviate(12) clean

       wellbeing   bonotonin   exercise
  1.    71.73816    196.5467    Control
  2.    68.66573    195.8572   Exercise
  3.    71.05155    228.6035   Exercise
  4.    69.44469    206.6651   Exercise
  5.    75.62035    261.6855   Exercise
```

Demonstration: Stata's `mediate` command

`mediate` (*ovar* [*omvarlist*, *omodel* `noconstant`])
(*mvar* [*mmvarlist*, *mmodel* `noconstant`])
(*tvar* [, `continuous`(*numlist*)]) [*if*] [*in*] [*weight*] [, *stat options*]

ovar is a continuous, binary, or count outcome of interest.

omvarlist specifies the covariates in the outcome model.

mvar is the mediator variable and may be continuous, binary, or count.

mmvarlist specifies the covariates in the mediator model.

tvar is the treatment variable and may be binary, multivalued, or continuous.

Demonstration: Estimation

```
. mediate (wellbeing) (bonotonin) (exercise)

Final EE criterion =  2.04e-28

Causal mediation analysis                              Number of obs = 2,000

Outcome model:     Linear
Mediator model:    Linear
Mediator variable: bonotonin
Treatment type:    Binary
```

wellbeing	Coefficient	Robust std. err.	z	P>|z|	[95% conf.	interval]
NIE						
exercise						
(Exercise vs Control)	9.799821	.3943251	24.85	0.000	9.026958	10.57268
NDE						
exercise						
(Exercise vs Control)	2.891453	.2304278	12.55	0.000	2.439823	3.343083
TE						
exercise						
(Exercise vs Control)	12.69127	.4005941	31.68	0.000	11.90612	13.47642

Note: Outcome equation includes treatment-mediator interaction.

Demonstration: Estimation

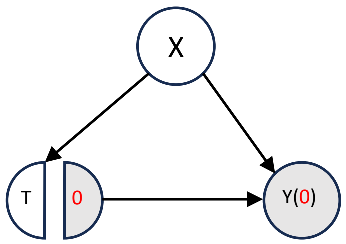
- **`mediate`** uses a method of moments estimator, also known as an estimating equations estimator, to estimate all auxiliary and effect parameters as well as their variance–covariance matrix.
- To report the auxiliary parameters:

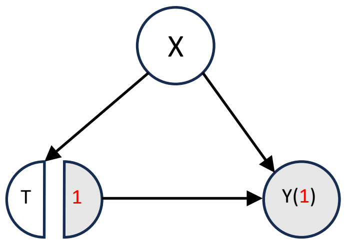
```
. mediate, aequations
 < output omitted >
──────────────────────────┬──────────────────────────────────────────────────────────────
wellbeing                 │
                  exercise│
                 Exercise │   2.065871    .8723559     2.37   0.018      .3560846    3.775657
                 bonotonin│   .2130222    .0034547    61.66   0.000      .2062512    .2197932
                          │
     exercise#c.bonotonin │
                 Exercise │   .0051424    .0046954     1.10   0.273     -.0040604    .0143452
                          │
                     _cons│   22.91374    .5633648    40.67   0.000      21.80956    24.01791
──────────────────────────┼──────────────────────────────────────────────────────────────
bonotonin                 │
                  exercise│
                 Exercise │   44.91939    1.641668    27.36   0.000      41.70178      48.137
                     _cons│    160.544    1.142508   140.52   0.000      158.3047    162.7832
──────────────────────────┴──────────────────────────────────────────────────────────────
```

Demonstration: Estimation without interaction

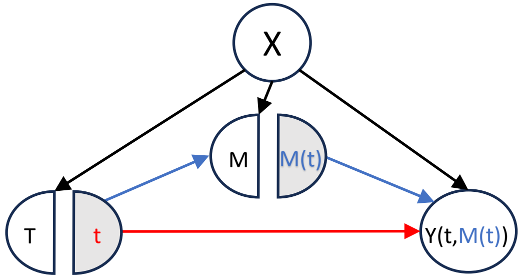
```
. mediate (wellbeing) (bonotonin) (exercise), aequations nointeract
 < output omitted >
─────────────────────────┬──────────────────────────────────────────────────────────────────
wellbeing                │
             exercise    │
            Exercise     │   2.996658    .2109357    14.21   0.000     2.583231    3.410084
            bonotonin    │   .2158225    .0023412    92.18   0.000     .2112338    .2204113
                _cons    │   22.46416    .3929094    57.17   0.000     21.69407    23.23425
─────────────────────────┼──────────────────────────────────────────────────────────────────
bonotonin                │
             exercise    │
            Exercise     │   44.91939    1.641668    27.36   0.000     41.70178      48.137
                _cons    │    160.544    1.142508   140.52   0.000     158.3047    162.7832
─────────────────────────┴──────────────────────────────────────────────────────────────────
Note: Outcome equation does not include treatment-mediator interaction.
```

Taking a step back: Preparing for causal identification





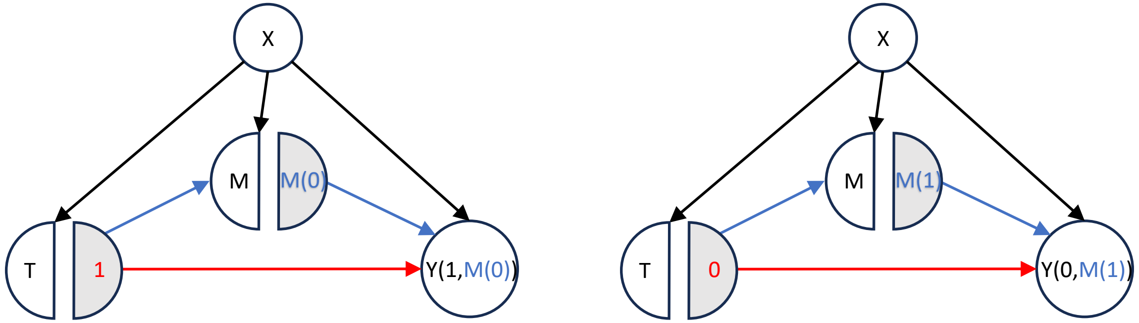
- Recall that our interest is in the contrast **Y(1) - Y(0)**
- For mediation, the idea is to split the contrast **Y(1) - Y(0)** into two other contrasts using a third potential outcome **M(t).**



- We introduce a new type of outcome **Y(t,m)**, which corresponds to the potential outcome when we set **T = t** and **M=m**.

- Note the familiar **Y(1) = Y[1, M(1)]** and **Y(0) = Y[0,M(0)]**.

Four potential outcomes



- Now, we have two new cross-world potential outcomes **Y[t, M(t')].**
- **Y[1,M(0)]** and **Y[0,M(1)]** are never observed (Fundamental problem of causal inference).
- These correspond to the unobserved worlds where treatment is set to **t** and the mediator is set to the value it would have taken under exposure **t'**.
- We use these **four potential outcomes** to define total effects, direct effects, and indirect effects

Potential-outcome means with `mediate`

```
. mediate (wellbeing) (bonotonin) (exercise), pomeans

Final EE criterion =  1.71e-28

Causal mediation analysis                                 Number of obs = 2,000

Outcome model:     Linear
Mediator model:    Linear
Mediator variable: bonotonin
Treatment type:    Binary
```

wellbeing	Coefficient	Robust std. err.	z	P>|z|	[95% conf.	interval]
POmeans						
Y0M0	57.11317	.2753201	207.44	0.000	56.57355	57.65278
Y1M0	60.00462	.3157888	190.02	0.000	59.38569	60.62356
Y0M1	66.68199	.3258477	204.64	0.000	66.04334	67.32064
Y1M1	69.80444	.2898927	240.79	0.000	69.23626	70.37262

Note: Outcome equation includes treatment–mediator interaction.

Different treatment effects

- The average total effect:

$$\tau = E[Y_i(1)] - E[Y_i(0)] = E\big[Y_i\big(1, M_i(1)\big)\big] - E\big[Y_i\big(0, M_i(0)\big)\big]$$

Effect decomposition

- The average total effect:

$$\tau = E[Y_i(1)] - E[Y_i(0)] = E\left[Y_i\left(1, M_i(1)\right)\right] - E\left[Y_i\left(0, M_i(0)\right)\right]$$

- The effect of the treatment on the outcome through the mediator is the indirect effect:

$$\delta(t) = E\left[Y_i\left(t, M_i(1)\right)\right] - E\left[Y_i\left(t, M_i(0)\right)\right], \qquad t \in \{0,1\}$$

Effect decomposition

- The average total effect:

$$\tau = E[Y_i(1)] - E[Y_i(0)] = E\left[Y_i\left(1, M_i(1)\right)\right] - E\left[Y_i\left(0, M_i(0)\right)\right]$$

- The effect of the treatment on the outcome through the mediator is the indirect effect:

$$\delta(t) = E\left[Y_i\left(t, M_i(1)\right)\right] - E\left[Y_i\left(t, M_i(0)\right)\right], \qquad t \in \{0,1\}$$

- The direct effect of the treatment on the outcome

$$\zeta(t) = E\left[Y_i\left(1, M_i(t)\right)\right] - E\left[Y_i\left(0, M_i(t)\right)\right], \qquad t \in \{0,1\}$$

Effect decomposition

- The average total effect:

$$\tau = E[Y_i(1)] - E[Y_i(0)] = E\left[Y_i\left(1, M_i(1)\right)\right] - E\left[Y_i\left(0, M_i(0)\right)\right]$$

- The effect of the treatment on the outcome through the mediator is the indirect effect:

$$\delta(t) = E\left[Y_i\left(t, M_i(1)\right)\right] - E\left[Y_i\left(t, M_i(0)\right)\right], \qquad t \in \{0,1\}$$

- The direct effect of the treatment on the outcome

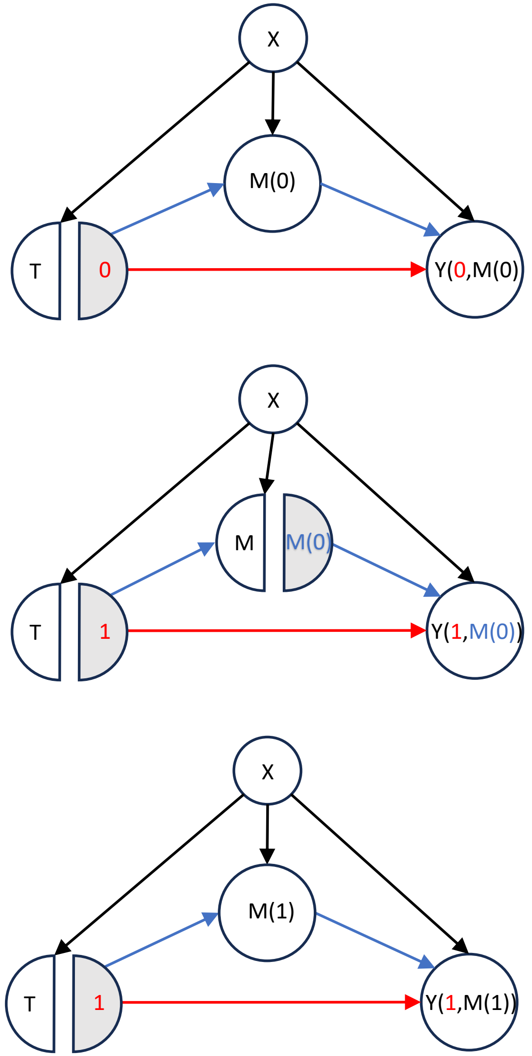
$$\zeta(t) = E\left[Y_i\left(1, M_i(t)\right)\right] - E\left[Y_i\left(0, M_i(t)\right)\right], \qquad t \in \{0,1\}$$

- The average total effect can be written as two decomposition of the sum of direct and indirect effect

$$\tau = \delta(1) + \zeta(0) = \delta(0) + \zeta(1)$$

Three different worlds

- The decomposition $\tau = \delta(1) + \zeta(0)$ contains 3 different worlds:



This difference is $\zeta\,(0)$

This difference is $\delta(1)$

*See Nguyen et al. (2020) for details on which decomposition should be used for specific analysis.

Different treatment effects

- Denoting E[Y(t,M(t'))] as $Y_{tM_{t'}}$, we define the following treatment effects of interest:

(Total) natural indirect effect (**NIE**)	$Y_{1M_1} - Y_{1M_0}$	δ(1)
(Pure) natural direct effect (**NDE**)	$Y_{1M_0} - Y_{0M_0}$	ζ (0)
(Pure) natural indirect effect (**PNIE**)	$Y_{0M_1} - Y_{0M_0}$	δ(0)
(Total) natural direct effect (**TNDE**)	$Y_{1M_1} - Y_{0M_1}$	ζ (1)
Total effect (**TE**)	$Y_{1M_1} - Y_{0M_0}$	τ

Alternative decompositions with mediate

```
. mediate (wellbeing) (bonotonin) (exercise), all
```

wellbeing	Coefficient	Robust std. err.	z	P>|z|	[95% conf.	interval]
POmeans						
Y0M0	57.11317	.2753201	207.44	0.000	56.57355	57.65278
Y1M0	60.00462	.3157888	190.02	0.000	59.38569	60.62356
Y0M1	66.68199	.3258477	204.64	0.000	66.04334	67.32064
Y1M1	69.80444	.2898927	240.79	0.000	69.23626	70.37262
NIE						
exercise						
(Exercise vs Control)	9.799821	.3943251	24.85	0.000	9.026958	10.57268
NDE						
exercise						
(Exercise vs Control)	2.891453	.2304278	12.55	0.000	2.439823	3.343083
PNIE						
exercise						
(Exercise vs Control)	9.568827	.3884522	24.63	0.000	8.807475	10.33018
TNDE						
exercise						
(Exercise vs Control)	3.122447	.2418591	12.91	0.000	2.648412	3.596482
TE						
exercise						
(Exercise vs Control)	12.69127	.4005941	31.68	0.000	11.90612	13.47642

Note: Outcome equation includes treatment-mediator interaction.

Which decomposition?

- Practical question remains: For a specific analysis, which decomposition should be used?
  $\tau = \delta(1) + \zeta(0)$ or $\tau = \delta(0) + \zeta(1)$

- Or should both be used?

- We follow Nguyen et al. (2020) and propose three answers for three cases.

Case 1: Is there a mediated effect? Or, is the causal effect partly mediated by this mediator?

Case 1: Is there a mediated effect? Or, is the causal effect partly mediated by this mediator?

- We propose using $\tau = \delta(1) + \zeta(0)$ decomposition (NIE and NDE)
- **Rational:** Here, we are not questioning the existence of a direct effect.
- We are researching the possibility of a mediated effect to the direct effect.
- If there is no mediated effect, then the total effect $\tau = \zeta(0)$ is the direct effect.

Case 2: In addition to the mediated effect, is there a direct effect?

Case 2: In addition to the mediated effect, is there a direct effect?

- We propose using $\tau = \delta(0) + \zeta(1)$ decomposition (PNIE and TNDE).
- This is a mirror image of the Case 1.
- **Rational:** Here, we are not questioning the existence of a mediator effect.
- We are researching the possibility of treatment affecting the outcome through other mechanisms.
- If there is no direct effect, then the total effect $\tau = \delta(0)$ is the indirect effect.

Case 3: No prior assumption or preferred question about either direct or indirect effect

Case 3: No prior assumption or preferred question about either direct or indirect effect

- We propose reporting both $\tau = \delta(1) + \zeta(0)$ and $\tau = \delta(0) + \zeta(1)$ decompositions.

- **Rational:** If the purpose is to describe all we can learn, there is no reason to prefer wither decomposition over the other.

Causal identification

- After defining different treatment effects, we are interested in causal assumptions that identify those effects

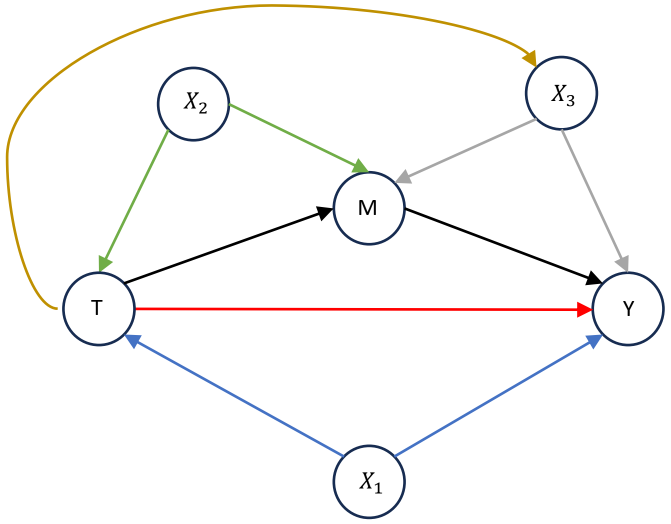
- That is, we are interested in assumptions such that

$$E_M[Y_i(t, M_i(t'))|X_i = x] = \int E[Y_i|M_i = m, T_i = t, X_i = x] df[m|T_i = t', X_i = x]$$

  - LHS is the causal estimand and cannot be estimated from the data
  - RHS is a conditional distribution that can be learned from the data

- This formula is often referred to as the “mediation formula” and is nonparametric.

Causal identification: Assumptions



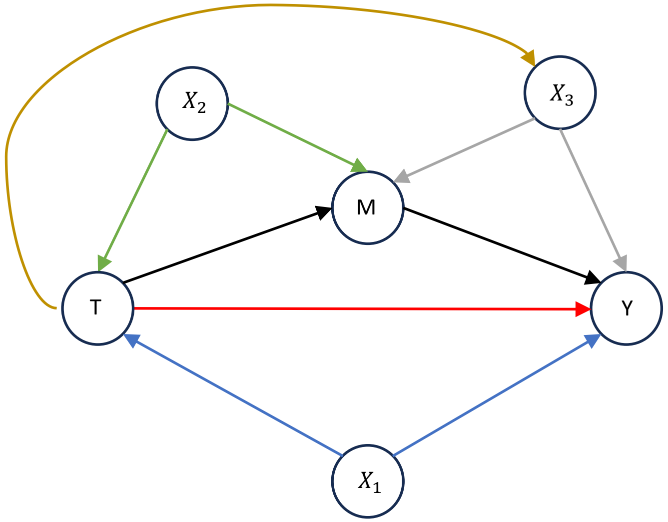
Sequential ignorabilty

1. No unobserved confounding in the treatment-outcome relationship.
2. No unobserved confounding in the mediator-outcome relationship.
3. No unmeasured confounding in the treatment-mediator relationship
4. No (observed) confounders in the mediator-outcome relationship that are caused by the treatment.

- In addition to **sequential ignorability**, we need **SUTVA** and **overlap** assumptions.

Returning to our example: Adding confounders

- Step 1: Hypothetical modeling



- **T** – exercise
- **M** – bonotonin
- **Y** – well-being
- $\boldsymbol{X_1} \cup \boldsymbol{X_3}$ – {age, gender, hstatus, basewell}
- $\boldsymbol{X_3} \cup \boldsymbol{X_2}$– {age, gender, hstatus, basebono}

- Step 2: Causal identification

- Step 3: Estimation in Stata

Estimation in Stata

```
. mediate (wellbeing basewell age gender hstatus)
>         (bonotonin basebono age gender hstatus)
>         (exercise)
```

Causal mediation analysis   Number of obs = 2,000

Outcome model: Linear
Mediator model: Linear
Mediator variable: bonotonin
Treatment type: Binary

wellbeing	Coefficient	Robust std. err.	z	P>|z|	[95% conf.	interval]
NIE						
exercise						
(Exercise vs Control)	10.02204	.2256812	44.41	0.000	9.579717	10.46437
NDE						
exercise						
(Exercise vs Control)	3.085412	.168631	18.30	0.000	2.754901	3.415922
TE						
exercise						
(Exercise vs Control)	13.10746	.2304752	56.87	0.000	12.65573	13.55918

Note: Outcome equation includes treatment-mediator interaction.

Postestimation in Stata

```
. estat proportion

Proportion mediated                                        Number of obs = 2,000
```

wellbeing	Proportion	Robust std. err.	z	P>|z|	[95% conf. interval]	
exercise (Exercise vs Control)	.7646064	.0118613	64.46	0.000	.7413587	.787854

Binary outcome and mediator

```
. mediate (bwellbeing basewell age gender hstatus, logit)
>         (bbonotonin basebono age gender hstatus, logit)
>         (exercise)

Causal mediation analysis                              Number of obs = 2,000

Outcome model:     Logit
Mediator model:    Logit
Mediator variable: bbonotonin
Treatment type:    Binary
```

bwellbeing	Coefficient	Robust std. err.	z	P>|z|	[95% conf.	interval]
NIE						
exercise						
(Exercise vs Control)	.1060631	.0171798	6.17	0.000	.0723914	.1397348
NDE						
exercise						
(Exercise vs Control)	.1521532	.0208609	7.29	0.000	.1112665	.1930399
TE						
exercise						
(Exercise vs Control)	.2582163	.0143273	18.02	0.000	.2301353	.2862973

Note: Outcome equation includes treatment-mediator interaction.

Risk ratios

- If the outcome is binary, and if the outcome model is either logit or probit, we can express the treatment effects as risk ratios or odds ratios.

- The treatment effects on risk-ratio are ratios of potential-outcome means:

$$NIE^{RR} = \frac{Y_{1M_1}}{Y_{1M_0}}$$

$$NDE^{RR} = \frac{Y_{1M_0}}{Y_{0M_0}}$$

$$PNIE^{RR} = \frac{Y_{0M_1}}{Y_{0M_0}}$$

$$TNDE^{RR} = \frac{Y_{1M_1}}{Y_{0M_1}}$$

$$TE^{RR} = \frac{Y_{1M_1}}{Y_{0M_0}}$$

Risk ratios

```
. estat rr
estat rr requires potential-outcome means; refitting model ...

Transformed treatment effects                         Number of obs = 2,000
```

bwellbeing	Risk ratio	Robust std. err.	z	P>|z|	[95% conf.	interval]
NIE						
exercise						
(Exercise vs Control)	1.231901	.0461649	5.57	0.000	1.144663	1.325789
NDE						
exercise						
(Exercise vs Control)	1.49852	.0768679	7.89	0.000	1.355188	1.657013
TE						
exercise						
(Exercise vs Control)	1.84603	.0707466	16.00	0.000	1.712449	1.990031

Odds-ratio

- For logit and probit outcome models, Y_{tM_t}, are probabilities, and so the treatment effects on the odds-ratio scale are

$$NIE^{OR} = \frac{Y_{1M_1}/(1-Y_{1M_1})}{Y_{1M_0}/(1-Y_{1M_0})}$$

$$NDE^{OR} = \frac{Y_{1M_0}(1-Y_{1M_0})}{Y_{0M_0}(1-Y_{0M_0})}$$

$$PNIE^{OR} = \frac{Y_{0M_1}(1-Y_{0M_1})}{Y_{0M_0}(1-Y_{0M_0})}$$

$$TNDE^{OR} = \frac{Y_{1M_1}(1-Y_{1M_1})}{Y_{0M_1}(1-Y_{0M_1})}$$

$$TE^{OR} = \frac{Y_{1M_1}(1-Y_{1M_1})}{Y_{0M_0}(1-Y_{0M_0})}$$

- Similar to risk-ratio, the total effect is the **product** of direct and indirect effect.

Odds-ratio

```
. estat or
estat or requires potential-outcome means; refitting model ...

Transformed treatment effects                          Number of obs = 2,000
```

bwellbeing	Odds ratio	Robust std. err.	z	P>|z|	[95% conf.	interval]
NIE						
exercise (Exercise vs Control)	1.531185	.1060583	6.15	0.000	1.336807	1.753826
NDE						
exercise (Exercise vs Control)	1.918699	.1669374	7.49	0.000	1.617885	2.275444
TE						
exercise (Exercise vs Control)	2.937883	.1843901	17.17	0.000	2.597829	3.322449

Multivalued treatment

```
. mediate (bwellbeing basewell age gender hstatus, logit)
>         (bbonotonin basebono age gender hstatus, logit)
>         (mexercise)
```

bwellbeing	Coefficient	Robust std. err.	z	P>|z|	[95% conf.	interval]
NIE						
mexercise						
(45 minutes vs Control)	.0917518	.0105476	8.70	0.000	.0710789	.1124248
(90 minutes vs Control)	.1082241	.0209235	5.17	0.000	.0672149	.1492334
NDE						
mexercise						
(45 minutes vs Control)	.0245277	.0176917	1.39	0.166	-.0101473	.0592027
(90 minutes vs Control)	.1411745	.0256778	5.50	0.000	.0908468	.1915021
TE						
mexercise						
(45 minutes vs Control)	.1162796	.0175356	6.63	0.000	.0819105	.1506487
(90 minutes vs Control)	.2493986	.0173024	14.41	0.000	.2154864	.2833108

Note: Outcome equation includes treatment-mediator interaction.

Continuous treatment

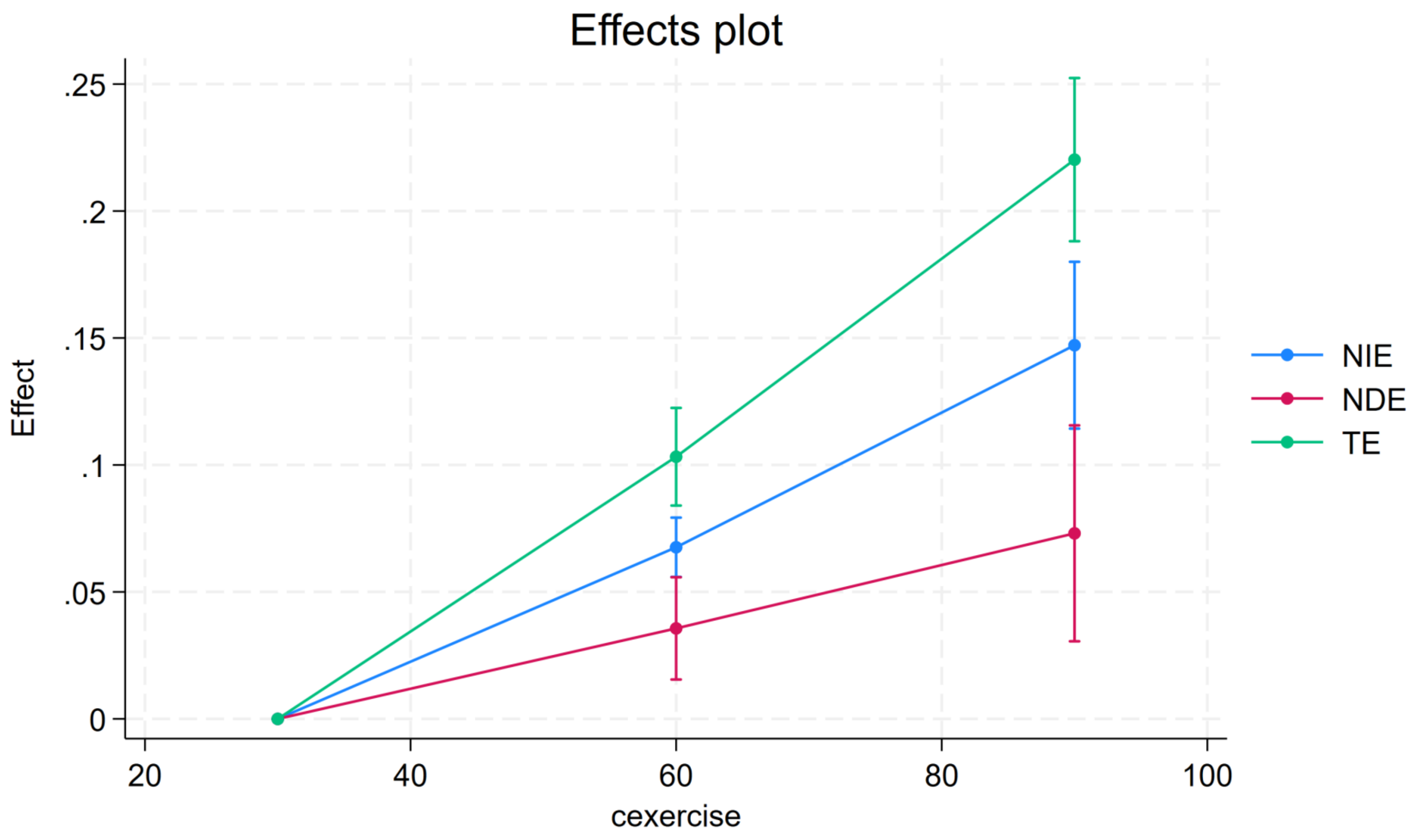
```
. mediate (bwellbeing basewell age gender hstatus, logit)
>         (bbonotonin basebono age gender hstatus, logit)
>         (cexercise, continuous (30 60 90))

Continuous treatment levels:
  0: cexercise = 30 (control)
  1: cexercise = 60
  2: cexercise = 90
```

bwellbeing	Coefficient	Robust std. err.	z	P>|z|	[95% conf.	interval]
NIE						
cexercise						
(1 vs 0)	.0675928	.0059601	11.34	0.000	.0559113	.0792743
(2 vs 0)	.1471622	.0167598	8.78	0.000	.1143136	.1800108
NDE						
cexercise						
(1 vs 0)	.0356447	.0102753	3.47	0.001	.0155056	.0557839
(2 vs 0)	.073076	.0216909	3.37	0.001	.0305626	.1155894
TE						
cexercise						
(1 vs 0)	.1032376	.0098062	10.53	0.000	.0840177	.1224574
(2 vs 0)	.2202382	.0163999	13.43	0.000	.1880949	.2523815

Note: Outcome equation includes treatment-mediator interaction.

Continuous treatment

```
. estat effectplot
```



Count mediator

- We consider the sample that includes women who gave birth to a child.

- We wish to find out whether the socioeconomic status and education of the mother affect the child’s health.

  - **Y -** the birthweight of the baby (**bweight**)
  - **T** - whether or not the mother has a college degree (**college**).
  - **M** -the number of cigarettes smoked per day during pregnancy (**ncigs**).
  - **X** – social economic status of parents (**sespar**)

- The hypothesis is that women with a higher educational degree are likely to smoke fewer cigarettes and that smoking during pregnancy has negative effects on birthweight.

Count mediator

```
. webuse birthweight
(Fictional birthweight data)

. list in 1/5, clean

       id   bweight   lbweight   ncigs   college        ses     sespar   age
  1.    1      3621         No       1        No     5.3581   3.308523    29
  2.    2      3278         No       0       Yes   9.556957   4.376035    38
  3.    3      3073         No       1        No   3.980829   6.580275    39
  4.    4      3306         No       0       Yes   11.17643   12.12075    30
  5.    5      4517         No       0       Yes   9.026146   4.738766    28
```

Count mediator

```
. mediate (bweight sespar c.age##c.age, expmean)
>         (ncigs sespar c.age##c.age, poisson)
>         (college), nointeract

Causal mediation analysis                           Number of obs = 2,000

Outcome model:     Exponential mean
Mediator model:    Poisson
Mediator variable: ncigs
Treatment type:    Binary
```

bweight	Coefficient	Robust std. err.	z	P>|z|	[95% conf.	interval]
NIE						
college						
(Yes vs No)	198.978	23.53279	8.46	0.000	152.8546	245.1014
NDE						
college						
(Yes vs No)	320.3318	34.47792	9.29	0.000	252.7563	387.9072
TE						
college						
(Yes vs No)	519.3098	28.70435	18.09	0.000	463.0503	575.5693

Note: Outcome equation does not include treatment-mediator interaction.

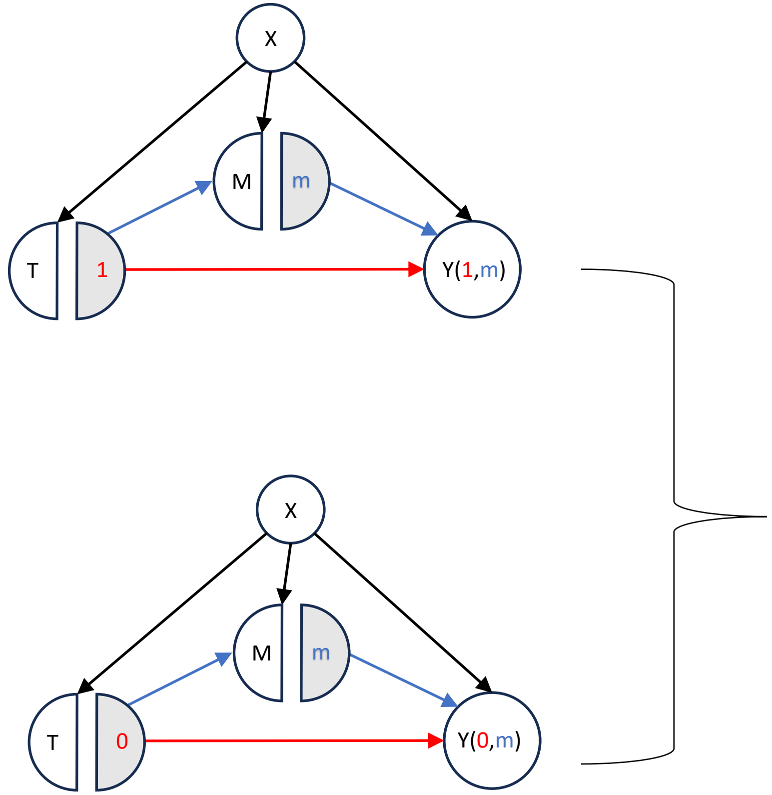
Incidence-rate-ratio

- The same formula as risk-ratio.
- Allowed when the outcome model is Poisson/exponential mean.

```
. estat irr
estat irr requires potential-outcome means; refitting model ...

Transformed treatment effects                         Number of obs = 2,000
```

bweight	IRR	Robust std. err.	z	P>|z|	[95% conf.	interval]
NIE						
college (Yes vs No)	1.057819	.0072037	8.25	0.000	1.043794	1.072033
NDE						
college (Yes vs No)	1.102636	.0113921	9.46	0.000	1.080533	1.125192
TE						
college (Yes vs No)	1.16639	.009948	18.05	0.000	1.147055	1.186052

Controlled direct effects (CDE)



- CDE is the effect of the treatment if the mediator were controlled, i.e., set to **a specific level (M =m)** for everyone.
- CDE may vary across individuals, and within an individual may vary depending on the mediator control value **m**.
- For binary treatment

$$CDE(m) = Y_{1m} - Y_{10}$$

$$CDE(m)^{RR} = \frac{Y_{1m}}{Y_{0m}}$$

$$CDE(m)^{IRR} = \frac{Y_{1m}}{Y_{0m}}$$

$$CDE(m)^{OR} = \frac{Y_{1m}/(1 - Y_{1m})}{Y_{0m}/(1 - Y_{0m})}$$

Controlled direct effects (CDE)

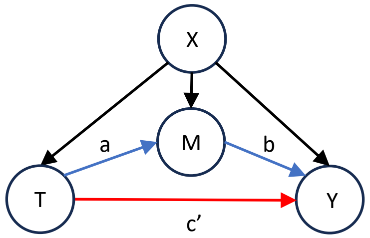
```
. estat cde, mvalue(0 1)

Controlled direct effect                                Number of obs = 2,000

Mediator variable: ncigs
Mediator values:
  1._at: ncigs = 0
  2._at: ncigs = 1
```

	CDE	Delta-method std. err.	z	P>|z|	[95% conf.	interval]
college@_at						
(Yes vs No) 1	341.955	35.26807	9.70	0.000	272.8308	411.0791
(Yes vs No) 2	332.6419	34.94916	9.52	0.000	264.1428	401.141

Controlled direct effects (CDE)

```
. estat cde, mvalue(0 1) contrast

Controlled direct effect                                Number of obs = 2,000

Mediator variable: ncigs
Mediator values:
  1._at: ncigs = 0
  2._at: ncigs = 1

------------------------------------------------------------------------------------
                      |             Delta-method
                      |        CDE   std. err.      z    P>|z|     [95% conf. interval]
----------------------+-------------------------------------------------------------
         _at#college  |
(2 vs 1) (Yes vs No)  |  -9.313066   .9748033    -9.55   0.000    -11.22365   -7.402487
------------------------------------------------------------------------------------
```

Final remarks: Traditional vs Causal Mediation Analysis



- Traditional approach uses a model-based definition

$$Y_i = \alpha_1 + c'T_i + bM_i + \varepsilon_{Y_i}$$
$$M_i = \alpha_2 + aT_i + \varepsilon_{M_i}$$

- Here, the indirect effect := ab and direct effect := c'
- Key differences between traditional and causal mediation analysis:

Traditional	Causal
Indirect and direct effects are mathematical objects that do not exist without the model	Effects are defined in a model-free manner, based on reasoning about what fits the notion of causal effect
No separation of the definition of an effect and its estimation method	Separates the definition of an effect, and its identification from estimation

Learn more:

- https://www.stata.com/manuals/causalmediate.pdf

References

1. Neal, B. (2020). Introduction to causal inference.

2. Nguyen, T., Schmid, I., and Stuart, E. (2020). Clarifying causal mediation analysis for the applied researcher: Defining effects based on what we want to learn. Psychological Methods, 26.

3. Nguyen, Trang Quynh, Schmid, Ian, Ogburn, Elizabeth L. and Stuart, Elizabeth A.. "Clarifying causal mediation analysis: Effect identification via three assumptions and five potential outcomes" *Journal of Causal Inference*, vol. 10, no. 1, 2022, pp. 246-279.

4. VanderWeele, T. and Shpitser, I. (2013). On the definition of a confounder. The Annals of Statistics, 41.

STATA® 18