Download Paper

The Identification and Estimation of Direct and Indirect Effects in A/B Tests through Causal Mediation Analysis

Xuan Yin [1] Liangjie Hong [2]

[1]xyin@etsy.com [2]lhong@etsy.com

What is the research about?

We disentangle the causal relationships between online products through the identification and estimation of direct and indirect effects in A/B tests where multiple unmeasured causally-dependent mediators invalidate Causal Mediation Analysis (**CMA**) in the literature. We propose new direct and indirect effects, the generalization of the effects in the literature, and prove their identification under some assumptions and show how to estimate and test them.

Why is it relevant?

What do we observe?

Induced Change: A change in one product would induce users to change their behaviors in others. Here are two examples:

Table 1: Recommendation Module A/B Test

Number of clicks on recommendation modules	Signficant increase
Number of clicks on organic search results	Significant decrease
Sitewide conversion/Gross merchandise value (GMV)	Insignificant change

Table 2: Promoted Listing A/B Test

Promoted listing	
click-through-rate	Significant increase
number of clicks	Significant increase
advertising revenue	Significant increase
Number of clicks on organic search results	Significant decrease
Sitewide conversion/Gross merchandise value (GMV)	Insignificant change

How does the observation challenge decision making?

- Different products may compete for users attention. How to quantify the competition?
- Although the decision whether to keep the new feature is typically based on sitewide conversion/GMV in A/B tests, it is difficult to make a decision when user engagement metrics of the tested product and conversion/GMV move differently.
- To optimize users' overall experiences, the *induced changes* of user engagement in other products cannot be ignored.

The challenges ask for more information to assist decision making.

- What is the effect of the *induced change* of user engagement in other products on conversion/GMV (indirect effect)?
- Leaving aside the induced change, what is the *direct effect* of the tested product change on conversion/GMV?
- Is the observed change of conversion/GMV the *synthesis* of the two effects?

A/B tests cannot answer above questions.

- It only shows how the change in recommendation modules or promoted listings impacts organic search and sitewide conversion/GMV (i.e., the identification of *Average Treatment Effect* (**ATE**)).

Introduction to causal inference

We conduct **CMA** within Potential Outcome Framework.
In an A/B test, a user i is randomly assigned to either treatment group ($T_i = 1$) or control group ($T_i = 0$). The treatment T impacts the outcome Y (partially) through the mediator M. Let $M_i(t)$ denote her potential mediator under treatment t. Let $Y_i(t, m)$ denote her potential outcome under the treatment t and the mediator m. Only one of potential mediators and only one of potential outcomes can be observed for each user. In Table 1, the recommendation module is T, the number of clicks on search results is M, and the conversion is Y. Examples:

- $M_i(1)$ ($M_i(0)$) is her numbers of clicks on search results if she was presented with the new (old) recommendation module.
- $Y_i(1, M_i(0))$ is her conversion status if she was presented with the new recommendation module and clicked on search results as if she had been presented with the old one.

The Fundamental Research Issue in Causal Inference Is Identification.

Causal effects are usually defined as the differences between potential outcomes. In order to estimate causal effects from population data, what kind of assumptions are required? (*Assumptions* ⇒ Causal Effects)

Rubin Causal Model

Strong Ignorability and *SUTVA* ⇒ ATE

- *Strong Ignorability*: $\{Y_i(0), Y_i(1)\} \perp\!\!\!\perp T_i$ and $0 < \mathbb{P}(T_i = t) < 1$
  Guaranteed by randomization in A/B tests.
- *SUTVA*: *Stable Unit-Treatment-Value Assumption*
  No multiple versions of treatment and no interference between users.
- ATE $:= \mathbb{E}(Y_i(1, M_i(1))) - \mathbb{E}(Y_i(0, M_i(0)))$ or $\mathbb{E}(M_i(1)) - \mathbb{E}(M_i(0))$

Introduction to CMA

CMA splits **ATE** into **ACME** (indirect effect) and **ADE** (direct effect).
Sequential Ignorability (SI) and *SUTVA* ⇒ ACME and ADE

- *SI*: add two extra conditions to *Strong Ignorability*, $Y_i(t', m) \perp\!\!\!\perp M_i(t) | T_i = t$ and $0 < \mathbb{P}(M_i(t) = m | T_i = t) < 1$.
  They are unverifiable. They mean, conditional on the treatment, each potential mediator behaves like the treatment and is ignorable to any potential outcomes.
- ACME$(t) := \mathbb{E}(Y_i(t, M_i(1))) - \mathbb{E}(Y_i(t, M_i(0)))$: *Average Causal Mediation Effect*. Because treatment is fixed at t, the difference between the two potential outcomes can only be attributed to the two different potential mediators, which are *induced* by different treatments.
  ACME(1) is the average effect of the *induced change* in organic search clicks upon conversion given users were presented with the new recommendation module all the time.
- ADE$(t) := \mathbb{E}(Y_i(1, M_i(t))) - \mathbb{E}(Y_i(0, M_i(t)))$: *Average Direct Effect*. Because mediator is fixed at $M(t)$, the difference between the two potential outcomes can only be attributed to the two different treatments. ADE(0) is the direct effect of the recommendation module change on conversion leaving aside the *induced change*.

Why can't we use CMA directly on A/B tests?

Multiple unmeasured causally-dependent mediators in A/B tests break ***SI*** and invalidates the identification.

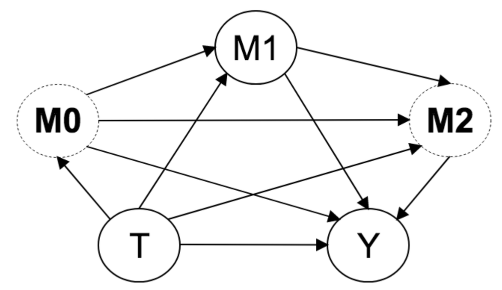

Figure 1: Multiple Unmeasured Causally-Dependent Mediators

Suppose M_1 is organic search clicks. Let $\boldsymbol{M}_0$ denote web pages/modules that inspire users to search (upstream mediators), e.g., "recently viewed." Let $\boldsymbol{M}_2$ denote web pages/modules that could play in the relationship between search and purchase (downstream mediators). In the real analysis, it is difficult and costly to know and to measure exactly upstream or downstream mediators of search because website has hundreds of web pages/modules.

What do we do?

We take care of multiple unmeasured causally-dependent mediators.
We prove that *Generalized SI* and *LSEM* ⇒ GADE and GACME

Generalized SI and *LSEM*

- *Generalized SI*: Each potential mediator, conditional on the treatment and its upstream mediators, behave like the treatment and is ignorable to all the potential outcomes and all the potential downstream mediators.
- *LSEM*: *Linear Structural Equation Model*. Potential mediators, potential outcomes, and treatment have linear relationships.

GACME and GADE

- GACME: *Generalized Average Causal Mediation Effect*. It captures the causal effect of the treatment T_i that goes through all the channels that have M_{i1}: $T \to M_1 \to Y$, $T \to \boldsymbol{M}_0 \to M_1 \to Y$, $T \to M_1 \to \boldsymbol{M}_2 \to Y$, and $T \to \boldsymbol{M}_0 \to M_1 \to \boldsymbol{M}_2 \to Y$.
- GADE: *Generalized Average Direct Effect*. It captures the causal effect of the treatment T_i that goes through all the channels that do not have M_{i1}: $T \to Y$, $T \to \boldsymbol{M}_0 \to Y$, $T \to \boldsymbol{M}_0 \to \boldsymbol{M}_2 \to Y$, $T \to \boldsymbol{M}_2 \to Y$

We estimate **GACME** and **GADE** by General Method of Moments.
We estimate the asymptotic variances of estimators by Delta method.
We test H_0: **GADE** $= 0$ and H_0: **GACME** $= 0$ based on asymptotic normality.

How is its application on A/B tests?

- **Recommendation Module A/B Tests**: The effect of induced reduction of user engagement of search on conversion is significantly negative. The direct effect on conversion is significantly positive. The indirect effect partially offsets the direct effect, which makes the total effect insignificant.
- **Promoted Listing A/B Tests**: The effect of induced reduction of user engagement of search on conversion is significantly negative. The direct effect on conversion is insignificant.