announcements

- Problem Set 7 is due right about now
- Problem Set 8 will be assigned later and due on Monday
- Drill will be happening tomorrow as usual

1

my inbox is full of bad news

Interacting With Curves: How to Validly Test and Probe Interactions in the Real (Nonlinear) World

Uri Simonsohn View all authors and affiliations

All Articles | https://doi.org/10.1177/25152459231207787

Contents | PDF / ePub | More

Abstract

Hypotheses involving interactions in which one variable modifies the association between another two are very common. They are typically tested relying on models that assume effects are linear, for example, with a regression like $y = \mathbf{a} + \mathbf{b}x + \mathbf{c}z + \mathbf{d}x \times z$. In the real world, however, few effects are linear, invalidating inferences about interactions. For instance, in realistic situations, the false-positive rate can be 100% for detecting an interaction, and a probed interaction can reliably produce estimated effects of the wrong sign. In this article, I propose a revised toolbox for studying interactions in a curvilinear-robust manner, giving correct answers "even" when effects are not linear. It is applicable to most study designs and produces results that are analogous to those of current—often invalid—practices. The presentation combines statistical intuition, demonstrations with published results, and simulations.

2

mediation analysis

March 27, 2024

3

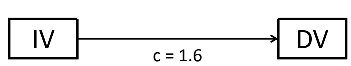
an example

- adolescents diagnosed with bipolar disorder are randomly assigned to a treatment group (a family counseling intervention + the usual pharmaceutical regimen) or a control group (only the pharmaceutical)
- the outcome is a measure of symptoms taken at 8 weeks after treatment begins
- we suspect that the counseling will be effective by reducing criticism; this is measured at 7 weeks

4

model 1: symptoms ~ treatment

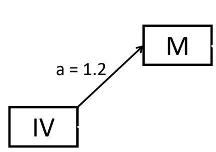
```
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   5.1000     0.2739  18.623 3.29e-13
X             1.6000     0.5477   2.921  0.00912
```



5

model 2: criticism ~ treatment

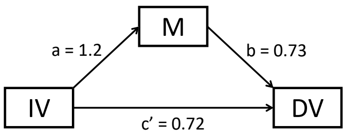
```
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   3.1000     0.1871  16.570 2.41e-12
X             1.2000     0.3742   3.207  0.00489
```



6

model 3: symptoms ~ tx + crit

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	2.8365	0.9847	2.881	0.0104
x	0.7238	0.6123	1.182	0.2535
c	0.7302	0.3077	2.373	0.0297



$a*b = 1.2*0.73 = 0.876 = c - c' = 1.6 - 0.724 = 0.876$ ✓

7

alternative approach: Sobel test

- test the *ab* path's significance using the Sobel test

$$z = \frac{ab}{\sqrt{a^2 SE_a^2 + b^2 SE_b^2}}$$

- works best with large samples
- doesn't require *a* and *b* to both be significant
- alternative denominator formulas exist
- note that this is also a test of $c - c'$

8

alternative approach: bootstrapping

- the sampling distribution of *ab* tends to be non-normal
- the original data is sampled (with replacement) at random
- this provides estimates of *ab* assuming H_0 (no mediation) to be true
- do this many times (1000s, at least) to generate an empirical sampling distribution, allowing the generation of a CI

9

please visit
quantpsy.org/medn.htm

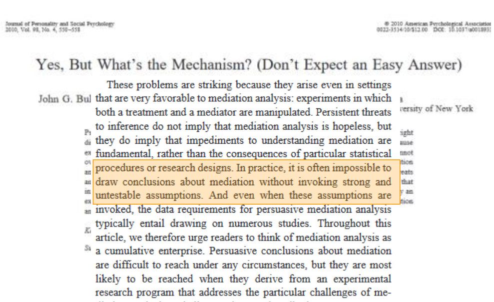
10

thinking about mediation

complexities of third-variable control & other considerations

11

drawing conclusions about mediation is hard


Journal of Personality and Social Psychology
2010, Vol. 98, No. 4, 550–558

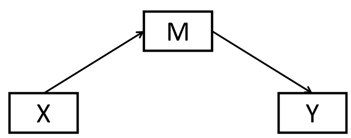
Yes, But What's the Mechanism? (Don't Expect an Easy Answer)

These problems are striking because they arise even in settings that are very favorable to mediation analysis: experiments in which both a treatment and a mediator are manipulated. Persistent threats to inference do not imply that mediation analysis is hopeless, but they do imply that impediments to understanding mediation are fundamental, rather than the consequences of particular statistical procedures or research designs. In practice, it is often impossible to draw conclusions about mediation without invoking strong and untestable assumptions. And even when these assumptions are invoked, the data requirements for persuasive mediation analysis typically entail drawing on numerous studies. Throughout this article, we therefore urge readers to think of mediation analysis as a cumulative enterprise. Persuasive conclusions about mediation are difficult to reach under any circumstances, but they are most likely to be reached when they derive from an experimental research program that addresses the particular challenges of mediation analysis—challenges that we describe here.

12

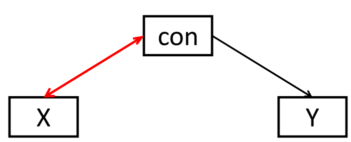
mediation: what we think is happening

- if one variable influences another through an intervening variable, the intervening variable is typically called a *mediator*



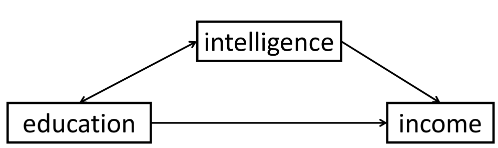
13

confounds: equivalent mathematically



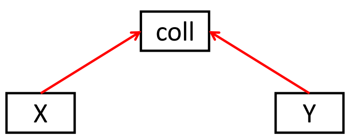
14

an example of a (probable?) confound



15

colliders:
similar mathematically



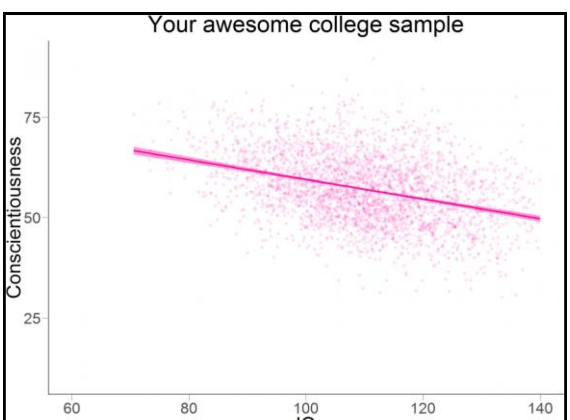
16

an example of collider bias

- imagine you are interested in the relationship between intelligence (indexed by IQ) and conscientiousness
- you find a large sample of college students and find … *r* = -.37
- what?!

http://www.the100.ci/2017/03/14/that-one-weird-third-variable-problem-nobody-ever-mentions-conditioning-on-a-collider/

17

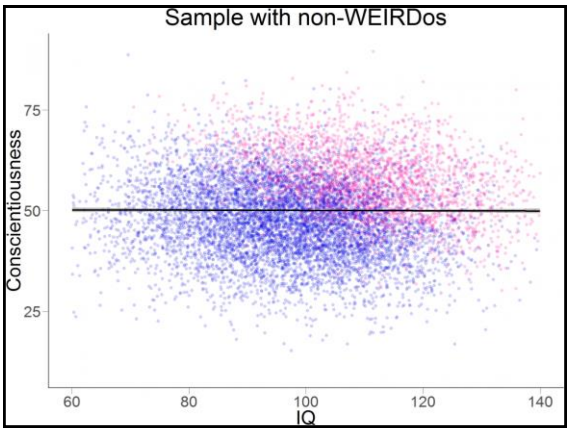


18

what's going on?

- college students tend to be higher IQ that the general population; they also tend to be higher with respect to conscientiousness
- that is, both of these variables are predictors of college-student membership; they "collide"
- so selecting from the college-student population "conditions on a collider", creating a (strange) relationship that doesn't exist in the whole population

19



20

third-variable patterns (problems?), cataloged

- confounds
- colliders
- suppressors
- mediators
- covariates

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2819361/
https://www.ncbi.nlm.nih.gov/pubmed/28575894

21