Testing Mediation Effects using Logic of Boolean Matrices (JASA)

Chengchun Shi [1] and Lexin Li [2]

[1]London School of Economics and Political Science

[2]University of California, Berkeley

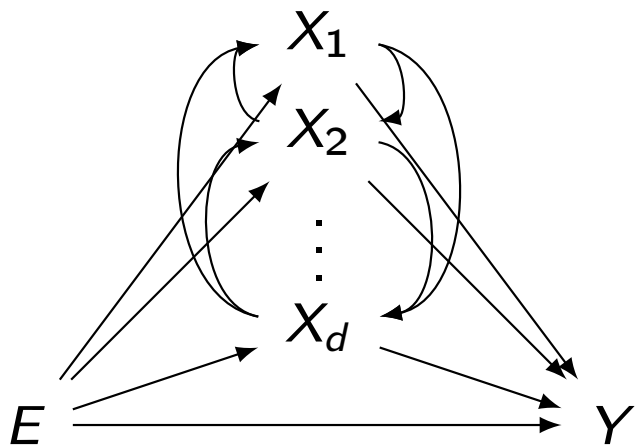
- Alzheimer's disease (AD) and normal aging:
  - AD is an irreversible neurodegenerative disorder, characterized by progressive impairment of cognitive and memory functions, then loss of independent living, and ultimately death
  - the leading form of dementia, and currently affecting 5.8 million American adults aged 65 years or older
  - prevalence continues to grow; projected to reach 13.8 million by 2050
  - there is no effective treatment
- **scientific questions of interest**:
  - neurodegeneration measure, often captured as grey matter cortical atrophy, is a well-known biomarker associated with AD
  - **question**: how age affects cortical thickness then cognitive outcome

- **high dimensional mediation analysis**:
  - to identify and explain the mechanism, or pathway, that underlies an observed relationship between an **exposure** and an **outcome** variable, through the inclusion of an intermediary variable, known as a **mediator**
  - question: infer the **significance** of individual mediators?

    $H_0(q)$: There is no path from E to Y that passes through X_q,

    $H_1(q)$: There is a path from E to Y that passes through X_q.

  - challenge: the total number of potential paths that go through any mediator is **super-exponential** in the number of mediators



- **what we propose** (in a nutshell):
  - propose a new testing procedure to evaluate the individual mediation effect, while allowing directed paths among the mediators
  - construct the test statistic using the **logic of Boolean matrices** $\rightarrow$ establish the proper limiting distribution under the null $\rightarrow$ the asymptotics of the test statistic built on regular matrix operations are difficult to establish
  - can be naturally coupled with a **screening** procedure $\rightarrow$ help scale down the number of potential paths to a moderate level $\rightarrow$ reduce the variance of the test statistic $\rightarrow$ enhance the power of the test
  - use a **data splitting** strategy $\rightarrow$ control type-I error
  - devise a **decorrelated estimator** to reduce potential bias induced by high-dimensional mediators
  - employ **multiplier bootstrap** to obtain the critical values
  - couple with a **multiple testing** procedure for FDR control
  - establish the **asymptotic size, power, and FDR control**, while allowing the number of mediators to **diverge** to ∞

- mediation inference:
  - **exposure**: age; **outcome**: PACC score; **mediators**: gray matter cortical thickness of $d = 68$ brain regions-of-interest (ROIs)
  - $n = 389$ subjects
  - set FDR level at 10%
- findings:

amyloid negative group	
l-entorhinal	l-precuneus
l-superiortemporal	r-inferiorparietal
r-superiorfrontal	r-superiortemporal

  - entorhinal cortex functions as a hub in a widespread network for memory, navigation and the perception of time; one of the most heavily damaged cortices in AD
  - precuneus is involved with episodic memory, visuospatial processing, reflections upon self, and aspects of consciousness, and is found to be an AD-signature region

Thank You!

Preprint https://arxiv.org/pdf/2006.02615.pdf,

Python code LOGAN https://github.com/callmespring/LOGAN