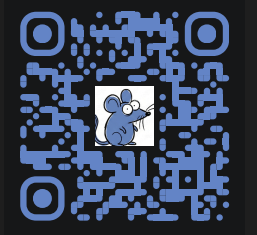


Cargo Cult Causal Inference

Understanding Derivation and Proof is Part of the Whole

(and the Whole $>$ \sum Parts)



Ashley I Naimi, PhD | Dept of Epidemiology
Professor | Emory University

✉ ashley.naimi@emory.edu

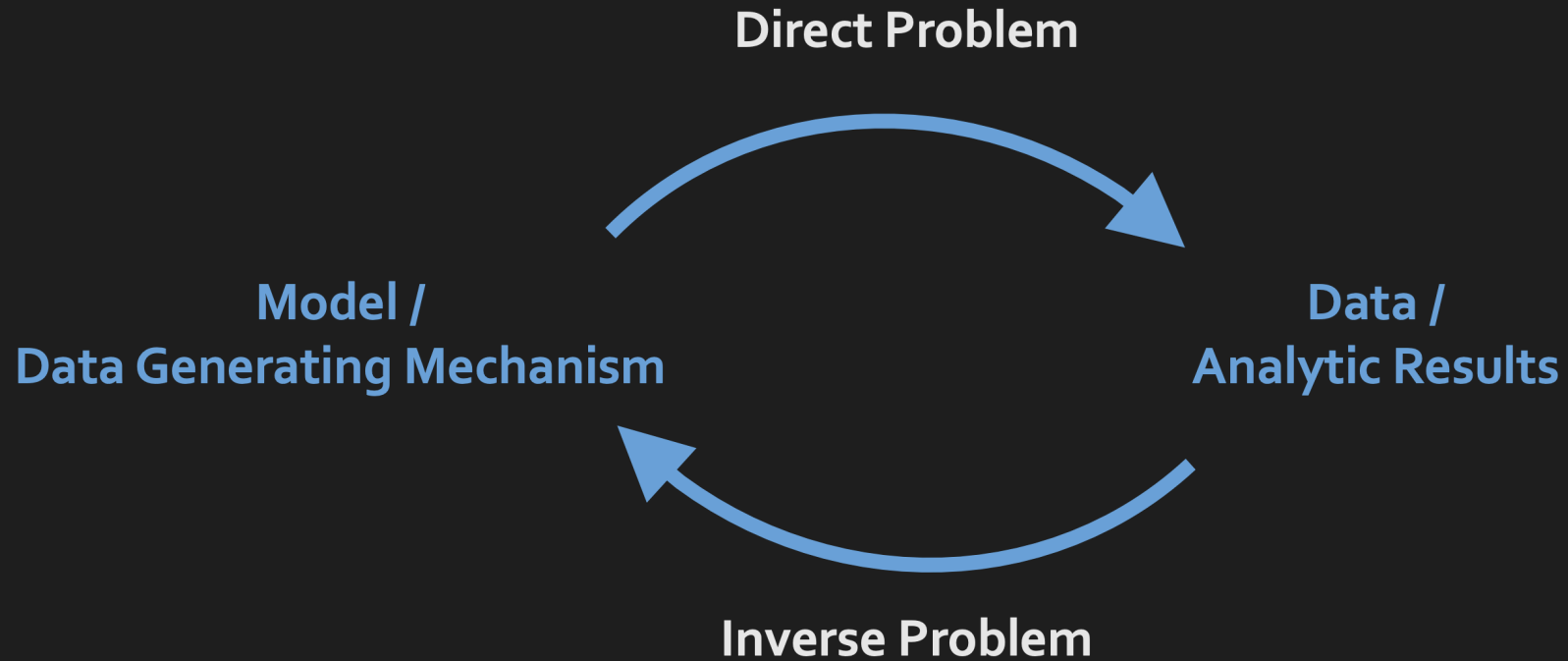
🌐 <https://ainaimi.github.io/>



EMORY

ROLLINS
SCHOOL OF
PUBLIC
HEALTH

Direct and Inverse Problems



The Direct Problem: Proof & Derivation

A container of 500,000 red and 500,000 blue marbles, uniformly mixed.

If you draw 4, what is $P(\text{exactly 2 red})$?

$$P(Y = 2 \mid \theta = 0.5, N = 4) = \binom{4}{2} 0.5^2 (1 - 0.5)^2 = 0.375$$

This is a **direct problem**

A **deductive statement**: if the premises are true, the conclusion is guaranteed.

The Inverse Problem: Inference

Same container with **unknown** number of red and blue marbles, uniformly mixed.

You draw 4 and observe 2 red.

What proportion of marbles in the container are red?

This is an **inverse problem**

An **inductive problem**: no guarantees (problem of induction).

Prone to *affirming the consequent*. Many containers could yield this sample.

Identifying the ATE Is a Direct Problem

We want the unmeasurable average potential outcome $E(Y^a)$.

Can observed data recover it?

Yes:

$$E(Y^a) = E\left(E(Y \mid A = a, W) \right)$$

If we can assume:

- **Causal Consistency:** if $A = a$, then $Y = Y^a$
- **Conditional Exchangeability:** $E(Y^a \mid A, W) = E(Y^a \mid W)$
- **Positivity:** $\epsilon \leq P(A = 1 \mid W) \leq 1 - \epsilon$

Estimating the ATE with Data: Inverse Problem

In applied causal inference, distinctions between statements of *proof* (or derivation) versus statements of *empirical investigation* (the analyses we actually do) are often blurred:

Mimicking statements about the conditions needed to solve the direct problem (i.e., the causal inference assumptions) when, in fact, we are solving an inverse problem.

Richard Feynman called this mimicry a form of **Cargo Cult Science**.

Cargo Cult Science

... they've arranged to make things like runways, to put fires along the sides of the runways, to make a wooden hut for a man to sit in, with two wooden pieces on his head like headphones and bars of bamboo sticking out like antennas—he's the controller—and they wait for the airplanes to land. They're doing everything right. The form is perfect. It looks exactly the way it looked before. But it doesn't work. No airplanes land. So I call these things Cargo Cult Science, because they follow all the apparent precepts and forms of scientific investigation, but they're missing something essential, because the planes don't land.

Cargo Cult Causal Inference

Target trial emulation in nutritional epidemiology: Guo et al (2025) *AJE* state that they "assumed randomization" after adjustment for baseline covariates, to establish identification assumptions.

But food-frequency questionnaires quietly threaten every assumption needed to identify the ATE:

- FFQs aggregate heterogeneous non-standardized foods. **causal consistency.**
- With FFQs, there is no clear assignment event: **immortal time bias.**
- FFQs cannot disentangle dietary components: **conditional exchangeability**

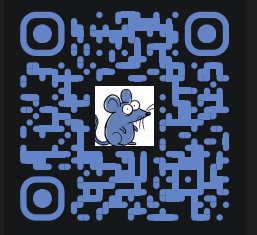
The Solution?: Questions, Not Incantations

Interrogate, don't Assume

Cargo Cult Causal Inference

Understanding Derivation and Proof is Part of the Whole

(and the Whole $>$ \sum Parts)



Ashley I Naimi, PhD | Dept of Epidemiology
Professor | Emory University

✉ ashley.naimi@emory.edu

🌐 <https://ainaimi.github.io/>



EMORY

ROLLINS
SCHOOL OF
PUBLIC
HEALTH