fundamental issues:

UNDERSTANDING DIFFERENCES BETWEEN OBSERVED AND UNOBSERVED CONFOUNDING

why do statistics?

- If you strip away all the artifice we're really just trying to convince one another to believe something.
- In stats, we put together rigorous arguments using data.
- The reason statistics is a discipline is that data tend to behave in ways that are repeatable.
- If you have *certain kinds* of data they behave in ways that have been well-studied and we have very precise ways of understanding what's happening (a.k.a. "backing into" what's going on, or "inferring" what the underlying structure may be); we have developed very powerful mathematical and computational machinery to help you.

understanding statistics

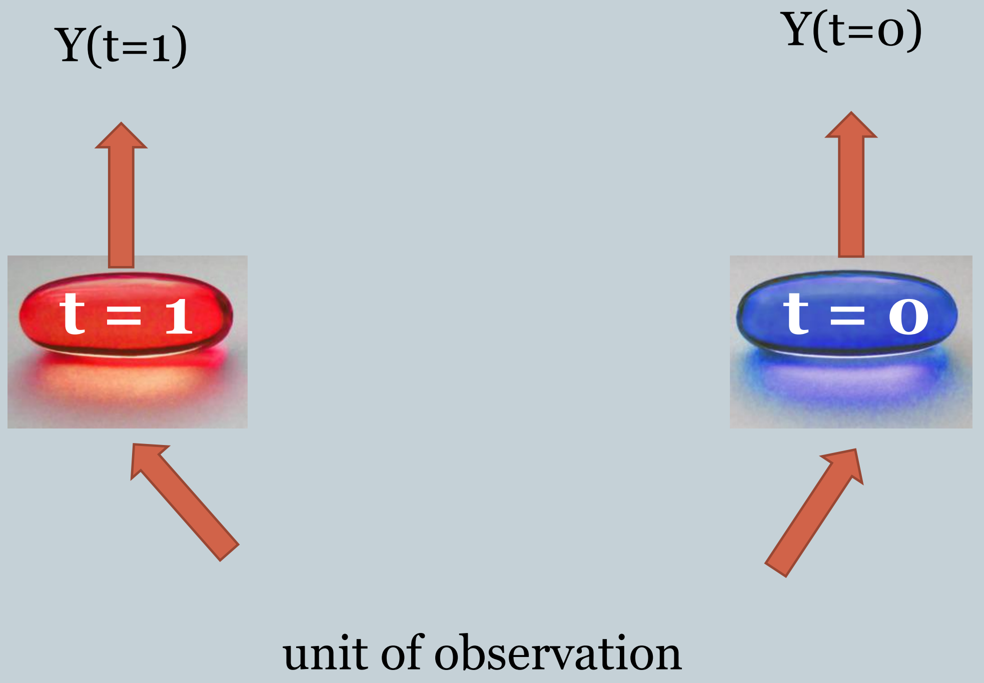
- Even these two days at CIMPOD are woefully inadequate.
- We focus on empirical measurements:
  (i) how to generate measurements and
  (ii) how to combine measurements in the most meaningful ways.

causal inference

- Our goal is to figure out what will happen what the change in the outcome will be for a person if we change from the control to the treatment:

$$Y_i(t = 1) - Y_i(t = 0) = \delta_i$$

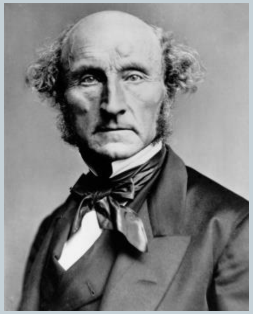
causal inference



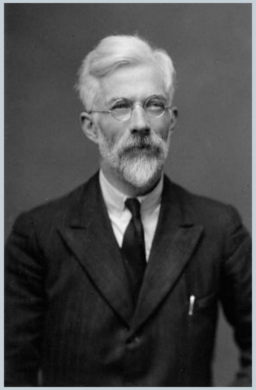
<u>Fundamental problem of causality</u>:
We cannot observe both $Y_i(t = 1)$ and $Y_i(t = 0)$ at the same time.

two approaches

- John Stuart Mill
  - Philosopher, economist, early feminist and civil servant.
  - Estimate effect through "method of differences."



- Sir Ronald Fisher
  - Statistician and biologist.
  - Estimate effect through "a controlled & random process."



[Heterogeneity and Causality (Rosenbaum 2005)]

terminology

- **Unit of observation** – the element in the study for which the intervention can be applied to or withheld from.
  - In our working example: people
  - You can imagine that we could talk about different levels of aggregation being "units of observations": doctors who treat patients, clinics, health systems, etc.

terminology

- **Covariate** – a variable, distinct from the intervention and outcome, that can change from unit of observation to unit of observation
  - In our working example: baseline weight, gender, BMI, age, hair color, favorite color…
  - Not all covariates are equally “important.” We’ll revisit this notion when we discuss the concept of *confounding*.

method of difference

- In 1864, in his *System of Logic: Principles of Evidence and Methods of Scientific Investigation*, Mill proposed four methods of experimental inquiry, including the "method of difference:"

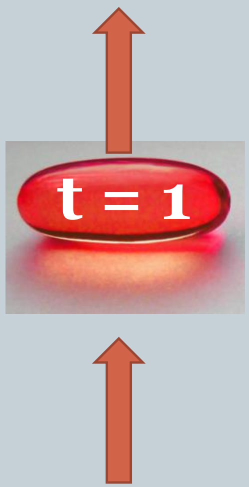
  If an instance in which the phenomenon … occurs and an instance in which it does not … have every circumstance save one in common … [then] the circumstance [in] which alone the two instances differ is the … cause or a necessary part of the cause (III, sec. 8)

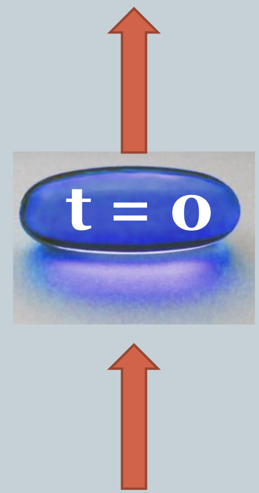
- For Mill, homogeneity and sound causal inference were closely linked: he wanted "two instances … exactly similar in all circumstances except the one" under study.

causal inference

$Y(t=1) = f(t=1, X=x)$

$Y(t=0) = f(t=0, X=x')$





x

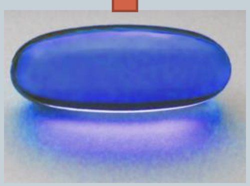
x'

The only difference

$Y(t{=}1) = f(t = 1, X = \boldsymbol{x})$

$Y(t{=}0) = f(t = 0, X = \boldsymbol{x})$

$\boldsymbol{x} =$ 

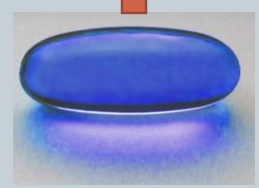


$\boldsymbol{x} =$ 

$\mathrm{Y(t{=}1)} = f(t = 1, X = x)$

$x =$

$\mathrm{Y(t{=}0)} = f(t = 0, X = x')$

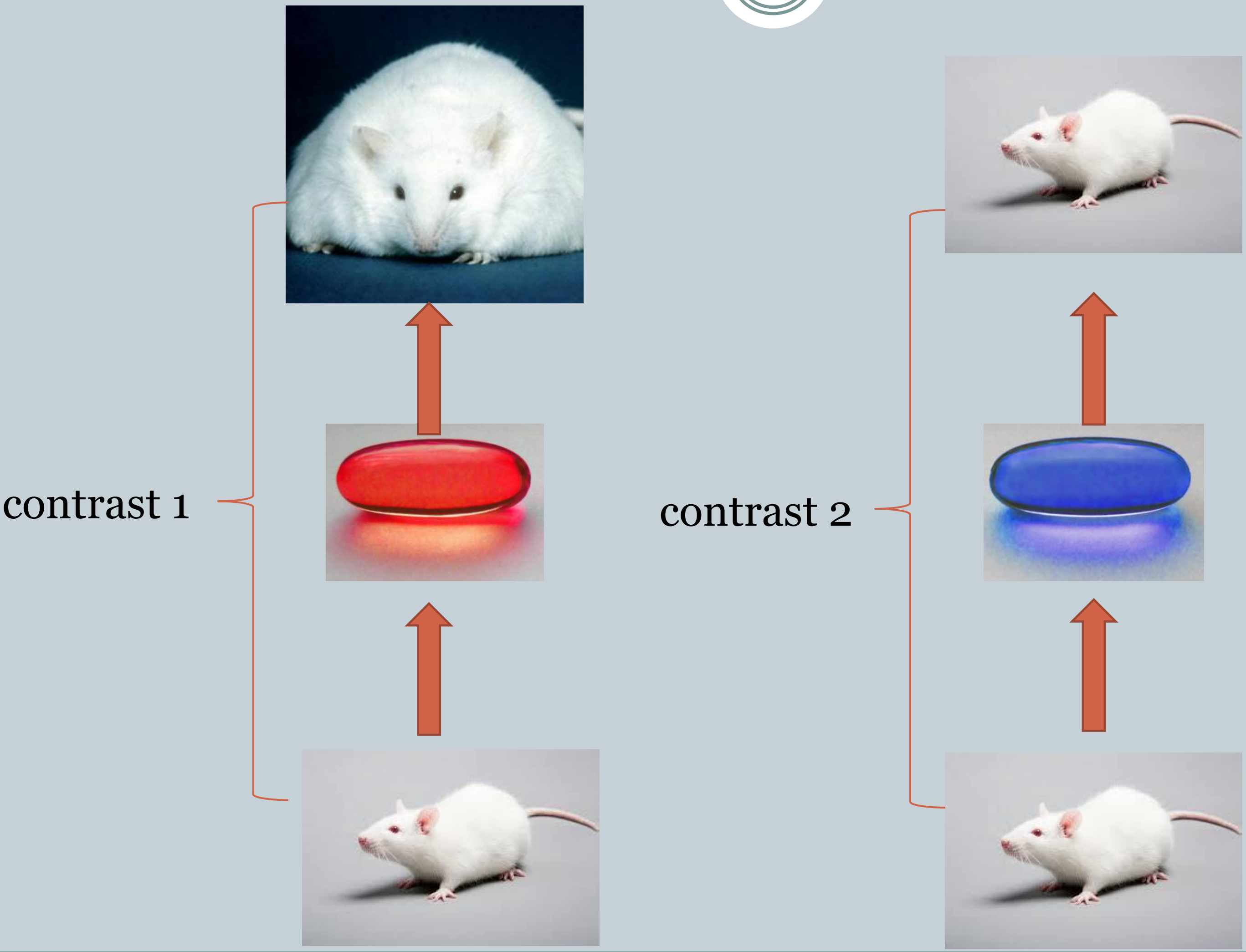


$x' =$

terminology

- **Confounding** – when something (usually a covariate or a set of covariates) makes your estimate of the causal effect biased (*loosely speaking*: makes your estimate – on average – report a number different than the number it should be).
  - In our working example: baseline weight, gender, BMI, age, ~~hair color~~, ~~favorite color~~...
  - Confounding is usually thought of as covariates that cause variation in the outcome as well as the treatment.
  - The way I like thinking about it: If the treatment group would have been different than the control group, even if we never applied any form of intervention, then we're almost surely going to experience confounding.

(contrast 1) – (contrast 2) = difference-in-differences



fisher: a deep insight

- Fisher's randomization
- IF we control the randomization process then we can describe, with mathematical certainty, how the data will behave.
- Armed with this understanding of data's behavior we can then make statements, with varying levels of certainty, about the state of the world.

fisher: a lady tasting tea

- In his 1935 groundbreaking book, *Design of Experiments*, he discusses an (apocryphal?) encounter he had with a lady at a gathering.
- She contended she could taste the difference between tea which had had its milk poured in first versus tea which had had milk poured in after the tea.
- Fisher thought this was hogwash and proceeded to develop a "test" of her claim.
- Interestingly, he discusses some of the reasoning that led him to this particular test.

fisher: a lady tasting tea

- In in Chapter 2 (p. 18) he wrote:

  It is not sufficient remedy to insist that "all the cups must be exactly alike" in every respect except that to be tested. For this is a totally impossible requirement in our example, and equally in all other forms of experimentation … These are only examples of the differences probably present; it would be impossible to present an exhaustive list of such possible differences … because [they] … are always strictly innumerable. When any such cause is named, it is usually perceived that, by increased labor and expense, it could be largely eliminated. Too frequently it is assumed that such refinements constitute improvements to the experiment …

confounding

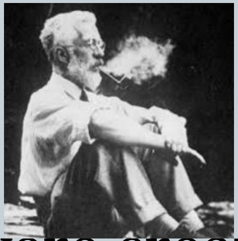
- Confounding comes in two flavors:
(i) observed confounding (the covariates are in your data set and we can probably do something), and
(ii) unobserved confounding (you're going to have a *really* rough time…)
- We'll come back to this… but there are quite different tools based on whether or not you can justify that there is no unobserved confounding.
- Assume you have unobserved confounding, until proven otherwise.

fisher: a lady tasting tea

- His point: You will always have confounding. It will be annoying. Let's move past that.
- His proposal?
- Propose a treatment assignment process that is well-described mathematically and random
- Propose a hypothesis that will explain how the data should look in general
  - This is really important, the theory here should contain information about how the intervention interacts with the outcome,
  - The theory should guide you in which confounders are most impactful, and how to measure the outcome(s).
- Run the experiment and compare the observed data to the actual way the world worked

beyond RCTs

- Randomized controlled trials (RCTs) are excellent
  - The "controlled" part addresses Mill's ideas of minimizing differences at baseline
  - The "randomized" part addresses Fisher's ideas of understanding what-else-could-have-happened
- Unfortunately, RCTs can't be implemented in all situations:
  - Does smoking cause cancer?
  - Are higher level NICUs better?
  - Expensive? Feasibility? Useful?



- There are study designs that were created to work "in the real world," and they follow many of these ideas…

beyond RCTs

- The world of “observational studies” is kind of hard to get into because it grew up in several distinct, but overlapping, disciplines:
  - Epidemiology
  - Demography
  - Economics (econometrics)
  - Political Science
  - Sociology
  - Biostatistics
  - Statistics
  - Psychology (psychometrics)
  - Computer Science

another way to look at confounding

- Economists created "structural equation models" to back out the confounding. These describe the mathematical essence of the real-world process.

$$t = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \overbrace{\alpha_p x_p + \varepsilon_t}^{\boldsymbol{\varepsilon_t^*}}$$

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \underbrace{\beta_p x_p + \eta_t}_{\boldsymbol{\eta_t^*}}$$

Sometimes called: *omitted variable bias*

study designs for high quality inference

- The best:
  - Randomized controlled trials
- IF you're dealing with observational studies (e.g., data that were not generated from an RCT)
  - Difference-in-differences
  - Regression discontinuity
  - Instrumental variables

You can point to the randomness.

- IF you're in a very peculiar situation, and have measured all covariates, then:

Hoping what's leftover is random.

  - Propensity score matching
  - Inverse probability weighting
  - Possibly even a regression (depending on the level of theory)

unobserved confounding

There are more things in heaven and earth, Horatio,
Than are dreamt of in your philosophy.
- *Hamlet* (1.5.167-8)

addressing unobserved confounding

- Negative control outcomes ([Eric Tchetgen Tchetgen, AJE 2014])
  - An outcome is said to be a valid negative control variable to the extent that it is influenced by unobserved confounders of the exposure effects on the outcome in view, although not directly influenced by the exposure.
  - Thus, a negative control outcome found to be empirically associated with the exposure after adjustment for observed confounders indicates that unobserved confounding may be present.

addressing unobserved confounding

- Sensitivity analysis (Paul Rosenbaum - [Sensitivity Analysis in Observational Studies])
  - A sensitivity analysis… asks what the unmeasured covariate[s] would have to be like to alter the conclusions of the study.
  - Observational studies vary markedly in their sensitivity to hidden bias: some are sensitive to very small biases, while others are insensitive to quit large biases.