Empirical Research on Economic Inequality
Experiments to test for discrimination in hiring

Maximilian Kasy

Harvard University, fall 2015

Inequality between groups

- We observe large economic inequalities along dimensions such as race and gender.
- Why?
- Many channels through which they might be created!

Possible channels

Differences in

1. early childhood influences
2. neighborhoods of growing up
3. access to / quality of
   primary, middle, and high school education
4. chance of being hired when applying for a job
5. wages conditional on being hired
6. chance of being promoted or fired in a given job
7. treatment by customers or clients
8. treatment by police and courts
9. ...

4. Chance of being hired when applying for a job

Decomposes further into

a. chance of being invited to an interview
b. chance of being hired given an interview

a. Chance of being invited to an interview

Bertrand, M. and Mullainathan, S. (2004). Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination. American Economic Review*, 94(4):991–1013.*

- Chance might depend on
  1. the (perceived) race and gender of an applicant,
  2. neighborhood of residence,
  3. the high school attended, ...
- Bertrand and Mullainathan (2004):
  What is the causal effect of perceived race
  on the chance of being invited to an interview,
  for *otherwise identical* CVs?

What is a causal effect?

- Potential outcome framework: answer to “what if” questions
- Two “treatments”: $D = 0$ or $D = 1$
- e.g. “black name” vs. “white name” on the CV
- Y_i: CV i’s outcome
  e.g. being invited for an interview
- Potential outcome Y_i^0:
  what if CV i had a “black name” (treatment 0)
- Potential outcome Y_i^1:
  what if CV i had a “white name” (treatment 1)

Questions for you

Does the “what if” question make sense?
After all, we can never observe what would have happened!

- Causal effect / treatment effect for CV i :
  $Y_i^1 - Y_i^0$.
- Average causal effect / average treatment effect:

$$ATE = E[Y^1 - Y^0],$$

- Expectation averages over the population of interest.

The fundamental problem of causal inference

- **We never observe both Y^0 and Y^1 at the same time.**
- One of the potential outcomes is always missing from the data.
- Treatment D determines which of the two we observe.
- Formally:

$$Y = D \cdot Y^1 + (1 - D) \cdot Y^0.$$

Selection problem

- Distribution of Y^1 among those with $D = 1$
  need not be the same as the distribution of Y^1 among everyone.
- In particular

$$E[Y|D=1] = E[Y^1|D=1] \neq E[Y^1]$$
$$E[Y|D=0] = E[Y^0|D=0] \neq E[Y^0]$$
$$E[Y|D=1] - E[Y|D=0] \neq E[Y^1 - Y^0] = ATE.$$

- e.g., for real job applicants, race correlates with neighborhood, school, etc. ...

Randomization

- No selection $\Leftrightarrow$ D is random

$$(Y^0, Y^1) \perp D.$$

- In this case,

$$E[Y|D=1] = E[Y^1|D=1] = E[Y^1]$$
$$E[Y|D=0] = E[Y^0|D=0] = E[Y^0]$$
$$E[Y|D=1] - E[Y|D=0] = E[Y^1 - Y^0] = ATE.$$

- Can ensure this by actually randomly assigning D.
- Independence $\Rightarrow$ comparing treatment and control actually compares "apples with apples."
- This gives **empirical content to** the "metaphysical" notion of **potential outcomes**!

Estimation

- Easy for randomized experiments
- Recall

$$ATE = E[Y_1 - Y_0] = E[Y|D=1] - E[Y|D=0].$$

- Estimator:

$$\widehat{\alpha} = \overline{Y}_1 - \overline{Y}_0,$$

where

$$\overline{Y}_1 = \frac{\sum Y_i \cdot D_i}{\sum D_i} = \frac{1}{N_1} \sum_{D_i=1} Y_i$$

$$\overline{Y}_0 = \frac{\sum Y_i \cdot (1 - D_i)}{\sum (1 - D_i)} = \frac{1}{N_0} \sum_{D_i=0} Y_i.$$

Questions for you

Show that

$$E[\widehat{\alpha}] = ATE$$

if $(Y^0, Y^1) \perp D$.

Inference

- Range of likely values for ATE?
- t-statistic:

$$t = \frac{\widehat{\alpha} - \alpha_{ATE}}{\widehat{\sigma}_{\alpha}}$$

where

$$\widehat{\sigma}_{\alpha} = \sqrt{\frac{\widehat{\sigma}_1^2}{N_1} + \frac{\widehat{\sigma}_0^2}{N_0}}$$

and

$$\widehat{\sigma}_1^2 = \frac{1}{N_1 - 1} \sum_{D_i = 1} (Y_i - \overline{Y}_1)^2.$$

- $\widehat{\sigma}_0^2$ is analogously defined.

Confidence interval

- t-statistic is approximately standard normal distributed (for samples of a reasonable size),

$$t \sim^{approx} N(0,1).$$

- 95% confidence interval:

$$CI = [\widehat{\alpha} - 1.96 \cdot \widehat{\sigma}_{\alpha}, \widehat{\alpha} + 1.96 \cdot \widehat{\sigma}_{\alpha}] \, .$$

Questions for you

Show that

$$P(\alpha \in CI) \approx 0.95.$$

(Homework)

Note that α is fixed, while CI is random!