# Causality

POST 8000 – Foundations of Social Science Research for Public Policy

## Steven V. Miller

Department of Political Science

# Goal for Today

*Introduce students to causality, and distinguishing causality from association.*

# The Problem, in Quotes

- "That correlation is not causation is perhaps the first thing that must be said." - Barnard, 1982 (p. 387)
- "If statistics cannot relate cause effect, they add to the rhetoric." - Smith, 1980 (p. 1000 [stylized by me])

# Associational Inference

A set of tools to understand how a response variable corresponds with some attribute. Tools include:

- Probability distributions (conditional, joint)
- Correlation
- Regression(?)

"Associational inference consists of [estimates, tests, posterior distributions, etc.] about the associational paramters relating $Y$ and $A$ [from units in $U$]. In this sense, associational inference is simply descriptive statistics." - Holland, 1986 (p. 946)

# Probability Distributions

**Joint probability**, in the event *A* and *B* are independent from each other:

$$p(A, B) = p(A) * p(B)$$

**Conditional probability**, in the event that *A* depends on *B* having already occurred:

$$p(A \mid B) = \frac{p(A, B)}{p(B)}$$

# Correlation (via Pearson's *r*)

$$\Sigma \frac{\left(\frac{x_i - \overline{x}}{s_x}\right)\left(\frac{y_i - \overline{y}}{s_y}\right)}{n-1}$$

...where:

- $x_i$, $y_i$ = individual observations of *x* or *y*, respectively.
- $\overline{x}$, $\overline{y}$ = sample means of *x* and *y*, respectively.
- $s_x$, $s_y$ = sample standard deviations of *x* and *y*, respectively.
- *n* = number of observations in the sample.

# Properties of Pearsons *r*

1. Pearson's *r* is symmetrical.
2. Pearson's *r* is bound between -1 and 1.
3. Pearson's *r* is standardized.

# Standardization

$$z = \frac{\text{Deviation from the mean}}{\text{Standard unit}} \tag{1}$$

The standard unit will vary, contingent on what you want.

- If you're working with just one random sample, it's the standard deviation.
- If you're comparing sample means across multiple random samples, it's the standard error.

# Standardization

Larger *z* values indicate greater difference from the mean.

- When *z* = 0, there is no deviation from the mean (obviously).

Standardization has a lot of cool properties you'll see through the semester.

- For now: it's a way to express a variable's scale.

# Causal Inference

Causal inference owes much to Rubin's "potential outcomes framework.

# The Problem in a Nutshell

An individual ($i$) who is offered a treatment ($Z_i$ = 1) with two potential outcomes:

- An outcome to be revealed if s/he accepts the treatment ($T_i$ = 1): $Y_i(T_i = 1 | Z_i = 1)$
- An outcome to be revealed if untreated ($T_i$ = 0): $Y_i(T_i = 0 | Z_i = 1)$

This is a missing data problem of a kind.

- We can only observe one.
- No perfect counterfactuals.
- Unicorns don't exist.

## The Solution

For $T_i = 0$ and $T_i = 1$, given both offered treatment ($Z_i = 1$):

Individual Treatment Effect for $i = Y_i(T_i = 1|Z_i = 1) - Y_i(T_i = 0|Z_i = 1)$

Think in terms of population averages.

- Per Rubin, there is an important population parameter to estimate.
- Hence why we (and he) referred to it as "effect of the treatment on the treated." (i.e. TOT)
- Also: the "average treatment effect" (i.e. ATE)

# The Importance of Random Assignment

Random assignment (to treatment/control) helps us with ATE because it's tough to imagine cases where ($Z_i = 1$ and $T_i = 0$).

- Per random assignment: participants assigned to treatment/control must be same on average in the population ("equal in expectation").
- i.e. $E[Y_i(T_i = 0|Z_i = 1)]$ must be equal to $E[Y_i(T_i = 0|Z_i = 0)]$

By substitution:

$$TOT = E[Y_i(T_i = 1|Z_i = 1)] - E[Y_i(T_i = 0|Z_i = 0)]$$

When unbiased, a difference in sample means is sufficient:

$$\hat{TOT} = \frac{\sum_{i=1}^{n_1} Y_i}{n_1} - \frac{\sum_{i=1}^{n_0} Y_i}{n_0}$$

# Some Other Important Assumptions

- Exogeneity (worth reiterating)
- Unit homogeneity
- Conditional independence
- SUTVA

# Criteria for Evaluating Causal Arguments

- Falsifiability
- Internal consistency
- Careful selection of DV
- Concreteness
- "Encompassibility" (sic)

# Table of Contents