# Gov 99r: Lecture 6

Ashley Anderson

Harvard University

October 23, 2014

# Outline

# Experiments

Experiments are about finding the *average treatment effect* that is the difference, on average, in outcomes between those subjects who are treated and those who are not.

# Experiments

Experiments are about finding the *average treatment effect* that is the difference, on average, in outcomes between those subjects who are treated and those who are not.

Mathematically this means we want:

# Experiments

Experiments are about finding the *average treatment effect* that is the difference, on average, in outcomes between those subjects who are treated and those who are not.

Mathematically this means we want:

$$\alpha_{ATE} = E[Y_1 - Y_0] = E[Y_1] - E[Y_0]$$

# Experiments: An Example

Consider a question: Does smoking cause individuals to die faster than non-smokers?

# Experiments: An Example

Consider a question: Does smoking cause individuals to die faster than non-smokers?

How would you answer this?

# Experiments: An Example

Consider a question: Does smoking cause individuals to die faster than non-smokers?

How would you answer this?

- Find a group of smokers (treated) and non-smokers (control)

# Experiments: An Example

Consider a question: Does smoking cause individuals to die faster than non-smokers?

How would you answer this?

- Find a group of smokers (treated) and non-smokers (control)
- Determine the average age of death for the two groups

# Experiments: An Example

Consider a question: Does smoking cause individuals to die faster than non-smokers?

How would you answer this?

- Find a group of smokers (treated) and non-smokers (control)
- Determine the average age of death for the two groups
- Subtract to find average treatment effect

# Experiments: An Example

Consider a question: Does smoking cause individuals to die faster than non-smokers?

How would you answer this?

- Find a group of smokers (treated) and non-smokers (control)
- Determine the average age of death for the two groups
- Subtract to find average treatment effect

What is wrong with this picture?

# Experiments: An Example

What is wrong with this picture?

- Smokers and non-smokers different

# Experiments: An Example

What is wrong with this picture?

- Smokers and non-smokers different
- Smokers may be different in ways that also affect life expectancy

# Experiments: An Example

What is wrong with this picture?

- Smokers and non-smokers different
- Smokers may be different in ways that also affect life expectancy
- Treatment process has selection bias

# Experiments: An Example

What is wrong with this picture?

- Smokers and non-smokers different
- Smokers may be different in ways that also affect life expectancy
- Treatment process has selection bias

$$E[Y_0|D=1] - E[Y_0|D=0]$$

## Experiments: The Ideal

Ideally to eliminate bias we would test the effect of treatment on each individual by seeing the outcome for the individual if they were treated and if they were not treated.

# Experiments: The Ideal

Ideally to eliminate bias we would test the effect of treatment on each individual by seeing the outcome for the individual if they were treated and if they were not treated.

$$\alpha - i = Y - 1i - Y_{0i}$$

# Experiments: The Ideal

Ideally to eliminate bias we would test the effect of treatment on each individual by seeing the outcome for the individual if they were treated and if they were not treated.

$$\alpha - i = Y - 1i - Y_{0i}$$

Yet this is impossible...

# Experiments: The Ideal

Ideally to eliminate bias we would test the effect of treatment on each individual by seeing the outcome for the individual if they were treated and if they were not treated.

$$\alpha - i = Y - 1i - Y_{0i}$$

Yet this is impossible...

# Assignment mechanisms: The Cure to All Ills

To assess and reduce bias in our estimates we must focus on the assignment mechanism

# Assignment mechanisms: The Cure to All Ills

To assess and reduce bias in our estimates we must focus on the assignment mechanism

- Random assignment

# Assignment mechanisms: The Cure to All Ills

To assess and reduce bias in our estimates we must focus on the assignment mechanism

- Random assignment
- Selection based on observable characteristics

# Assignment mechanisms: The Cure to All Ills

To assess and reduce bias in our estimates we must focus on the assignment mechanism

- Random assignment
- Selection based on observable characteristics
- Selection based on unobservable characteristics

# Assignment mechanisms: The Cure to All Ills

To assess and reduce bias in our estimates we must focus on the assignment mechanism

- Random assignment
- Selection based on observable characteristics
- Selection based on unobservable characteristics

The goal is to get at causality by identifying a set of observable outcomes (for the treated) that can stand in for the unobserved potential outcomes (for treated units as if they did not receive treatment)

# Regression Basics

In most cases we cannot perform an experiment to estimate causal effects. This can happen for a number of reasons:

## Regression Basics

In most cases we cannot perform an experiment to estimate causal effects. This can happen for a number of reasons:

- Humans are hard to control

# Regression Basics

In most cases we cannot perform an experiment to estimate causal effects. This can happen for a number of reasons:

- Humans are hard to control
- Unethical to assign "inconvenient treatments " like war, bad institutions, crisis

# Regression Basics

In most cases we cannot perform an experiment to estimate causal effects. This can happen for a number of reasons:

- Humans are hard to control
- Unethical to assign "inconvenient treatments " like war, bad institutions, crisis
- Things that happened in the past are interesting to study

# Regression Basics

In most cases we cannot perform an experiment to estimate causal effects. This can happen for a number of reasons:

- Humans are hard to control
- Unethical to assign "inconvenient treatments " like war, bad institutions, crisis
- Things that happened in the past are interesting to study

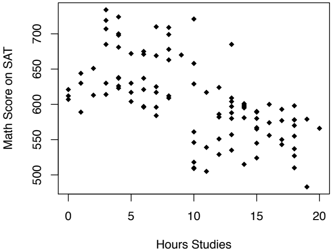
Here regression can be helpful to estimate "causal" effects.

# Regression Basics

At the most base level, regression is finding the line of best fit for a set of data

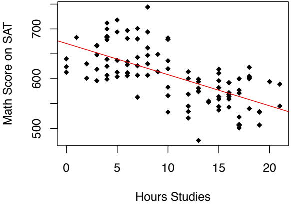
# Regression Basics

At the most base level, regression is finding the line of best fit for a set of data



# Regression Basics

At the most base level, regression is finding the line of best fit for a set of data



# Regression Basics

While all regressions find a line of best fit, not all regressions provide causal evidence

# Regression Basics

While all regressions find a line of best fit, not all regressions provide causal evidence

- Regression Discontinuity

# Regression Basics

While all regressions find a line of best fit, not all regressions provide causal evidence

- Regression Discontinuity
- Difference in Difference

# Regression Basics

While all regressions find a line of best fit, not all regressions provide causal evidence

- Regression Discontinuity
- Difference in Difference
- Matching

# Regression Basics

While all regressions find a line of best fit, not all regressions provide causal evidence

- Regression Discontinuity
- Difference in Difference
- Matching

Most of you will be doing linear regression so we will focus on that.

# Regression Basics

While all regressions find a line of best fit, not all regressions provide causal evidence

- Regression Discontinuity
- Difference in Difference
- Matching

Most of you will be doing linear regression so we will focus on that.

# Regression Basics

- Linear regression is trying to find the best $\beta$'s for the equation $y = \beta_0 + \beta_1 * x$

## Regression Basics

- Linear regression is trying to find the best $\beta$'s for the equation $y = \beta_0 + \beta_1 * x$
- What is best?

# Regression Basics

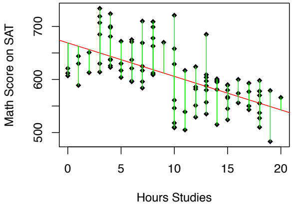
- Linear regression is trying to find the best $\beta$'s for the equation $y = \beta_0 + \beta_1 * x$
- What is best?
- This means picking $\beta$'s to minimize the sum of $(y = \beta_0 + \beta_1 * x)^2$

# Regression Basics

- Linear regression is trying to find the best $\beta$'s for the equation $y = \beta_0 + \beta_1 * x$
- What is best?
- This means picking $\beta$'s to minimize the sum of $(y = \beta_0 + \beta_1 * x)^2$
- Minimizing how far the points are from the line

# Regression Basics

- Linear regression is trying to find the best $\beta$'s for the equation $y = \beta_0 + \beta_1 * x$
- What is best?
- This means picking $\beta$'s to minimize the sum of $(y = \beta_0 + \beta_1 * x)^2$
- Minimizing how far the points are from the line



# Multivariable Regression

This is simple when regressing one variable onto an outcome, but harder with more variables. When multiple variables are used we need the following to be satisfied:

# Multivariable Regression

This is simple when regressing one variable onto an outcome, but harder with more variables. When multiple variables are used we need the following to be satisfied:

- Independence assumption

# Multivariable Regression

This is simple when regressing one variable onto an outcome, but harder with more variables. When multiple variables are used we need the following to be satisfied:

- Independence assumption
- Constant variance

# Multivariable Regression

This is simple when regressing one variable onto an outcome, but harder with more variables. When multiple variables are used we need the following to be satisfied:

- Independence assumption
- Constant variance
- Knowledge of data generating process

# Multivariable Regression

This is simple when regressing one variable onto an outcome, but harder with more variables. When multiple variables are used we need the following to be satisfied:

- Independence assumption
- Constant variance
- Knowledge of data generating process

Satisfying these assumptions fully will get you close to causal language, but for the most part you will be identifying correlation not causation

# Models for Regression

Different regression models do different things.

# Models for Regression

Different regression models do different things.

- Logit analysis: binary outcome, log distribution

# Models for Regression

Different regression models do different things.

- Logit analysis: binary outcome, log distribution
- Probit analysis: binary outcome, normal distribution

# Models for Regression

Different regression models do different things.

- Logit analysis: binary outcome, log distribution
- Probit analysis: binary outcome, normal distribution
- Orderded probit/logit: more than two mutually exclusive outcomes

# Models for Regression

Different regression models do different things.

- Logit analysis: binary outcome, log distribution
- Probit analysis: binary outcome, normal distribution
- Orderded probit/logit: more than two mutually exclusive outcomes

# Case Selection Basics

In doing causal inference in qualitative studies, case study selection is crucial. Yet, there are many ways to choose cases:

# Case Selection Basics

In doing causal inference in qualitative studies, case study selection is crucial. Yet, there are many ways to choose cases:

- Most similar

# Case Selection Basics

In doing causal inference in qualitative studies, case study selection is crucial. Yet, there are many ways to choose cases:

- Most similar
- Most different

# Case Selection Basics

In doing causal inference in qualitative studies, case study selection is crucial. Yet, there are many ways to choose cases:

- Most similar
- Most different
- Typical

# Case Selection Basics

In doing causal inference in qualitative studies, case study selection is crucial. Yet, there are many ways to choose cases:

- Most similar
- Most different
- Typical
- Diverse

# Case Selection Basics

In doing causal inference in qualitative studies, case study selection is crucial. Yet, there are many ways to choose cases:

- Most similar
- Most different
- Typical
- Diverse
- Extreme

# Case Selection Basics

In doing causal inference in qualitative studies, case study selection is crucial. Yet, there are many ways to choose cases:

- Most similar
- Most different
- Typical
- Diverse
- Extreme
- Deviant

# Case Selection Basics

In doing causal inference in qualitative studies, case study selection is crucial. Yet, there are many ways to choose cases:

- Most similar
- Most different
- Typical
- Diverse
- Extreme
- Deviant
- Influential

# Case Selection Basics

In doing causal inference in qualitative studies, case study selection is crucial. Yet, there are many ways to choose cases:

- Most similar
- Most different
- Typical
- Diverse
- Extreme
- Deviant
- Influential
- Crucial

## Case Selection Basics

In doing causal inference in qualitative studies, case study selection is crucial. Yet, there are many ways to choose cases:

- Most similar
- Most different
- Typical
- Diverse
- Extreme
- Deviant
- Influential
- Crucial

# Case Selection Basics

Diverse, extreme, typical, influential, deviant, and crucial cases all include selection bias. We are selecting these cases *particularly because* they are odd or tell us something special about the data.

# Case Selection Basics

Diverse, extreme, typical, influential, deviant, and crucial cases all include selection bias. We are selecting these cases *particularly because* they are odd or tell us something special about the data.

You might think of this as "selecting on the dependent variable"

# Case Selection Basics

Diverse, extreme, typical, influential, deviant, and crucial cases all include selection bias. We are selecting these cases *particularly because* they are odd or tell us something special about the data.

You might think of this as "selecting on the dependent variable"

To get closer to experimentation and to causal inference, we might employ most similar or most different cases.

# Case Selection Basics

Diverse, extreme, typical, influential, deviant, and crucial cases all include selection bias. We are selecting these cases *particularly because* they are odd or tell us something special about the data.

You might think of this as "selecting on the dependent variable"

To get closer to experimentation and to causal inference, we might employ most similar or most different cases.

# Most Similar Design

In most similar design, cases are selected because they are (reasonably) identical in terms of their observable characteristics.

# Most Similar Design

In most similar design, cases are selected because they are (reasonably) identical in terms of their observable characteristics.

- Mimics an experiment (independence assumption)

# Most Similar Design

In most similar design, cases are selected because they are (reasonably) identical in terms of their observable characteristics.

- Mimics an experiment (independence assumption)
- All IVs of interest have the same or very similar values

# Most Similar Design

In most similar design, cases are selected because they are (reasonably) identical in terms of their observable characteristics.

- Mimics an experiment (independence assumption)
- All IVs of interest have the same or very similar values
- Bias still exists in unobservables, confounders must be exogenous

# Most Similar Design

In most similar design, cases are selected because they are (reasonably) identical in terms of their observable characteristics.

- Mimics an experiment (independence assumption)
- All IVs of interest have the same or very similar values
- Bias still exists in unobservables, confounders must be exogenous

This helps us to isolate our variable of interest (treatment) and see the effects

# Most Different Design

In most similar design, cases are selected because they are (reasonably) different in terms of their observable characteristics.

# Most Different Design

In most similar design, cases are selected because they are (reasonably) different in terms of their observable characteristics.

Very different from an experiment because IVs have opposite values, and bias still exists.

# Most Different Design

In most similar design, cases are selected because they are (reasonably) different in terms of their observable characteristics.

Very different from an experiment because IVs have opposite values, and bias still exists.

In an inverse way, this helps us to isolate our variable of interest (treatment) and see the effects on producing a similar outcome **despite** differences.

## Most Different Design

In most similar design, cases are selected because they are (reasonably) different in terms of their observable characteristics.

Very different from an experiment because IVs have opposite values, and bias still exists.

In an inverse way, this helps us to isolate our variable of interest (treatment) and see the effects on producing a similar outcome **despite** differences.