# 9.S918: Statistical Inference in Brain and Cognitive Sciences

## Week 1 Day 2: Introduction to causal inference


Roger Levy
Dept. of Brain & Cognitive Sciences
Massachusetts Institute of Technology


Feb 5, 2025

# A tiny bit of statistics

- On Monday we reviewed basics of **probability**: the logical calculus of uncertainty–a branch of mathematics
- The primary focus of this class is **statistics**: the mathematics, science, craft, and art of drawing inferences from data
- The two fields are fundamentally different
- But, probability is used extensively throughout statistics

# Perhaps the simplest probability **distribution**

# Perhaps the simplest probability **distribution**

- Consider a binary random variable $Y$ with two possible outcomes: 0 and 1

# Perhaps the simplest probability **distribution**

- Consider a binary random variable $Y$ with two possible outcomes: 0 and 1
- $Y$ is a **Bernoulli random variable** with **parameter** $P(\text{heads}) = \pi$, where $0 \leq \pi \leq 1$

# Perhaps the simplest probability **distribution**

- Consider a binary random variable $Y$ with two possible outcomes: 0 and 1
- $Y$ is a **Bernoulli random variable** with **parameter** $P(\text{heads}) = \pi$, where $0 \leq \pi \leq 1$
- Figuring out from observed data what the weighting is likely to be is **parameter estimation**

# Perhaps the simplest probability **distribution**

- Consider a binary random variable $Y$ with two possible outcomes: 0 and 1
- $Y$ is a **Bernoulli random variable** with **parameter** $P(\text{heads}) = \pi$, where $0 \leq \pi \leq 1$
- Figuring out from observed data what the weighting is likely to be is **parameter estimation**
- In general, we will use $\mathbf{y}$ to refer to observed-outcome **data** and $\theta$ to refer to the model parameters to be estimated

# Statistical estimators

# Statistical estimators

- **Estimator:** a procedure for guessing a quantity of interest within a population from a sample from that population

# Statistical estimators

- **Estimator:** a procedure for guessing a quantity of interest within a population from a sample from that population
- For example, the **relative frequency estimator:** if we observe $r$ instances of heads in $n$ coin flips,

$$\hat{\pi} = \frac{r}{n}$$

# Statistical estimators

- **Estimator:** a procedure for guessing a quantity of interest within a population from a sample from that population
- For example, the **relative frequency estimator:** if we observe $r$ instances of heads in $n$ coin flips,

"this is an estimator"

$$\boxed{\hat{\phantom{\pi}}}\!\!\!\!\pi = \frac{r}{n}$$

# Statistical estimators

- **Estimator:** a procedure for guessing a quantity of interest within a population from a sample from that population
- For example, the **relative frequency estimator:** if we observe $r$ instances of heads in $n$ coin flips,

"this is an estimator"

$$\hat{\pi} = \frac{r}{n}$$

- Data are stochastic, so estimators give random variables!

# Statistical estimators

- **Estimator:** a procedure for guessing a quantity of interest within a population from a sample from that population
- For example, the **relative frequency estimator:** if we observe $r$ instances of heads in $n$ coin flips,

"this is an estimator"

$$\boxed{\widehat{\phantom{\pi}}}\!\!\!\!\pi = \frac{r}{n}$$

- Data are stochastic, so estimators give random variables!
- **Bias** of an estimator is $E[\widehat{\theta}] - \theta$

# Statistical estimators

- **Estimator:** a procedure for guessing a quantity of interest within a population from a sample from that population
- For example, the **relative frequency estimator:** if we observe $r$ instances of heads in $n$ coin flips,

"this is an estimator"

$$\boxed{\widehat{\ }}\pi = \frac{r}{n}$$

- Data are stochastic, so estimators give random variables!
- **Bias** of an estimator is $E[\widehat{\theta}] - \theta$

$$E[\widehat{\pi}] = E[\frac{r}{n}] = \frac{1}{n}E[r] = \frac{1}{n}\sum_{i=1}^{n} E[Y_i] = \frac{1}{n}n\pi = \pi$$

# Statistical estimators

- **Estimator:** a procedure for guessing a quantity of interest within a population from a sample from that population
- For example, the **relative frequency estimator:** if we observe $r$ instances of heads in $n$ coin flips,

"this is an estimator"

$$\boxed{\hat{\ }}\!\pi = \frac{r}{n}$$

- Data are stochastic, so estimators give random variables!
- **Bias** of an estimator is $E[\hat{\theta}] - \theta$

Here we used **linearity of the expectation**

$$E[\hat{\pi}] = E[\frac{r}{n}] = \frac{1}{n}E[r] = \frac{1}{n}\sum_{i=1}^{n} E[Y_i] = \frac{1}{n}n\pi = \pi$$

# Statistical estimators

- **Estimator:** a procedure for guessing a quantity of interest within a population from a sample from that population
- For example, the **relative frequency estimator:** if we observe $r$ instances of heads in $n$ coin flips,

"this is an estimator"

$$\boxed{\widehat{\phantom{\pi}}}\!\!\!\pi = \frac{r}{n}$$

- Data are stochastic, so estimators give random variables!
- **Bias** of an estimator is $E[\widehat{\theta}] - \theta$

Here we used **linearity of the expectation**

$$E[\widehat{\pi}] = E[\frac{r}{n}] = \frac{1}{n}E[r] = \frac{1}{n}\sum_{i=1}^{n} E[Y_i] = \frac{1}{n}n\pi = \pi$$

...so $\widehat{\pi}$ is **unbiased**

# Statistical estimators

- **Estimator:** a procedure for guessing a quantity of interest within a population from a sample from that population
- For example, the **relative frequency estimator:** if we observe $r$ instances of heads in $n$ coin flips,

"this is an estimator"

$$\boxed{\hat{\phantom{\pi}}}\!\!\pi = \frac{r}{n}$$

- Data are stochastic, so estimators give random variables!
- **Bias** of an estimator is $E[\hat{\theta}] - \theta$

Here we used **linearity of the expectation**

$$E[\hat{\pi}] = E[\frac{r}{n}] = \frac{1}{n}E[r] = \frac{1}{n}\sum_{i=1}^{n} E[Y_i] = \frac{1}{n}n\pi = \pi$$

...so $\hat{\pi}$ is **unbiased**

- **Variance** of an estimator is ordinary variance

# Statistical estimators

- **Estimator:** a procedure for guessing a quantity of interest within a population from a sample from that population
- For example, the **relative frequency estimator:** if we observe $r$ instances of heads in $n$ coin flips,

"this is an estimator"

$$\boxed{\widehat{\ }}\pi = \frac{r}{n}$$

- Data are stochastic, so estimators give random variables!
- **Bias** of an estimator is $E[\widehat{\theta}] - \theta$

Here we used **linearity of the expectation**

$$E[\widehat{\pi}] = E[\frac{r}{n}] = \frac{1}{n}E[r] = \frac{1}{n}\sum_{i=1}^{n} E[Y_i] = \frac{1}{n}n\pi = \pi$$

...so $\widehat{\pi}$ is **unbiased**

- **Variance** of an estimator is ordinary variance

$$\text{Var}(X) \equiv E[(X - E[X])^2]$$

# Statistical estimators

- **Estimator:** a procedure for guessing a quantity of interest within a population from a sample from that population
- For example, the **relative frequency estimator:** if we observe $r$ instances of heads in $n$ coin flips,

"this is an estimator"

$$\boxed{\hat{\pi}} = \frac{r}{n}$$

- Data are stochastic, so estimators give random variables!
- **Bias** of an estimator is $E[\hat{\theta}] - \theta$

Here we used **linearity of the expectation**

$$E[\hat{\pi}] = E[\frac{r}{n}] = \frac{1}{n}E[r] = \frac{1}{n}\sum_{i=1}^{n} E[Y_i] = \frac{1}{n}n\pi = \pi$$

...so $\hat{\pi}$ is **unbiased**

- **Variance** of an estimator is ordinary variance

$$\text{Var}(X) \equiv E[(X - E[X])^2] \qquad \text{Var}(\hat{\pi}) = \frac{\pi(1-\pi)}{n}$$

(see reading materials)

# Statistical estimators

- **Estimator:** a procedure for guessing a quantity of interest within a population from a sample from that population
- For example, the **relative frequency estimator:** if we observe $r$ instances of heads in $n$ coin flips,

"this is an estimator"

$$\boxed{\hat{\ }}\pi = \frac{r}{n}$$

- Data are stochastic, so estimators give random variables!
- **Bias** of an estimator is $E[\hat{\theta}] - \theta$

Here we used **linearity of the expectation**

$$E[\hat{\pi}] = E[\frac{r}{n}] = \frac{1}{n}E[r] = \frac{1}{n}\sum_{i=1}^{n} E[Y_i] = \frac{1}{n}n\pi = \pi$$

...so $\hat{\pi}$ is **unbiased**

- **Variance** of an estimator is ordinary variance

$$\text{Var}(X) \equiv E[(X - E[X])^2] \qquad \text{Var}(\hat{\pi}) = \frac{\pi(1-\pi)}{n}$$

(see reading materials)

- Good estimators have favorable **bias–variance** tradeoff

# Maximum likelihood estimation

$$\mathrm{Lik}(\boldsymbol{\theta}; \boldsymbol{y}) \equiv P(\boldsymbol{y}|\boldsymbol{\theta}) \qquad \hat{\boldsymbol{\theta}}_{MLE} \stackrel{\text{def}}{=} \arg\max_{\boldsymbol{\theta}} \mathrm{Lik}(\boldsymbol{\theta}; \boldsymbol{y})$$

| ***i*** | $\boldsymbol{y_i}$ |
|---|---|
| 1 | T |
| 2 | T |
| 3 | H |
| 4 | T |

*(repeat slide from lecture 3)*

# Maximum likelihood estimation

$$\mathrm{Lik}(\boldsymbol{\theta}; \boldsymbol{y}) \equiv P(\boldsymbol{y}|\boldsymbol{\theta}) \qquad \hat{\boldsymbol{\theta}}_{MLE} \stackrel{\text{def}}{=} \arg\max_{\boldsymbol{\theta}} \mathrm{Lik}(\boldsymbol{\theta}; \boldsymbol{y})$$

| $i$ | $y_i$ |
|---|---|
| 1 | T |
| 2 | T |
| 3 | H |
| 4 | T |

- *p* refers to the value of P(coin toss$_i$ = Heads)
- Likelihood for the following dataset

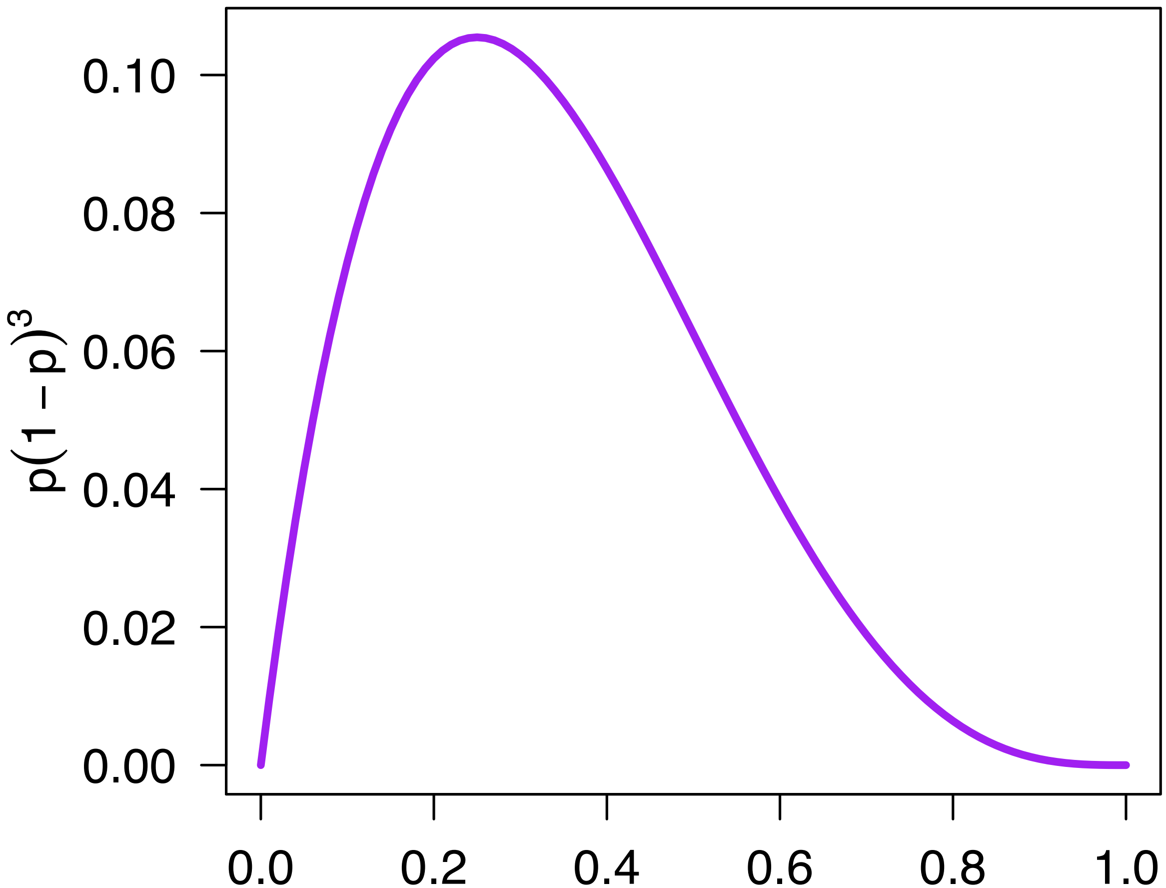
*(repeat slide from lecture 3)*

# Maximum likelihood estimation

$$\mathrm{Lik}(\boldsymbol{\theta}; \boldsymbol{y}) \equiv P(\boldsymbol{y}|\boldsymbol{\theta}) \qquad \hat{\boldsymbol{\theta}}_{MLE} \stackrel{\mathrm{def}}{=} \arg\max_{\boldsymbol{\theta}} \mathrm{Lik}(\boldsymbol{\theta}; \boldsymbol{y})$$

| $i$ | $y_i$ |
|---|---|
| 1 | T |
| 2 | T |
| 3 | H |
| 4 | T |

- *p* refers to the value of P(coin toss$_i$ = Heads)
- Likelihood for the following dataset



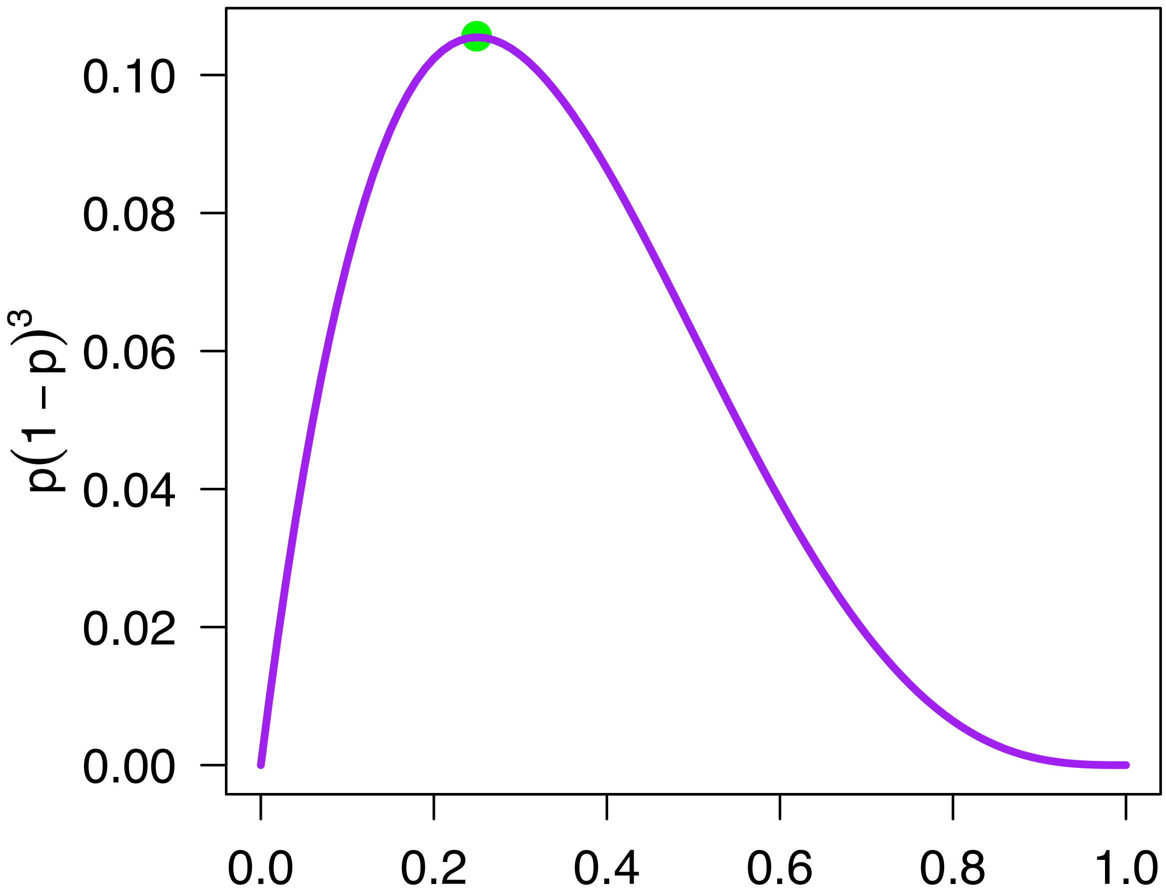

*(repeat slide from lecture 3)*

# Maximum likelihood estimation

$$\text{Lik}(\boldsymbol{\theta}; \boldsymbol{y}) \equiv P(\boldsymbol{y}|\boldsymbol{\theta}) \qquad \hat{\boldsymbol{\theta}}_{MLE} \stackrel{\text{def}}{=} \underset{\boldsymbol{\theta}}{\arg\max}\ \text{Lik}(\boldsymbol{\theta}; \boldsymbol{y})$$

| ***i*** | $y_i$ |
|---|---|
| 1 | T |
| 2 | T |
| 3 | H |
| 4 | T |

- *p* refers to the value of P(coin toss$_i$ = Heads)
- Likelihood for the following dataset



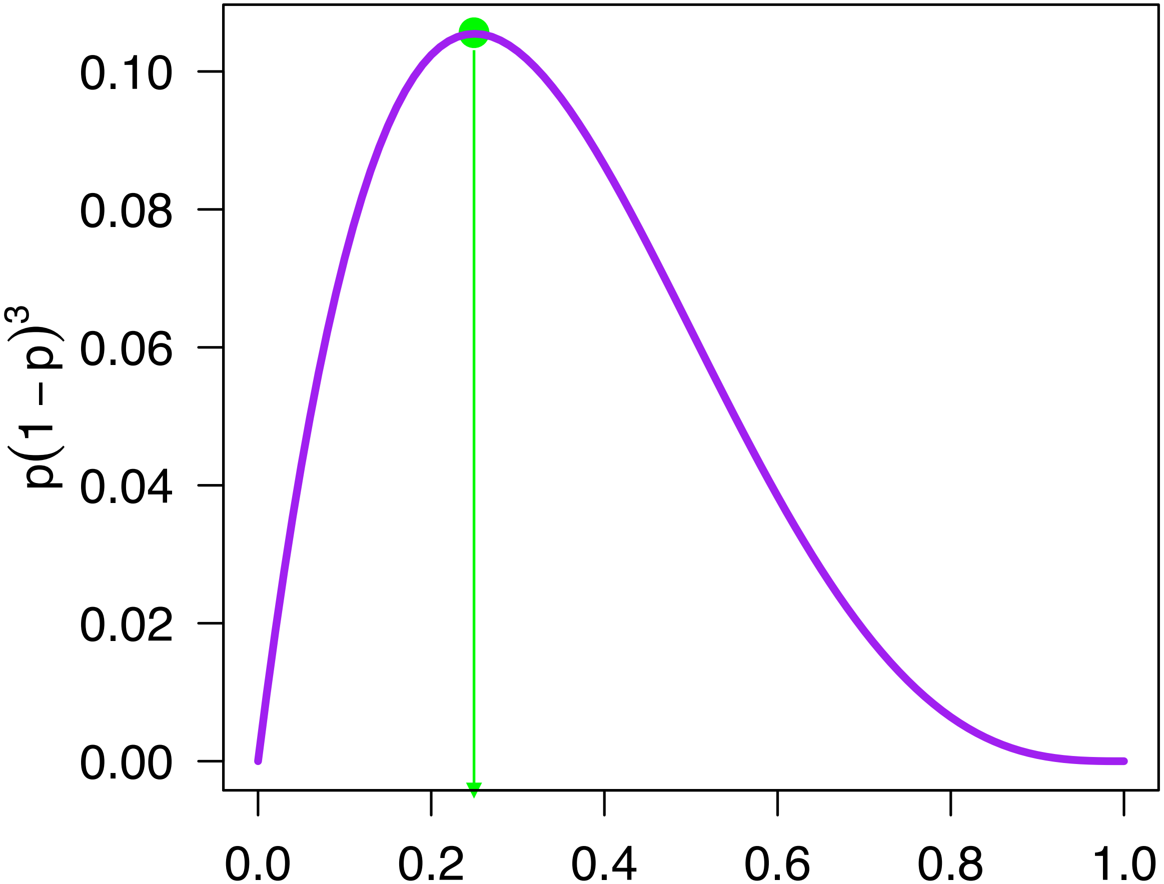

*(repeat slide from lecture 3)*

# Maximum likelihood estimation

$$\mathrm{Lik}(\boldsymbol{\theta}; \boldsymbol{y}) \equiv P(\boldsymbol{y}|\boldsymbol{\theta}) \qquad \hat{\boldsymbol{\theta}}_{MLE} \stackrel{\mathrm{def}}{=} \arg\max_{\boldsymbol{\theta}} \mathrm{Lik}(\boldsymbol{\theta}; \boldsymbol{y})$$

| $i$ | $y_i$ |
|---|---|
| 1 | T |
| 2 | T |
| 3 | H |
| 4 | T |

- $p$ refers to the value of P(coin toss$_i$ = Heads)
- Likelihood for the following dataset



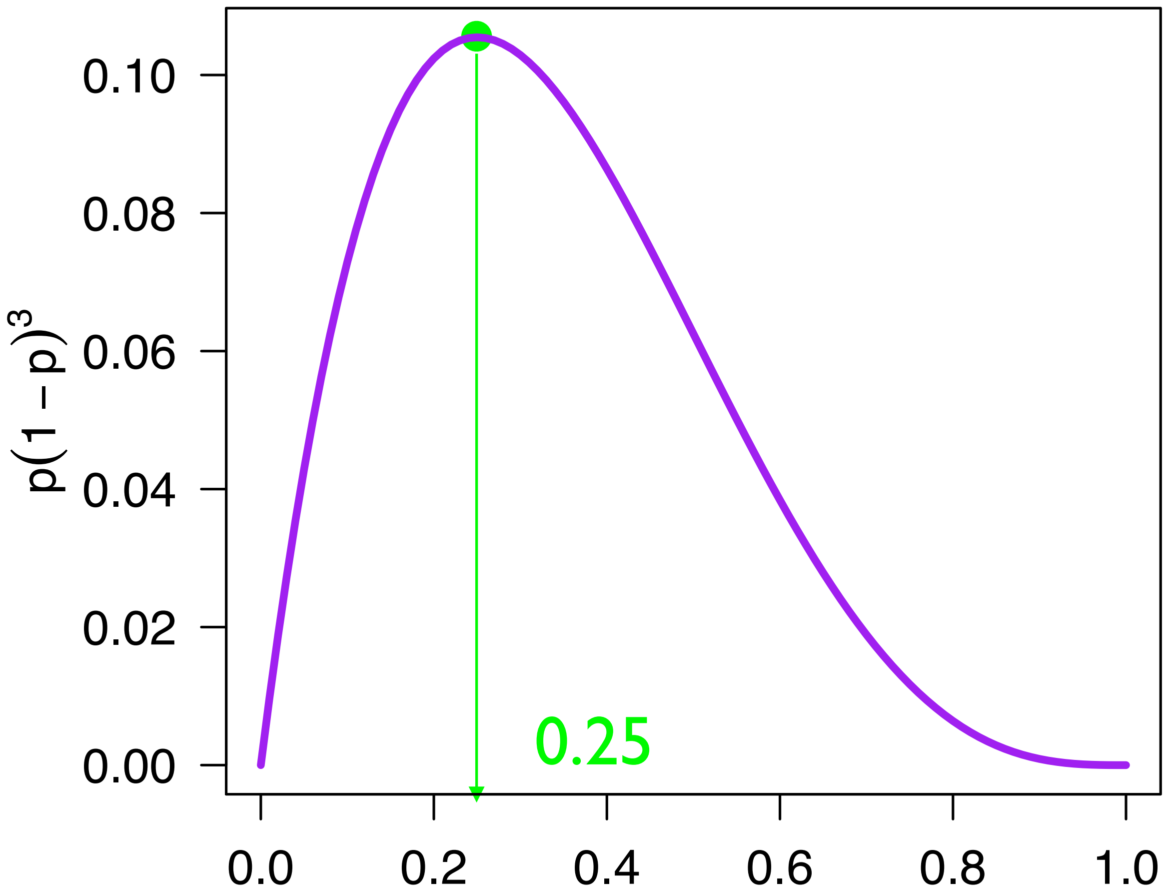

*(repeat slide from lecture 3)*

# Maximum likelihood estimation

$$\mathrm{Lik}(\boldsymbol{\theta}; \boldsymbol{y}) \equiv P(\boldsymbol{y}|\boldsymbol{\theta}) \qquad \hat{\boldsymbol{\theta}}_{MLE} \stackrel{\text{def}}{=} \arg\max_{\boldsymbol{\theta}} \mathrm{Lik}(\boldsymbol{\theta}; \boldsymbol{y})$$

| $i$ | $y_i$ |
|---|---|
| 1 | T |
| 2 | T |
| 3 | H |
| 4 | T |

- *p* refers to the value of P(coin toss$_i$ = Heads)
- Likelihood for the following dataset



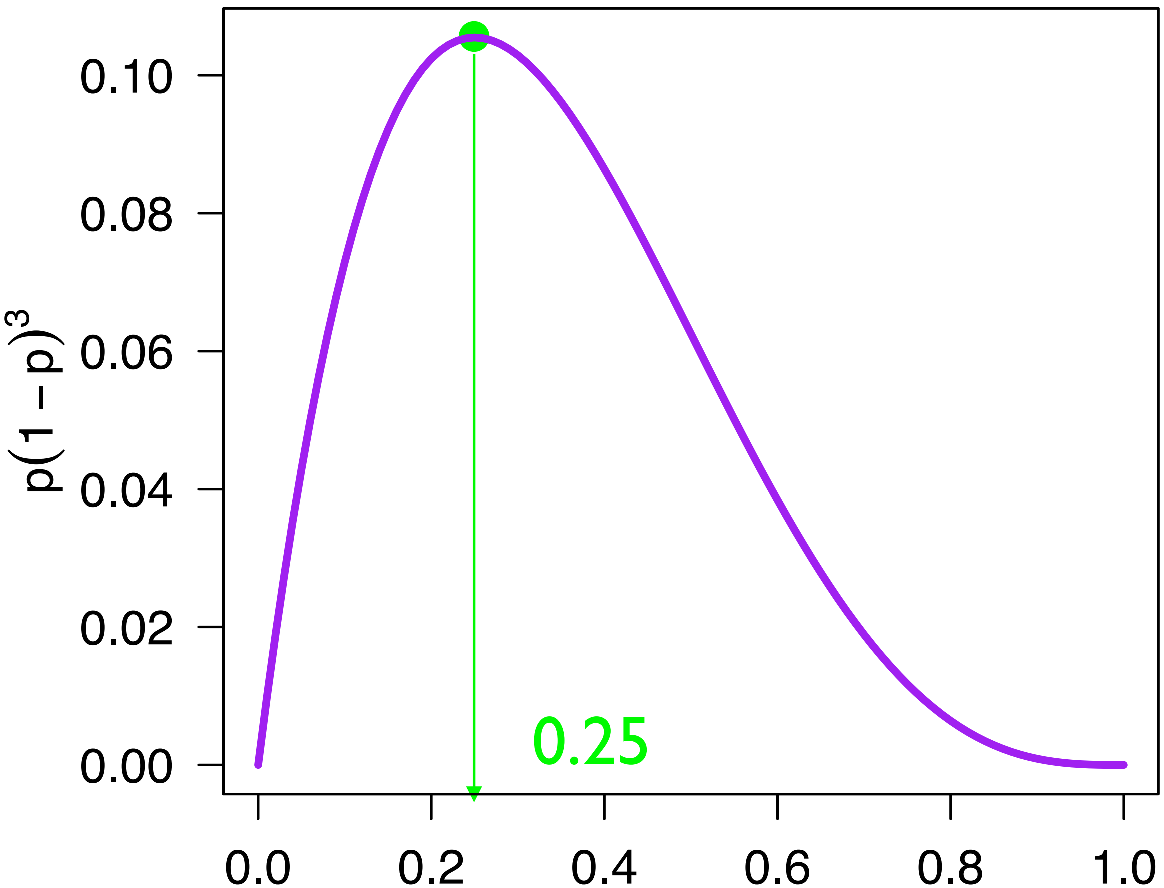

*(repeat slide from lecture 3)*

# Maximum likelihood estimation

$$\mathrm{Lik}(\boldsymbol{\theta};\boldsymbol{y}) \equiv P(\boldsymbol{y}|\boldsymbol{\theta}) \qquad \hat{\boldsymbol{\theta}}_{MLE} \stackrel{\mathrm{def}}{=} \arg\max_{\boldsymbol{\theta}} \mathrm{Lik}(\boldsymbol{\theta};\boldsymbol{y})$$

| ***i*** | $y_i$ |
|---|---|
| 1 | T |
| 2 | T |
| 3 | H |
| 4 | T |

- *p* refers to the value of P(coin toss$_i$ = Heads)
- Likelihood for the following dataset




This is choosing the *maximum likelihood estimate* (**MLE**)

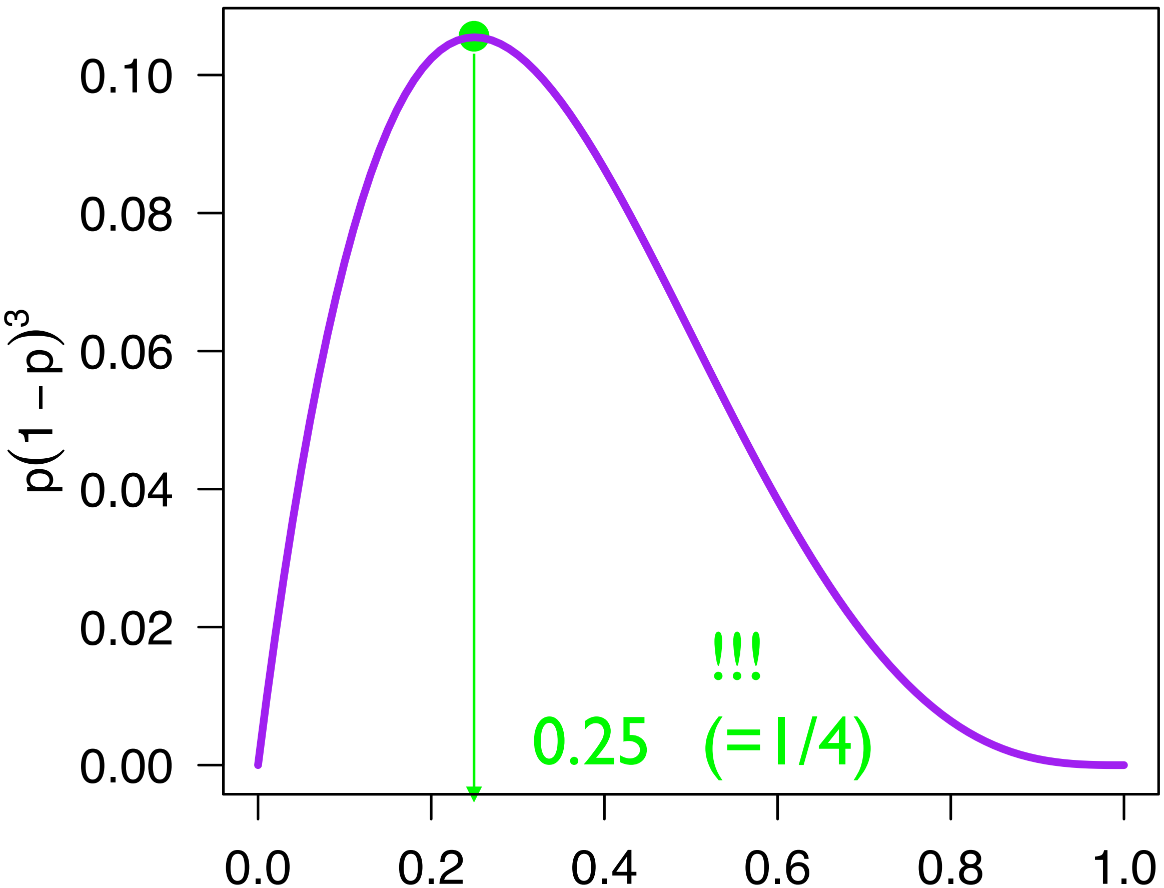
*(repeat slide from lecture 3)*

# Maximum likelihood estimation

$$\text{Lik}(\boldsymbol{\theta};\boldsymbol{y}) \equiv P(\boldsymbol{y}|\boldsymbol{\theta}) \qquad \hat{\boldsymbol{\theta}}_{MLE} \stackrel{\text{def}}{=} \arg\max_{\boldsymbol{\theta}} \text{Lik}(\boldsymbol{\theta};\boldsymbol{y})$$

| $i$ | $y_i$ |
|---|---|
| 1 | T |
| 2 | T |
| 3 | H |
| 4 | T |

- *p* refers to the value of P(coin toss$_i$ = Heads)
- Likelihood for the following dataset




This is choosing the *maximum likelihood estimate* (**MLE**)

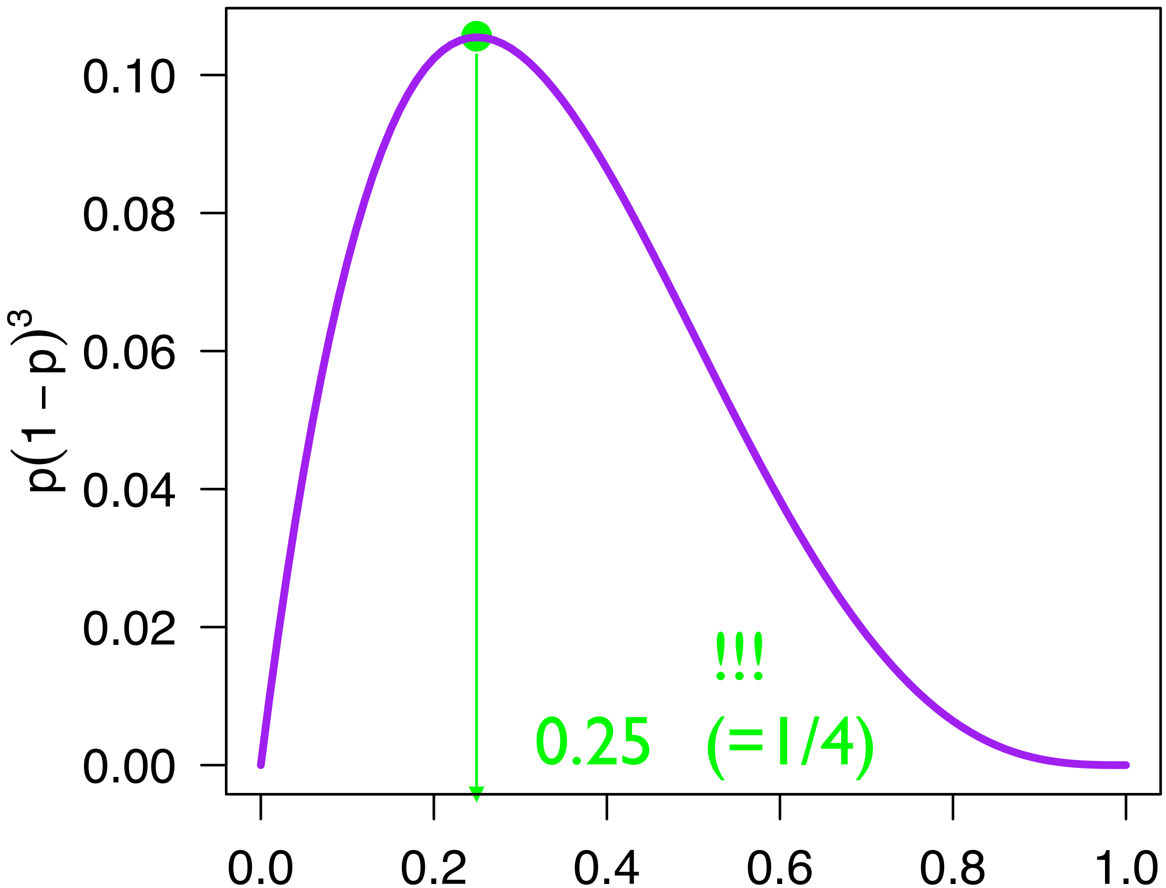
*(repeat slide from lecture 3)*

# Maximum likelihood estimation

$$\mathrm{Lik}(\boldsymbol{\theta}; \boldsymbol{y}) \equiv P(\boldsymbol{y}|\boldsymbol{\theta}) \qquad \hat{\boldsymbol{\theta}}_{MLE} \stackrel{\text{def}}{=} \arg\max_{\boldsymbol{\theta}} \mathrm{Lik}(\boldsymbol{\theta}; \boldsymbol{y})$$

| $i$ | $y_i$ |
|---|---|
| 1 | T |
| 2 | T |
| 3 | H |
| 4 | T |

- $p$ refers to the value of P(coin toss$_i$ = Heads)
- Likelihood for the following dataset




This is choosing the *maximum likelihood estimate* (**MLE**)

The **MLE** also turns out to be the *relative frequency estimate* (RFE)

*(repeat slide from lecture 3)*

# Introductory causal inference

- You have probably had previous exposure to both probability and statistics
- You are less likely to have had exposure to **causal inference**
- Causal inference uses probability and statistics, but it is something separate from the traditional construal of those two fields
- You can think of causal inference as being a framework extending more traditional statistics by:
  - Adding new probability-based mathematical constructs; and,
  - Developing a set of practice for statistical inference based on those constructs
- Two causal inference frameworks:
  - The **potential outcomes** framework
  - The **causal graphical models** framework

# The potential-outcomes framework

- In epidemiology and many other areas of statistics, causal inference was developed out of the idea of **potential outcomes** (Neyman 1923, Rubin 1974)
- Consider an outcome, $Y$, and a potential **treatment** $A$
- **Example:**

$Y$: an individual survives to the end of the year (0: no, 1: yes)

$A$: an individual with heart disease receives a heart transplant (0: no, 1: yes)

# Potential-outcome random variables

# Potential-outcome random variables

- Suppose that $A$ is discrete; for this case, $A \in \{0,1\}$

# Potential-outcome random variables

- Suppose that $A$ is discrete; for this case, $A \in \{0,1\}$
- The **potential outcomes**, or **counterfactual outcomes**, are random variables for $Y$ for each potential value of $A$

| | |
|---|---|
| $Y^{a=0}$ | The value that $Y$ would take if $A$ were 0 |
| $Y^{a=1}$ | The value that $Y$ would take if $A$ were 1 |

# Potential-outcome random variables

- Suppose that $A$ is discrete; for this case, $A \in \{0,1\}$
- The **potential outcomes**, or **counterfactual outcomes**, are random variables for $Y$ for each potential value of $A$

| | |
|---|---|
| $Y^{a=0}$ | The value that $Y$ would take if $A$ were 0 |
| $Y^{a=1}$ | The value that $Y$ would take if $A$ were 1 |

- **Counterfactual risk** is the **expected value** of each counterfactual-outcome random variable:

$$E[Y^{a=0}] \qquad\qquad E[Y^{a=1}]$$

# Potential-outcome random variables

- Suppose that $A$ is discrete; for this case, $A \in \{0,1\}$
- The **potential outcomes**, or **counterfactual outcomes**, are random variables for $Y$ for each potential value of $A$

| | |
|---|---|
| $Y^{a=0}$ | The value that $Y$ would take if $A$ were 0 |
| $Y^{a=1}$ | The value that $Y$ would take if $A$ were 1 |

- **Counterfactual risk** is the **expected value** of each counterfactual-outcome random variable:

$$E[Y^{a=0}] \qquad\qquad E[Y^{a=1}]$$

- Expected value, or expectation, is defined as follows:

$$E[X] = \sum_x xP(X = x)$$

# Potential-outcome random variables

- Suppose that $A$ is discrete; for this case, $A \in \{0,1\}$
- The **potential outcomes**, or **counterfactual outcomes**, are random variables for $Y$ for each potential value of $A$

| | |
|---|---|
| $Y^{a=0}$ | The value that $Y$ would take if $A$ were 0 |
| $Y^{a=1}$ | The value that $Y$ would take if $A$ were 1 |

- **Counterfactual risk** is the **expected value** of each counterfactual-outcome random variable:

$$E[Y^{a=0}] \qquad\qquad E[Y^{a=1}]$$

- Expected value, or expectation, is defined as follows:

$$E[X] = \sum_x xP(X = x)$$

- So we are interested in (and likewise for $Y^{a=1}$):

$$E[Y^{a=0}] = \sum_y yP(Y^{a=0} = y) = 0 \times P(Y^{a=0} = 0) + 1 \times P(Y^{a=0} = 1) = P(Y^{a=0} = 1)$$

# Potential-outcome random variables

- Suppose that $A$ is discrete; for this case, $A \in \{0,1\}$
- The **potential outcomes**, or **counterfactual outcomes**, are random variables for $Y$ for each potential value of $A$

| | |
|---|---|
| $Y^{a=0}$ | The value that $Y$ would take if $A$ were 0 |
| $Y^{a=1}$ | The value that $Y$ would take if $A$ were 1 |

- **Counterfactual risk** is the **expected value** of each counterfactual-outcome random variable:

$$E[Y^{a=0}] \qquad\qquad E[Y^{a=1}]$$

- Expected value, or expectation, is defined as follows:

$$E[X] = \sum_x xP(X = x)$$

- So we are interested in (and likewise for $Y^{a=1}$):

$$E[Y^{a=0}] = \sum_y yP(Y^{a=0} = y) = 0 \times P(Y^{a=0} = 0) + 1 \times P(Y^{a=0} = 1) = \boxed{P(Y^{a=0} = 1)}$$

# Counterfactual data and causal effects

(Hernan & Robins, 2020, Table 1.1)

| | $Y^{a=0}$ | $Y^{a=1}$ |
|---|---|---|
| Rheia | 0 | 1 |
| Kronos | 1 | 0 |
| Demeter | 0 | 0 |
| Hades | 0 | 0 |
| Hestia | 0 | 0 |
| Poseidon | 1 | 0 |
| Hera | 0 | 0 |
| Zeus | 0 | 1 |
| Artemis | 1 | 1 |
| Apollo | 1 | 0 |
| Leto | 0 | 1 |
| Ares | 1 | 1 |
| Athena | 1 | 1 |
| Hephaestus | 0 | 1 |
| Aphrodite | 0 | 1 |
| Cyclope | 0 | 1 |
| Persephone | 1 | 1 |
| Hermes | 1 | 0 |
| Hebe | 1 | 0 |
| Dionysus | 1 | 0 |
| $P(Y^{a=*}) = 1$ | 0.5 | 0.5 |

# Counterfactual data and causal effects

- Suppose we knew **what would happen** for each individual in the population under each value of the treatment

(Hernan & Robins, 2020, Table 1.1)

| | $Y^{a=0}$ | $Y^{a=1}$ |
|---|---|---|
| Rheia | 0 | 1 |
| Kronos | 1 | 0 |
| Demeter | 0 | 0 |
| Hades | 0 | 0 |
| Hestia | 0 | 0 |
| Poseidon | 1 | 0 |
| Hera | 0 | 0 |
| Zeus | 0 | 1 |
| Artemis | 1 | 1 |
| Apollo | 1 | 0 |
| Leto | 0 | 1 |
| Ares | 1 | 1 |
| Athena | 1 | 1 |
| Hephaestus | 0 | 1 |
| Aphrodite | 0 | 1 |
| Cyclope | 0 | 1 |
| Persephone | 1 | 1 |
| Hermes | 1 | 0 |
| Hebe | 1 | 0 |
| Dionysus | 1 | 0 |
| $P(Y^{a=*}) = 1$ | 0.5 | 0.5 |

# Counterfactual data and causal effects

- Suppose we knew **what would happen** for each individual in the population under each value of the treatment
- Then we could compute the counterfactual risks:

$$E[Y^{a=0}] = 0.5 \qquad E[Y^{a=1}] = 0.5$$

(Hernan & Robins, 2020, Table 1.1)

| | $Y^{a=0}$ | $Y^{a=1}$ |
|---|---|---|
| Rheia | 0 | 1 |
| Kronos | 1 | 0 |
| Demeter | 0 | 0 |
| Hades | 0 | 0 |
| Hestia | 0 | 0 |
| Poseidon | 1 | 0 |
| Hera | 0 | 0 |
| Zeus | 0 | 1 |
| Artemis | 1 | 1 |
| Apollo | 1 | 0 |
| Leto | 0 | 1 |
| Ares | 1 | 1 |
| Athena | 1 | 1 |
| Hephaestus | 0 | 1 |
| Aphrodite | 0 | 1 |
| Cyclope | 0 | 1 |
| Persephone | 1 | 1 |
| Hermes | 1 | 0 |
| Hebe | 1 | 0 |
| Dionysus | 1 | 0 |
| $P(Y^{a=*}) = 1$ | 0.5 | 0.5 |

# Counterfactual data and causal effects

- Suppose we knew **what would happen** for each individual in the population under each value of the treatment
- Then we could compute the counterfactual risks:

$$E[Y^{a=0}] = 0.5 \qquad E[Y^{a=1}] = 0.5$$

- The **average causal effect** of treatment $A$ is defined as the difference of counterfactual risks:

$$E[Y^{a=1}] - E[Y^{a=0}] = 0$$

(Hernan & Robins, 2020, Table 1.1)

| | $Y^{a=0}$ | $Y^{a=1}$ |
|---|---|---|
| Rheia | 0 | 1 |
| Kronos | 1 | 0 |
| Demeter | 0 | 0 |
| Hades | 0 | 0 |
| Hestia | 0 | 0 |
| Poseidon | 1 | 0 |
| Hera | 0 | 0 |
| Zeus | 0 | 1 |
| Artemis | 1 | 1 |
| Apollo | 1 | 0 |
| Leto | 0 | 1 |
| Ares | 1 | 1 |
| Athena | 1 | 1 |
| Hephaestus | 0 | 1 |
| Aphrodite | 0 | 1 |
| Cyclope | 0 | 1 |
| Persephone | 1 | 1 |
| Hermes | 1 | 0 |
| Hebe | 1 | 0 |
| Dionysus | 1 | 0 |
| $P(Y^{a=*}) = 1$ | 0.5 | 0.5 |

# Counterfactual data and causal effects

- Suppose we knew **what would happen** for each individual in the population under each value of the treatment
- Then we could compute the counterfactual risks:

$$E[Y^{a=0}] = 0.5 \qquad E[Y^{a=1}] = 0.5$$

- The **average causal effect** of treatment $A$ is defined as the difference of counterfactual risks:

$$E[Y^{a=1}] - E[Y^{a=0}] = 0$$

- Here, treatment is **ineffective**

(Hernan & Robins, 2020, Table 1.1)

| | $Y^{a=0}$ | $Y^{a=1}$ |
|---|---|---|
| Rheia | 0 | 1 |
| Kronos | 1 | 0 |
| Demeter | 0 | 0 |
| Hades | 0 | 0 |
| Hestia | 0 | 0 |
| Poseidon | 1 | 0 |
| Hera | 0 | 0 |
| Zeus | 0 | 1 |
| Artemis | 1 | 1 |
| Apollo | 1 | 0 |
| Leto | 0 | 1 |
| Ares | 1 | 1 |
| Athena | 1 | 1 |
| Hephaestus | 0 | 1 |
| Aphrodite | 0 | 1 |
| Cyclope | 0 | 1 |
| Persephone | 1 | 1 |
| Hermes | 1 | 0 |
| Hebe | 1 | 0 |
| Dionysus | 1 | 0 |
| $P(Y^{a=*}) = 1$ | 0.5 | 0.5 |

# Estimating causal effects

Remember, $E[Y^{a=i}] = P(Y^{a=i} = 1)$

| | $L$ | $A$ | $Y$ | $Y^0$ | $Y^1$ |
|---|---|---|---|---|---|
| Rheia | 0 | 0 | 0 | 0 | ? |
| Kronos | 0 | 0 | 1 | 1 | ? |
| Demeter | 0 | 0 | 0 | 0 | ? |
| Hades | 0 | 0 | 0 | 0 | ? |
| Hestia | 0 | 1 | 0 | ? | 0 |
| Poseidon | 0 | 1 | 0 | ? | 0 |
| Hera | 0 | 1 | 0 | ? | 0 |
| Zeus | 0 | 1 | 1 | ? | 1 |
| Artemis | 1 | 0 | 1 | 1 | ? |
| Apollo | 1 | 0 | 1 | 1 | ? |
| Leto | 1 | 0 | 0 | 0 | ? |
| Ares | 1 | 1 | 1 | ? | 1 |
| Athena | 1 | 1 | 1 | ? | 1 |
| Hephaestus | 1 | 1 | 1 | ? | 1 |
| Aphrodite | 1 | 1 | 1 | ? | 1 |
| Polyphemus | 1 | 1 | 1 | ? | 1 |
| Persephone | 1 | 1 | 1 | ? | 1 |
| Hermes | 1 | 1 | 0 | ? | 0 |
| Hebe | 1 | 1 | 0 | ? | 0 |
| Dionysus | 1 | 1 | 0 | ? | 0 |

# Estimating causal effects

Remember, $E[Y^{a=i}] = P(Y^{a=i} = 1)$

| | $L$ | $A$ | $Y$ | $Y^0$ | $Y^1$ |
|---|---|---|---|---|---|
| Rheia | 0 | 0 | 0 | 0 | ? |
| Kronos | 0 | 0 | 1 | 1 | ? |
| Demeter | 0 | 0 | 0 | 0 | ? |
| Hades | 0 | 0 | 0 | 0 | ? |
| Hestia | 0 | 1 | 0 | ? | 0 |
| Poseidon | 0 | 1 | 0 | ? | 0 |
| Hera | 0 | 1 | 0 | ? | 0 |
| Zeus | 0 | 1 | 1 | ? | 1 |
| Artemis | 1 | 0 | 1 | 1 | ? |
| Apollo | 1 | 0 | 1 | 1 | ? |
| Leto | 1 | 0 | 0 | 0 | ? |
| Ares | 1 | 1 | 1 | ? | 1 |
| Athena | 1 | 1 | 1 | ? | 1 |
| Hephaestus | 1 | 1 | 1 | ? | 1 |
| Aphrodite | 1 | 1 | 1 | ? | 1 |
| Polyphemus | 1 | 1 | 1 | ? | 1 |
| Persephone | 1 | 1 | 1 | ? | 1 |
| Hermes | 1 | 1 | 0 | ? | 0 |
| Hebe | 1 | 1 | 0 | ? | 0 |
| Dionysus | 1 | 1 | 0 | ? | 0 |

- Naively, we might estimate the counterfactual risks $P(Y^{a=i} = 1)$ directly from observed $A$ and $Y$:

# Estimating causal effects

Remember, $E[Y^{a=i}] = P(Y^{a=i} = 1)$

| | $L$ | $A$ | $Y$ | $Y^0$ | $Y^1$ |
|---|---|---|---|---|---|
| Rheia | 0 | 0 | 0 | 0 | ? |
| Kronos | 0 | 0 | 1 | 1 | ? |
| Demeter | 0 | 0 | 0 | 0 | ? |
| Hades | 0 | 0 | 0 | 0 | ? |
| Hestia | 0 | 1 | 0 | ? | 0 |
| Poseidon | 0 | 1 | 0 | ? | 0 |
| Hera | 0 | 1 | 0 | ? | 0 |
| Zeus | 0 | 1 | 1 | ? | 1 |
| Artemis | 1 | 0 | 1 | 1 | ? |
| Apollo | 1 | 0 | 1 | 1 | ? |
| Leto | 1 | 0 | 0 | 0 | ? |
| Ares | 1 | 1 | 1 | ? | 1 |
| Athena | 1 | 1 | 1 | ? | 1 |
| Hephaestus | 1 | 1 | 1 | ? | 1 |
| Aphrodite | 1 | 1 | 1 | ? | 1 |
| Polyphemus | 1 | 1 | 1 | ? | 1 |
| Persephone | 1 | 1 | 1 | ? | 1 |
| Hermes | 1 | 1 | 0 | ? | 0 |
| Hebe | 1 | 1 | 0 | ? | 0 |
| Dionysus | 1 | 1 | 0 | ? | 0 |

- Naively, we might estimate the counterfactual risks $P(Y^{a=i} = 1)$ directly from observed $A$ and $Y$:

$$\hat{P}_{MLE}(Y = 1 \,|\, A = 0) = \frac{3}{7} \quad \hat{P}_{MLE}(Y = 1 \,|\, A = 1) = \frac{7}{13}$$

# Estimating causal effects

Remember, $E[Y^{a=i}] = P(Y^{a=i} = 1)$

| | $L$ | $A$ | $Y$ | $Y^0$ | $Y^1$ |
|---|---|---|---|---|---|
| Rheia | 0 | 0 | 0 | 0 | ? |
| Kronos | 0 | 0 | 1 | 1 | ? |
| Demeter | 0 | 0 | 0 | 0 | ? |
| Hades | 0 | 0 | 0 | 0 | ? |
| Hestia | 0 | 1 | 0 | ? | 0 |
| Poseidon | 0 | 1 | 0 | ? | 0 |
| Hera | 0 | 1 | 0 | ? | 0 |
| Zeus | 0 | 1 | 1 | ? | 1 |
| Artemis | 1 | 0 | 1 | 1 | ? |
| Apollo | 1 | 0 | 1 | 1 | ? |
| Leto | 1 | 0 | 0 | 0 | ? |
| Ares | 1 | 1 | 1 | ? | 1 |
| Athena | 1 | 1 | 1 | ? | 1 |
| Hephaestus | 1 | 1 | 1 | ? | 1 |
| Aphrodite | 1 | 1 | 1 | ? | 1 |
| Polyphemus | 1 | 1 | 1 | ? | 1 |
| Persephone | 1 | 1 | 1 | ? | 1 |
| Hermes | 1 | 1 | 0 | ? | 0 |
| Hebe | 1 | 1 | 0 | ? | 0 |
| Dionysus | 1 | 1 | 0 | ? | 0 |

- Naively, we might estimate the counterfactual risks $P(Y^{a=i} = 1)$ directly from observed $A$ and $Y$:

$$\hat{P}_{MLE}(Y = 1 \,|\, A = 0) = \frac{3}{7} \quad \hat{P}_{MLE}(Y = 1 \,|\, A = 1) = \frac{7}{13}$$

- But under what circumstances $\hat{P}_{MLE}(Y \,|\, A = i) = \hat{P}_{MLE}(Y^{a=i} = 1)$?

# Estimating causal effects

Remember, $E[Y^{a=i}] = P(Y^{a=i} = 1)$

| | $L$ | $A$ | $Y$ | $Y^0$ | $Y^1$ |
|---|---|---|---|---|---|
| Rheia | 0 | 0 | 0 | 0 | ? |
| Kronos | 0 | 0 | 1 | 1 | ? |
| Demeter | 0 | 0 | 0 | 0 | ? |
| Hades | 0 | 0 | 0 | 0 | ? |
| Hestia | 0 | 1 | 0 | ? | 0 |
| Poseidon | 0 | 1 | 0 | ? | 0 |
| Hera | 0 | 1 | 0 | ? | 0 |
| Zeus | 0 | 1 | 1 | ? | 1 |
| Artemis | 1 | 0 | 1 | 1 | ? |
| Apollo | 1 | 0 | 1 | 1 | ? |
| Leto | 1 | 0 | 0 | 0 | ? |
| Ares | 1 | 1 | 1 | ? | 1 |
| Athena | 1 | 1 | 1 | ? | 1 |
| Hephaestus | 1 | 1 | 1 | ? | 1 |
| Aphrodite | 1 | 1 | 1 | ? | 1 |
| Polyphemus | 1 | 1 | 1 | ? | 1 |
| Persephone | 1 | 1 | 1 | ? | 1 |
| Hermes | 1 | 1 | 0 | ? | 0 |
| Hebe | 1 | 1 | 0 | ? | 0 |
| Dionysus | 1 | 1 | 0 | ? | 0 |

- Naively, we might estimate the counterfactual risks $P(Y^{a=i} = 1)$ directly from observed $A$ and $Y$:

$$\hat{P}_{MLE}(Y = 1 \,|\, A = 0) = \frac{3}{7} \quad \hat{P}_{MLE}(Y = 1 \,|\, A = 1) = \frac{7}{13}$$

- But under what circumstances $\hat{P}_{MLE}(Y \,|\, A = i) = \hat{P}_{MLE}(Y^{a=i} = 1)$?
- The following is certainly true:

# Estimating causal effects

Remember, $E[Y^{a=i}] = P(Y^{a=i} = 1)$

| | $L$ | $A$ | $Y$ | $Y^0$ | $Y^1$ |
|---|---|---|---|---|---|
| Rheia | 0 | 0 | 0 | 0 | ? |
| Kronos | 0 | 0 | 1 | 1 | ? |
| Demeter | 0 | 0 | 0 | 0 | ? |
| Hades | 0 | 0 | 0 | 0 | ? |
| Hestia | 0 | 1 | 0 | ? | 0 |
| Poseidon | 0 | 1 | 0 | ? | 0 |
| Hera | 0 | 1 | 0 | ? | 0 |
| Zeus | 0 | 1 | 1 | ? | 1 |
| Artemis | 1 | 0 | 1 | 1 | ? |
| Apollo | 1 | 0 | 1 | 1 | ? |
| Leto | 1 | 0 | 0 | 0 | ? |
| Ares | 1 | 1 | 1 | ? | 1 |
| Athena | 1 | 1 | 1 | ? | 1 |
| Hephaestus | 1 | 1 | 1 | ? | 1 |
| Aphrodite | 1 | 1 | 1 | ? | 1 |
| Polyphemus | 1 | 1 | 1 | ? | 1 |
| Persephone | 1 | 1 | 1 | ? | 1 |
| Hermes | 1 | 1 | 0 | ? | 0 |
| Hebe | 1 | 1 | 0 | ? | 0 |
| Dionysus | 1 | 1 | 0 | ? | 0 |

- Naively, we might estimate the counterfactual risks $P(Y^{a=i} = 1)$ directly from observed $A$ and $Y$:

$$\hat{P}_{MLE}(Y = 1 \,|\, A = 0) = \frac{3}{7} \quad \hat{P}_{MLE}(Y = 1 \,|\, A = 1) = \frac{7}{13}$$

- But under what circumstances $\hat{P}_{MLE}(Y \,|\, A = i) = \hat{P}_{MLE}(Y^{a=i} = 1)$?
- The following is certainly true:

$$\hat{P}_{MLE}(Y = 1 \,|\, A = i) = \frac{\text{Count}(Y = 1 \wedge A = i)}{\text{Count}(A = i)}$$

# Estimating causal effects

Remember, $E[Y^{a=i}] = P(Y^{a=i} = 1)$

| | $L$ | $A$ | $Y$ | $Y^0$ | $Y^1$ |
|---|---|---|---|---|---|
| Rheia | 0 | 0 | 0 | 0 | ? |
| Kronos | 0 | 0 | 1 | 1 | ? |
| Demeter | 0 | 0 | 0 | 0 | ? |
| Hades | 0 | 0 | 0 | 0 | ? |
| Hestia | 0 | 1 | 0 | ? | 0 |
| Poseidon | 0 | 1 | 0 | ? | 0 |
| Hera | 0 | 1 | 0 | ? | 0 |
| Zeus | 0 | 1 | 1 | ? | 1 |
| Artemis | 1 | 0 | 1 | 1 | ? |
| Apollo | 1 | 0 | 1 | 1 | ? |
| Leto | 1 | 0 | 0 | 0 | ? |
| Ares | 1 | 1 | 1 | ? | 1 |
| Athena | 1 | 1 | 1 | ? | 1 |
| Hephaestus | 1 | 1 | 1 | ? | 1 |
| Aphrodite | 1 | 1 | 1 | ? | 1 |
| Polyphemus | 1 | 1 | 1 | ? | 1 |
| Persephone | 1 | 1 | 1 | ? | 1 |
| Hermes | 1 | 1 | 0 | ? | 0 |
| Hebe | 1 | 1 | 0 | ? | 0 |
| Dionysus | 1 | 1 | 0 | ? | 0 |

- Naively, we might estimate the counterfactual risks $P(Y^{a=i} = 1)$ directly from observed $A$ and $Y$:

$$\hat{P}_{MLE}(Y = 1 \,|\, A = 0) = \frac{3}{7} \quad \hat{P}_{MLE}(Y = 1 \,|\, A = 1) = \frac{7}{13}$$

- But under what circumstances $\hat{P}_{MLE}(Y \,|\, A = i) = \hat{P}_{MLE}(Y^{a=i} = 1)$?
- The following is certainly true:

$$\hat{P}_{MLE}(Y = 1 \,|\, A = i) = \frac{\text{Count}(Y = 1 \wedge A = i)}{\text{Count}(A = i)}$$

**Consistency:** when $A = i$, $Y = Y^{a=i}$

# Estimating causal effects

Remember, $E[Y^{a=i}] = P(Y^{a=i} = 1)$

| | $L$ | $A$ | $Y$ | $Y^0$ | $Y^1$ |
|---|---|---|---|---|---|
| Rheia | 0 | 0 | 0 | 0 | ? |
| Kronos | 0 | 0 | 1 | 1 | ? |
| Demeter | 0 | 0 | 0 | 0 | ? |
| Hades | 0 | 0 | 0 | 0 | ? |
| Hestia | 0 | 1 | 0 | ? | 0 |
| Poseidon | 0 | 1 | 0 | ? | 0 |
| Hera | 0 | 1 | 0 | ? | 0 |
| Zeus | 0 | 1 | 1 | ? | 1 |
| Artemis | 1 | 0 | 1 | 1 | ? |
| Apollo | 1 | 0 | 1 | 1 | ? |
| Leto | 1 | 0 | 0 | 0 | ? |
| Ares | 1 | 1 | 1 | ? | 1 |
| Athena | 1 | 1 | 1 | ? | 1 |
| Hephaestus | 1 | 1 | 1 | ? | 1 |
| Aphrodite | 1 | 1 | 1 | ? | 1 |
| Polyphemus | 1 | 1 | 1 | ? | 1 |
| Persephone | 1 | 1 | 1 | ? | 1 |
| Hermes | 1 | 1 | 0 | ? | 0 |
| Hebe | 1 | 1 | 0 | ? | 0 |
| Dionysus | 1 | 1 | 0 | ? | 0 |

- Naively, we might estimate the counterfactual risks $P(Y^{a=i} = 1)$ directly from observed $A$ and $Y$:

$$\hat{P}_{MLE}(Y = 1 \,|\, A = 0) = \frac{3}{7} \quad \hat{P}_{MLE}(Y = 1 \,|\, A = 1) = \frac{7}{13}$$

- But under what circumstances $\hat{P}_{MLE}(Y \,|\, A = i) = \hat{P}_{MLE}(Y^{a=i} = 1)$?
- The following is certainly true:

$$\hat{P}_{MLE}(Y = 1 \,|\, A = i) = \frac{\text{Count}(Y = 1 \wedge A = i)}{\text{Count}(A = i)}$$

**Consistency:** when $A = i,\ Y = Y^{a=i}$

$$= \frac{\text{Count}(Y^{a=1} = 1 \wedge A = i)}{\text{Count}(A = i)}$$

*Crucial step; make sure you understand it!*

# Estimating causal effects

Remember, $E[Y^{a=i}] = P(Y^{a=i} = 1)$

| | $L$ | $A$ | $Y$ | $Y^0$ | $Y^1$ |
|---|---|---|---|---|---|
| Rheia | 0 | 0 | 0 | 0 | ? |
| Kronos | 0 | 0 | 1 | 1 | ? |
| Demeter | 0 | 0 | 0 | 0 | ? |
| Hades | 0 | 0 | 0 | 0 | ? |
| Hestia | 0 | 1 | 0 | ? | 0 |
| Poseidon | 0 | 1 | 0 | ? | 0 |
| Hera | 0 | 1 | 0 | ? | 0 |
| Zeus | 0 | 1 | 1 | ? | 1 |
| Artemis | 1 | 0 | 1 | 1 | ? |
| Apollo | 1 | 0 | 1 | 1 | ? |
| Leto | 1 | 0 | 0 | 0 | ? |
| Ares | 1 | 1 | 1 | ? | 1 |
| Athena | 1 | 1 | 1 | ? | 1 |
| Hephaestus | 1 | 1 | 1 | ? | 1 |
| Aphrodite | 1 | 1 | 1 | ? | 1 |
| Polyphemus | 1 | 1 | 1 | ? | 1 |
| Persephone | 1 | 1 | 1 | ? | 1 |
| Hermes | 1 | 1 | 0 | ? | 0 |
| Hebe | 1 | 1 | 0 | ? | 0 |
| Dionysus | 1 | 1 | 0 | ? | 0 |

- Naively, we might estimate the counterfactual risks $P(Y^{a=i} = 1)$ directly from observed $A$ and $Y$:

$$\hat{P}_{MLE}(Y = 1 \mid A = 0) = \frac{3}{7} \quad \hat{P}_{MLE}(Y = 1 \mid A = 1) = \frac{7}{13}$$

- But under what circumstances $\hat{P}_{MLE}(Y \mid A = i) = \hat{P}_{MLE}(Y^{a=i} = 1)$?
- The following is certainly true:

$$\hat{P}_{MLE}(Y = 1 \mid A = i) = \frac{\text{Count}(Y = 1 \wedge A = i)}{\text{Count}(A = i)}$$

**Consistency:** when $A = i$, $Y = Y^{a=i}$

$$= \frac{\text{Count}(Y^{a=1} = 1 \wedge A = i)}{\text{Count}(A = i)}$$

*Crucial step; make sure you understand it!*

$$= \hat{P}_{MLE}(Y^{a=i} = 1 \mid A = i)$$

# Estimating causal effects

Remember, $E[Y^{a=i}] = P(Y^{a=i} = 1)$

| | $L$ | $A$ | $Y$ | $Y^0$ | $Y^1$ |
|---|---|---|---|---|---|
| Rheia | 0 | 0 | 0 | 0 | ? |
| Kronos | 0 | 0 | 1 | 1 | ? |
| Demeter | 0 | 0 | 0 | 0 | ? |
| Hades | 0 | 0 | 0 | 0 | ? |
| Hestia | 0 | 1 | 0 | ? | 0 |
| Poseidon | 0 | 1 | 0 | ? | 0 |
| Hera | 0 | 1 | 0 | ? | 0 |
| Zeus | 0 | 1 | 1 | ? | 1 |
| Artemis | 1 | 0 | 1 | 1 | ? |
| Apollo | 1 | 0 | 1 | 1 | ? |
| Leto | 1 | 0 | 0 | 0 | ? |
| Ares | 1 | 1 | 1 | ? | 1 |
| Athena | 1 | 1 | 1 | ? | 1 |
| Hephaestus | 1 | 1 | 1 | ? | 1 |
| Aphrodite | 1 | 1 | 1 | ? | 1 |
| Polyphemus | 1 | 1 | 1 | ? | 1 |
| Persephone | 1 | 1 | 1 | ? | 1 |
| Hermes | 1 | 1 | 0 | ? | 0 |
| Hebe | 1 | 1 | 0 | ? | 0 |
| Dionysus | 1 | 1 | 0 | ? | 0 |

- Naively, we might estimate the counterfactual risks $P(Y^{a=i} = 1)$ directly from observed $A$ and $Y$:

$$\hat{P}_{MLE}(Y = 1 \mid A = 0) = \frac{3}{7} \quad \hat{P}_{MLE}(Y = 1 \mid A = 1) = \frac{7}{13}$$

- But under what circumstances $\hat{P}_{MLE}(Y \mid A = i) = \hat{P}_{MLE}(Y^{a=i} = 1)$?
- The following is certainly true:

$$\hat{P}_{MLE}(Y = 1 \mid A = i) = \frac{\text{Count}(Y = 1 \wedge A = i)}{\text{Count}(A = i)}$$

**CONSISTENCY:** when $A = i,\ Y = Y^{a=i}$

$$= \frac{\text{Count}(Y^{a=1} = 1 \wedge A = i)}{\text{Count}(A = i)}$$

*Crucial step; make sure you understand it!*

$$= \hat{P}_{MLE}(Y^{a=i} = 1 \mid A = i)$$

- So, the following condition suffices:

$$P(Y^{a=i} = 1 \mid A = i) = P(Y^{a=i} = 1)$$

# Estimating causal effects

Remember, $E[Y^{a=i}] = P(Y^{a=i} = 1)$

| | $L$ | $A$ | $Y$ | $Y^0$ | $Y^1$ |
|---|---|---|---|---|---|
| Rheia | 0 | 0 | 0 | 0 | ? |
| Kronos | 0 | 0 | 1 | 1 | ? |
| Demeter | 0 | 0 | 0 | 0 | ? |
| Hades | 0 | 0 | 0 | 0 | ? |
| Hestia | 0 | 1 | 0 | ? | 0 |
| Poseidon | 0 | 1 | 0 | ? | 0 |
| Hera | 0 | 1 | 0 | ? | 0 |
| Zeus | 0 | 1 | 1 | ? | 1 |
| Artemis | 1 | 0 | 1 | 1 | ? |
| Apollo | 1 | 0 | 1 | 1 | ? |
| Leto | 1 | 0 | 0 | 0 | ? |
| Ares | 1 | 1 | 1 | ? | 1 |
| Athena | 1 | 1 | 1 | ? | 1 |
| Hephaestus | 1 | 1 | 1 | ? | 1 |
| Aphrodite | 1 | 1 | 1 | ? | 1 |
| Polyphemus | 1 | 1 | 1 | ? | 1 |
| Persephone | 1 | 1 | 1 | ? | 1 |
| Hermes | 1 | 1 | 0 | ? | 0 |
| Hebe | 1 | 1 | 0 | ? | 0 |
| Dionysus | 1 | 1 | 0 | ? | 0 |

- Naively, we might estimate the counterfactual risks $P(Y^{a=i} = 1)$ directly from observed $A$ and $Y$:

$$\hat{P}_{MLE}(Y = 1 \,|\, A = 0) = \frac{3}{7} \quad \hat{P}_{MLE}(Y = 1 \,|\, A = 1) = \frac{7}{13}$$

- But under what circumstances $\hat{P}_{MLE}(Y \,|\, A = i) = \hat{P}_{MLE}(Y^{a=i} = 1)$?
- The following is certainly true:

$$\hat{P}_{MLE}(Y = 1 \,|\, A = i) = \frac{\text{Count}(Y = 1 \wedge A = i)}{\text{Count}(A = i)}$$

**Consistency:** when $A = i,\, Y = Y^{a=i}$

$$= \frac{\text{Count}(Y^{a=1} = 1 \wedge A = i)}{\text{Count}(A = i)}$$

*Crucial step; make sure you understand it!*

$$= \hat{P}_{MLE}(Y^{a=i} = 1 \,|\, A = i)$$

- So, the following condition suffices:

$$P(Y^{a=i} = 1 \,|\, A = i) = P(Y^{a=i} = 1)$$

- This is called **Exchangeability**:

$$Y^a \perp A \,|\, \{\}$$

# Exchangeability and randomization

**Goal:** $\hat{P}(Y^a = 1)$

| | $Y^{a=0}$ | $Y^{a=1}$ |
|---|---|---|
| Rheia | 0 | 1 |
| Kronos | 1 | 0 |
| Demeter | 0 | 0 |
| Hades | 0 | 0 |
| Hestia | 0 | 0 |
| Poseidon | 1 | 0 |
| Hera | 0 | 0 |
| Zeus | 0 | 1 |
| Artemis | 1 | 1 |
| Apollo | 1 | 0 |
| Leto | 0 | 1 |
| Ares | 1 | 1 |
| Athena | 1 | 1 |
| Hephaestus | 0 | 1 |
| Aphrodite | 0 | 1 |
| Polyphemus | 0 | 1 |
| Persephone | 1 | 1 |
| Hermes | 1 | 0 |
| Hebe | 1 | 0 |
| Dionysus | 1 | 0 |

# Exchangeability and randomization

**Goal:** $\hat{P}(Y^a = 1)$

- Why is a randomized experiment so powerful?

| | $Y^{a=0}$ | $Y^{a=1}$ |
|---|---|---|
| Rheia | 0 | 1 |
| Kronos | 1 | 0 |
| Demeter | 0 | 0 |
| Hades | 0 | 0 |
| Hestia | 0 | 0 |
| Poseidon | 1 | 0 |
| Hera | 0 | 0 |
| Zeus | 0 | 1 |
| Artemis | 1 | 1 |
| Apollo | 1 | 0 |
| Leto | 0 | 1 |
| Ares | 1 | 1 |
| Athena | 1 | 1 |
| Hephaestus | 0 | 1 |
| Aphrodite | 0 | 1 |
| Polyphemus | 0 | 1 |
| Persephone | 1 | 1 |
| Hermes | 1 | 0 |
| Hebe | 1 | 0 |
| Dionysus | 1 | 0 |

# Exchangeability and randomization

**Goal:** $\hat{P}(Y^a = 1)$

- Why is a randomized experiment so powerful?
- Recap of exchangeability criterion:

$$Y^a \perp A \mid \{\}$$

| | $Y^{a=0}$ | $Y^{a=1}$ |
|---|---|---|
| Rheia | 0 | 1 |
| Kronos | 1 | 0 |
| Demeter | 0 | 0 |
| Hades | 0 | 0 |
| Hestia | 0 | 0 |
| Poseidon | 1 | 0 |
| Hera | 0 | 0 |
| Zeus | 0 | 1 |
| Artemis | 1 | 1 |
| Apollo | 1 | 0 |
| Leto | 0 | 1 |
| Ares | 1 | 1 |
| Athena | 1 | 1 |
| Hephaestus | 0 | 1 |
| Aphrodite | 0 | 1 |
| Polyphemus | 0 | 1 |
| Persephone | 1 | 1 |
| Hermes | 1 | 0 |
| Hebe | 1 | 0 |
| Dionysus | 1 | 0 |

# Exchangeability and randomization

**Goal:** $\hat{P}(Y^a = 1)$

- Why is a randomized experiment so powerful?
- Recap of exchangeability criterion:

$$Y^a \perp A \mid \{\}$$

- If we ourselves determine $A$ in a way that is *truly blind to* $Y^a$, it **imposes** exchangeability!

| | $Y^{a=0}$ | $Y^{a=1}$ |
|---|---|---|
| Rheia | 0 | 1 |
| Kronos | 1 | 0 |
| Demeter | 0 | 0 |
| Hades | 0 | 0 |
| Hestia | 0 | 0 |
| Poseidon | 1 | 0 |
| Hera | 0 | 0 |
| Zeus | 0 | 1 |
| Artemis | 1 | 1 |
| Apollo | 1 | 0 |
| Leto | 0 | 1 |
| Ares | 1 | 1 |
| Athena | 1 | 1 |
| Hephaestus | 0 | 1 |
| Aphrodite | 0 | 1 |
| Polyphemus | 0 | 1 |
| Persephone | 1 | 1 |
| Hermes | 1 | 0 |
| Hebe | 1 | 0 |
| Dionysus | 1 | 0 |

# Exchangeability and randomization

**Goal:** $\hat{P}(Y^a = 1)$

- Why is a randomized experiment so powerful?
- Recap of exchangeability criterion:

$$Y^a \perp A \,|\, \{\}$$

- If we ourselves determine $A$ in a way that is *truly blind to* $Y^a$, it **imposes** exchangeability!

| |
|---|
| Rheia |
| Kronos |
| Demeter |
| Hades |
| Hestia |
| Poseidon |
| Hera |
| Zeus |
| Artemis |
| Apollo |
| Leto |
| Ares |
| Athena |
| Hephaestus |
| Aphrodite |
| Polyphemus |
| Persephone |
| Hermes |
| Hebe |
| Dionysus |

# Exchangeability and randomization

**Goal:** $\hat{P}(Y^a = 1)$

- Why is a randomized experiment so powerful?
- Recap of exchangeability criterion:

$$Y^a \perp A \mid \{\}$$

- If we ourselves determine $A$ in a way that is *truly blind to* $Y^a$, it **imposes** exchangeability!

| | $A$ |
|---|---|
| Rheia | 0 |
| Kronos | 0 |
| Demeter | 0 |
| Hades | 0 |
| Hestia | 1 |
| Poseidon | 1 |
| Hera | 1 |
| Zeus | 1 |
| Artemis | 0 |
| Apollo | 0 |
| Leto | 0 |
| Ares | 1 |
| Athena | 1 |
| Hephaestus | 1 |
| Aphrodite | 1 |
| Polyphemus | 1 |
| Persephone | 1 |
| Hermes | 1 |
| Hebe | 1 |
| Dionysus | 1 |

# Exchangeability and randomization

**Goal:** $\hat{P}(Y^a = 1)$

- Why is a randomized experiment so powerful?
- Recap of exchangeability criterion:

$$Y^a \perp A \mid \{\}$$

- If we ourselves determine $A$ in a way that is *truly blind to* $Y^a$, it **imposes** exchangeability!

| | $A$ | $Y$ |
|---|---|---|
| Rheia | 0 | 0 |
| Kronos | 0 | 1 |
| Demeter | 0 | 0 |
| Hades | 0 | 0 |
| Hestia | 1 | 0 |
| Poseidon | 1 | 0 |
| Hera | 1 | 0 |
| Zeus | 1 | 1 |
| Artemis | 0 | 1 |
| Apollo | 0 | 1 |
| Leto | 0 | 0 |
| Ares | 1 | 1 |
| Athena | 1 | 1 |
| Hephaestus | 1 | 1 |
| Aphrodite | 1 | 1 |
| Polyphemus | 1 | 1 |
| Persephone | 1 | 1 |
| Hermes | 1 | 0 |
| Hebe | 1 | 0 |
| Dionysus | 1 | 0 |

# Exchangeability and randomization

**Goal: $\hat{P}(Y^a = 1)$**

- Why is a randomized experiment so powerful?
- Recap of exchangeability criterion:

$$Y^a \perp A \,|\, \{\}$$

- If we ourselves determine $A$ in a way that is *truly blind to* $Y^a$, it **imposes** exchangeability!

| | $Y^{a=0}$ | $Y^{a=1}$ | $A$ | $Y$ |
|---|---|---|---|---|
| Rheia | 0 | 1 | 0 | 0 |
| Kronos | 1 | 0 | 0 | 1 |
| Demeter | 0 | 0 | 0 | 0 |
| Hades | 0 | 0 | 0 | 0 |
| Hestia | 0 | 0 | 1 | 0 |
| Poseidon | 1 | 0 | 1 | 0 |
| Hera | 0 | 0 | 1 | 0 |
| Zeus | 0 | 1 | 1 | 1 |
| Artemis | 1 | 1 | 0 | 1 |
| Apollo | 1 | 0 | 0 | 1 |
| Leto | 0 | 1 | 0 | 0 |
| Ares | 1 | 1 | 1 | 1 |
| Athena | 1 | 1 | 1 | 1 |
| Hephaestus | 0 | 1 | 1 | 1 |
| Aphrodite | 0 | 1 | 1 | 1 |
| Polyphemus | 0 | 1 | 1 | 1 |
| Persephone | 1 | 1 | 1 | 1 |
| Hermes | 1 | 0 | 1 | 0 |
| Hebe | 1 | 0 | 1 | 0 |
| Dionysus | 1 | 0 | 1 | 0 |

# Exchangeability and randomization

**Goal:** $\hat{P}(Y^a = 1)$

- Why is a randomized experiment so powerful?
- Recap of exchangeability criterion:

$$Y^a \perp A \,|\, \{\,\}$$

- If we ourselves determine $A$ in a way that is *truly blind to* $Y^a$, it **imposes** exchangeability!
- We can now go ahead and estimate

$$\hat{P}(Y^{a=i} = 1) = \hat{P}(Y = 1 \,|\, A = i)$$

| | $Y^{a=0}$ | $Y^{a=1}$ | $A$ | $Y$ |
|---|---|---|---|---|
| Rheia | 0 | 1 | 0 | 0 |
| Kronos | 1 | 0 | 0 | 1 |
| Demeter | 0 | 0 | 0 | 0 |
| Hades | 0 | 0 | 0 | 0 |
| Hestia | 0 | 0 | 1 | 0 |
| Poseidon | 1 | 0 | 1 | 0 |
| Hera | 0 | 0 | 1 | 0 |
| Zeus | 0 | 1 | 1 | 1 |
| Artemis | 1 | 1 | 0 | 1 |
| Apollo | 1 | 0 | 0 | 1 |
| Leto | 0 | 1 | 0 | 0 |
| Ares | 1 | 1 | 1 | 1 |
| Athena | 1 | 1 | 1 | 1 |
| Hephaestus | 0 | 1 | 1 | 1 |
| Aphrodite | 0 | 1 | 1 | 1 |
| Polyphemus | 0 | 1 | 1 | 1 |
| Persephone | 1 | 1 | 1 | 1 |
| Hermes | 1 | 0 | 1 | 0 |
| Hebe | 1 | 0 | 1 | 0 |
| Dionysus | 1 | 0 | 1 | 0 |

# Exchangeability and randomization

**Goal:** $\hat{P}(Y^a = 1)$

- Why is a randomized experiment so powerful?
- Recap of exchangeability criterion:

$$Y^a \perp A \,|\, \{\}$$

- If we ourselves determine $A$ in a way that is *truly blind to* $Y^a$, it **imposes** exchangeability!
- We can now go ahead and estimate

$$\hat{P}(Y^{a=i} = 1) = \hat{P}(Y = 1 \,|\, A = i)$$

- Hooray!!!

| | $Y^{a=0}$ | $Y^{a=1}$ | $A$ | $Y$ |
|---|---|---|---|---|
| Rheia | 0 | 1 | 0 | 0 |
| Kronos | 1 | 0 | 0 | 1 |
| Demeter | 0 | 0 | 0 | 0 |
| Hades | 0 | 0 | 0 | 0 |
| Hestia | 0 | 0 | 1 | 0 |
| Poseidon | 1 | 0 | 1 | 0 |
| Hera | 0 | 0 | 1 | 0 |
| Zeus | 0 | 1 | 1 | 1 |
| Artemis | 1 | 1 | 0 | 1 |
| Apollo | 1 | 0 | 0 | 1 |
| Leto | 0 | 1 | 0 | 0 |
| Ares | 1 | 1 | 1 | 1 |
| Athena | 1 | 1 | 1 | 1 |
| Hephaestus | 0 | 1 | 1 | 1 |
| Aphrodite | 0 | 1 | 1 | 1 |
| Polyphemus | 0 | 1 | 1 | 1 |
| Persephone | 1 | 1 | 1 | 1 |
| Hermes | 1 | 0 | 1 | 0 |
| Hebe | 1 | 0 | 1 | 0 |
| Dionysus | 1 | 0 | 1 | 0 |

# Does loss of randomization make things hopeless?

# Does loss of randomization make things hopeless?

- In the real world, many datasets are ***not*** randomized this way

# Does loss of randomization make things hopeless?

- In the real world, many datasets are ***not*** randomized this way
- **Example:** let's imagine some other variable that might affect whether treatment $A$ is applied; e.g., $L =$ whether the patient was in critical condition (1=yes, 0=no)

# Does loss of randomization make things hopeless?

- In the real world, many datasets are ***not*** randomized this way
- **Example:** let's imagine some other variable that might affect whether treatment $A$ is applied; e.g., $L =$ whether the patient was in critical condition (1=yes, 0=no)

| | $L$ |
|---|---|
| Rheia | 0 |
| Kronos | 0 |
| Demeter | 0 |
| Hades | 0 |
| Hestia | 0 |
| Poseidon | 0 |
| Hera | 0 |
| Zeus | 0 |
| Artemis | 1 |
| Apollo | 1 |
| Leto | 1 |
| Ares | 1 |
| Athena | 1 |
| Hephaestus | 1 |
| Aphrodite | 1 |
| Polyphemus | 1 |
| Persephone | 1 |
| Hermes | 1 |
| Hebe | 1 |
| Dionysus | 1 |

# Does loss of randomization make things hopeless?

- In the real world, many datasets are ***not*** randomized this way
- **Example:** let's imagine some other variable that might affect whether treatment $A$ is applied; e.g., $L =$ whether the patient was in critical condition (1=yes, 0=no)

| | $L$ | $A$ |
|---|---|---|
| Rheia | 0 | 0 |
| Kronos | 0 | 0 |
| Demeter | 0 | 0 |
| Hades | 0 | 0 |
| Hestia | 0 | 1 |
| Poseidon | 0 | 1 |
| Hera | 0 | 1 |
| Zeus | 0 | 1 |
| Artemis | 1 | 0 |
| Apollo | 1 | 0 |
| Leto | 1 | 0 |
| Ares | 1 | 1 |
| Athena | 1 | 1 |
| Hephaestus | 1 | 1 |
| Aphrodite | 1 | 1 |
| Polyphemus | 1 | 1 |
| Persephone | 1 | 1 |
| Hermes | 1 | 1 |
| Hebe | 1 | 1 |
| Dionysus | 1 | 1 |

# Does loss of randomization make things hopeless?

- In the real world, many datasets are ***not*** randomized this way
- **Example:** let's imagine some other variable that might affect whether treatment $A$ is applied; e.g., $L =$ whether the patient was in critical condition (1=yes, 0=no)

| | $L$ | $A$ |
|---|---|---|
| Rheia | 0 | 0 |
| Kronos | 0 | 0 |
| Demeter | 0 | 0 |
| Hades | 0 | 0 |
| Hestia | 0 | 1 |
| Poseidon | 0 | 1 |
| Hera | 0 | 1 |
| Zeus | 0 | 1 |
| Artemis | 1 | 0 |
| Apollo | 1 | 0 |
| Leto | 1 | 0 |
| Ares | 1 | 1 |
| Athena | 1 | 1 |
| Hephaestus | 1 | 1 |
| Aphrodite | 1 | 1 |
| Polyphemus | 1 | 1 |
| Persephone | 1 | 1 |
| Hermes | 1 | 1 |
| Hebe | 1 | 1 |
| Dionysus | 1 | 1 |

# Does loss of randomization make things hopeless?

- In the real world, many datasets are ***not*** randomized this way
- **Example:** let's imagine some other variable that might affect whether treatment $A$ is applied; e.g., $L =$ whether the patient was in critical condition (1=yes, 0=no)

| | $L$ | $A$ | $Y^{a=0}$ | $Y^{a=1}$ |
|---|---|---|---|---|
| Rheia | 0 | 0 | 0 | 1 |
| Kronos | 0 | 0 | 1 | 0 |
| Demeter | 0 | 0 | 0 | 0 |
| Hades | 0 | 0 | 0 | 0 |
| Hestia | 0 | 1 | 0 | 0 |
| Poseidon | 0 | 1 | 1 | 0 |
| Hera | 0 | 1 | 0 | 0 |
| Zeus | 0 | 1 | 0 | 1 |
| Artemis | 1 | 0 | 1 | 1 |
| Apollo | 1 | 0 | 1 | 0 |
| Leto | 1 | 0 | 0 | 1 |
| Ares | 1 | 1 | 1 | 1 |
| Athena | 1 | 1 | 1 | 1 |
| Hephaestus | 1 | 1 | 0 | 1 |
| Aphrodite | 1 | 1 | 0 | 1 |
| Polyphemus | 1 | 1 | 0 | 1 |
| Persephone | 1 | 1 | 1 | 1 |
| Hermes | 1 | 1 | 1 | 0 |
| Hebe | 1 | 1 | 1 | 0 |
| Dionysus | 1 | 1 | 1 | 0 |

# Does loss of randomization make things hopeless?

- In the real world, many datasets are ***not*** randomized this way
- **Example:** let's imagine some other variable that might affect whether treatment $A$ is applied; e.g., $L =$ whether the patient was in critical condition (1=yes, 0=no)
- In general, $L$ will be related to $Y^a$

| | $L$ | $A$ | $Y^{a=0}$ | $Y^{a=1}$ |
|---|---|---|---|---|
| Rheia | 0 | 0 | 0 | 1 |
| Kronos | 0 | 0 | 1 | 0 |
| Demeter | 0 | 0 | 0 | 0 |
| Hades | 0 | 0 | 0 | 0 |
| Hestia | 0 | 1 | 0 | 0 |
| Poseidon | 0 | 1 | 1 | 0 |
| Hera | 0 | 1 | 0 | 0 |
| Zeus | 0 | 1 | 0 | 1 |
| Artemis | 1 | 0 | 1 | 1 |
| Apollo | 1 | 0 | 1 | 0 |
| Leto | 1 | 0 | 0 | 1 |
| Ares | 1 | 1 | 1 | 1 |
| Athena | 1 | 1 | 1 | 1 |
| Hephaestus | 1 | 1 | 0 | 1 |
| Aphrodite | 1 | 1 | 0 | 1 |
| Polyphemus | 1 | 1 | 0 | 1 |
| Persephone | 1 | 1 | 1 | 1 |
| Hermes | 1 | 1 | 1 | 0 |
| Hebe | 1 | 1 | 1 | 0 |
| Dionysus | 1 | 1 | 1 | 0 |

# Does loss of randomization make things hopeless?

- In the real world, many datasets are ***not*** randomized this way
- **Example:** let's imagine some other variable that might affect whether treatment $A$ is applied; e.g., $L =$ whether the patient was in critical condition (1=yes, 0=no)
- In general, $L$ will be related to $Y^a$

| | $L$ | $A$ | $Y^{a=0}$ | $Y^{a=1}$ |
|---|---|---|---|---|
| Rheia | 0 | 0 | 0 | 1 |
| Kronos | 0 | 0 | 1 | 0 |
| Demeter | 0 | 0 | 0 | 0 |
| Hades | 0 | 0 | 0 | 0 |
| Hestia | 0 | 1 | 0 | 0 |
| Poseidon | 0 | 1 | 1 | 0 |
| Hera | 0 | 1 | 0 | 0 |
| Zeus | 0 | 1 | 0 | 1 |
| Artemis | 1 | 0 | 1 | 1 |
| Apollo | 1 | 0 | 1 | 0 |
| Leto | 1 | 0 | 0 | 1 |
| Ares | 1 | 1 | 1 | 1 |
| Athena | 1 | 1 | 1 | 1 |
| Hephaestus | 1 | 1 | 0 | 1 |
| Aphrodite | 1 | 1 | 0 | 1 |
| Polyphemus | 1 | 1 | 0 | 1 |
| Persephone | 1 | 1 | 1 | 1 |
| Hermes | 1 | 1 | 1 | 0 |
| Hebe | 1 | 1 | 1 | 0 |
| Dionysus | 1 | 1 | 1 | 0 |

# Does loss of randomization make things hopeless?

- In the real world, many datasets are ***not*** randomized this way
- **Example:** let's imagine some other variable that might affect whether treatment $A$ is applied; e.g., $L =$ whether the patient was in critical condition (1=yes, 0=no)
- In general, $L$ will be related to $Y^a$

| | $L$ | $A$ | $Y^{a=0}$ | $Y^{a=1}$ | $Y$ |
|---|---|---|---|---|---|
| Rheia | 0 | 0 | 0 | 1 | 0 |
| Kronos | 0 | 0 | 1 | 0 | 1 |
| Demeter | 0 | 0 | 0 | 0 | 0 |
| Hades | 0 | 0 | 0 | 0 | 0 |
| Hestia | 0 | 1 | 0 | 0 | 0 |
| Poseidon | 0 | 1 | 1 | 0 | 0 |
| Hera | 0 | 1 | 0 | 0 | 0 |
| Zeus | 0 | 1 | 0 | 1 | 1 |
| Artemis | 1 | 0 | 1 | 1 | 1 |
| Apollo | 1 | 0 | 1 | 0 | 1 |
| Leto | 1 | 0 | 0 | 1 | 0 |
| Ares | 1 | 1 | 1 | 1 | 1 |
| Athena | 1 | 1 | 1 | 1 | 1 |
| Hephaestus | 1 | 1 | 0 | 1 | 1 |
| Aphrodite | 1 | 1 | 0 | 1 | 1 |
| Polyphemus | 1 | 1 | 0 | 1 | 1 |
| Persephone | 1 | 1 | 1 | 1 | 1 |
| Hermes | 1 | 1 | 1 | 0 | 0 |
| Hebe | 1 | 1 | 1 | 0 | 0 |
| Dionysus | 1 | 1 | 1 | 0 | 0 |

# Does loss of randomization make things hopeless?

- In the real world, many datasets are ***not*** randomized this way
- **Example:** let's imagine some other variable that might affect whether treatment $A$ is applied; e.g., $L =$ whether the patient was in critical condition (1=yes, 0=no)
- In general, $L$ will be related to $Y^a$

| | $L$ | $A$ | $Y^{a=0}$ | $Y^{a=1}$ | $Y$ |
|---|---|---|---|---|---|
| Rheia | 0 | 0 | 0 | 1 | 0 |
| Kronos | 0 | 0 | 1 | 0 | 1 |
| Demeter | 0 | 0 | 0 | 0 | 0 |
| Hades | 0 | 0 | 0 | 0 | 0 |
| Hestia | 0 | 1 | 0 | 0 | 0 |
| Poseidon | 0 | 1 | 1 | 0 | 0 |
| Hera | 0 | 1 | 0 | 0 | 0 |
| Zeus | 0 | 1 | 0 | 1 | 1 |
| Artemis | 1 | 0 | 1 | 1 | 1 |
| Apollo | 1 | 0 | 1 | 0 | 1 |
| Leto | 1 | 0 | 0 | 1 | 0 |
| Ares | 1 | 1 | 1 | 1 | 1 |
| Athena | 1 | 1 | 1 | 1 | 1 |
| Hephaestus | 1 | 1 | 0 | 1 | 1 |
| Aphrodite | 1 | 1 | 0 | 1 | 1 |
| Polyphemus | 1 | 1 | 0 | 1 | 1 |
| Persephone | 1 | 1 | 1 | 1 | 1 |
| Hermes | 1 | 1 | 1 | 0 | 0 |
| Hebe | 1 | 1 | 1 | 0 | 0 |
| Dionysus | 1 | 1 | 1 | 0 | 0 |

# Does loss of randomization make things hopeless?

- In the real world, many datasets are ***not*** randomized this way
- **Example:** let's imagine some other variable that might affect whether treatment $A$ is applied; e.g., $L =$ whether the patient was in critical condition (1=yes, 0=no)
- In general, $L$ will be related to $Y^a$
  - E.g., in this example, patients in critical condition are surely more likely to die overall!

| | $L$ | $A$ | $Y^{a=0}$ | $Y^{a=1}$ | $Y$ |
|---|---|---|---|---|---|
| Rheia | 0 | 0 | 0 | 1 | 0 |
| Kronos | 0 | 0 | 1 | 0 | 1 |
| Demeter | 0 | 0 | 0 | 0 | 0 |
| Hades | 0 | 0 | 0 | 0 | 0 |
| Hestia | 0 | 1 | 0 | 0 | 0 |
| Poseidon | 0 | 1 | 1 | 0 | 0 |
| Hera | 0 | 1 | 0 | 0 | 0 |
| Zeus | 0 | 1 | 0 | 1 | 1 |
| Artemis | 1 | 0 | 1 | 1 | 1 |
| Apollo | 1 | 0 | 1 | 0 | 1 |
| Leto | 1 | 0 | 0 | 1 | 0 |
| Ares | 1 | 1 | 1 | 1 | 1 |
| Athena | 1 | 1 | 1 | 1 | 1 |
| Hephaestus | 1 | 1 | 0 | 1 | 1 |
| Aphrodite | 1 | 1 | 0 | 1 | 1 |
| Polyphemus | 1 | 1 | 0 | 1 | 1 |
| Persephone | 1 | 1 | 1 | 1 | 1 |
| Hermes | 1 | 1 | 1 | 0 | 0 |
| Hebe | 1 | 1 | 1 | 0 | 0 |
| Dionysus | 1 | 1 | 1 | 0 | 0 |

# Does loss of randomization make things hopeless?

- In the real world, many datasets are ***not*** randomized this way
- **Example:** let's imagine some other variable that might affect whether treatment $A$ is applied; e.g., $L =$ whether the patient was in critical condition (1=yes, 0=no)
- In general, $L$ will be related to $Y^a$
  - E.g., in this example, patients in critical condition are surely more likely to die overall!

$$A \perp Y^a \mid \{\}$$

| | $L$ | $A$ | $Y^{a=0}$ | $Y^{a=1}$ | $Y$ |
|---|---|---|---|---|---|
| Rheia | 0 | 0 | 0 | 1 | 0 |
| Kronos | 0 | 0 | 1 | 0 | 1 |
| Demeter | 0 | 0 | 0 | 0 | 0 |
| Hades | 0 | 0 | 0 | 0 | 0 |
| Hestia | 0 | 1 | 0 | 0 | 0 |
| Poseidon | 0 | 1 | 1 | 0 | 0 |
| Hera | 0 | 1 | 0 | 0 | 0 |
| Zeus | 0 | 1 | 0 | 1 | 1 |
| Artemis | 1 | 0 | 1 | 1 | 1 |
| Apollo | 1 | 0 | 1 | 0 | 1 |
| Leto | 1 | 0 | 0 | 1 | 0 |
| Ares | 1 | 1 | 1 | 1 | 1 |
| Athena | 1 | 1 | 1 | 1 | 1 |
| Hephaestus | 1 | 1 | 0 | 1 | 1 |
| Aphrodite | 1 | 1 | 0 | 1 | 1 |
| Polyphemus | 1 | 1 | 0 | 1 | 1 |
| Persephone | 1 | 1 | 1 | 1 | 1 |
| Hermes | 1 | 1 | 1 | 0 | 0 |
| Hebe | 1 | 1 | 1 | 0 | 0 |
| Dionysus | 1 | 1 | 1 | 0 | 0 |

# Does loss of randomization make things hopeless?

- In the real world, many datasets are ***not*** randomized this way
- **Example:** let's imagine some other variable that might affect whether treatment $A$ is applied; e.g., $L =$ whether the patient was in critical condition (1=yes, 0=no)
- In general, $L$ will be related to $Y^a$
    - E.g., in this example, patients in critical condition are surely more likely to die overall!

$$\cancel{A \perp\!\!\!\perp Y^a \mid \{\}}$$

| | $L$ | $A$ | $Y^{a=0}$ | $Y^{a=1}$ | $Y$ |
|---|---|---|---|---|---|
| Rheia | 0 | 0 | 0 | 1 | 0 |
| Kronos | 0 | 0 | 1 | 0 | 1 |
| Demeter | 0 | 0 | 0 | 0 | 0 |
| Hades | 0 | 0 | 0 | 0 | 0 |
| Hestia | 0 | 1 | 0 | 0 | 0 |
| Poseidon | 0 | 1 | 1 | 0 | 0 |
| Hera | 0 | 1 | 0 | 0 | 0 |
| Zeus | 0 | 1 | 0 | 1 | 1 |
| Artemis | 1 | 0 | 1 | 1 | 1 |
| Apollo | 1 | 0 | 1 | 0 | 1 |
| Leto | 1 | 0 | 0 | 1 | 0 |
| Ares | 1 | 1 | 1 | 1 | 1 |
| Athena | 1 | 1 | 1 | 1 | 1 |
| Hephaestus | 1 | 1 | 0 | 1 | 1 |
| Aphrodite | 1 | 1 | 0 | 1 | 1 |
| Polyphemus | 1 | 1 | 0 | 1 | 1 |
| Persephone | 1 | 1 | 1 | 1 | 1 |
| Hermes | 1 | 1 | 1 | 0 | 0 |
| Hebe | 1 | 1 | 1 | 0 | 0 |
| Dionysus | 1 | 1 | 1 | 0 | 0 |

# Conditional exchangeability

| | $L$ | $A$ | $Y^{a=0}$ | $Y^{a=1}$ | $Y$ |
|---|---|---|---|---|---|
| Rheia | 0 | 0 | 0 | 1 | 0 |
| Kronos | 0 | 0 | 1 | 0 | 1 |
| Demeter | 0 | 0 | 0 | 0 | 0 |
| Hades | 0 | 0 | 0 | 0 | 0 |
| Hestia | 0 | 1 | 0 | 0 | 0 |
| Poseidon | 0 | 1 | 1 | 0 | 0 |
| Hera | 0 | 1 | 0 | 0 | 0 |
| Zeus | 0 | 1 | 0 | 1 | 1 |
| Artemis | 1 | 0 | 1 | 1 | 1 |
| Apollo | 1 | 0 | 1 | 0 | 1 |
| Leto | 1 | 0 | 0 | 1 | 0 |
| Ares | 1 | 1 | 1 | 1 | 1 |
| Athena | 1 | 1 | 1 | 1 | 1 |
| Hephaestus | 1 | 1 | 0 | 1 | 1 |
| Aphrodite | 1 | 1 | 0 | 1 | 1 |
| Polyphemus | 1 | 1 | 0 | 1 | 1 |
| Persephone | 1 | 1 | 1 | 1 | 1 |
| Hermes | 1 | 1 | 1 | 0 | 0 |
| Hebe | 1 | 1 | 1 | 0 | 0 |
| Dionysus | 1 | 1 | 1 | 0 | 0 |

# Conditional exchangeability

- But now suppose we have observed (i.e., it's in our dataset) the factor $L$ that affected whether the treatment $A$ was applied

| | $L$ | $A$ | $Y^{a=0}$ | $Y^{a=1}$ | $Y$ |
|---|---|---|---|---|---|
| Rheia | 0 | 0 | 0 | 1 | 0 |
| Kronos | 0 | 0 | 1 | 0 | 1 |
| Demeter | 0 | 0 | 0 | 0 | 0 |
| Hades | 0 | 0 | 0 | 0 | 0 |
| Hestia | 0 | 1 | 0 | 0 | 0 |
| Poseidon | 0 | 1 | 1 | 0 | 0 |
| Hera | 0 | 1 | 0 | 0 | 0 |
| Zeus | 0 | 1 | 0 | 1 | 1 |
| Artemis | 1 | 0 | 1 | 1 | 1 |
| Apollo | 1 | 0 | 1 | 0 | 1 |
| Leto | 1 | 0 | 0 | 1 | 0 |
| Ares | 1 | 1 | 1 | 1 | 1 |
| Athena | 1 | 1 | 1 | 1 | 1 |
| Hephaestus | 1 | 1 | 0 | 1 | 1 |
| Aphrodite | 1 | 1 | 0 | 1 | 1 |
| Polyphemus | 1 | 1 | 0 | 1 | 1 |
| Persephone | 1 | 1 | 1 | 1 | 1 |
| Hermes | 1 | 1 | 1 | 0 | 0 |
| Hebe | 1 | 1 | 1 | 0 | 0 |
| Dionysus | 1 | 1 | 1 | 0 | 0 |

# Conditional exchangeability

- But now suppose we have observed (i.e., it's in our dataset) the factor $L$ that affected whether the treatment $A$ was applied
- If the following condition holds, it can help us estimate the counterfactual risks $P(Y^a)$:

$$A \perp Y^a \,|\, L$$

| | $L$ | $A$ | $Y^{a=0}$ | $Y^{a=1}$ | $Y$ |
|---|---|---|---|---|---|
| Rheia | 0 | 0 | 0 | 1 | 0 |
| Kronos | 0 | 0 | 1 | 0 | 1 |
| Demeter | 0 | 0 | 0 | 0 | 0 |
| Hades | 0 | 0 | 0 | 0 | 0 |
| Hestia | 0 | 1 | 0 | 0 | 0 |
| Poseidon | 0 | 1 | 1 | 0 | 0 |
| Hera | 0 | 1 | 0 | 0 | 0 |
| Zeus | 0 | 1 | 0 | 1 | 1 |
| Artemis | 1 | 0 | 1 | 1 | 1 |
| Apollo | 1 | 0 | 1 | 0 | 1 |
| Leto | 1 | 0 | 0 | 1 | 0 |
| Ares | 1 | 1 | 1 | 1 | 1 |
| Athena | 1 | 1 | 1 | 1 | 1 |
| Hephaestus | 1 | 1 | 0 | 1 | 1 |
| Aphrodite | 1 | 1 | 0 | 1 | 1 |
| Polyphemus | 1 | 1 | 0 | 1 | 1 |
| Persephone | 1 | 1 | 1 | 1 | 1 |
| Hermes | 1 | 1 | 1 | 0 | 0 |
| Hebe | 1 | 1 | 1 | 0 | 0 |
| Dionysus | 1 | 1 | 1 | 0 | 0 |

# Conditional exchangeability

- But now suppose we have observed (i.e., it's in our dataset) the factor $L$ that affected whether the treatment $A$ was applied
- If the following condition holds, it can help us estimate the counterfactual risks $P(Y^a)$:

$$A \perp Y^a \mid L$$

- That is, $L$ captures all the information available in $A$ that is relevant to all $Y^a$

| | $L$ | $A$ | $Y^{a=0}$ | $Y^{a=1}$ | $Y$ |
|---|---|---|---|---|---|
| Rheia | 0 | 0 | 0 | 1 | 0 |
| Kronos | 0 | 0 | 1 | 0 | 1 |
| Demeter | 0 | 0 | 0 | 0 | 0 |
| Hades | 0 | 0 | 0 | 0 | 0 |
| Hestia | 0 | 1 | 0 | 0 | 0 |
| Poseidon | 0 | 1 | 1 | 0 | 0 |
| Hera | 0 | 1 | 0 | 0 | 0 |
| Zeus | 0 | 1 | 0 | 1 | 1 |
| Artemis | 1 | 0 | 1 | 1 | 1 |
| Apollo | 1 | 0 | 1 | 0 | 1 |
| Leto | 1 | 0 | 0 | 1 | 0 |
| Ares | 1 | 1 | 1 | 1 | 1 |
| Athena | 1 | 1 | 1 | 1 | 1 |
| Hephaestus | 1 | 1 | 0 | 1 | 1 |
| Aphrodite | 1 | 1 | 0 | 1 | 1 |
| Polyphemus | 1 | 1 | 0 | 1 | 1 |
| Persephone | 1 | 1 | 1 | 1 | 1 |
| Hermes | 1 | 1 | 1 | 0 | 0 |
| Hebe | 1 | 1 | 1 | 0 | 0 |
| Dionysus | 1 | 1 | 1 | 0 | 0 |

# Conditional exchangeability

- But now suppose we have observed (i.e., it's in our dataset) the factor $L$ that affected whether the treatment $A$ was applied
- If the following condition holds, it can help us estimate the counterfactual risks $P(Y^a)$:

$$A \perp Y^a \mid L$$

- That is, $L$ captures all the information available in $A$ that is relevant to all $Y^a$
- This is called CONDITIONAL EXCHANGEABILITY

| | $L$ | $A$ | $Y^{a=0}$ | $Y^{a=1}$ | $Y$ |
|---|---|---|---|---|---|
| Rheia | 0 | 0 | 0 | 1 | 0 |
| Kronos | 0 | 0 | 1 | 0 | 1 |
| Demeter | 0 | 0 | 0 | 0 | 0 |
| Hades | 0 | 0 | 0 | 0 | 0 |
| Hestia | 0 | 1 | 0 | 0 | 0 |
| Poseidon | 0 | 1 | 1 | 0 | 0 |
| Hera | 0 | 1 | 0 | 0 | 0 |
| Zeus | 0 | 1 | 0 | 1 | 1 |
| Artemis | 1 | 0 | 1 | 1 | 1 |
| Apollo | 1 | 0 | 1 | 0 | 1 |
| Leto | 1 | 0 | 0 | 1 | 0 |
| Ares | 1 | 1 | 1 | 1 | 1 |
| Athena | 1 | 1 | 1 | 1 | 1 |
| Hephaestus | 1 | 1 | 0 | 1 | 1 |
| Aphrodite | 1 | 1 | 0 | 1 | 1 |
| Polyphemus | 1 | 1 | 0 | 1 | 1 |
| Persephone | 1 | 1 | 1 | 1 | 1 |
| Hermes | 1 | 1 | 1 | 0 | 0 |
| Hebe | 1 | 1 | 1 | 0 | 0 |
| Dionysus | 1 | 1 | 1 | 0 | 0 |

# Conditional exchangeability

| | $L$ | $A$ | $Y$ |
|---|---|---|---|
| Rheia | 0 | 0 | 0 |
| Kronos | 0 | 0 | 1 |
| Demeter | 0 | 0 | 0 |
| Hades | 0 | 0 | 0 |
| Hestia | 0 | 1 | 0 |
| Poseidon | 0 | 1 | 0 |
| Hera | 0 | 1 | 0 |
| Zeus | 0 | 1 | 1 |
| Artemis | 1 | 0 | 1 |
| Apollo | 1 | 0 | 1 |
| Leto | 1 | 0 | 0 |
| Ares | 1 | 1 | 1 |
| Athena | 1 | 1 | 1 |
| Hephaestus | 1 | 1 | 1 |
| Aphrodite | 1 | 1 | 1 |
| Polyphemus | 1 | 1 | 1 |
| Persephone | 1 | 1 | 1 |
| Hermes | 1 | 1 | 0 |
| Hebe | 1 | 1 | 0 |
| Dionysus | 1 | 1 | 0 |

- Suppose we observe some other potentially relevant factor in our data, such as $L =$ whether the patient was in critical condition (1=yes, 0=no)
- CONDITIONAL EXCHANGEABILITY holds iff the counterfactual outcomes are conditionally independent of treatment, GIVEN this other relevant factor $L$:

$$A \perp Y^a \mid L$$

- Conditional exchangeability intuitively seems to give us an inroad into estimating the average causal effect of $A$ on $Y$

# IDENTIFIABILITY of causal effects

# IDENTIFIABILITY of causal effects

- IDENTIFIABILITY means, our assumptions allow the causal effect we are interested in to be ***uniquely estimated*** from the available data (set of observed/measured variables)

# IDENTIFIABILITY of causal effects

- IDENTIFIABILITY means, our assumptions allow the causal effect we are interested in to be ***uniquely estimated*** from the available data (set of observed/measured variables)
  - "Uniquely estimate": if we had an arbitrary large quantity of data, we could estimate the causal effect with arbitrarily high accuracy and precision

# IDENTIFIABILITY of causal effects

- IDENTIFIABILITY means, our assumptions allow the causal effect we are interested in to be ***uniquely estimated*** from the available data (set of observed/measured variables)
  - "Uniquely estimate": if we had an arbitrary large quantity of data, we could estimate the causal effect with arbitrarily high accuracy and precision
- Simple case of unidentifiability: Hernan & Robins's heart transplant example, if $L$ (severity of disease) affects probability of a heart transplant and we don't measure it

# IDENTIFIABILITY of causal effects

- IDENTIFIABILITY means, our assumptions allow the causal effect we are interested in to be ***uniquely estimated*** from the available data (set of observed/measured variables)
  - "Uniquely estimate": if we had an arbitrary large quantity of data, we could estimate the causal effect with arbitrarily high accuracy and precision
- Simple case of unidentifiability: Hernan & Robins's heart transplant example, if $L$ (severity of disease) affects probability of a heart transplant and we don't measure it
  - Suppose that people with transplants have lower survival rates: $\hat{P}_{MLE}(Y|A=1) < \hat{P}_{MLE}(Y|A=0)$

# IDENTIFIABILITY of causal effects

- IDENTIFIABILITY means, our assumptions allow the causal effect we are interested in to be ***uniquely estimated*** from the available data (set of observed/measured variables)
  - "Uniquely estimate": if we had an arbitrary large quantity of data, we could estimate the causal effect with arbitrarily high accuracy and precision
- Simple case of unidentifiability: Hernan & Robins's heart transplant example, if $L$ (severity of disease) affects probability of a heart transplant and we don't measure it
  - Suppose that people with transplants have lower survival rates: $\hat{P}_{MLE}(Y|A=1) < \hat{P}_{MLE}(Y|A=0)$
    - Could be because heart transplants are dangerous

# IDENTIFIABILITY of causal effects

- IDENTIFIABILITY means, our assumptions allow the causal effect we are interested in to be ***uniquely estimated*** from the available data (set of observed/measured variables)
  - "Uniquely estimate": if we had an arbitrary large quantity of data, we could estimate the causal effect with arbitrarily high accuracy and precision
- Simple case of unidentifiability: Hernan & Robins's heart transplant example, if $L$ (severity of disease) affects probability of a heart transplant and we don't measure it
  - Suppose that people with transplants have lower survival rates: $\hat{P}_{MLE}(Y|A=1) < \hat{P}_{MLE}(Y|A=0)$
    - Could be because heart transplants are dangerous
    - Or: sicker people are more likely to get transplants!

# IDENTIFIABILITY of causal effects

- IDENTIFIABILITY means, our assumptions allow the causal effect we are interested in to be ***uniquely estimated*** from the available data (set of observed/measured variables)
  - "Uniquely estimate": if we had an arbitrary large quantity of data, we could estimate the causal effect with arbitrarily high accuracy and precision
- Simple case of unidentifiability: Hernan & Robins's heart transplant example, if $L$ (severity of disease) affects probability of a heart transplant and we don't measure it
  - Suppose that people with transplants have lower survival rates: $\hat{P}_{MLE}(Y|A=1) < \hat{P}_{MLE}(Y|A=0)$
    - Could be because heart transplants are dangerous
    - Or: sicker people are more likely to get transplants!

***Not an issue of sample size—no amount of data would help!***

# The three criteria for identifiability

- **Consistency**: $Y = Y^{a=i}$ whenever $A = i$
  - Consequence: different individuals' outcomes don't affect each other
  - Consequence: there can be no "multiple versions" of the same treatment $A$ in terms of their influence on $Y$
- **Exchangeability**: for all $i$, $Y^{a=i} \perp A \,|\, Z$ for some set of observed variables $Z$
  - Consequence: there can be no "hidden common causes" or "hidden mediators" of $A$ and $Y^{a=i}$
- **Positivity**: for all $i$ and all values of $Z$, $P(A = i \,|\, Z) > 0$
  - e.g., in our example, it can't be the case that individuals with heart disease are *always* given transplants

# What to do with conditional exchangeability

- One approach is to

| | $L$ | $A$ | $Y$ |
|---|---|---|---|
| Rheia | 0 | 0 | 0 |
| Kronos | 0 | 0 | 1 |
| Demeter | 0 | 0 | 0 |
| Hades | 0 | 0 | 0 |
| Hestia | 0 | 1 | 0 |
| Poseidon | 0 | 1 | 0 |
| Hera | 0 | 1 | 0 |
| Zeus | 0 | 1 | 1 |
| Artemis | 1 | 0 | 1 |
| Apollo | 1 | 0 | 1 |
| Leto | 1 | 0 | 0 |
| Ares | 1 | 1 | 1 |
| Athena | 1 | 1 | 1 |
| Hephaestus | 1 | 1 | 1 |
| Aphrodite | 1 | 1 | 1 |
| Polyphemus | 1 | 1 | 1 |
| Persephone | 1 | 1 | 1 |
| Hermes | 1 | 1 | 0 |
| Hebe | 1 | 1 | 0 |
| Dionysus | 1 | 1 | 0 |

# Why causal inference is challenging

(Hernan & Robins, 2020, Table 1.1)

| | $Y^{a=0}$ | $Y^{a=1}$ |
|---|---|---|
| Rheia | 0 | 1 |
| Kronos | 1 | 0 |
| Demeter | 0 | 0 |
| Hades | 0 | 0 |
| Hestia | 0 | 0 |
| Poseidon | 1 | 0 |
| Hera | 0 | 0 |
| Zeus | 0 | 1 |
| Artemis | 1 | 1 |
| Apollo | 1 | 0 |
| Leto | 0 | 1 |
| Ares | 1 | 1 |
| Athena | 1 | 1 |
| Hephaestus | 0 | 1 |
| Aphrodite | 0 | 1 |
| Cyclope | 0 | 1 |
| Persephone | 1 | 1 |
| Hermes | 1 | 0 |
| Hebe | 1 | 0 |
| Dionysus | 1 | 0 |

# Why causal inference is challenging

- But, we don't get to see individual-specific counterfactual RV values!

(Hernan & Robins, 2020, Table 1.1)

| | $Y^{a=0}$ | $Y^{a=1}$ |
|---|---|---|
| Rheia | 0 | 1 |
| Kronos | 1 | 0 |
| Demeter | 0 | 0 |
| Hades | 0 | 0 |
| Hestia | 0 | 0 |
| Poseidon | 1 | 0 |
| Hera | 0 | 0 |
| Zeus | 0 | 1 |
| Artemis | 1 | 1 |
| Apollo | 1 | 0 |
| Leto | 0 | 1 |
| Ares | 1 | 1 |
| Athena | 1 | 1 |
| Hephaestus | 0 | 1 |
| Aphrodite | 0 | 1 |
| Cyclope | 0 | 1 |
| Persephone | 1 | 1 |
| Hermes | 1 | 0 |
| Hebe | 1 | 0 |
| Dionysus | 1 | 0 |

# Why causal inference is challenging

- But, we don't get to see individual-specific counterfactual RV values!
- At most we see one **actual** outcome

(Hernan & Robins, 2020, Table 1.1)

| | $Y^{a=0}$ | $Y^{a=1}$ |
|---|---|---|
| Rheia | 0 | 1 |
| Kronos | 1 | 0 |
| Demeter | 0 | 0 |
| Hades | 0 | 0 |
| Hestia | 0 | 0 |
| Poseidon | 1 | 0 |
| Hera | 0 | 0 |
| Zeus | 0 | 1 |
| Artemis | 1 | 1 |
| Apollo | 1 | 0 |
| Leto | 0 | 1 |
| Ares | 1 | 1 |
| Athena | 1 | 1 |
| Hephaestus | 0 | 1 |
| Aphrodite | 0 | 1 |
| Cyclope | 0 | 1 |
| Persephone | 1 | 1 |
| Hermes | 1 | 0 |
| Hebe | 1 | 0 |
| Dionysus | 1 | 0 |

Typical observed $Y$

# Why causal inference is challenging

- But, we don't get to see individual-specific counterfactual RV values!
- At most we see one **actual** outcome
- Let us bundle up everything we don't know about the process by which actual $A$ and $Y$ were determined into the RV $U$

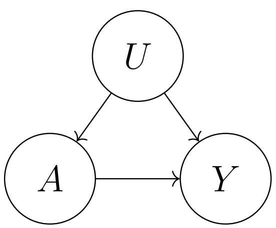
(Hernan & Robins, 2020, Table 1.1)

| | $Y^{a=0}$ | $Y^{a=1}$ |
|---|---|---|
| Rheia | 0 | 1 |
| Kronos | 1 | 0 |
| Demeter | 0 | 0 |
| Hades | 0 | 0 |
| Hestia | 0 | 0 |
| Poseidon | 1 | 0 |
| Hera | 0 | 0 |
| Zeus | 0 | 1 |
| Artemis | 1 | 1 |
| Apollo | 1 | 0 |
| Leto | 0 | 1 |
| Ares | 1 | 1 |
| Athena | 1 | 1 |
| Hephaestus | 0 | 1 |
| Aphrodite | 0 | 1 |
| Cyclope | 0 | 1 |
| Persephone | 1 | 1 |
| Hermes | 1 | 0 |
| Hebe | 1 | 0 |
| Dionysus | 1 | 0 |

Typical observed $Y$

# Why causal inference is challenging

- But, we don't get to see individual-specific counterfactual RV values!
- At most we see one **actual** outcome
- Let us bundle up everything we don't know about the process by which actual $A$ and $Y$ were determined into the RV $U$



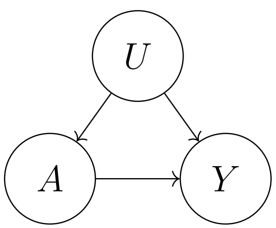

(Hernan & Robins, 2020, Table 1.1)

| | $Y^{a=0}$ | $Y^{a=1}$ |
|---|---|---|
| Rheia | 0 | 1 |
| Kronos | 1 | 0 |
| Demeter | 0 | 0 |
| Hades | 0 | 0 |
| Hestia | 0 | 0 |
| Poseidon | 1 | 0 |
| Hera | 0 | 0 |
| Zeus | 0 | 1 |
| Artemis | 1 | 1 |
| Apollo | 1 | 0 |
| Leto | 0 | 1 |
| Ares | 1 | 1 |
| Athena | 1 | 1 |
| Hephaestus | 0 | 1 |
| Aphrodite | 0 | 1 |
| Cyclope | 0 | 1 |
| Persephone | 1 | 1 |
| Hermes | 1 | 0 |
| Hebe | 1 | 0 |
| Dionysus | 1 | 0 |

Typical observed $Y$

# Why causal inference is challenging

- But, we don't get to see individual-specific counterfactual RV values!
- At most we see one **actual** outcome
- Let us bundle up everything we don't know about the process by which actual $A$ and $Y$ were determined into the RV $U$




Confounding!

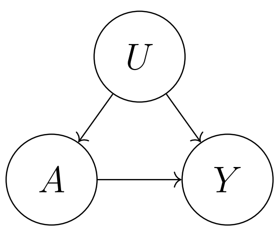
(Hernan & Robins, 2020, Table 1.1)

| | $Y^{a=0}$ | $Y^{a=1}$ |
|---|---|---|
| Rheia | 0 | 1 |
| Kronos | 1 | 0 |
| Demeter | 0 | 0 |
| Hades | 0 | 0 |
| Hestia | 0 | 0 |
| Poseidon | 1 | 0 |
| Hera | 0 | 0 |
| Zeus | 0 | 1 |
| Artemis | 1 | 1 |
| Apollo | 1 | 0 |
| Leto | 0 | 1 |
| Ares | 1 | 1 |
| Athena | 1 | 1 |
| Hephaestus | 0 | 1 |
| Aphrodite | 0 | 1 |
| Cyclope | 0 | 1 |
| Persephone | 1 | 1 |
| Hermes | 1 | 0 |
| Hebe | 1 | 0 |
| Dionysus | 1 | 0 |

Typical observed $Y$

# Why causal inference is challenging

- But, we don't get to see individual-specific counterfactual RV values!
- At most we see one **actual** outcome
- Let us bundle up everything we don't know about the process by which actual $A$ and $Y$ were determined into the RV $U$




Confounding!

- This is why randomized experiments are so powerful!

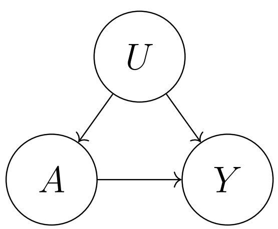
(Hernan & Robins, 2020, Table 1.1)

| | $Y^{a=0}$ | $Y^{a=1}$ |
|---|---|---|
| Rheia | 0 | 1 |
| Kronos | 1 | 0 |
| Demeter | 0 | 0 |
| Hades | 0 | 0 |
| Hestia | 0 | 0 |
| Poseidon | 1 | 0 |
| Hera | 0 | 0 |
| Zeus | 0 | 1 |
| Artemis | 1 | 1 |
| Apollo | 1 | 0 |
| Leto | 0 | 1 |
| Ares | 1 | 1 |
| Athena | 1 | 1 |
| Hephaestus | 0 | 1 |
| Aphrodite | 0 | 1 |
| Cyclope | 0 | 1 |
| Persephone | 1 | 1 |
| Hermes | 1 | 0 |
| Hebe | 1 | 0 |
| Dionysus | 1 | 0 |

Typical observed $Y$

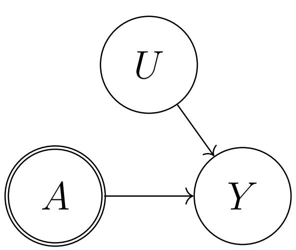
# Why causal inference is challenging

- But, we don't get to see individual-specific counterfactual RV values!
- At most we see one **actual** outcome
- Let us bundle up everything we don't know about the process by which actual $A$ and $Y$ were determined into the RV $U$




Confounding!

- This is why randomized experiments are so powerful!




(Hernan & Robins, 2020, Table 1.1)

| | $Y^{a=0}$ | $Y^{a=1}$ |
|---|---|---|
| Rheia | 0 | 1 |
| Kronos | 1 | 0 |
| Demeter | 0 | 0 |
| Hades | 0 | 0 |
| Hestia | 0 | 0 |
| Poseidon | 1 | 0 |
| Hera | 0 | 0 |
| Zeus | 0 | 1 |
| Artemis | 1 | 1 |
| Apollo | 1 | 0 |
| Leto | 0 | 1 |
| Ares | 1 | 1 |
| Athena | 1 | 1 |
| Hephaestus | 0 | 1 |
| Aphrodite | 0 | 1 |
| Cyclope | 0 | 1 |
| Persephone | 1 | 1 |
| Hermes | 1 | 0 |
| Hebe | 1 | 0 |
| Dionysus | 1 | 0 |

Typical observed $Y$

# Pearl's functional causal models

(Pearl 2009, section 1.4)

# Pearl's functional causal models

- Examples like these highlight the value of explicitly representing the **unknowns** – parts of the causal system that we will not in practice be able to observe

(Pearl 2009, section 1.4)

# Pearl's functional causal models

- Examples like these highlight the value of explicitly representing the **unknowns** – parts of the causal system that we will not in practice be able to observe
- For a system of RVs $V = \{X_i\}$, each gets an unseen "error term" $\{U_i\}$ (**not** necessarily independent)

(Pearl 2009, section 1.4)

# Pearl's functional causal models

- Examples like these highlight the value of explicitly representing the **unknowns** – parts of the causal system that we will not in practice be able to observe
- For a system of RVs $V = \{X_i\}$, each gets an unseen "error term" $\{U_i\}$ (**not** necessarily independent)
- A functional causal model then consists of a **deterministic** system of equations of the form

$$x_i = f_i(\mathrm{pa}_i, u_i)$$

(Pearl 2009, section 1.4)

# Pearl's functional causal models

- Examples like these highlight the value of explicitly representing the **unknowns** – parts of the causal system that we will not in practice be able to observe
- For a system of RVs $V = \{X_i\}$, each gets an unseen "error term" $\{U_i\}$ (**not** necessarily independent)
- A functional causal model then consists of a **deterministic** system of equations of the form

$$x_i = f_i(\mathrm{pa}_i, u_i)$$

- where $\mathrm{pa}_i$ is the subset of $V$ that are parents of $X_i$

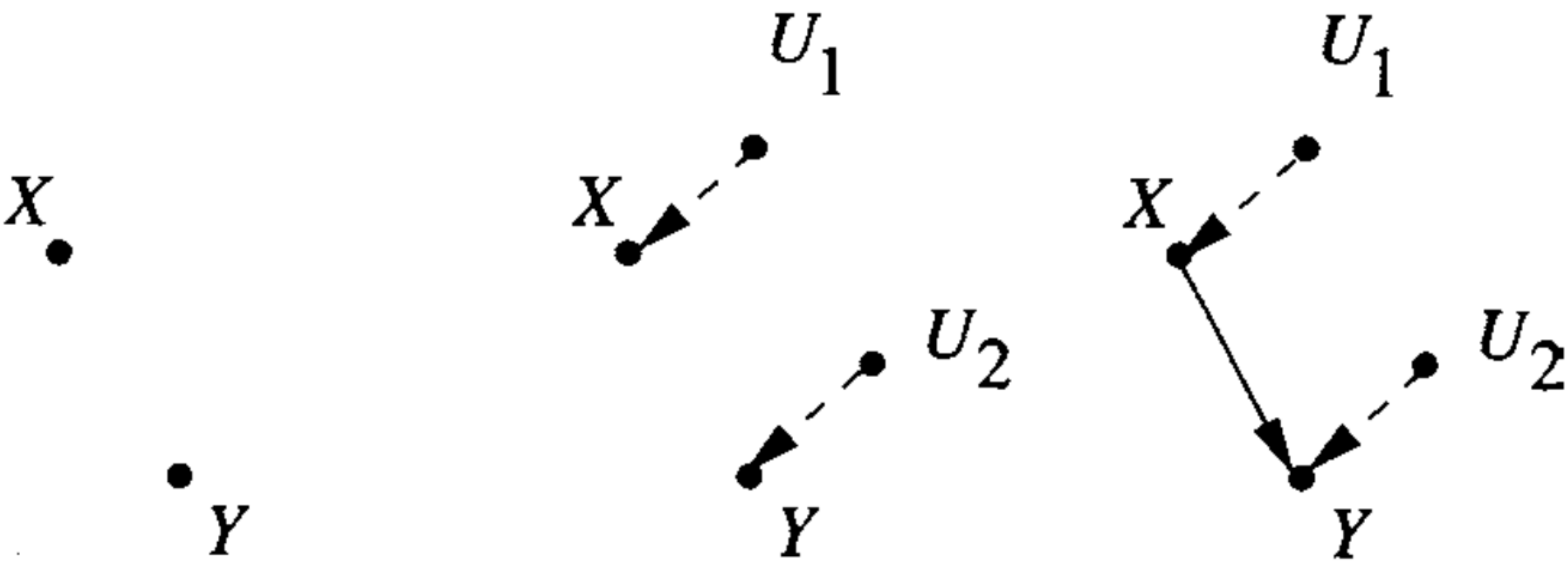
(Pearl 2009, section 1.4)

# Bayes Nets & functional causal models

**Bayes Net**

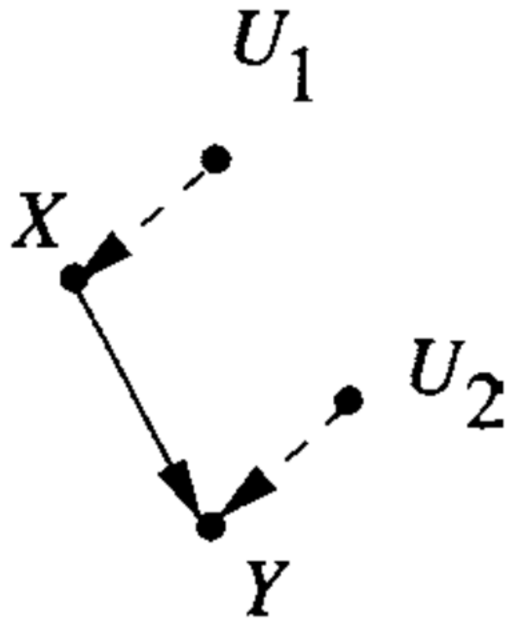
**Corresponding functional causal model**

**Functional causal model with $X$ causally upstream of $Y$**




(Pearl 2009, Figure 1.6)

# Example model & structural equations




$x = f_1(u_1)$ $\qquad\qquad$ $y = f_2(x, u_2)$

- An example structural equation:

$$y = a + bx + u_2$$

**(Could be linear regression!)**

(Pearl 2009, Figure 1.6)