Machine learning and causal inference (randomized experiments)

Robert Castelo

robert.castelo@upf.edu

@robertclab

Dept. of Medicine and Life Sciences
Universitat Pompeu Fabra

upf.

Barcelona

Barcelona School of Economics
Data Science Methodology Program
Winter Term 2024-25

BSE
Barcelona School of Economics

Introductory concepts and definitions

- Causal inference is a field of study that (1) attempts to identify factors that cause an effect on an outcome of interest and (2) estimates the magnitude of those effects.

Barcelona School of Economics

Introductory concepts and definitions

- Causal inference is a field of study that (1) attempts to identify factors that cause an effect on an outcome of interest and (2) estimates the magnitude of those effects.
- Causality can be studied from a philosophical (metaphysical) viewpoint (Aristotle, Hume, Kant, etc., see https://en.wikipedia.org/wiki/Causality). Our treatment will be more pragmatic: math, statistics, computation .. data science!

Barcelona School of Economics

Introductory concepts and definitions

- Causal inference is a field of study that (1) attempts to identify factors that cause an effect on an outcome of interest and (2) estimates the magnitude of those effects.
- Causality can be studied from a philosophical (metaphysical) viewpoint (Aristotle, Hume, Kant, etc., see https://en.wikipedia.org/wiki/Causality). Our treatment will be more pragmatic: math, statistics, computation .. data science!
- The term *inference* here doesn't have the same meaning as in statistical *inference*.

Barcelona School of Economics

Introductory concepts and definitions

- Causal inference is a field of study that (1) attempts to identify factors that cause an effect on an outcome of interest and (2) estimates the magnitude of those effects.
- Causality can be studied from a philosophical (metaphysical) viewpoint (Aristotle, Hume, Kant, etc., see https://en.wikipedia.org/wiki/Causality). Our treatment will be more pragmatic: math, statistics, computation .. data science!
- The term *inference* here doesn't have the same meaning as in statistical *inference*.
- Upfront, association-correlation-prediction is not causation.

Barcelona School of Economics

Introductory concepts and definitions

- Causal inference is a field of study that (1) attempts to identify factors that cause an effect on an outcome of interest and (2) estimates the magnitude of those effects.
- Causality can be studied from a philosophical (metaphysical) viewpoint (Aristotle, Hume, Kant, etc., see https://en.wikipedia.org/wiki/Causality). Our treatment will be more pragmatic: math, statistics, computation .. data science!
- The term *inference* here doesn't have the same meaning as in statistical *inference*.
- Upfront, association-correlation-prediction is not causation.
- Methods in causal inference will help us to decide when an association is causal $\Longrightarrow$ it can improve our decision making.

Barcelona School of Economics

Introductory concepts and definitions

Causal inference is used in many different fields such as:

- **Epidemiology:** Ballon, M. et al. Which modifiable prenatal factors mediate the relation between socio-economic position and a child's weight and length at birth? *Matern. Child. Nutr.* (2019); https://doi.org/10.1111/mcn.12878.

- **Economics:** Varian H.R. Causal inference in economics and marketing. *PNAS* (2016); https://doi.org/10.1073/pnas.1510479113.

- **Biology:** Meinshausen et al. Methods for causal inference from gene perturbation experiments and validation. *PNAS* (2016); https://doi.org/10.1073/pnas.1510493113.

- **Sociology:** Knox, D. and Mummolo, J. Toward a general causal framework for the study of racial bias in policing. *J. Pol. Inst. and Pol. Econ.* (2020); https://doi.org/10.1561/113.00000018.

- **Medicine:** Polack, F.P. et al. Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. *NEJM* (2020); https://doi.org/10.1056/NEJMoa2034577.

More examples in section 1.3 from Morgan and Winship (2015).

Barcelona School of Economics

Introductory concepts and definitions

- **Unit:** basic object of study in an investigation, e.g., human subjects, living cells, laboratory animals, households, economic indicators, etc.

Barcelona School of Economics

Introductory concepts and definitions

- **Unit:** basic object of study in an investigation, e.g., human subjects, living cells, laboratory animals, households, economic indicators, etc.

- **Treatment:** an intervention or exposure. Let X be a variable indicating the treatment, i.e., the cause to which each unit is exposed. For instance, $X = 1$ or $X = t$ denotes *treatment*, while $X = 0$ or $X = c$ denotes *no-treatment*, often referred to as *control*.

Barcelona School of Economics

Introductory concepts and definitions

- **Unit:** basic object of study in an investigation, e.g., human subjects, living cells, laboratory animals, households, economic indicators, etc.

- **Treatment:** an intervention or exposure. Let X be a variable indicating the treatment, i.e., the cause to which each unit is exposed. For instance, $X = 1$ or $X = t$ denotes *treatment*, while $X = 0$ or $X = c$ denotes *no-treatment*, often referred to as *control*.

- **Observed outcomes:** Values of the effect of the treatment, denoted by the variable Y_i for the value of the unit i. They need to occur after the treatment. Sometimes they are referred to as the observed *responses*.

Barcelona School of Economics

Introductory concepts and definitions

- The observed outcomes in Y_i depend on the treatment X. To faithfully represent this setting we actually need **two** variables.

Introductory concepts and definitions

- The observed outcomes in Y_i depend on the treatment X. To faithfully represent this setting we actually need **two** variables.

- **Potential outcomes:** outcomes under the treatment and control, denoted in different ways by $Y_i(0)$, $Y_i(1)$, $Y_i(c)$, $Y_i(t)$ and $Y_i(x)$ for a given treatment $X = x$, i.e., $Y_i(x = 1) \equiv Y_i(1) \equiv Y_i(t)$.

observed outcome Y i when treatment is 1 or zero (x=1/0)

Barcelona School of Economics

Introductory concepts and definitions

- The observed outcomes in Y_i depend on the treatment X. To faithfully represent this setting we actually need **two** variables.

- **Potential outcomes:** outcomes under the treatment and control, denoted in different ways by $Y_i(0)$, $Y_i(1)$, $Y_i(c)$, $Y_i(t)$ and $Y_i(x)$ for a given treatment $X = x$, i.e., $Y_i(x = 1) \equiv Y_i(1) \equiv Y_i(t)$.

- Crawling literature: sometimes the unit subscript is dropped, $Y(x)$; sometimes unit and treatment are exchanged in the notation, $Y_t(i)$; sometimes they are referred to as the *potential responses*.

Barcelona School of Economics

Introductory concepts and definitions

- The observed outcomes in Y_i depend on the treatment X. To faithfully represent this setting we actually need **two** variables.

- **Potential outcomes:** outcomes under the treatment and control, denoted in different ways by $Y_i(0)$, $Y_i(1)$, $Y_i(c)$, $Y_i(t)$ and $Y_i(x)$ for a given treatment $X = x$, i.e., $Y_i(x = 1) \equiv Y_i(1) \equiv Y_i(t)$.

- Crawling literature: sometimes the unit subscript is dropped, $Y(x)$; sometimes unit and treatment are exchanged in the notation, $Y_t(i)$; sometimes they are referred to as the *potential responses*.

- For each unit i we have a set of potential outcomes $\{Y_i(0), Y_i(1)\}$, e.g.:
  $Y_i(0) =$ *lung cancer 5-year survival without immunotherapy,*
  $Y_i(1) =$ *lung cancer 5-year survival with immunotherapy.*

Barcelona School of Economics

Introductory concepts and definitions

- The potential outcomes framework was originally proposed by Jerzy Neyman in the context of randomized experiments and was extended by Donald Rubin to both observational and experimental studies (Rubin, 2005)[1]. Sometimes called Neyman-Rubin framework.

[1] Rubin, D.B. Causal inference using potential outcomes. *JASA*, 2005. https://doi.org/10.1198/016214504000001880

[2] Holland, P.W., Statistics and Causal Inference, *JASA*, 1986. https://doi.org/10.1080/01621459.1986.10478354

Barcelona School of Economics

Introductory concepts and definitions

- The potential outcomes framework was originally proposed by Jerzy Neyman in the context of randomized experiments and was extended by Donald Rubin to both observational and experimental studies (Rubin, 2005)[1]. Sometimes called Neyman-Rubin framework.

- **Goal:** assess the effect of treatment compared to control. Holland (1986)[2]:

  - "The emphasis here will be on *measuring the effects of causes* because this seems to be a place where statistics, which is concerned with measurement, has contributions to make".

  - "Everything has a cause, but not everything can be cause" $\Longrightarrow$ "No causation without manipulation", e.g., "Do COVID19 vaccines reduce COVID19 mortality rates?".

[1] Rubin, D.B. Causal inference using potential outcomes. *JASA*, 2005. https://doi.org/10.1198/016214504000001880

[2] Holland, P.W., Statistics and Causal Inference, *JASA*, 1986. https://doi.org/10.1080/01621459.1986.10478354

Barcelona School of Economics

- Individual treatment effect (ITE): difference between potential outcomes in a specific unit i.

$$\delta_i := Y_i(1) - Y_i(0)\,.$$

Introductory concepts and definitions

- Individual treatment effect (ITE): difference between potential outcomes in a specific unit i.

$$\delta_i := Y_i(1) - Y_i(0)\,.$$

- If treatment has no ITE on unit i, then $Y_i(0) = Y_i(1)$ and $\delta_i = 0$.

- Individual treatment effect (ITE): difference between potential outcomes in a specific unit i.

$$\delta_i := Y_i(1) - Y_i(0)\,.$$

- If treatment has no ITE on unit i, then $Y_i(0) = Y_i(1)$ and $\delta_i = 0$.

- Observed outcomes can be expressed in terms of potential outcomes:

$$Y_i = X_i \cdot Y_i(1) + (1 - X_i) \cdot Y_i(0)\,,$$

  which implies that $Y_i = Y_i(1)$ if $X_i = 1$ and $Y_i = Y_i(0)$ if $X_i = 0$.

- Observed outcomes can be expressed in terms of potential outcomes:

$$Y_i = X_i \cdot Y_i(1) + (1 - X_i) \cdot Y_i(0) \,.$$

- Toy example with data:

i	X_i	$Y_i(1)$	$Y_i(0)$	$\delta_i = Y_i(1) - Y_i(0)$	Y_i
1	0	7	2	5	2
2	1	5	1	4	5
3	1	6	1	5	6
4	0	4	2	2	2

- Observed outcomes can be expressed in terms of potential outcomes:

$$Y_i = X_i \cdot Y_i(1) + (1 - X_i) \cdot Y_i(0)\,.$$

- Toy example with data:

i	X_i	$Y_i(1)$	$Y_i(0)$	$\delta_i = Y_i(1) - Y_i(0)$	Y_i
1	0	7	2	5	2
2	1	5	1	4	5
3	1	6	1	5	6
4	0	4	2	2	2

- Do we actually have the previous table in real life?

No --> fundamental problem of causal inference

Introductory concepts and definitions

- Holland (1986):

  Fundamental Problem of Causal Inference. It is impossible to observe all potential outcomes on the same unit.

Introductory concepts and definitions

- Holland (1986):

  Fundamental Problem of Causal Inference. It is impossible to observe all potential outcomes on the same unit.

The Matrix, 1999. See https://en.wikipedia.org/wiki/The_Matrix

Introductory concepts and definitions

- Holland (1986):

  Fundamental Problem of Causal Inference. It is impossible to observe all potential outcomes on the same unit.

The Matrix, 1999. See https://en.wikipedia.org/wiki/The_Matrix

- A chunk of literature refers to potential outcomes as *counterfactuals* because they resemble *what if ..* or *had I ..* type of statements.

- Toy example revisited, what we actually see in real life:

i	X_i	$Y_i(1)$	$Y_i(0)$	$\delta_i = Y_i(1) - Y_i(0)$	Y_i
1	0	? (cf)	2	?	2
2	1	5	? (cf)	?	5
3	1	6	? (cf)	?	6
4	0	? (cf)	2	?	2

- Toy example revisited, what we actually see in real life:

i	X_i	$Y_i(1)$	$Y_i(0)$	$\delta_i = Y_i(1) - Y_i(0)$	Y_i
1	0	? (cf)	2	?	2
2	1	5	? (cf)	?	5
3	1	6	? (cf)	?	6
4	0	? (cf)	2	?	2

- Causal inference is a missing data problem. The complete table, i.e., without missing data, is referred to in the literature as the *nature table* or the *science table*.

Introductory concepts and definitions

- Toy example revisited, what we actually see in real life:

i	X_i	$Y_i(1)$	$Y_i(0)$	$\delta_i = Y_i(1) - Y_i(0)$	Y_i
1	0	? (cf)	2	?	2
2	1	5	? (cf)	?	5
3	1	6	? (cf)	?	6
4	0	? (cf)	2	?	2

- Causal inference is a missing data problem. The complete table, i.e., without missing data, is referred to in the literature as the *nature table* or the *science table*.

- How can we calculate $\delta_i = Y_i(1) - Y_i(0)$?

Introductory concepts and definitions

- "Typical unit-level" (Rubin, 2005)[3] or average treatment effect (ATE):

$$ATE := \frac{1}{n}\sum_{i=1}^{n}\{Y_i(1) - Y_i(0)\}.$$

[3] Rubin, D.B. Causal inference using potential outcomes. *JASA*, 2005. https://doi.org/10.1198/016214504000001880

Barcelona School of Economics

Introductory concepts and definitions

- "Typical unit-level" (Rubin, 2005)[3] or average treatment effect (ATE):

$$ATE := \frac{1}{n}\sum_{i=1}^{n}\{\, Y_i(1) - Y_i(0)\}\,.$$

- Other summaries of causal effects can be obtained using the median:

$$\text{median}\{\, Y_i(1) - Y_i(0), i = 1, \ldots, n\}\,,$$

or the margin, e.g., difference in median causal effects:

$$\text{median}\{\, Y_i(1), i = 1, \ldots, n\} - \text{median}\{\, Y_i(0), i = 1, \ldots, n\}\,,$$

and difference in mean causal effects, which is equal to typical unit-level:

$$ATE := \frac{1}{n}\sum_{i=1}^{n}\{\, Y_i(1) - Y_i(0)\} = \frac{\sum_{i=1}^{n} Y_i(1)}{n} - \frac{\sum_{i=1}^{n} Y_i(0)}{n} = \bar{Y}(1) - \bar{Y}(0).$$

[3] Rubin, D.B. Causal inference using potential outcomes. *JASA*, 2005. https://doi.org/10.1198/016214504000001880

Introductory concepts and definitions

- "Typical unit-level" (Rubin, 2005)[3] or average treatment effect (ATE):

$$ATE := \frac{1}{n}\sum_{i=1}^{n}\{\, Y_i(1) - Y_i(0)\}\,.$$

- Other summaries of causal effects can be obtained using the median:

$$\text{median}\{\, Y_i(1) - Y_i(0), i = 1, \ldots, n\}\,,$$

or the margin, e.g., difference in median causal effects:

$$\text{median}\{\, Y_i(1), i = 1, \ldots, n\} - \text{median}\{\, Y_i(0), i = 1, \ldots, n\}\,,$$

and difference in mean causal effects, which is equal to typical unit-level:

$$ATE := \frac{1}{n}\sum_{i=1}^{n}\{\, Y_i(1) - Y_i(0)\} = \frac{\sum_{i=1}^{n} Y_i(1)}{n} - \frac{\sum_{i=1}^{n} Y_i(0)}{n} = \bar{Y}(1) - \bar{Y}(0).$$

- **Critical requirement:** to estimate causal effects, comparisons between $Y_i(1)$ and $Y_i(0)$ must be made on the same set of units $i = 1, \ldots, n$.

[3] Rubin, D.B. Causal inference using potential outcomes. *JASA*, 2005. https://doi.org/10.1198/016214504000001880

Assumptions: additivity

- Rubin (2005): "Nothing is wrong with making assumptions; causal inference is impossible without making assumptions, and they are the strands that link statistics to science".

Assumptions: additivity

- Rubin (2005): "Nothing is wrong with making assumptions; causal inference is impossible without making assumptions, and they are the strands that link statistics to science".

- **Additivity**: the ITE is *additive* if the treatment adds a fixed amount to each control value. More specifically, given $\delta_i = Y_i(1) - Y_i(0)$, then $\delta_i = \delta$ for $i = 1, \ldots, n$ and some constant δ, i.e., the causal effect is identical for all units i.

i	X_i	$Y_i(1)$	$Y_i(0)$	$Y_i(1) - Y_i(0)$
1	0	7	2	5
2	1	5	0	5
3	1	6	1	5
4	0	8	3	5
mean		6.5	1.5	5

Assumptions: SUTVA

- Rubin (1980)[4]: stable unit treatment value assumption (SUTVA), also known as *stability assumption*.

- Imbens and Rubin (2015, pg.10)[5] SUTVA definition:

  The potential outcomes for any unit do not vary with the treatments assigned to other units, and, for each unit, there are no different forms of versions of each treatment level, which lead to different potential outcomes.

[4] Rubin, D.B. Discussion of "Randomization analysis of experimental data in the Fisher randomization test" by Basu, *JASA*, 1980. https://doi.org/10.2307/2287653.

[5] Imbens, G.W. and Rubin, D.B. *Causal Inference for Statistics, Social and Biomedical Sciences*, Cambridge University Press, 2015.

Assumptions: SUTVA

- Rubin (1980)[4]: stable unit treatment value assumption (SUTVA), also known as *stability assumption*.
- Imbens and Rubin (2015, pg.10)[5] SUTVA definition:

  The potential outcomes for any unit do not vary with the treatments assigned to other units, and, for each unit, there are no different forms of versions of each treatment level, which lead to different potential outcomes.

- The goal of the stability assumption is to enable us to exploit the presence of multiple units for estimating causal effects.

[4] Rubin, D.B. Discussion of "Randomization analysis of experimental data in the Fisher randomization test" by Basu, *JASA*, 1980. https://doi.org/10.2307/2287653.

[5] Imbens, G.W. and Rubin, D.B. *Causal Inference for Statistics, Social and Biomedical Sciences*, Cambridge University Press, 2015.

Assumptions: SUTVA

- Rubin (1980)[4]: stable unit treatment value assumption (SUTVA), also known as *stability assumption*.
- Imbens and Rubin (2015, pg.10)[5] SUTVA definition:

  The potential outcomes for any unit do not vary with the treatments assigned to other units, and, for each unit, there are no different forms of versions of each treatment level, which lead to different potential outcomes.

- The goal of the stability assumption is to enable us to exploit the presence of multiple units for estimating causal effects.
- Without SUTVA, causal inference using potential outcomes is far more complicated.

[4] Rubin, D.B. Discussion of "Randomization analysis of experimental data in the Fisher randomization test" by Basu, *JASA*, 1980. https://doi.org/10.2307/2287653.

[5] Imbens, G.W. and Rubin, D.B. *Causal Inference for Statistics, Social and Biomedical Sciences*, Cambridge University Press, 2015.

Assumptions: SUTVA

- "The potential outcomes for any unit do not vary with the treatments assigned to other units": **no intereference between units**, i.e., neither $Y_i(1)$ nor $Y_i(0)$ is affected by the treatment on other units j with $j \neq i$.

e.g. vaccinations: one person being vaccinated affects the other person, --> intereferenc

Assumptions: SUTVA

- "The potential outcomes for any unit do not vary with the treatments assigned to other units": **no intereference between units**, i.e., neither $Y_i(1)$ nor $Y_i(0)$ is affected by the treatment on other units j with $j \neq i$.

- No interference may be reasonable in some context (e.g., therapies for noncommunicable diseases), but not in others (e.g., vaccines for communicable -infectious- diseases).

Barcelona School of Economics

Assumptions: SUTVA

- "The potential outcomes for any unit do not vary with the treatments assigned to other units": **no intereference between units**, i.e., neither $Y_i(1)$ nor $Y_i(0)$ is affected by the treatment on other units j with $j \neq i$.

- No interference may be reasonable in some context (e.g., therapies for noncommunicable diseases), but not in others (e.g., vaccines for communicable -infectious- diseases).

- We need to rely on assumed existing knowledge of the current subject matter to assert that some treatments do not affect outcomes for some units. Ruling out certain causal effects, also known as exclusion restriction.

Barcelona School of Economics

Assumptions: SUTVA

- "The potential outcomes for any unit do not vary with the treatments assigned to other units": **no intereference between units**, i.e., neither $Y_i(1)$ nor $Y_i(0)$ is affected by the treatment on other units j with $j \neq i$.

- No interference may be reasonable in some context (e.g., therapies for noncommunicable diseases), but not in others (e.g., vaccines for communicable -infectious- diseases).

- We need to rely on assumed existing knowledge of the current subject matter to assert that some treatments do not affect outcomes for some units. Ruling out certain causal effects, also known as exclusion restriction.

- Motivated by the problem of measuring vaccine causal effects, approaches have been developed to perform causal inference in the presence of interference.

Barcelona School of Economics

Assumptions: SUTVA

- "For each unit, there are no different forms of versions of each treatment level, which lead to different potential outcomes": **causal consistency**, i.e., the potential outcome $y_i(x)$ of a unit i receiving a specific treatment x is the same no matter how unit i is exposed to treatment x.

Assumptions: SUTVA

- "For each unit, there are no different forms of versions of each treatment level, which lead to different potential outcomes": **causal consistency**, i.e., the potential outcome $y_i(x)$ of a unit i receiving a specific treatment x is the same no matter how unit i is exposed to treatment x.

- For instance, consider treatment/exposure to be smoking, a health-related outcome (e.g., COPD, lung cancer, etc.) may differ depending on the type of smoker (sporadic vs. chain smokers).

Barcelona School of Economics

Assumptions: SUTVA

- "For each unit, there are no different forms of versions of each treatment level, which lead to different potential outcomes": **causal consistency**, i.e., the potential outcome $y_i(x)$ of a unit i receiving a specific treatment x is the same no matter how unit i is exposed to treatment x.

- For instance, consider treatment/exposure to be smoking, a health-related outcome (e.g., COPD, lung cancer, etc.) may differ depending on the type of smoker (sporadic vs. chain smokers).

- Causal consistency is a different concept from statistical consistency, where an estimator is consistent if it converges in probability to the value that it is designed to estimate.

Barcelona School of Economics

Assumptions: the treatment assignment mechanism

- **Assignment mechanism:** how units receive treatment, the process by which some units receive treatment/exposure and other don't (control).

Assumptions: the treatment assignment mechanism

- **Assignment mechanism:** how units receive treatment, the process by which some units receive treatment/exposure and other don't (control).

- Challenge 1: assigning treatment may or may not be under our control, e.g., randomized clinical trials vs. observational studies.

Assumptions: the treatment assignment mechanism

- **Assignment mechanism:** how units receive treatment, the process by which some units receive treatment/exposure and other don't (control).

- Challenge 1: assigning treatment may or may not be under our control, e.g., randomized clinical trials vs. observational studies.

- Challenge 2: even if we can assign treatment, some units may actually not receive it, a setting known as *non-compliance*. For instance, treatment invividuals in a clinical trial for a vaccine mistakenly received placebo.

Barcelona School of Economics

Assumptions: the treatment assignment mechanism

- Randomized assignment mechanism: treatments in X are randomly assigned.

Assumptions: the treatment assignment mechanism

- Randomized assignment mechanism: treatments in X are randomly assigned.

- If units randomly receive treatment according to X, then by causal consistency,

$$Y_i = X_i \cdot Y_i(1) + (1 - X_i) \cdot Y_i(0) \,.$$

Assumptions: the treatment assignment mechanism

- Randomized assignment mechanism: treatments in X are randomly assigned.

- If units randomly receive treatment according to X, then by causal consistency,

$$Y_i = X_i \cdot Y_i(1) + (1 - X_i) \cdot Y_i(0)\,.$$

- Let $\mathbf{X} := (X_1, \ldots, X_n)$ and $\mathbf{x} := (x_1, \ldots, x_n)$ be actual values in $\mathbf{X}$. In a completely randomized experiment, where k of n units receive treatment equally likely,

$$\Pr(\mathbf{X} = \mathbf{x}) = \begin{cases} 1/\binom{n}{k} & \text{if } \sum_i x_i = k \\ 0 & \text{otherwise} \end{cases}$$

Assumptions: the treatment assignment mechanism

- Randomized assignment mechanism: treatments in X are randomly assigned.

- If units randomly receive treatment according to X, then by causal consistency,

$$Y_i = X_i \cdot Y_i(1) + (1 - X_i) \cdot Y_i(0)\,.$$

- Let $\mathbf{X} := (X_1, \ldots, X_n)$ and $\mathbf{x} := (x_1, \ldots, x_n)$ be actual values in $\mathbf{X}$. In a completely randomized experiment, where k of n units receive treatment equally likely,

$$\Pr(\mathbf{X} = \mathbf{x}) = \begin{cases} 1/\binom{n}{k} & \text{if } \sum_i x_i = k \\ 0 & \text{otherwise} \end{cases}$$

- Toy example: for $n = 4$ units, there are $2^4 = 16$ possible treatment/control assignment allocations. Let's say we want to assign treatment to $k = 2$ units. There are $\binom{4}{2} = 6$ ways to assign treatment $x = 1$ to them and those assignments will be made with equal probability $1/6$.

Randomization-based causal inference

- **Goal:** given data such as

i	X_i	$Y_i(1)$	$Y_i(0)$	Y_i
1	0	?	2	2
2	1	5	?	5
3	1	6	?	6
4	0	?	3	3

  estimate the average treatment effect (ATE).

- Randomization-based causal inference: classical statistical approaches based on results from Neyman and Fisher.

- Assumptions: SUTVA and randomized treatment assignment. No distributional assumptions on observed outcomes.

Barcelona School of Economics

Randomization-based causal inference (Neyman)

- Neyman (1923)[6] showed that given a completely randomized experiment, the difference in sample means:

$$\widehat{\mathrm{ATE}} := \frac{\sum_i^n Y_i X_i}{\sum_i^n X_i} - \frac{\sum_i^n Y_i (1 - X_i)}{\sum_i^n (1 - X_i)} \,.$$

  is unbiased estimator of ATE, i.e., $\mathrm{E}[\widehat{\mathrm{ATE}}] = \mathrm{ATE}$, taking the expectation over all possible randomizations.

[6] Splawa-Neyman, J. On the application of probability theory to agricultural experiments. Essay on principles. Section 9. *Annals of Agricultural Sciences*, 1923. Translated in *Statistical Science*, 5:465-472, 1990. https://www.jstor.org/stable/2245382.

Randomization-based causal inference (Neyman)

- Neyman (1923)[6] showed that given a completely randomized experiment, the difference in sample means:

$$\widehat{\mathrm{ATE}} := \frac{\sum_i^n Y_i X_i}{\sum_i^n X_i} - \frac{\sum_i^n Y_i (1 - X_i)}{\sum_i^n (1 - X_i)} .$$

  is unbiased estimator of ATE, i.e., $\mathrm{E}[\widehat{\mathrm{ATE}}] = \mathrm{ATE}$, taking the expectation over all possible randomizations.

- The potential outcomes $\{ Y_i(0), Y_i(1), i = 1, \ldots, n \}$ are considered fixed and only treatment assignments X_i are considered random.

[6] Splawa-Neyman, J. On the application of probability theory to agricultural experiments. Essay on principles. Section 9. *Annals of Agricultural Sciences*, 1923. Translated in *Statistical Science*, 5:465-472, 1990. https://www.jstor.org/stable/2245382.

Randomization-based causal inference (Neyman)

- Neyman (1923)[6] showed that given a completely randomized experiment, the difference in sample means:

$$\widehat{\mathrm{ATE}} := \frac{\sum_i^n Y_i X_i}{\sum_i^n X_i} - \frac{\sum_i^n Y_i(1-X_i)}{\sum_i^n (1-X_i)}\,.$$

  is unbiased estimator of ATE, i.e., $\mathrm{E}[\widehat{\mathrm{ATE}}] = \mathrm{ATE}$, taking the expectation over all possible randomizations.

- The potential outcomes $\{Y_i(0), Y_i(1), i = 1, \ldots, n\}$ are considered fixed and only treatment assignments X_i are considered random.

- In a completely randomized experiment, assuming SUTVA, **association is causation**.

[6] Splawa-Neyman, J. On the application of probability theory to agricultural experiments. Essay on principles. Section 9. *Annals of Agricultural Sciences*, 1923. Translated in *Statistical Science*, 5:465-472, 1990. https://www.jstor.org/stable/2245382.

Randomization-based causal inference (Neyman)

- Proof (1/4). In a completely randomized experiment, the number of units assigned to treatment k is fixed.

$$\mathrm{E}\left[\frac{\sum_i^n Y_i X_i}{\sum_i^n X_i}\right] = \frac{1}{k}\sum_i^n \mathrm{E}[Y_i X_i]\,.$$

Randomization-based causal inference (Neyman)

- Proof (1/4). In a completely randomized experiment, the number of units assigned to treatment k is fixed.

$$\mathrm{E}\left[\frac{\sum_i^n Y_i X_i}{\sum_i^n X_i}\right] = \frac{1}{k}\sum_i^n \mathrm{E}[Y_i X_i]\,.$$

- Proof (2/4). By causal consistency we can replace Y_i by $Y_i(1)$

$$\mathrm{E}\left[\frac{\sum_i^n Y_i X_i}{\sum_i^n X_i}\right] = \frac{1}{k}\sum_i^n \mathrm{E}[Y_i X_i] = \frac{1}{k}\sum_i^n Y_i(1)\mathrm{E}[X_i]\,.$$

Barcelona School of Economics

Randomization-based causal inference (Neyman)

- Proof (3/4). In a completely randomized experiment, potential outcomes $Y_i(1)$ are fixed and $\mathrm{E}[X_i] = k/n$.

$$\mathrm{E}\left[\frac{\sum_i^n Y_i X_i}{\sum_i^n X_i}\right] = \frac{1}{k}\sum_i^n \mathrm{E}[Y_i X_i] = \frac{1}{k}\sum_i^n Y_i(1)\mathrm{E}[X_i] = \frac{1}{n}\sum_i^n Y_i(1) = \bar{Y}(1)\,.$$

Randomization-based causal inference (Neyman)

- Proof (3/4). In a completely randomized experiment, potential outcomes $Y_i(1)$ are fixed and $\mathrm{E}[X_i] = k/n$.

$$\mathrm{E}\left[\frac{\sum_i^n Y_i X_i}{\sum_i^n X_i}\right] = \frac{1}{k}\sum_i^n \mathrm{E}[Y_i X_i] = \frac{1}{k}\sum_i^n Y_i(1)\mathrm{E}[X_i] = \frac{1}{n}\sum_i^n Y_i(1) = \bar{Y}(1)\,.$$

- Proof (4/4). Analogously for untreated assignments.

$$\mathrm{E}\left[\frac{\sum_i^n Y_i(1-X_i)}{\sum_i^n (1-X_i)}\right] = \bar{Y}(0)\,,$$

and therefore,

$$\mathrm{E}[\widehat{\mathrm{ATE}}] = \mathrm{ATE}\,.$$

Randomization-based causal inference (Neyman)

- Toy example with an additive $\text{ITE} = 5$.

i	$Y_i(1)$	$Y_i(0)$	$Y_i(1) - Y_i(0)$
1	7	2	5
2	5	0	5
3	6	1	5
4	8	3	5

- Completely randomized experiment with $k = 2$, calculate $\text{E}[\widehat{ATE}]$.

x	$\Pr(\mathbf{X} = \mathbf{x})$	$\widehat{\text{ATE}}$
0011	1/6	(6+8)/2 - (2+0)/2 = 7 - 1 = 6
0101	1/6	(5+8)/2 - (2+1)/2 = 6.5 - 1.5 = 5
0110	1/6	(5+6)/2 - (2+3)/2 = 5.5 - 2.5 = 3
1001	1/6	(7+8)/2 - (0+1)/2 = 7.5 - 0.5 = 7
1010	1/6	(7+6)/2 - (0+3)/2 = 6.5 - 1.5 = 5
1100	1/6	(7+5)/2 - (1+3)/2 = 6 - 2 = 4
$\text{E}[\widehat{ATE}]$		(6+5+3+7+5+4)/6 = 30/6 = **5**

normally we don't have 0011 as seen before, but we just want to show E[ATE] her

Randomization-based causal inference (Neyman)

- Given that $\widehat{ATE}$ is the difference in sample means, we can estimate its variance pooling the within-group sample variances:

$$\widehat{\mathrm{Var}}(\widehat{\mathrm{ATE}}) := \frac{s_1^2}{k} + \frac{s_0^2}{n-k},$$

where s_j^2 are the sample variances for $\{ Y_i : X_i = j,\ j = (0,1)\}$.

Randomization-based causal inference (Neyman)

- Given that $\widehat{ATE}$ is the difference in sample means, we can estimate its variance pooling the within-group sample variances:

$$\widehat{\mathrm{Var}}(\widehat{\mathrm{ATE}}) := \frac{s_1^2}{k} + \frac{s_0^2}{n-k},$$

where s_j^2 are the sample variances for $\{Y_i : X_i = j,\ j = (0,1)\}$.

- This is a positively biased estimate of the ATE true variance over all possible randomizations, unless the ITEs are additive. More concretely,

0 if additive

$$\mathrm{E}[\widehat{Var}(\widehat{ATE})] = \mathrm{Var}(\widehat{ATE}) + \frac{1}{n-1}\mathrm{Var}(ITE),$$

this only approaches the true variance if the ITEs are additive

where

$$\mathrm{Var}(ITE) := \frac{1}{n}\sum_i^n([Y_i(1) - Y_i(0)] - [\bar{Y}(1) - \bar{Y}(0)])^2.$$

Randomization-based causal inference (Neyman)

- Toy example with an additive $\mathrm{ITE} = 5$.

i	$Y_i(1)$	$Y_i(0)$	TE_i
1	7	2	5
2	5	0	5
3	6	1	5
4	8	3	5

additive, so ATE and var(ATE) should match

- Completely randomized experiment with $k = 2$.

var in R sample mean

x	$\Pr(\mathbf{X} = \mathbf{x})$	$\widehat{\mathrm{ATE}}$	$\widehat{\mathrm{Var}}(\widehat{ATE})$
0011	1/6	(6+8)/2 - (2+0)/2 = 6	2
0101	1/6	(5+8)/2 - (2+1)/2 = 5	2.5
0110	1/6	(5+6)/2 - (2+3)/2 = 3	0.5
1001	1/6	(7+8)/2 - (0+1)/2 = 7	0.5
1010	1/6	(7+6)/2 - (0+3)/2 = 5	2.5
1100	1/6	(7+5)/2 - (1+3)/2 = 4	2
$\mathrm{E}[\cdot]$		(6+5+3+7+5+4)/6 = 5	1.67
$\widehat{\mathrm{Var}}(\widehat{ATE})$		1.67	

n-1 * var()/n in R

Randomization-based causal inference (Neyman)

- Toy example with a non-additive ITE.

i	$Y_i(1)$	$Y_i(0)$	$Y_i(1) - Y_i(0)$
1	7	2	5
2	5	1	4
3	6	1	5
4	4	2	2

- Completely randomized experiment with $k = 2$.

$\mathbf{x}$	$\Pr(\mathbf{X} = \mathbf{x})$	$\widehat{\text{ATE}}$	$\widehat{\text{Var}}(\widehat{ATE})$
0011	1/6	3.50	1.25
0101	1/6	3.00	0.50
0110	1/6	3.50	0.25
1001	1/6	4.50	2.25
1010	1/6	5.00	0.50
1100	1/6	5.50	1.25
$\text{E}[\cdot]$		4.17	1.00
$\text{Var}(\widehat{ATE})$		0.81	

Expected var is larger than the true one because it's non-additive

Randomization-based causal inference (Neyman)

- By the finite population CLT and assuming additive ITEs, we can obtain the Wald-type statistic:

$$Z = \frac{\widehat{\mathrm{ATE}} - \mathrm{ATE}}{\sqrt{\widehat{\mathrm{Var}}(\widehat{ATE})}} \to N(0,1)\,,$$

as $n \to \infty$.

Wald test: Basically the same as normal CI e.g. Z test:Specifically used in causal inference to evaluate treatment effects.It ass

- For large samples ($n > 30$), we can obtain a valid $(1-\alpha)\%$ confidence interval as

$$\widehat{\mathrm{ATE}} \pm z_{1-\alpha/2}\sqrt{\widehat{\mathrm{Var}}(\widehat{\mathrm{ATE}})}\,.$$

Large-sample frequentist causal inference

- Assume the data from a randomized experiment $(Y_1, X_1) \ldots, (Y_n, X_n)$ are i.i.d. sampled from an infinite larger population.

- Here potential outcomes are now a random sample, but we can still link them to the observed outcomes assuming causal consistency:

$$Y = Y(1)X + Y(0)(1 - X)\,.$$

- Define ATE using expectations (instead of finite averages):

$$\text{ATE} := \text{E}[\,Y(1) - Y(0)]\,.$$

- Apply standard large sample frequentists methods to draw inference about ATE under additional assumptions.

Barcelona School of Economics

Large-sample frequentist causal inference

- When two random variables X, Y are independent, denoted $X \perp\!\!\!\perp Y$, then $P(X = x, Y = y) = P(X = x)P(Y = y)$, $P(X = x|Y = y) = P(X = x)$ and $P(Y = y|X = x) = P(Y = y)$.

Large-sample frequentist causal inference

- When two random variables X, Y are independent, denoted $X \perp\!\!\!\perp Y$, then $P(X = x, Y = y) = P(X = x)P(Y = y)$, $P(X = x|Y = y) = P(X = x)$ and $P(Y = y|X = x) = P(Y = y)$.

- Strong ignorability assumption on the assignment mechanism:

$$X \perp\!\!\!\perp \{Y(0), Y(1)\}.$$

The assumption of strong ignorability is crucial for identifying causal effects. It says that the treatment a

Large-sample frequentist causal inference

- When two random variables X, Y are independent, denoted $X \perp\!\!\!\perp Y$, then $P(X = x, Y = y) = P(X = x)P(Y = y)$, $P(X = x|Y = y) = P(X = x)$ and $P(Y = y|X = x) = P(Y = y)$.

- Strong ignorability assumption on the assignment mechanism:

$$X \perp\!\!\!\perp \{Y(0), Y(1)\}.$$

- Weak ignorability assumption on the assignment mechanism:

$$X \perp\!\!\!\perp Y(0) \text{ and } X \perp\!\!\!\perp Y(1).$$

Weak ignorability is a slightly relaxed assumption. It states that XY(0)XY(0) and XY(1)XY(1), which means that the treatment a

Large-sample frequentist causal inference

- When two random variables X, Y are independent, denoted $X \perp\!\!\!\perp Y$, then $P(X = x, Y = y) = P(X = x)P(Y = y)$, $P(X = x|Y = y) = P(X = x)$ and $P(Y = y|X = x) = P(Y = y)$.

- Strong ignorability assumption on the assignment mechanism:

$$X \perp\!\!\!\perp \{Y(0), Y(1)\}\,.$$

- Weak ignorability assumption on the assignment mechanism:

$$X \perp\!\!\!\perp Y(0) \;\text{ and }\; X \perp\!\!\!\perp Y(1)\,.$$

- Independence of treatment assignments from potential outcomes do **not** imply $X \perp\!\!\!\perp Y$, independence from observed outcomes, because Y depends on X:

$$Y := Y(1)X + Y(0)(1 - X)\,.$$

Large-sample frequentist causal inference

- Assuming weak ignorability, then $\widehat{\mathrm{ATE}} \xrightarrow{p} \mathrm{ATE}$, i.e., Neyman's $\widehat{\mathrm{ATE}}$ is a consistent (converges in probability) estimator of ATE.

Large-sample frequentist causal inference

- Assuming weak ignorability, then $\widehat{\mathrm{ATE}} \xrightarrow{p} \mathrm{ATE}$, i.e., Neyman's $\widehat{\mathrm{ATE}}$ is a consistent (converges in probability) estimator of ATE.
- Proof (1/4). By causal consistency we can replace Y_i by $Y_i(1)$:

$$\frac{\sum_i^n Y_i X_i}{n} = \frac{\sum_i^n Y_i(1) X_i}{n}.$$

Large-sample frequentist causal inference

- Assuming weak ignorability, then $\widehat{\mathrm{ATE}} \xrightarrow{p} \mathrm{ATE}$, i.e., Neyman's $\widehat{\mathrm{ATE}}$ is a consistent (converges in probability) estimator of ATE.
- Proof (1/4). By causal consistency we can replace Y_i by $Y_i(1)$:

$$\frac{\sum_i^n Y_i X_i}{n} = \frac{\sum_i^n Y_i(1) X_i}{n}\,.$$

- Proof (2/4). By the law of large numbers (LLN):

$$\frac{\sum_i^n Y_i(1) X_i}{n} \xrightarrow{p} \mathrm{E}[\,Y(1)X\,] \;\text{ and }\; \frac{\sum_i^n X_i}{n} \xrightarrow{p} \mathrm{E}[X]\,.$$

Large-sample frequentist causal inference

- Assuming weak ignorability, then $\widehat{\mathrm{ATE}} \xrightarrow{p} \mathrm{ATE}$, i.e., Neyman's $\widehat{\mathrm{ATE}}$ is a consistent (converges in probability) estimator of ATE.
- Proof (1/4). By causal consistency we can replace Y_i by $Y_i(1)$:

$$\frac{\sum_i^n Y_i X_i}{n} = \frac{\sum_i^n Y_i(1) X_i}{n} .$$

- Proof (2/4). By the law of large numbers (LLN):

$$\frac{\sum_i^n Y_i(1) X_i}{n} \xrightarrow{p} \mathrm{E}[Y(1)X] \text{ and } \frac{\sum_i^n X_i}{n} \xrightarrow{p} \mathrm{E}[X] .$$

- Proof (3/4). By weak ignorability:

$$\mathrm{E}[Y(1)X] = \mathrm{E}[Y(1)]\mathrm{E}[X] .$$

Large-sample frequentist causal inference

- Assuming weak ignorability, then $\widehat{\mathrm{ATE}} \xrightarrow{p} \mathrm{ATE}$, i.e., Neyman's $\widehat{\mathrm{ATE}}$ is a consistent (converges in probability) estimator of ATE.
- Proof (1/4). By causal consistency we can replace Y_i by $Y_i(1)$:

$$\frac{\sum_i^n Y_i X_i}{n} = \frac{\sum_i^n Y_i(1) X_i}{n} .$$

- Proof (2/4). By the law of large numbers (LLN):

$$\frac{\sum_i^n Y_i(1) X_i}{n} \xrightarrow{p} \mathrm{E}[Y(1)X] \text{ and } \frac{\sum_i^n X_i}{n} \xrightarrow{p} \mathrm{E}[X] .$$

- Proof (3/4). By weak ignorability:

$$\mathrm{E}[Y(1)X] = \mathrm{E}[Y(1)]\mathrm{E}[X] .$$

- Proof (4/4). By the Slutsky's theorem, assuming a randomized experiment:

$$\mathrm{E}\left[\frac{\sum_i^n Y_i X_i}{\sum_i^n X_i}\right] \xrightarrow{p} \frac{\mathrm{E}[Y(1)]\mathrm{E}[X]}{\mathrm{E}[X]} = \mathrm{E}[Y(1)] .$$

Large-sample frequentist causal inference

- Consider the system of two linear equations:

$$\mathrm{E}[\,Y(x)] = \beta_0 + \beta_1 x \quad \text{for} \quad x = \{0, 1\}\,,$$

such that $\beta_1 = \mathrm{ATE}$.

Large-sample frequentist causal inference

- Consider the system of two linear equations:

$$\mathrm{E}[\,Y(x)] = \beta_0 + \beta_1 x \;\; \text{for} \;\; x = \{0, 1\}\,,$$

  such that $\beta_1 = \mathrm{ATE}$.

- Given data from a randomized experiment, consider the regression model:

$$\mathrm{E}[\,Y|X = x] = \alpha_0 + \alpha_1 x\,.$$

Large-sample frequentist causal inference

- Consider the system of two linear equations:

$$\mathrm{E}[\,Y(x)] = \beta_0 + \beta_1 x \;\text{ for }\; x = \{0, 1\}\,,$$

  such that $\beta_1 = \mathrm{ATE}$.

- Given data from a randomized experiment, consider the regression model:

$$\mathrm{E}[\,Y|X = x] = \alpha_0 + \alpha_1 x\,.$$

- Assuming causal consistency and weak ignorability ($X \perp\!\!\!\perp Y(x)$), then for $x \in \{0, 1\}$,

$$\mathrm{E}[\,Y|X = x] = \mathrm{E}[\,Y(x)|X = x] = \mathrm{E}[\,Y(x)]$$

  implying that $\alpha_1 = \mathrm{ATE}$ and thus $\alpha_1 = \beta_1$.

Large-sample frequentist causal inference

- Consider the system of two linear equations:

$$\mathrm{E}[\,Y(x)] = \beta_0 + \beta_1 x \;\text{ for }\; x = \{0, 1\}\,,$$

  such that $\beta_1 = \mathrm{ATE}$.

- Given data from a randomized experiment, consider the regression model:

$$\mathrm{E}[\,Y|X = x] = \alpha_0 + \alpha_1 x\,.$$

- Assuming causal consistency and weak ignorability ($X \perp\!\!\!\perp Y(x)$), then for $x \in \{0, 1\}$,

$$\mathrm{E}[\,Y|X = x] = \mathrm{E}[\,Y(x)|X = x] = \mathrm{E}[\,Y(x)]$$

  implying that $\alpha_1 = \mathrm{ATE}$ and thus $\alpha_1 = \beta_1$.

- Simple linear regression with a binary explanatory variable on data from a randomized experiment: the estimated slope is the difference in sample means, corresponding to the Neyman's estimator for ATE.

Barcelona School of Economics

Conditional average treatment effects

- The $\text{ATE} := \text{E}[Y(1) - Y(0)]$ can be seen as an *unconditional* average treatment effect.

- Conditional ATEs may be also of interest, particularly the two obtained by conditioning on either assigning or not assigning treatment.

- Average treatment effect for the treated (ATT):

$$\text{ATT} := \text{E}[Y(1) - Y(0)|X = 1] = \text{E}[Y(1)|X = 1] - \text{E}[Y(0)|X = 1]\,.$$

X = 1 = people smoking

Y(1) what is the effect of quiting

Y(0) what is the effect of nott quitting

- Average treatment effect for the controls (ATC):

$$\text{ATC} := \text{E}[Y(1) - Y(0)|X = 0] = \text{E}[Y(1)|X = 0] - \text{E}[Y(0)|X = 0]\,.$$

- The ATT is often more interesting. For instance, what would be the effect of smoking in smoker individuals who quitted smoking.

Barcelona School of Economics

Concluding remarks

- Unit i, treatment/exposure X_i, observed outcomes Y_i, potential outcomes $\{Y_i(1), Y_i(0)\}$.

Concluding remarks

- Unit i, treatment/exposure X_i, observed outcomes Y_i, potential outcomes $\{Y_i(1), Y_i(0)\}$.

- Assumptions: additivity, SUTVA (no-interference, consistency), randomized assignment mechanism, ignorability.

when you do this and assume sutva, association = causation

Concluding remarks

- Unit i, treatment/exposure X_i, observed outcomes Y_i, potential outcomes $\{Y_i(1), Y_i(0)\}$.
- Assumptions: additivity, SUTVA (no-interference, consistency), randomized assignment mechanism, ignorability.
- $\text{ITE} := Y_i(1) - Y_i(0)$ and the fundamental problem of causal inference.

Concluding remarks

- Unit i, treatment/exposure X_i, observed outcomes Y_i, potential outcomes $\{Y_i(1), Y_i(0)\}$.
- Assumptions: additivity, SUTVA (no-interference, consistency), randomized assignment mechanism, ignorability.
- $\text{ITE} := Y_i(1) - Y_i(0)$ and the fundamental problem of causal inference.
- Finite population $\text{ATE} := 1/n \sum_{i=1}^{n} \{Y_i(1) - Y_i(0)\}$, sometimes known as sample average treatment effect (SATE).

Concluding remarks

- Unit i, treatment/exposure X_i, observed outcomes Y_i, potential outcomes $\{Y_i(1), Y_i(0)\}$.
- Assumptions: additivity, SUTVA (no-interference, consistency), randomized assignment mechanism, ignorability.
- $\text{ITE} := Y_i(1) - Y_i(0)$ and the fundamental problem of causal inference.
- Finite population $\text{ATE} := 1/n \sum_{i=1}^{n} \{Y_i(1) - Y_i(0)\}$, sometimes known as sample average treatment effect (SATE).
- Neyman's randomization-based causal inference estimates $\widehat{\text{ATE}}$ by difference in sample means between treated and control.

Barcelona School of Economics

Concluding remarks

- Unit i, treatment/exposure X_i, observed outcomes Y_i, potential outcomes $\{Y_i(1), Y_i(0)\}$.
- Assumptions: additivity, SUTVA (no-interference, consistency), randomized assignment mechanism, ignorability.
- $\text{ITE} := Y_i(1) - Y_i(0)$ and the fundamental problem of causal inference.
- Finite population $\text{ATE} := 1/n \sum_{i=1}^{n} \{Y_i(1) - Y_i(0)\}$, sometimes known as sample average treatment effect (SATE).
- Neyman's randomization-based causal inference estimates $\widehat{\text{ATE}}$ by difference in sample means between treated and control.
- Large-sample frequentist causal inference estimates $\text{ATE} := \text{E}[Y(1) - Y(0)]$, sometimes known as population average treatment effect (PATE). Slope in a simple linear regression with a binary explanatory variable. Conditional ATEs, ATT and ATC, also of interest.