

# ATE vs ATT Explanation



Fill in the blanks.  $T^\dagger = E[Y_1 - Y_0]$  is \_\_\_\_\_ while  $T^* = E[Y_1 - Y_0 | X = 1]$  is \_\_\_\_\_

Help me to solve the problem shown in the picture.

Here are the available options for the blank:

- A. ATT
- B. ATE



To solve this problem, let's break it down:

1.  $T^\dagger = E[Y_1 - Y_0]$ :

This is the **average treatment effect (ATE)**, which measures the expected difference in outcomes ( $Y_1$  and  $Y_0$ ) for the entire population, irrespective of whether they received the treatment or not. So, the blank here is:

**ATE.**

2.  $T^* = E[Y_1 - Y_0 | X = 1]$ :

This represents the **average treatment effect on the treated (ATT)**, which measures the expected difference in outcomes ( $Y_1$  and  $Y_0$ ) only for those who actually received the treatment ( $X = 1$ ). So, the blank here is:

**ATT.**

## Final Answer:

- $T^\dagger = \text{ATE}$
- $T^* = \text{ATT}$



$$E[Y_1 | X = 1] - E[Y_0 | X = 0] = \underbrace{E[Y_1 | X = 1] - E[Y_0 | X = 1]}_{\text{ATT}} + \underbrace{\{E[Y_0 | X = 1] - E[Y_0 | X = 0]\}}_{\text{bias}}$$

Fill in the blanks. On the right of the above equation, the first part is the \_\_\_\_\_ and the second part is the \_\_\_\_\_. Randomization is one of the methods to ensure treatment-control balance which is represented by \_\_\_\_\_ = 0.

Now please solve the problem shown in the picture.

Here are the available options for the blank:

- A. ATT
- B. ATE
- C. bias



Here's the solution based on the equation provided:

## Explanation:

The equation breaks down the difference between  $\mathbb{E}[Y_1|X = 1] - \mathbb{E}[Y_0|X = 0]$  into two parts:

1. **The first part** represents the **Average Treatment Effect on the Treated (ATT)**, which captures the treatment effect for the treated group.
2. **The second part** represents the **bias**, which arises because  $\mathbb{E}[Y_0|X = 1] \neq \mathbb{E}[Y_0|X = 0]$  in observational data. Randomization eliminates this bias, ensuring that the treatment and control groups are comparable.

Randomization ensures that the bias term equals zero, as it balances the treatment and control groups, making  $\mathbb{E}[Y_0|X = 1] = \mathbb{E}[Y_0|X = 0]$ .

## Fill-in-the-Blanks:

- **First part:** ATT
- **Second part:** bias
- **Randomization ensures treatment-control balance:** bias = 0

## Final Answer:

- The first part is the **ATT**.
- The second part is the **bias**.
- Randomization is one of the methods to ensure treatment-control balance, which is represented by **bias = 0**.



Why does the Vietnam Draft Lottery in 1972 allow us to learn about the effects of military service on subsequent civilian earnings?

- Because it created sequential dates of birth from January 1 to December 31 to determine the priority to be called up for military service.
- Because it randomized the dates of birth throughout the year to determine the priority to be called up for military service.
- Because the Vietnam War happened to end before half of the dates of birth were called up for military service.
- Because half of the dates of birth were called up for Vietnam War and the other half were called up for Korean War.



The correct answer is:

**B. Because it randomized the dates of birth throughout the year to determine the priority to be called up for military service.**

## Explanation:

The Vietnam Draft Lottery in 1972 used a randomization mechanism to assign priority for military service based on dates of birth. This random assignment creates a **quasi-experimental setting**, which helps researchers identify the causal effects of military service on subsequent civilian earnings. By randomizing birth dates, the draft lottery effectively eliminated confounding variables that might otherwise affect the relationship between military service and later earnings, making it possible to estimate causal effects.



Based on all of the problems above, please construct an engaging explanation that explains the fancy concepts from scratch behind them. In the explanation, please include the following:

- all necessary fancy math tools and symbols (if any),
- some quantitative tangible examples,
- some humor.



Alright, let's demystify the fancy concepts behind these problems with an engaging explanation, some math, relatable examples, and a touch of humor!

## Causal Inference and the Mystery of What-Ifs

Imagine you're debating whether eating chocolate before bedtime makes you sleep better or worse. One day you eat chocolate, and the next day you don't. You compare your sleep quality, and voilà, you conclude chocolate is magical!

But wait—what if the "no-chocolate day" just coincided with a barking dog keeping you awake? That's the core of causal inference: **how do we isolate the actual cause-and-effect relationship?**

Mathematically, we're chasing  $\mathbb{E}[Y_1 - Y_0]$ , which translates to: **What would happen to your sleep ( $Y_1$ ) if you ate chocolate, minus what would happen if you didn't ( $Y_0$ )?**

## ATE vs. ATT: The Chocolate Chronicles

### ATE (Average Treatment Effect):

This is the average effect of chocolate for *everyone*. Whether you're a chocolate lover or hater, we take your results and average them. Think of it as polling everyone's opinion on chocolate-induced sleep.

### ATT (Average Treatment Effect on the Treated):

This is the effect *specifically for people who ate chocolate*. If you're allergic to chocolate, we don't bother including you—no offense.

### Example:\*\*

If eating chocolate improves sleep by 2 hours for chocolate eaters ( $ATT = 2$ ) but only by 1 hour on average across all people ( $ATE = 1$ ), you know chocolate mainly helps the die-hard chocoholics.

## Vietnam Draft Lottery: The Original Netflix Series

Now, let's time-travel to 1972 and the Vietnam Draft Lottery. Birthdays were randomly drawn to decide who was called for military service. This randomization was like the Netflix algorithm suggesting a show: some people got drafted, others didn't, purely by chance.

This randomness is gold for researchers! Why? It's like having a controlled experiment without the expensive lab coats. The lottery created two groups:

1. **Treatment group:** Those drafted (Netflix suggested *Bridgerton*).
2. **Control group:** Those not drafted (Netflix suggested nothing).

By comparing their civilian earnings years later, researchers estimate how military service affected their income.

Mathematically:

$$\text{Effect} = \mathbb{E}[Y_1 \mid X = 1] - \mathbb{E}[Y_0 \mid X = 1],$$

where  $X = 1$  means drafted. Randomization ensures fairness—like flipping a coin, not a weighted die.

---

## Bias: The Sneaky Villain

Bias is like that one friend who insists pineapple pizza is the best without ever trying other toppings. It skews your results because you're not comparing apples to apples (or pizzas to pizzas).

Example: Suppose we compare chocolate lovers to broccoli lovers to see which group sleeps better. Broccoli lovers might already be health freaks who naturally sleep well, so attributing their good sleep to broccoli is biased.

In math terms:

$$\text{Bias} = \mathbb{E}[Y_0 \mid X = 1] - \mathbb{E}[Y_0 \mid X = 0].$$

Randomization kills bias—like deleting pineapple from the menu and forcing everyone to eat pepperoni instead.

---

## Why Randomization is Like the Ultimate Superhero

Randomization ensures everyone has an equal shot at being treated, whether it's eating chocolate, being drafted, or getting Netflix recommendations. It levels the playing field, making comparisons fair.

Here's the kicker: randomization eliminates bias, so you're left with the *pure effect*. For example, in the Vietnam Draft, we assume birthdays are randomly distributed (no "luckier" days), so any earnings difference between groups can be attributed to military service, not random quirks.

---

## Let's Laugh at Math (Sort of)

- Imagine bias as that one friend who always asks, "Do you like my outfit?" and you can't say no. Randomization? It's like throwing a dart at your closet—it's nobody's fault you wore pajamas to the party.
- ATT vs. ATE? Think of it as:

- ATE: "Does chocolate help everyone sleep?"
- ATT: "Does chocolate help the chocoholics?"

## Closing Example: Chocolate Therapy

Suppose we randomly assign 100 people to eat chocolate and 100 people to abstain. We find:

- Chocoholics sleep 2 hours longer on average ( $ATT = 2$ ).
- Across everyone, the sleep boost averages 1 hour ( $ATE = 1$ ).
- Bias? Zero! Because we randomized the chocolate assignment, no broccoli-loving health freaks were sneaking in.

With this, we've answered the *what-if*: **What if everyone ate chocolate? And what if only chocolate lovers ate it?**

And that, my friend, is how fancy math helps us uncover truths, whether it's about chocolate, military service, or Netflix recommendations. 🎉



To reinforce the understanding, please generate some similar problems and show the solutions.



## Problem 1: Effect of Online Courses on Exam Scores

A university wants to know if taking an online course improves student exam scores. Some students voluntarily take the course, and others do not. Researchers randomly assign **free course access** to a group of students to study the causal effect. The researchers define the following:

- **Treatment Group**: Students who were given free course access.
- **Control Group**: Students who were not given free course access.

Given the following:

- $\mathbb{E}[Y_1]$ : Average exam score if all students took the course.
- $\mathbb{E}[Y_0]$ : Average exam score if no students took the course.
- $\mathbb{E}[Y_1|X = 1]$ : Average exam score of students who were treated (took the course).
- $\mathbb{E}[Y_0|X = 0]$ : Average exam score of students who were not treated.

### Questions:

1. What does  $\mathbb{E}[Y_1 - Y_0]$  represent?
2. What does  $\mathbb{E}[Y_1 - Y_0|X = 1]$  represent?
3. If  $\mathbb{E}[Y_0|X = 1] - \mathbb{E}[Y_0|X = 0] \neq 0$ , what does this imply about the study?

### Solutions:

1.  $\mathbb{E}[Y_1 - Y_0]$ : This is the **Average Treatment Effect (ATE)**, the average impact of taking the course on all students.
2.  $\mathbb{E}[Y_1 - Y_0|X = 1]$ : This is the **Average Treatment Effect on the Treated (ATT)**, the impact of the course on students who were assigned free access.
3.  $\mathbb{E}[Y_0|X = 1] - \mathbb{E}[Y_0|X = 0] \neq 0$ : This implies **bias**. It means the treated and control groups are not directly comparable, possibly due to differences in inherent ability, motivation, or background. Randomization can eliminate this bias.

## Problem 2: Effect of Gym Membership on Weight Loss

A gym wants to understand if providing free memberships helps participants lose weight. Due to limited resources, they randomly assign free gym memberships to half of the applicants and leave the other half on the waitlist.

### Given:

- **Treatment Group:** People given free gym memberships.
- **Control Group:** People on the waitlist.

Data shows:

- $\mathbb{E}[Y_1|X = 1] = 5$ : Average weight loss for the treatment group (in kilograms).
- $\mathbb{E}[Y_0|X = 0] = 2$ : Average weight loss for the control group.
- $\mathbb{E}[Y_0|X = 1] = 3$ : Expected weight loss for the treatment group had they not been treated.

### Questions:

1. Calculate the ATT.
2. Calculate the bias.
3. Does randomization ensure  $\text{bias} = 0$ ? Why?

### Solutions:

1.  $ATT = \mathbb{E}[Y_1|X = 1] - \mathbb{E}[Y_0|X = 1] = 5 - 3 = 2$ .  
The gym helps treated individuals lose an additional 2 kg on average.
2.  $Bias = \mathbb{E}[Y_0|X = 1] - \mathbb{E}[Y_0|X = 0] = 3 - 2 = 1$ .  
Without randomization, the treatment and control groups differ by 1 kg in their expected weight loss.
3. Yes, randomization ensures  $\text{bias} = 0$ , because random assignment balances the two groups, making their expected outcomes comparable.

## Problem 3: Effect of a Job Training Program on Salaries

A job training program is offered to unemployed individuals. Researchers want to know the impact of the program on post-training salaries. They use randomization to assign applicants into two groups:

- **Treated:** Those who undergo the training.

- **Control:** Those who do not undergo the training.

#### Data:

- $\mathbb{E}[Y_1] = 50,000$ : Average salary if everyone undergoes training.
- $\mathbb{E}[Y_0] = 45,000$ : Average salary if no one undergoes training.
- $\mathbb{E}[Y_1|X = 1] = 52,000$ : Average salary for treated individuals.
- $\mathbb{E}[Y_0|X = 1] = 46,000$ : Counterfactual salary for treated individuals had they not undergone training.

#### Questions:

1. Calculate the ATE.
2. Calculate the ATT.
3. Interpret the difference between ATE and ATT in this case.

#### Solutions:

1.  $ATE = \mathbb{E}[Y_1] - \mathbb{E}[Y_0] = 50,000 - 45,000 = 5,000$ .  
On average, training increases salaries by \$5,000 for everyone.
2.  $ATT = \mathbb{E}[Y_1|X = 1] - \mathbb{E}[Y_0|X = 1] = 52,000 - 46,000 = 6,000$ .  
The program increases salaries by \$6,000 for treated individuals.
3. The ATT is higher than the ATE, indicating that the treated group benefits more from the training than the average person. This could imply that the training program is particularly effective for motivated or high-potential individuals.

### Problem 4: Randomization Humor—The Cat Experiment 🐱

You want to test if playing cat videos makes people happier. So you randomly assign participants to:

- **Treatment Group:** Watch 30 minutes of cat videos.
- **Control Group:** No videos (just stare at a blank wall).

#### Data:

- Happiness score for the treated group increases by 10 points.
- For the control group, it remains unchanged.
- $Bias = 0$ , because participants were randomly assigned.

#### Solution:

Randomization ensures fairness: no bias! Watching cat videos improves happiness by 10 points for the treated group, and that's all you need to know. The moral? Randomization is purr-fect! 🐱