Lecture: 04

Introduce to Hypothesis Testing

(PART01)

Topics we will cover

1. What is Hypothesis Testing
2. Null and Alternative Hypothesis
3. Level of Significance α
4. Collect Data using Different Sampling Methods
5. Type I and II Error

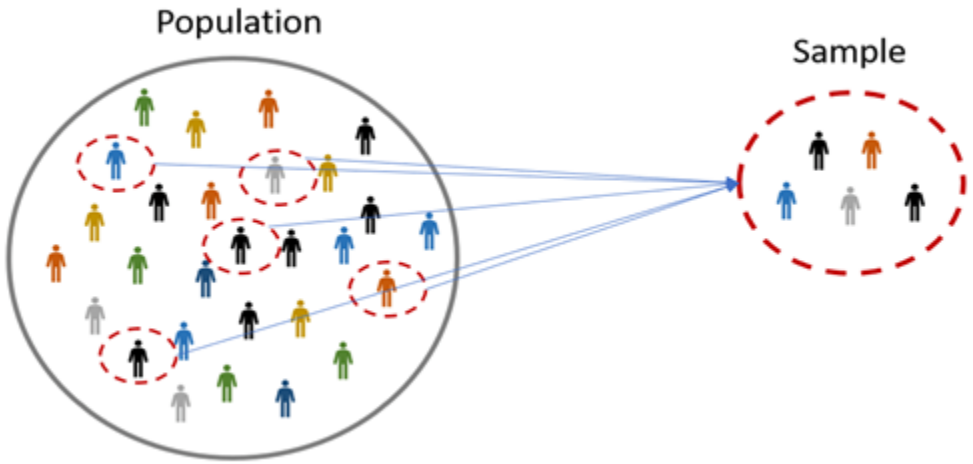
"Data alone is not interesting.
It is the interpretation of the data that we are really interested in."

Hypothesis Testing (Making Informed Decisions with Data)

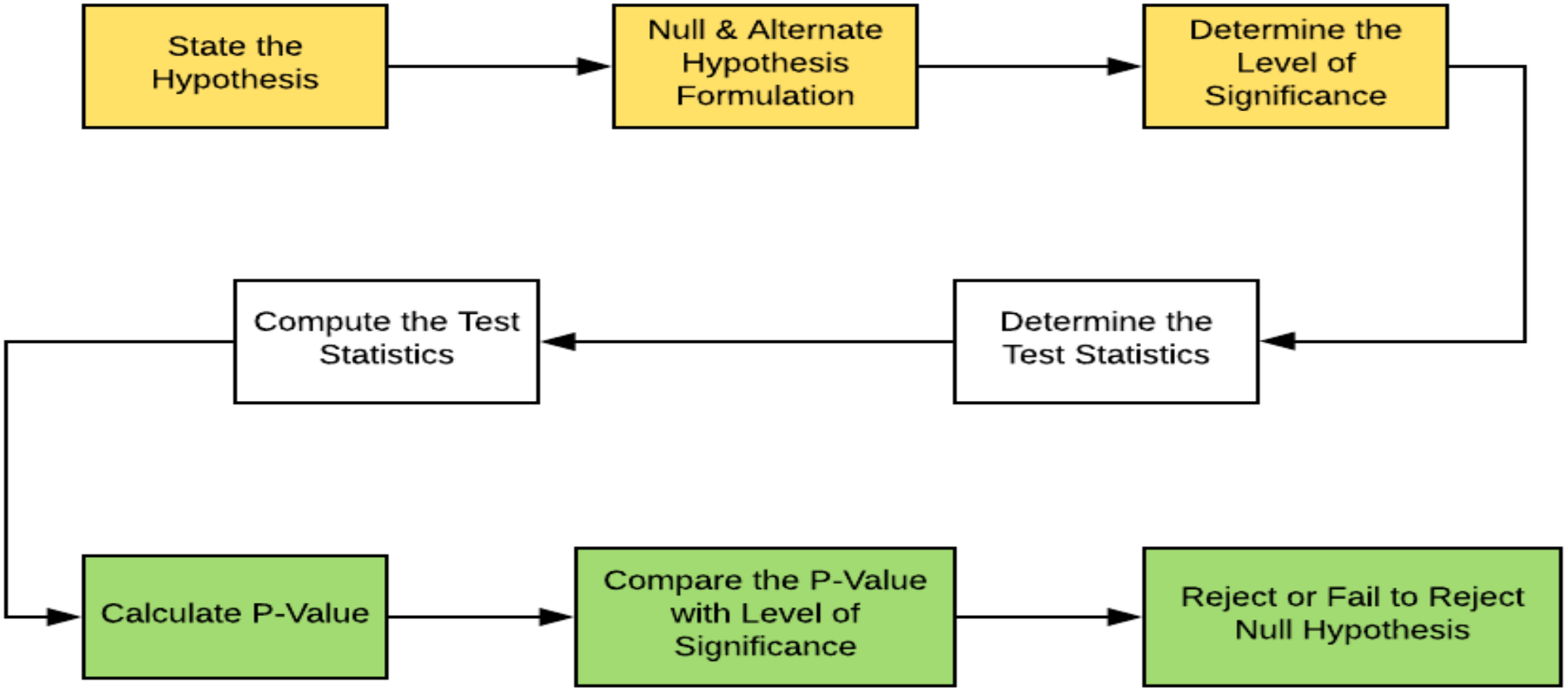
Hypothesis Testing

Hypothesis testing is a statistical method used to make informed decisions or draw conclusions about a *population* based on a *sample of data*.

RECAP:
Population: The entire group that you are interested in studying.
Sample: A subset of the population used for analysis.



Main Steps of Hypothesis Testing



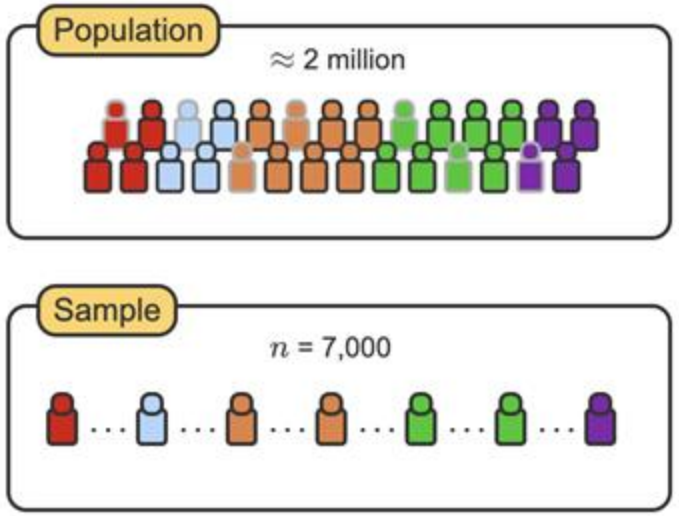
Hypothesis Testing Workflow

Main Steps of Hypothesis Testing

1. State the Null Hypothesis
2. State the Alternative Hypothesis
3. Pick a Level of Significance α
4. Choose a Test
5. Collect Data
6. Calculate a test statistic
7. Calculate P-Value and compare with α
8. Draw a Conclusion

Collect Data

Obtain a representative sample from the population.



Remember the importance of recognizing whether data is collected through an experimental design or observational study.

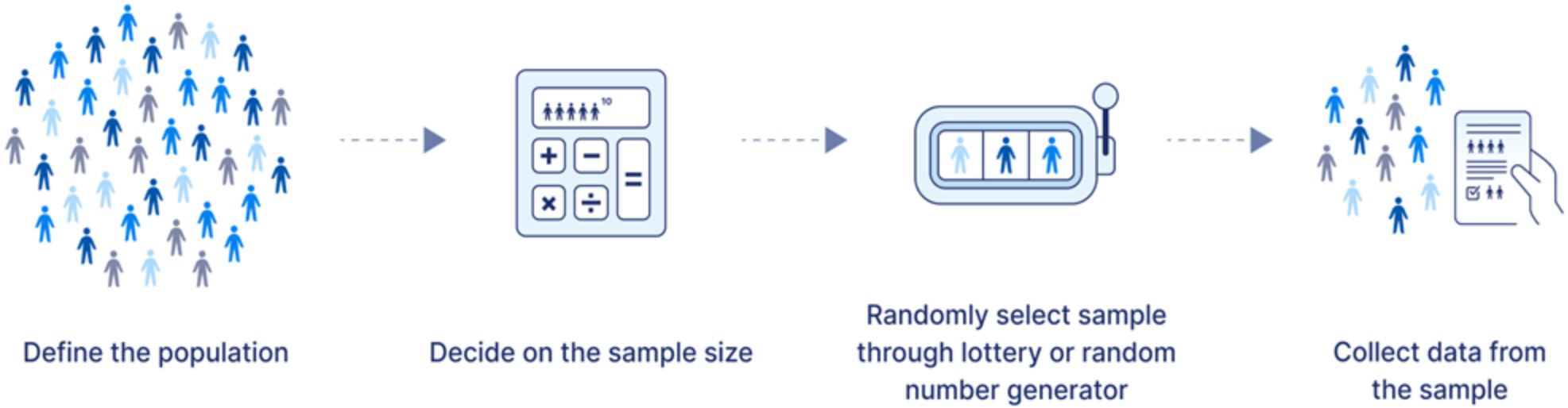
Sampling Methods

A sampling method is a process by which individual items/event (observational units) are selected from the population to be included in the sample. Common sampling methods:

- **Random sampling:** Choosing names randomly from a list for a survey
- **Stratified sampling:** Surveying transportation preferences involves dividing the population into three income strata—low, middle, and high-income households—and then randomly selecting households from each stratum in a city.
- **Cluster sampling:** Estimating the average income of households in a city by dividing the city into clusters based on neighborhoods or districts, then randomly selecting specific neighborhoods as clusters and surveying all households within the selected neighborhoods.
- **Systematic sampling:** Selecting every 10th person from a list of customers
- **Convenience sampling:** Surveying people in a shopping mall (ease of access, availability), Recruiting volunteers from a specific organization or community etc.

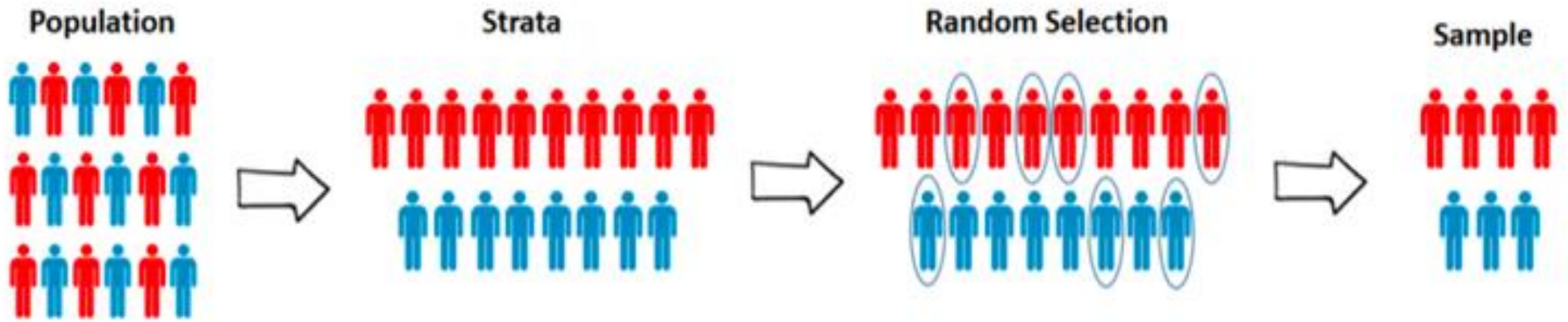
Different Sampling Methods

- **Random sampling:** Observational units are chosen completely at random from the entire population. Each unit has an equal chance of being selected, ensuring that the sample represents the population well.



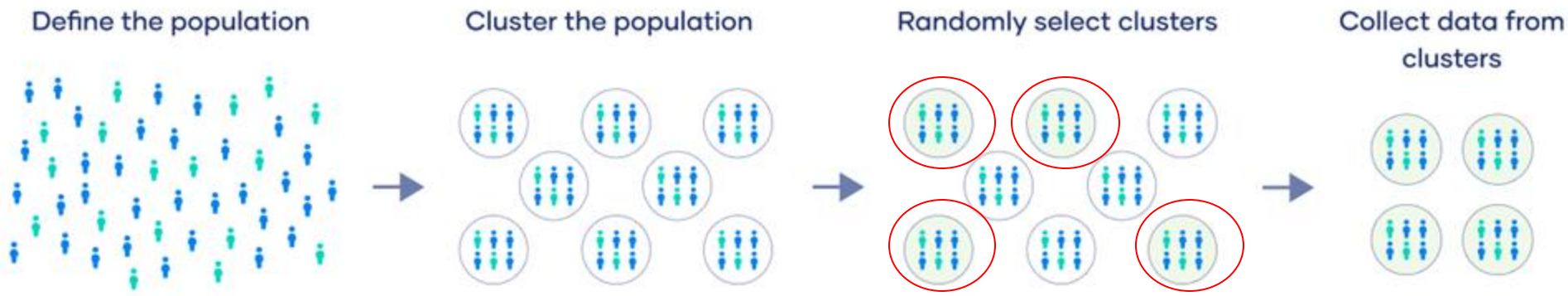
Different Sampling Methods

- **Stratified sampling**: The population is divided into distinct groups or strata based on specific characteristics (e.g., gender, age, race, education level, or income) that are relevant to the study. Then, random samples are taken from each stratum. This ensures representation from all important subgroups in the population, regardless of their individual similarities or differences.



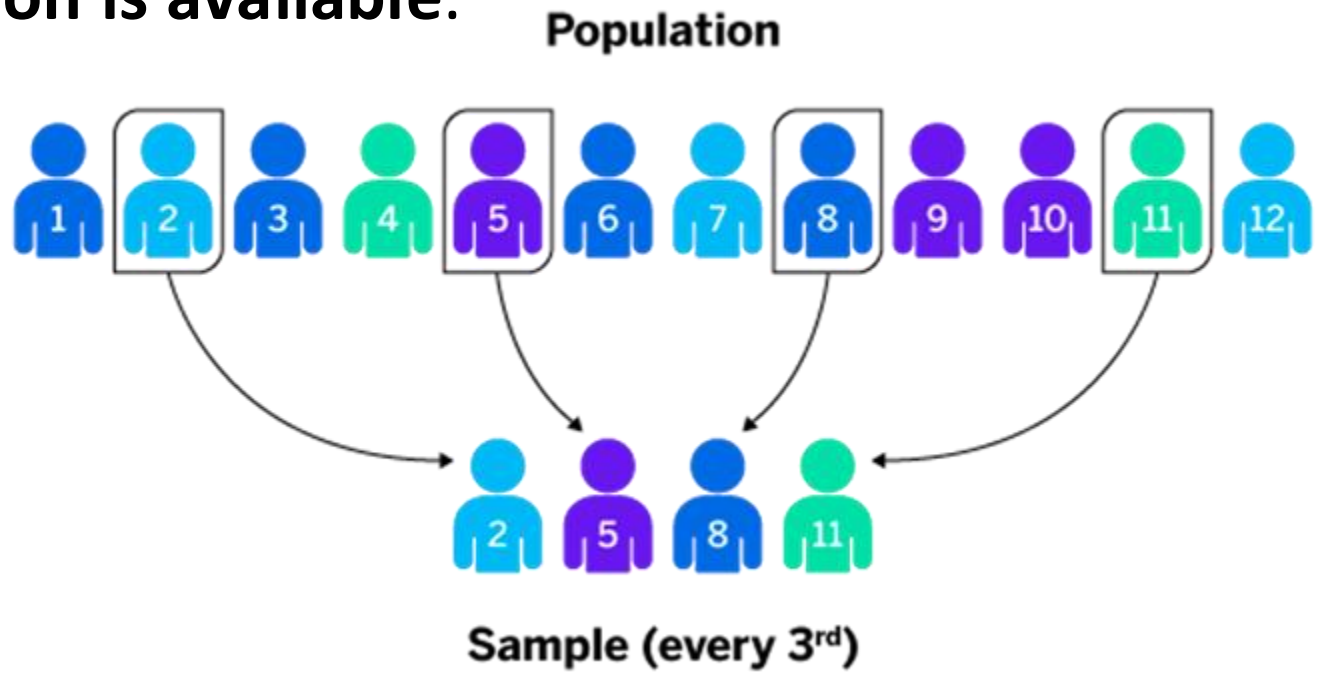
Different Sampling Methods

- **Cluster sampling:** The population is divided into clusters based on a certain characteristic (such as geographic location). Then a random selection of clusters is made, and all units (individuals) within the selected clusters are included in the sample. This method is particularly **useful when obtaining a complete list of the population is difficult**, as clusters provide a more manageable sampling frame.



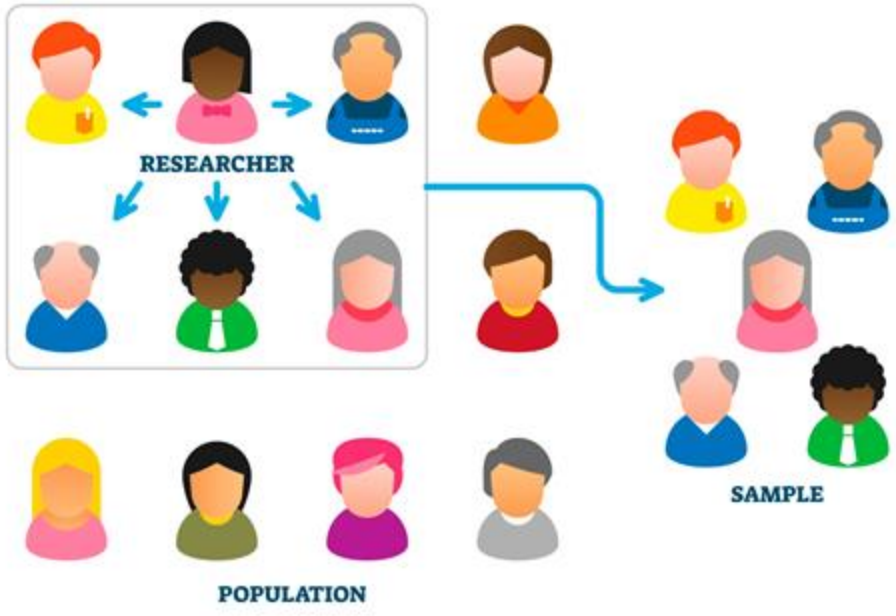
Different Sampling Methods

- **Systematic sampling:** Researchers select a random starting point in the population and then choose every kth unit thereafter until the desired sample size is reached. This method is efficient and easier to conduct compared to simple random sampling, especially **when a complete list of the population is available**.



Different Sampling Methods

- **Convenience sampling**: This method involves selecting units that are readily available and accessible to the researcher. It is a non-probabilistic sampling technique where researchers select individuals who are readily available and accessible to participate in a study.
    - This method involves choosing participants based on their convenience or accessibility, rather than using random selection procedures.
    - While convenient, this method may introduce bias because it may not represent the entire population accurately. It's often used in exploratory research or when other sampling methods are impractical.



STATE/ FORMULATE THE HYPOTHESIS

Hypothesis → A **premise** or **claim** that we want to test/investigate.

Hypothesis

Hypothesis → A **premise** or **claim** that we want to test/investigate.

Why do hypothesis testing? ***Sample mean may be different from the population mean***

It is a statement about one or more populations.

It is usually concerned with the parameters (such as the mean (average), variance (spread), or proportion (percentage)) of the population (help in understanding the overall characteristics of the population).

The hospital administrator may want to **test the hypothesis** that the *average length of stay* of patients admitted to the hospital is **5 days**

Two hypotheses involved in hypothesis testing

1. Null Hypothesis: This is the hypothesis to be tested
   a. Def: A statement of no effect or no difference (currently established).
   b. Represents currently accepted value for a parameter
   c. Suggests that the observed data does not deviate from the population
   d. Denoted as H_o

2. Alternative Hypothesis (also known as Research Hypothesis): Statement you will adopt if the evidence (data) is strong enough to reject the null hypothesis.

   a. Contradicts the null hypothesis.
   b. Suggests that the observed data is DIFFERENT than the population
   c. Represents what you think might be true
   d. Denoted as H_a

Example: H_o and H_a

Let's say, **It is believed** that a candy making machine makes chocolate bars that are on average **5 gram in weight**. A worker claims that the machine after maintenance no longer makes 5 gram bar. Write down H_0 and H_a.

Null Hypothesis H_0 : $\mu = 5$ gram (in avg).

The average weight of the chocolate bars is still 5 grams.

Alternative Hypothesis: H_a : $\mu \neq 5$ gram (in avg).

The average weight of the chocolate bars is **no longer** 5 grams

H_0 and H_a. are mathematical opposite (mutually exclusive)

Possible Outcomes of the Test

Let's say, It is believed that a candy making machine makes chocolate bars that are on average 5 gram in weight. A worker claims that the machine after maintenance no longer makes 5 gram bar. Write down **H0** and **Ha.**

Ho: μ = 5 gram. ; Ha : μ ≠ 5 gram.

This test can lead to TWO possible outcome of the test

Criteria: Based on evidence from sample data.

Possible Outcomes of this test:

Q: How do you do the testing? How do you actually decide to reject the null or not?

1. **Reject** the null hypothesis
2. **Fail to Reject** the null hypothesis→ *not enough evidence*

Do Test Statistics → this is <u>computed from sample data</u> (during the execution of a statistical test) used to decide possible outcome

PRACTICE

WRITING NULL AND ALTERNATIVE HYPOTHESIS

Doctors believe that the average teen sleeps on average of no longer than **10 hours per day**. A researcher believes that teens on average sleep longer.

Write down **H0** and **Ha.**

Choose a Statistical Test and calculate Test Statistics

Choose a Statistical Test:

Select a test based on your research question, hypothesis, data type (categorical or continuous), and number of groups. For example:

- Compare 2 means → *t-test*
- Compare > 2 means → *ANOVA*
- Compare proportions → *Chi-square test* (compare categorical data *i.e.* observed frequencies with expected frequencies)
- Relationship between variables → *Correlation/Regression*

Calculate a Test Statistic:

A numerical value computed from sample data that measures how much the observed results deviate from what is expected under the null hypothesis. It helps determine whether to reject the null hypothesis.

Ex: To compare two group means, we use the t-test to calculate a t-value that measures the difference relative to variability and how strongly the data deviate from the null hypothesis.

The test statistic basically tells you if your data agrees or disagrees with your starting idea (the null hypothesis). It shows how strong your evidence is either for or against your idea.

Q: How do you do the testing?

Let's say, It is believed that a candy making machine makes chocolate bars that are on average 5 gram in weight. A worker claims that the machine after maintenance no longer makes 5 gram bar. Write down **H0** and **Ha.**

Ho: μ = 5 gram. ; Ha : μ ≠ 5 gram.

Possible Outcomes of this test:

1. Reject the null hypothesis
2. Fail to Reject the null hypothesis

Do Test Statistics → this is calculated from sample data used to decide possible outcome

Example: We sample 50 chocolate bars and get average value of mass of the bar.

- Then we calculate test statistics (depends on what type of problem you have)
- It help to determine that the data you have statistically significant enough to reject this null hypothesis or not.

What step helps us **decide whether to reject** the null hypothesis or not, and **how confident** can we be in our decision?

Determine the Significance Level (Alpha):

The **significance level (α)** is a threshold (commonly 0.05) that defines how much evidence we require to reject the null hypothesis

Significance Level

Significance Level: (usually denoted as α) represents is a measure of the strength of the evidence used as a threshold **for deciding whether to reject or fail to reject the null hypothesis in a hypothesis test.**

It is the probability of rejecting the null hypothesis when it's actually true.

- Common values for **α are 0.05 or 0.01**.

The significance level helps you **determine how confident you need to be in your results before** you make a decision.

Significance Level: Example

Let's say you're testing a new treatment for headaches.

Your null hypothesis (H0) : the treatment has no effect, and

Your alternative hypothesis (H1) : the treatment reduces headaches.

- **You set your significance level (α) at 0.05**, meaning **you're willing to accept a 5% chance of mistakenly** concluding that the treatment works when it actually doesn't.

- After conducting your study, let's say you calculate a p-value of 0.03.
  - Since the **p-value is less than α (0.03 < 0.05)**, you reject the null hypothesis and
  - indicates that the observed data is statistically significant, providing enough evidence to support the alternative hypothesis.
    - conclude that the treatment likely reduces headaches.

Back to the previous example

Q: How confident you are with your decision?

Ho: μ = 5 gram. ; Ha : μ ≠ 5 gram.

Possible Outcomes of this test: Reject the null hypothesis or Fail to Reject the null hypothesis

Example: We sample 50 chocolate bars and get average value of mass of the bar.

- Then we calculate test statistics (depends on what type of problem you have)
- It help to determine that the data you have statistically significant enough to reject this null hypothesis or not.

Let's say:

Monday → Received average value of mass of the bar 5.12 gram

Wednesday → Received average value of mass of the bar 5.75 gram

Friday → Received average value of mass of the bar 7.82 gram

How much I am confident that I should reject the Ho or not! → we need a CONCRETE WAY to look at the Ho

Understanding Level of Confidence (C) or Significance (Alpha)

How confident are we in our decision?

We use variable **C (Confidence Level):** Represents the degree of certainty in our decision-making process; could be 95%, 99%.

Level of Significance (α): Indicates the likelihood of making a Type I error, which is rejecting the null hypothesis when it's true.

- **α** is a term used in scientific research to **describe when to reject the null hypothesis.** Often times it is .05 or .01.
- This gives us a **quick way of describing whether or not it is likely an effect exists.**

Understanding Level of Confidence (C) or Significance (Alpha)

How confident are we in our decision?

Interpreting Confidence Level:

- **High Confidence Level (e.g., 99%):** If we're 99% confident in our decision and reject the null hypothesis, we're highly certain that rejecting the null is the correct choice.
- **Low Confidence Level (e.g., 50%):** If our confidence level is only 50%, we're less convinced and might hesitate to reject the null hypothesis.

Understanding Level of Confidence (C) or Significance (Alpha)

How confident are we in our decision?

Remember: Level of significance; alpha=1-C (always sum up to 1)

If Level of Confidence=95% , C=0.95 , so alpha = 1-C =0.95 = 0.05

- C and α are related mathematically but not often used interchangeably.
- While both deal with certainty, they represent different aspects

How to decide C or Alpha value? depends on several factors, including the nature of the research, the consequences of making Type I and Type II errors, and the standard practices in the field. A general guide for choosing the value:

- Evaluate the impact of Type I and Type II errors on your study and adjust alpha or C based on which error is more critical.
- Check if your field typically follows specific alpha (e.g., α = 0.05). or C values, as preferences may vary based on disciplinary standards

Type I and Type II Errors

These errors are important to consider because they affect the validity of the conclusions drawn from hypothesis testing.

Hypothesis testing involves the risk of making errors.

- Type I Error (False Positive): Occurs when you incorrectly reject a true null hypothesis (α) (**We reject the null hypothesis when the it is true.**) The significance level (α) represents the probability of making a Type I error.
- Type II Error (False Negative): Occurs when you fail to reject a false null hypothesis (β). (**we accept the null hypothesis when it is not true**)

	Null Hypothesis is TRUE	Null Hypothesis is FALSE
Reject null hypothesis	Type I Error (False positive)	Correct Outcome! (True positive)
Fail to reject null hypothesis	Correct Outcome! (True negative)	Type II Error (False negative)

Hypothesis Test	TRUE	FALSE
REJECTED	Type I Error	Correct Decision
NOT REJECTED	Correct Decision	Type II Error

Example: What is Type I and II error?

Let's say you're testing whether a new drug is effective in reducing blood pressure.

Your null hypothesis (H0) : the drug has no effect on blood pressure, and

Your alternative hypothesis (H1) : the drug reduces blood pressure on avg.

Collect data from a sample of patients and conduct a statistical test.

Type I error → concluding that the drug is effective at reducing blood pressure when, in reality, it has no effect. (**a false positive result, indicating the drug works when it doesn't.**)

Type II error → failing to detect that the drug is effective at reducing blood pressure when, in fact, it does reduce blood pressure (**a false negative result, indicating the drug doesn't work when it actually does.)**

PRACTICE

WRITING NULL AND ALTERNATIVE HYPOTHESIS

A company stated that their straw machine makes straws that are 4mm in diameter. But a worker believes that the machine no longer makes straws of this size and samples 100 straws to perform a hypothesis test with 99% confidence

Write down **H0**, **H, N, C, alpha.**

Extra Reading Slides

Understanding Level of Confidence (C) or Significance (Alpha)

How confident are we in our decision? We use variable C → A variable C ; could be 95%, 99%.

Level of significance; alpha=1-C

The significance level (alpha)is a term used in scientific research to describe when to reject the null hypothesis. Often times it is .05 or .01. This gives us a quick way of describing whether or not it is likely an effect exists.

If Level of Confidence=95% , C=0.95 , so alpha = 1-C =0.95 = 0.05

WRITING NULL AND ALTERNATIVE HYPOTHESIS

A company stated that their straw machine makes straws that are 4mm in diameter. But a worker believes that the machine no longer makes straws of this size and samples 100 straws to perform a hypothesis test with 99% confidence

- Ho : $\mu=4mm$
- Ha ; $\mu\neq4mm$

- N=100
- C= 0.99
- α= 0.01 (α =1 - C)

WRITING NULL AND ALTERNATIVE HYPOTHESIS

The school board claims that atleast 60% of students bring a phone at school. A teacher believe that this number is too high and randomly samples 25 students to test at at level of significance of .02

$H_o : p \geq 0.60$
$H_a : p < 0.60$

$N=25$
$\alpha=0.02$
$C=0.98$ $(C=1- \alpha)$

Try

A grocery store assumes that the average shopper spends no more than $100 in the store. The new store manager claims that they spend more. He chooses 27 shoppers randomly and they spend an average of $104.93 with a standard deviation of$9.07. Test the manager claims at 0.05 significance level? In this question only use the critical value.

RECAP:

A hypothesis test is a method that evaluate population claims (or hypothesis) using sample data.

- The conclusion of a hypothesis test determines if the observed data support the claim or suggest an alternative explanation.
  - The null hypothesis typically represents the claim of no effect or no difference
    - E.g., The new drug has no effect on blood pressure.
  - The alternative hypothesis represents the claim of an effect or difference.
    - E.g., The new drug reduces blood pressure.
- Now to evaluate hypotheses about population parameters based on sample data, we apply different types of test statistics (TODAY's topic) which return us p-values or probability values.
- Draw a conclusion about the null hypothesis based on the statistical evidence provided by the p-value

The steps involved in the Hypothesis Testing