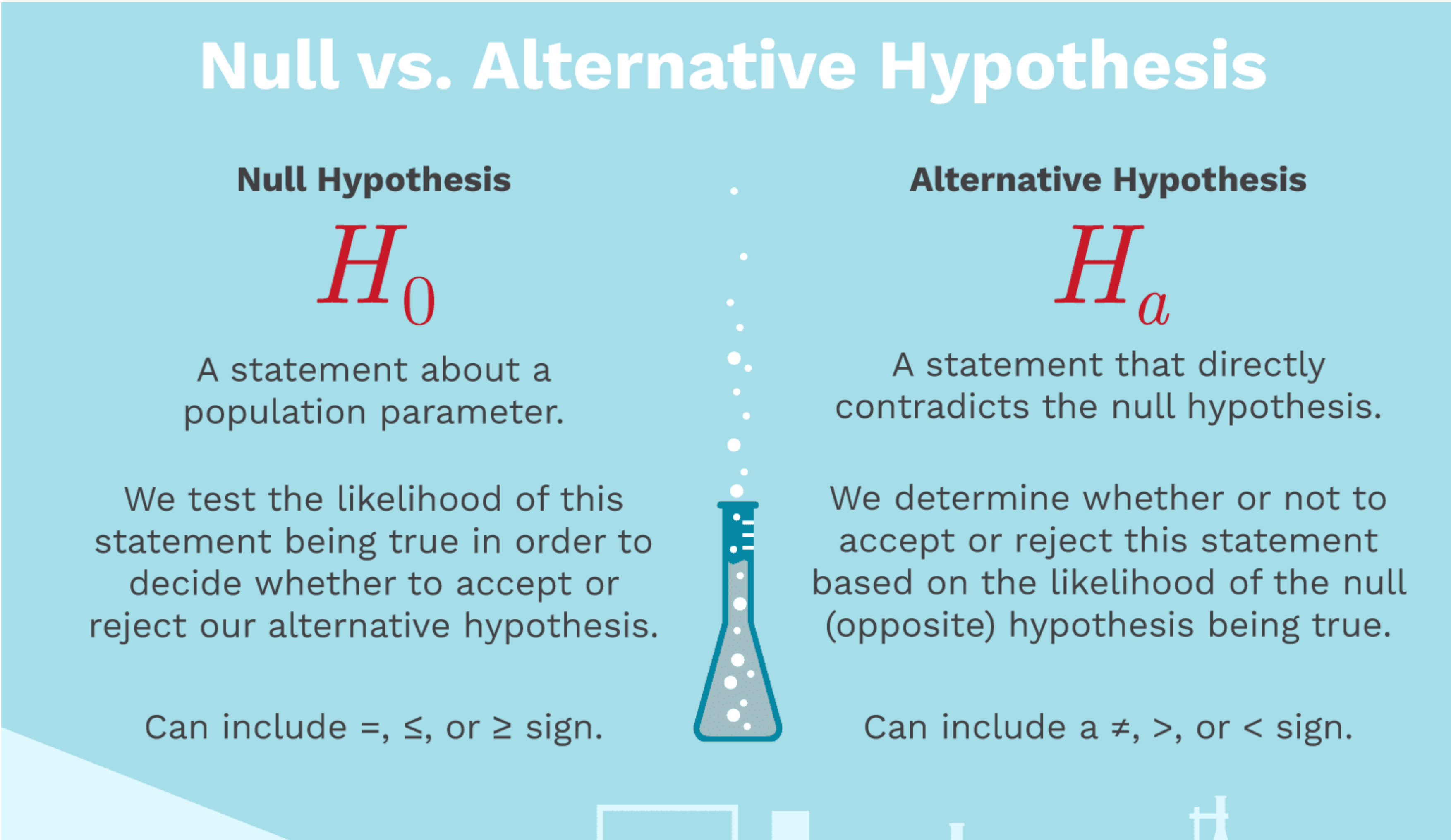
WHAT IS A HYPOTHESIS?

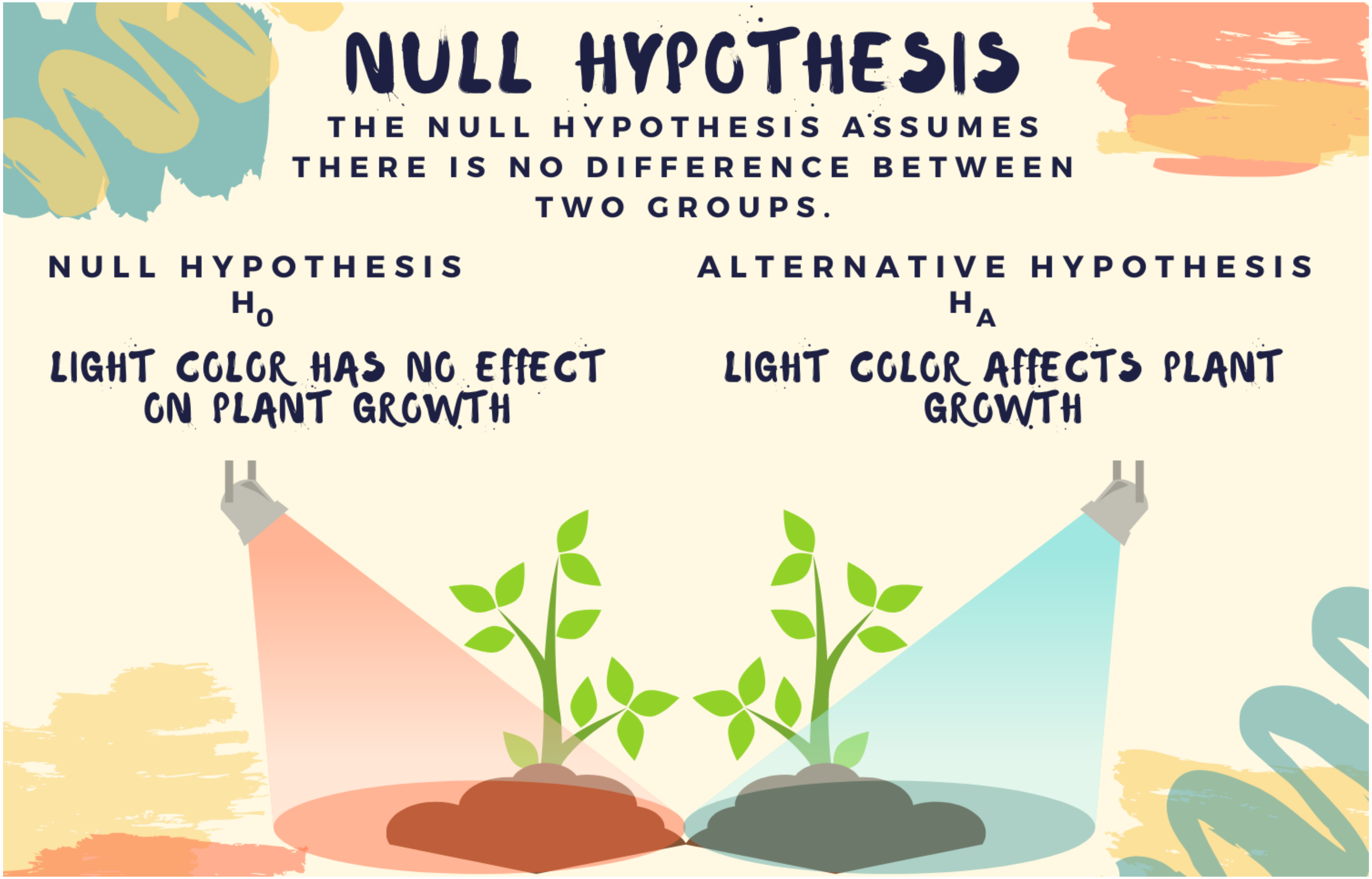
A hypothesis is a precise, testable statement of what the researcher predicts will be the outcome of the study. It is stated at the start of the study.

This usually involves proposing a possible relationship between **two variables**: the **independent variable** (what the researcher changes) and the **dependent variable** (what the research measures).

A fundamental requirement of a hypothesis is that is can be tested against reality, and can then be **supported or rejected**.


Null vs. Alternative Hypothesis
Null Hypothesis
H_0
A statement about a population parameter.
We test the likelihood of this statement being true in order to decide whether to accept or reject our alternative hypothesis.
Can include =, ≤, or ≥ sign.
Alternative Hypothesis
H_a
A statement that directly contradicts the null hypothesis.
We determine whether or not to accept or reject this statement based on the likelihood of the null (opposite) hypothesis being true.
Can include a ≠, >, or < sign.

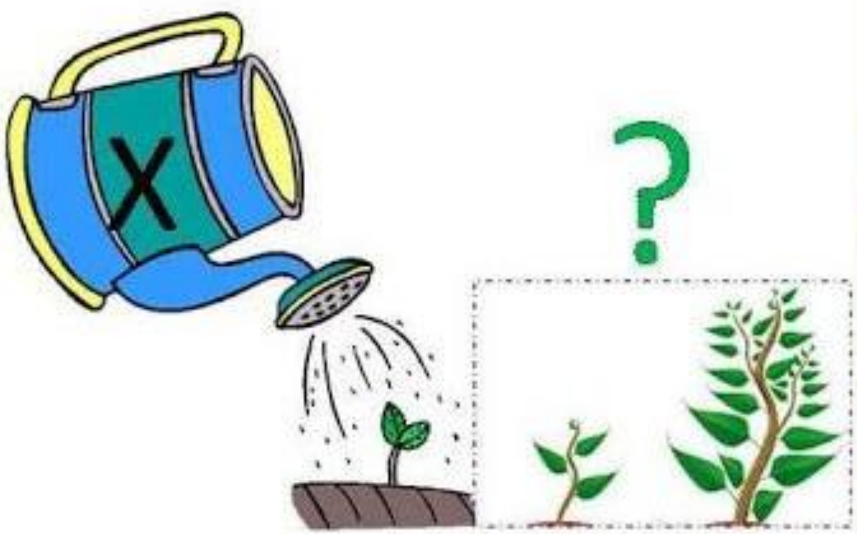
Unit –VI : Test of Hypothesis



Effect of Bio-fertilizer 'x' on Plant growth

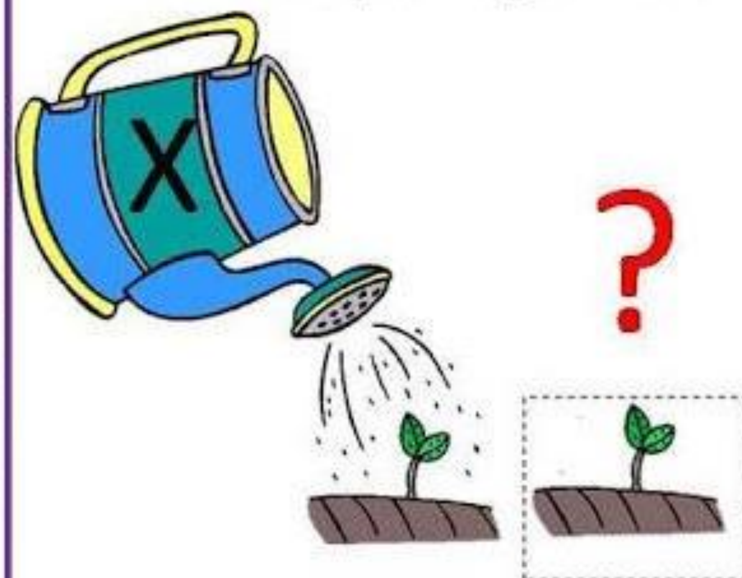
Alternative Hypothesis

H_1: Application of bio-fertilizer 'x' increase plant growth.



Null Hypothesis

H_0: Application of bio-fertilizer 'x' do not increase plant growth.



What is a Hypothesis Test?

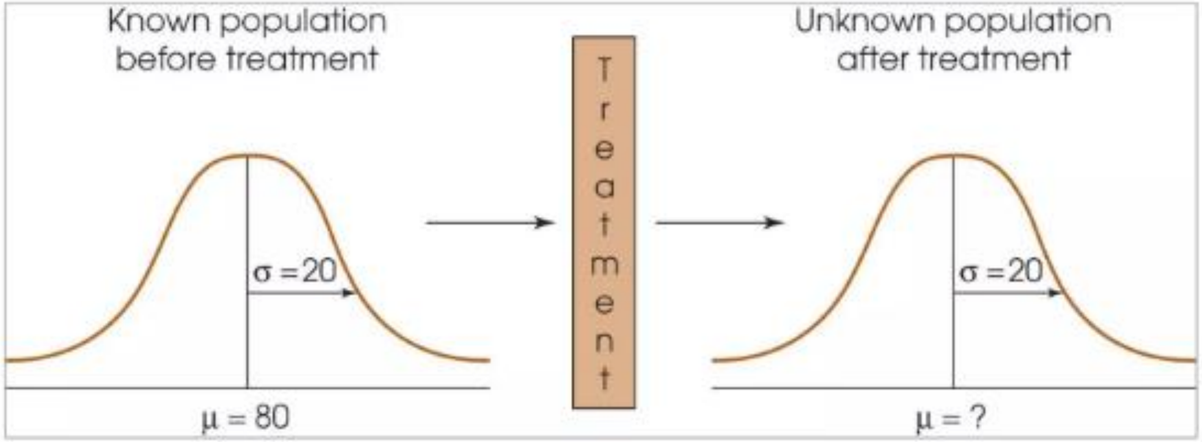
- A ***hypothesis test*** is a statistical method that uses sample data to evaluate a hypothesis about a population
  - A standardized methodology we use to answer a research question of interest
  - The most commonly utilized inferential procedure
  - Incorporates the concepts of z-scores, probability, and the distribution of sample means
- The general goal of the hypothesis test:
  - To rule out chance (sampling error) as a plausible explanation for the results from a research study

The Logic of Hypothesis Testing

1. We have a hypothesis about a population
   - Typically about a population parameter: $\mu = 50$
2. Based on our hypothesis, we predict the characteristics our sample *should* have
   - *M* should be *around* 50
3. Obtain a random sample from the population
   - Sample $n = 20$ from the population and compute their *M*
4. Compare the sample statistic with our hypothesis about the population
   - If the values are consistent, the hypothesis is reasonable
   - If there is a large discrepancy, we decide the hypothesis is unreasonable

The Unknown Population

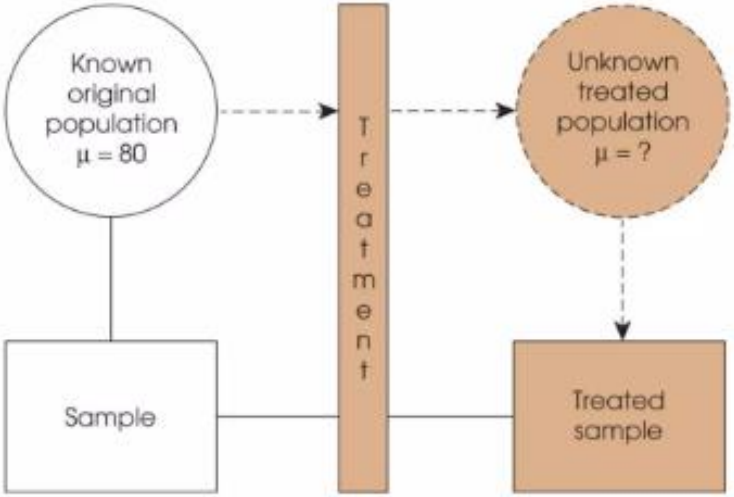
- Typically, research involves an ***unknown population***



- After we administer "tutoring treatment" to our sample of students, what is the new μ?
  - Remember, σ should remain the same
- The focus of our hypothesis is the *unknown population*

The Research Study Sample

- Theoretically:
  - We would treat the entire population, randomly sample from the treated population, and evaluate the *M*



- Realistically:
  - We sample from the original population, treat the sample only, and evaluate the *M*

The Purpose of the Hypothesis Test

To decide between two explanations:

1.	2.
There **does not** appear to be a treatment effect	There **does** appear to be a treatment effect
The difference between the sample and the population *can be explained by sampling error*	The difference between the sample and the population *is too large to be explained by sampling error*

The Hypothesis Test: Step 1

State the hypothesis about the unknown population

- The **null hypothesis** (H_0)
  - States that there is **no change** in the general population before and after an intervention.
  - In the context of an experiment, H_0 predicts that the independent variable (IV) will have **no effect** on the dependent variable
- The **scientific/alternative hypothesis** (H_1)
  - States that there **is** a change in the general population following an intervention
  - In the context of an experiment, H_1 predicts that the independent variable (IV) **will have an effect** on the dependent variable

SAT Example

- Population parameters of a normally distributed variable are $\mu = 500$ and $\sigma = 50$
- I believe my prep course will have some effect on this mean value
- My hypotheses:
  - Null (H_0): Even with treatment, the mean SAT score will ***remain at*** 500
  - Alternative (H_1): With treatment, the mean SAT score will ***be different from**** 500

$$H_0 : \mu_{SAT} = 500$$

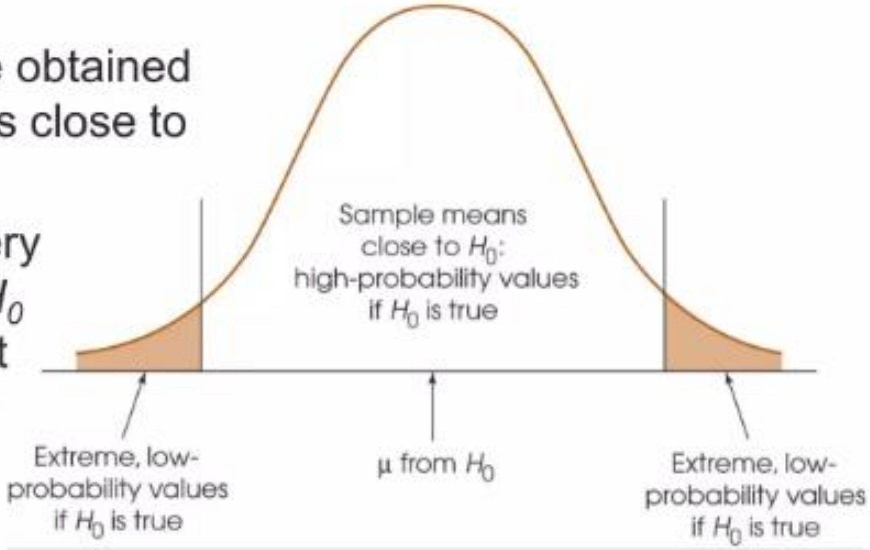
$$H_1 : \mu_{SAT} \neq 500$$

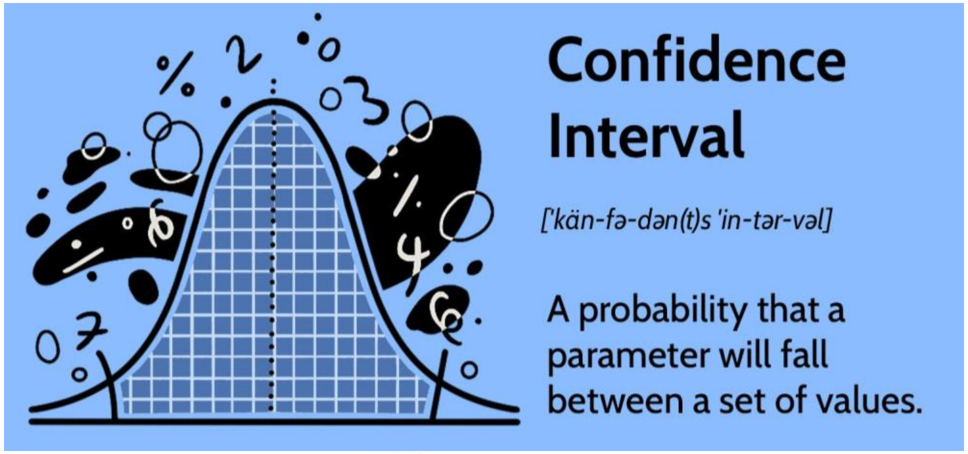
SURESH GYAN VIHAR UNIVERSITY
Accredited by NAAC with 'A' Grade

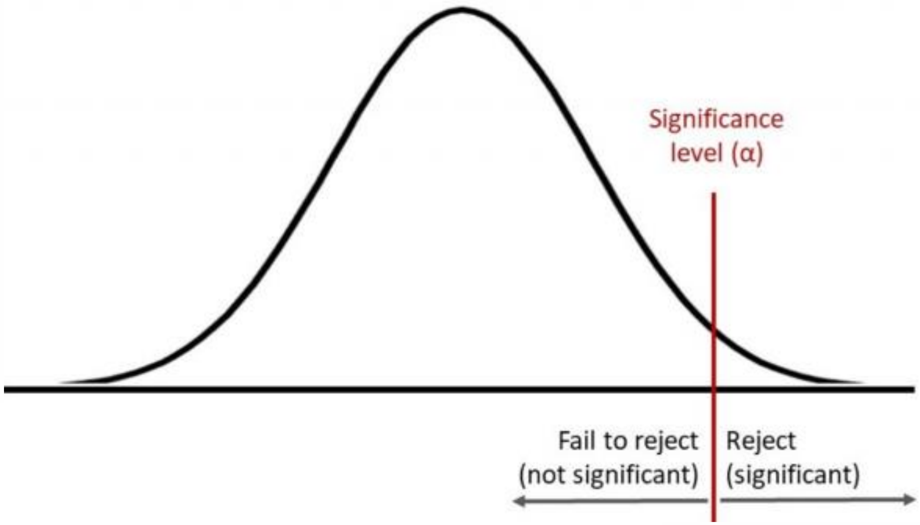
The Hypothesis Test: Step 2

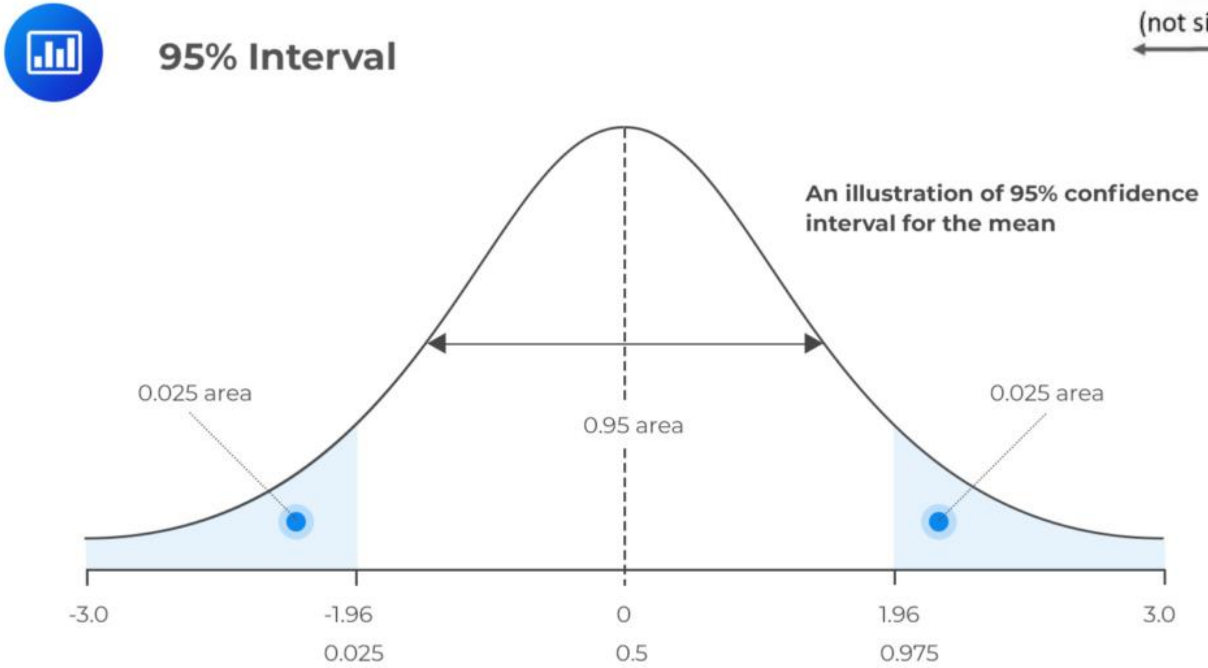
Set the criteria for a decision

- Determine before the experiment what sample means are inconsistent with the null hypothesis
- The sampling distribution of the mean for the population is divided into two parts:
  1. Sample means likely to be obtained if H_0 is true (sample means close to the null value)
  2. Sample means that are very unlikely to be obtained if H_0 is true (sample means that are very different from the null value)




Confidence Interval
['kän-fə-dən(t)s 'in-tər-vəl]
A probability that a parameter will fall between a set of values.

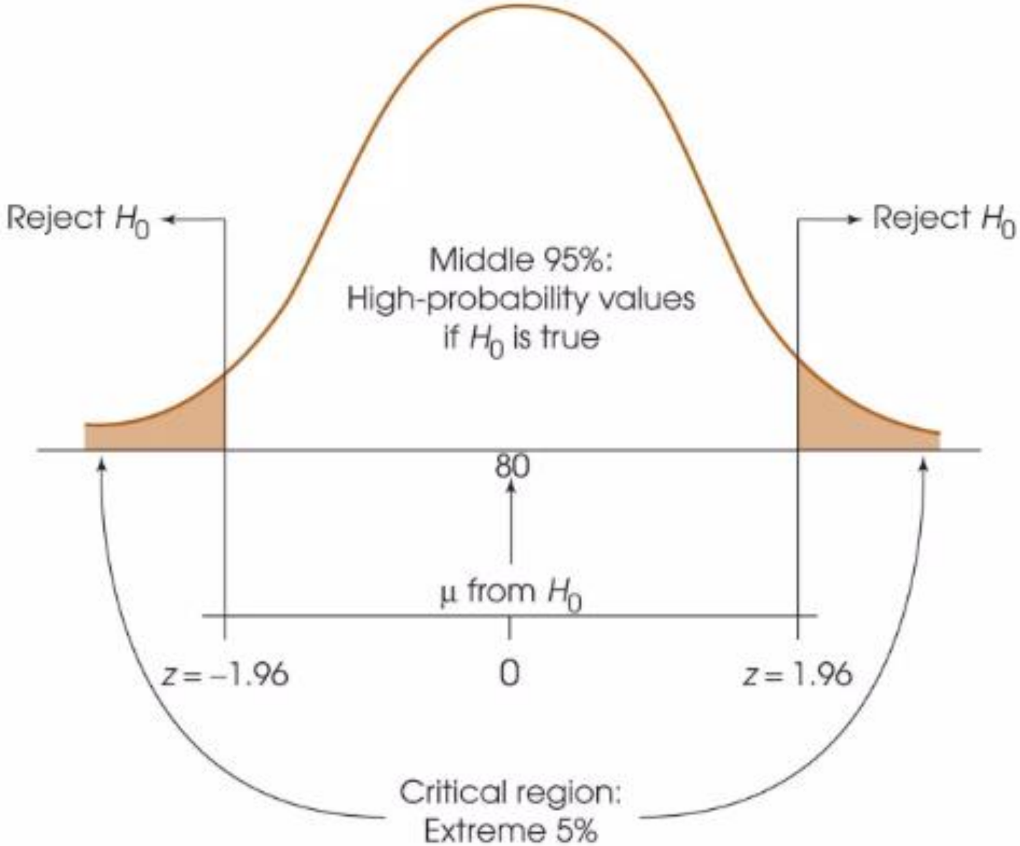

Significance level (α)
Fail to reject (not significant)
Reject (significant)


95% Interval
An illustration of 95% confidence interval for the mean
0.025 area
0.95 area
0.025 area
-3.0
-1.96
0
1.96
3.0
0.025
0.5
0.975

The Alpha (α) Level

a.k.a. the Level of Significance

- How we define which sample means have "high" or "low" probabilities
- Typically set at:
  - $\alpha = 0.05$,
  - $\alpha = 0.01$
  - $\alpha = 0.001$
- Separates the typical values from the extreme, unlikely-to-occur values (in the ***critical region***)



The α level establishes the criterion, or "cut-off", for making a decision about the null hypothesis. The alpha level also determines the risk of a Type I error (TBD in next section)

The Hypothesis Test: Step 3

Collect data and compute sample statistics

- *After* we state they hypotheses and establish our critical regions:
  - Randomly sample from the population
  - Give the sample the treatment/intervention
  - Summarize the sample with the appropriate statistic (e.g.: the *M*)
- Compute the test statistic

$$z = \frac{M - \mu}{\sigma_M}$$

The z-score (the test statistic) forms a ratio comparing the obtained difference between the sample mean and the hypothesized population mean versus the amount of difference we would expect without any treatment effect (the standard error)

The Hypothesis Test: Step 4

Make a decision

- Use the z-score obtained for the sample statistic and make a decision about the null hypothesis (H_0) based on the critical region. Either:
  - *Reject the null hypothesis*
    - The sample data fall in the critical region
    - This event is unlikely if the null hypothesis is true, so we conclude to reject the null
    - This does *NOT* mean that we have proven the alternative
  - *Fail to reject the null hypothesis* (retain the null)
    - The sample data do not fall in the critical region
    - We do not have enough evidence to show that the null hypothesis is incorrect (*NOT* that the null hypothesis is "true")

Making your decision: One mo' time!

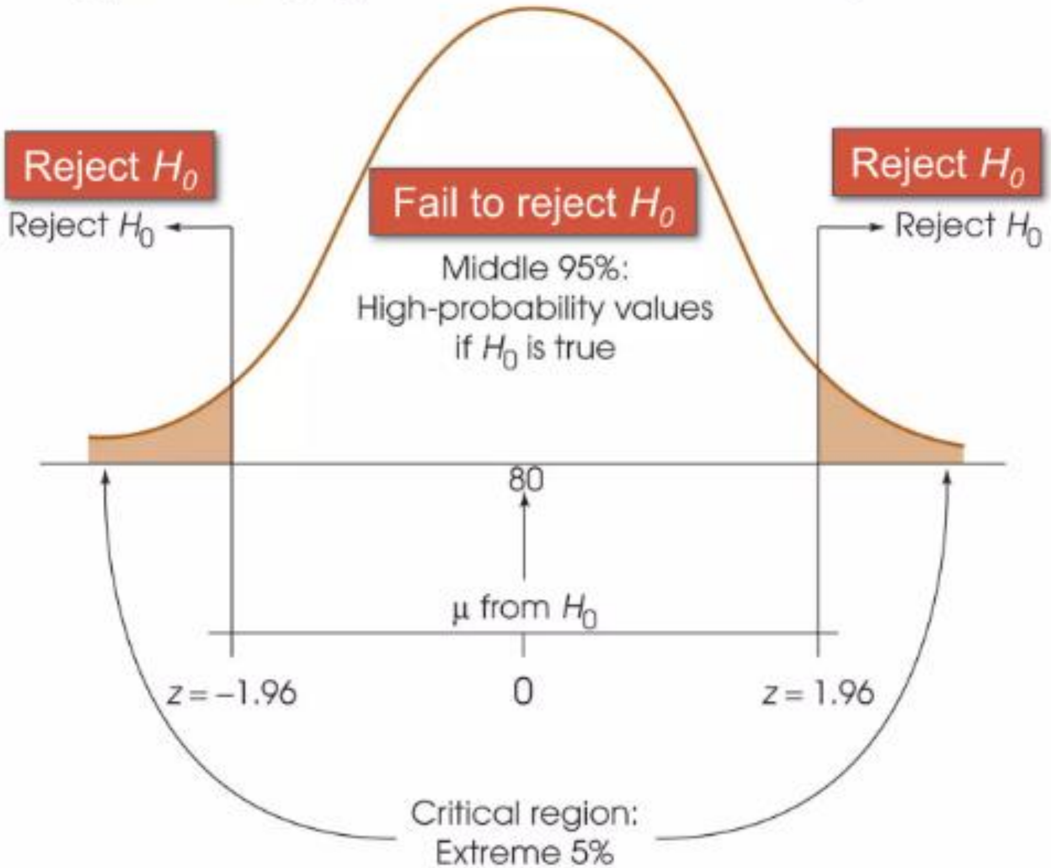
IF: My sample mean has a corresponding z = 0.87

- My treatment effect is not convincing. I do not have enough evidence and fail to reject the null.

IF: My sample mean has a corresponding z = -2.50

- This is an unlikely event if the null is true, so I reject the null. I have convincing evidence my treatment works and has an effect.

- Use the z-score obtained for the sample statistic and make a decision about the null hypothesis (H_0) based on the critical region



An Analogy for Hypothesis Testing

Research Study

1. H_0: There is no treatment effect
2. Researchers gather evidence to show that the treatment actually does have an effect

Jury Trial

1. H_0: Defendant did not commit a crime
2. Police gathers evidence to show that the defendant really did commit a crime

Both are trying to refute the null hypothesis (H_0)

Research Study

3. If there is enough evidence, the researchers reject H_0 and conclude that there is a treatment effect
4. If there is *not* enough evidence, the researcher fails to reject H_0

They are not concluding that there is no treatment effect; simply that there is not enough *evidence* to conclude there is an effect

Jury Trial

3. If there is enough evidence, the jury rejects H_0 and concludes that the defendant is guilty of the crime
4. If there is not enough evidence, the jury fails to find the defendant guilty

They are not concluding that the defendant is innocent; simply that there is not enough *evidence* for a guilty verdict

Statistical Significance

A result is statistically significant if it is unlikely to occur when the null hypothesis (H_0) is true
(Sufficient to reject the null hypothesis [H_0])

- An effect is significant if we decide to reject the null hypothesis after conducting a hypothesis test
  - $Z = 0.87$, $p > 0.05$
    does not fall in the critical region, fail to reject the null
  - $Z = -2.50$, $p < 0.05$
    falls in the critical region, reject the null

The z-Score as a Test Statistic

- **A Test Statistic** = A single, specific statistic (converted from sample data) that is used to test the hypothesis

$$z = \frac{M - \mu}{\sigma_M}$$

- **The z-score formula as a recipe**
  - Make a hypothesis about the value of μ. This is H_0.
  - Plug the hypothesized value in the formula along with the other values
  - If the results is a z-score near zero (where z-scores are supposed to be), fail to reject H_0.

SURESH GYAN VIHAR UNIVERSITY
Accredited by NAAC with 'A' Grade

The z-Score Formula as a Ratio

$$z = \frac{M - \mu}{\sigma_M} \Leftrightarrow \frac{\text{Sample mean – hypothesized population mean}}{\text{Standard error between } M \text{ and } \mu}$$

OR

$$\frac{\text{Obtained difference}}{\text{Difference due to chance}} = \frac{3}{1} = 3$$

If z is large, the obtained difference is larger than expected by chance

- For example:
  z = 3 indicates the obtained difference is three times larger than what would be expected by chance

Type I Errors (α)

- A **Type I error** occurs when the sample data appear to show a treatment effect when, in fact, there is none
  - We **reject** a null hypothesis (H_0) that is actually true
  - We falsely conclude that the treatment has an effect
- Caused by unusual, unrepresentative samples that fall into the critical region despite a null effect of treatment
- Type I errors are very unlikely

Type II Errors (β)

- A **Type II error** occurs when the sample does not appear to have been affected by the treatment when, in fact, it has (the treatment does have an effect)
  - We **fail to reject** the null hypothesis (H_0) that is actually false
  - We falsely conclude that the treatment had no effect

- Commonly the result of a very small treatment effect
  - Although the treatment does have an effect, it is not large enough to show up in our data

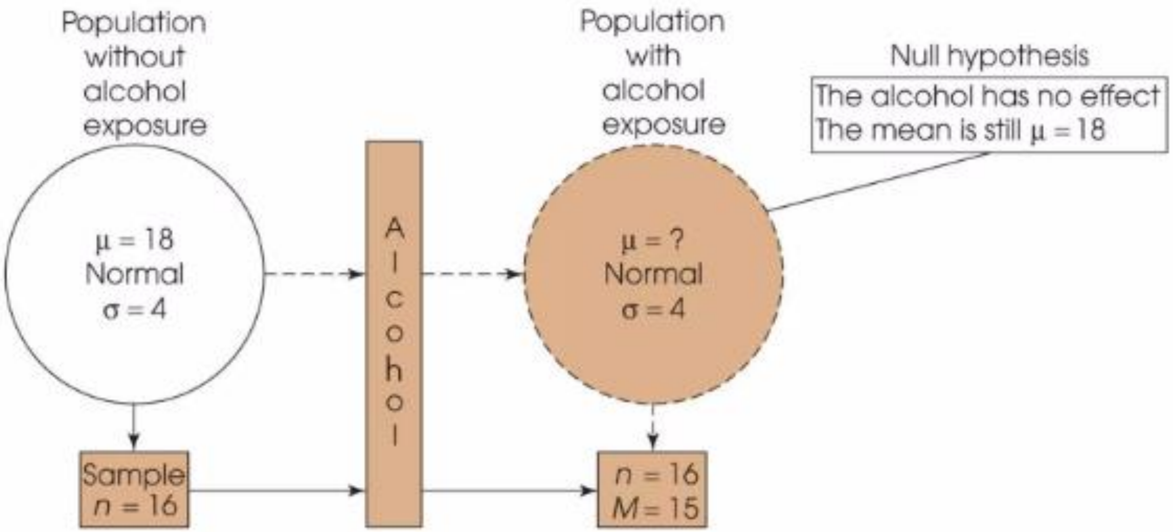
The Relationship Between α & Errors

		Actual Situation	
		No Effect, H_0 True	Effect Exists, H_0 False
Experimenter's Decision	Reject H_0	Type I error	Decision correct
	Retain H_0	Decision correct	Type II error

- The alpha (α) level is the probability that the test will lead to a Type I error
  - *If α = .05: The probability that the test will lead to a Type I error is 5%*
  - *If α = .01: The probability that the test will lead to a Type I error is 1%*
- How does changing α influence Type II errors?
  - As we make α smaller, we decrease the chance of making a Type I error.
  - However, we also make it harder to reject the null; thus increasing the risk of a Type II error

The Study:
Prenatal alcohol exposure on birth weights

A random sample of n = 16 pregnant rats is obtained. The mother rats are given a daily dose of alcohol. At birth, one pup is selected from each litter to produce a sample of n = 16 newborn rats. The average weight for the sample is M = 15 grams. It is known that regular newborn rats have an average weight of μ = 18 grams, with σ = 4



Step 1

State the hypotheses and select the alpha (α) level

$$H_0 : m_{alcohol\,\mathrm{exp}} = 18$$

Even with alcohol exposure, the rats still average 18 grams at birth

or

Alcohol exposure does not have an effect on birth weight

$$H_1 : m_{alcohol\,\mathrm{exp}} \;^{1}\; 18$$

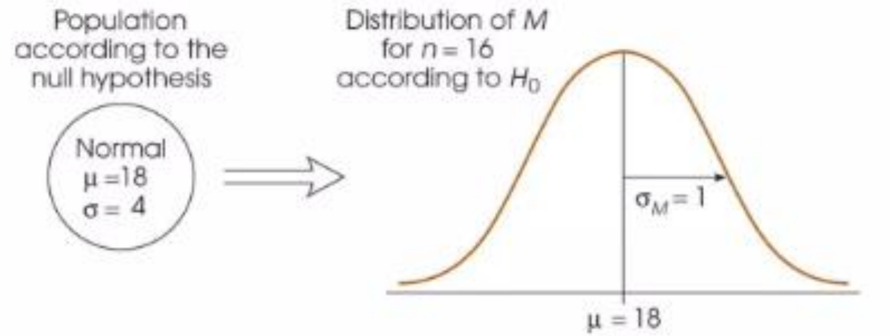
Alcohol exposure will change birth weight

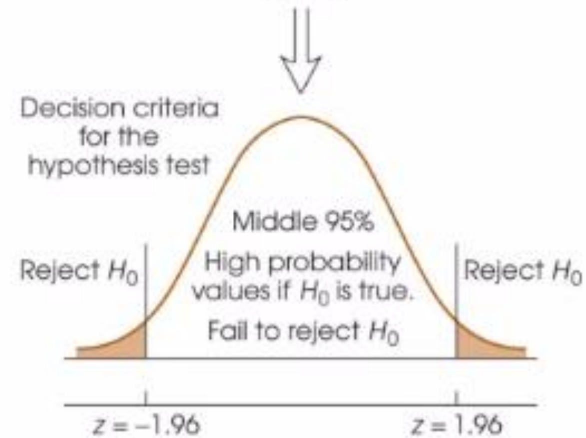
- For this test, we will use an alpha level of α = .05
  - We are taking a 5% risk of committing a Type I error

Step 2

Set the decision criteria by locating the critical region

- The 3-step process:





If we obtain a sample mean that is in the critical region, we conclude that the sample is not compatible with the null hypothesis and we reject H_0

Step 3

- Collect the data, and compute the test statistic
- In this case, the z-score:

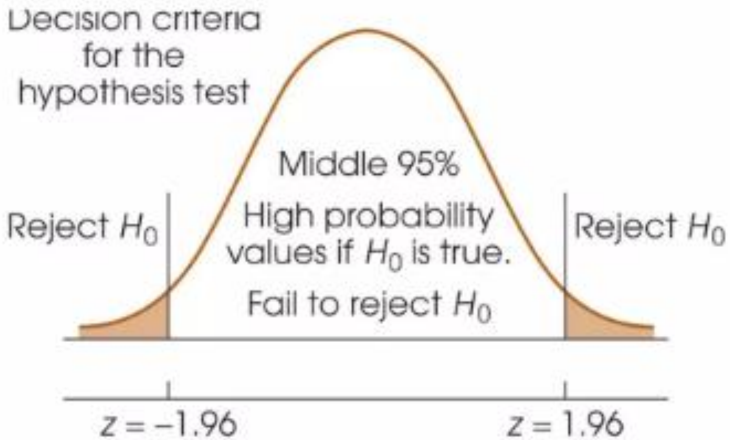
$$z = \frac{M - \mu}{\sigma_M} = \frac{15 - 18}{\left(\frac{4}{\sqrt{16}}\right)} = \frac{-3}{\left(\frac{4}{4}\right)} = \frac{-3}{1} = -3.00$$

Step 4

Make a decision

- Our z-score (-3.00) exceeds the critical value of +/-1.96
  - It is in the critical region
- Our interpretation:
  - This event is unlikely if the null is true

Reject the null hypothesis. Alcohol does have a significant effect on the birth weight of rats.



Things to Keep In Mind

- State the hypotheses:
  - Your hypotheses lays out the whole *point* of the test
  - Phrased in terms of the *population parameters*
- When setting critical values:
  - What are they?
  - Are they positive? Negative? Both?
- Results
  - What do we conclude about the null?
  - State your results in scientific language:
    The treatment with alcohol had a significant effect on the birth weight of newborn rats, $z = -3.00$, $p < .05$.
- Conclusions
  - What do we conclude about the point of the study?

Factors Influencing a Hypothesis Test

1. The size of difference between *M* and μ
   - Numerator of z-score
   - Bigger difference, more likely to reject

2. The variability of the scores (σ)
   - Denominator of z-score (standard error)
   - More variability, less accurate, less likely to reject

1. The number of scores in the sample (*n*)
   - Denominator of z-score (standard error)
   - Bigger sample, more accurate, more likely to reject