

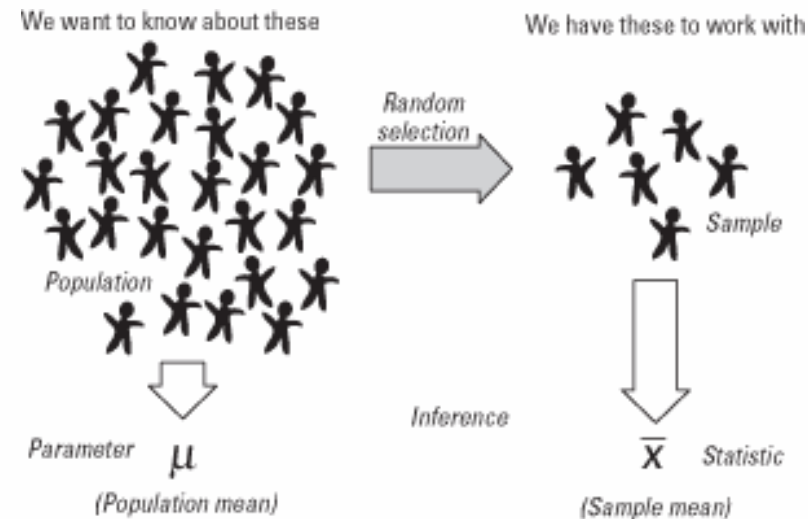
Hypothesis Testing

Statistical Hypothesis

- A hypothesis is a statement or assertion about the state of nature. e.g. The campaign was effective.
- Every hypothesis implies its contradiction or alternative. e.g. The campaign was not effective.
- Either of the hypothesis statement can be true or false.

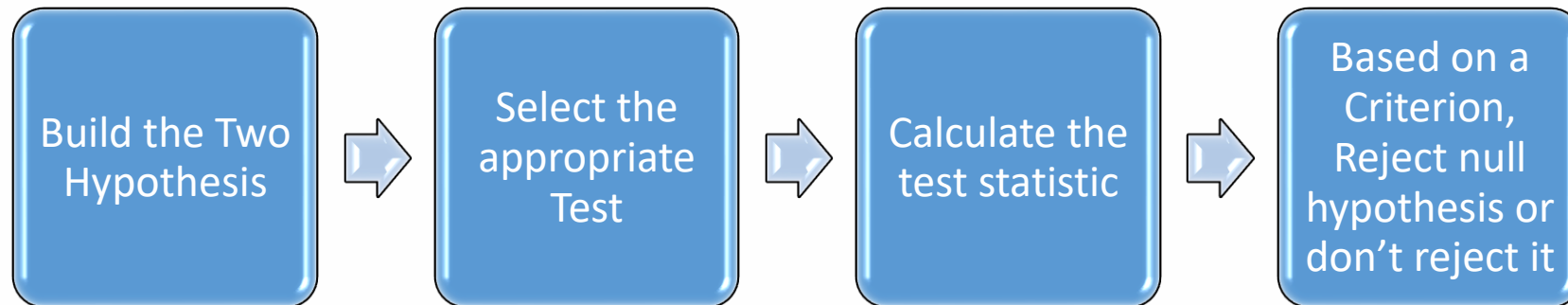
Testing the Hypothesis

- For testing any hypothesis we make use of the sample data.
- Based on the sample data, we reject either of the hypothesis statements.
- The decision to reject or failing to reject any of the hypothesis can be either correct or wrong.



Courtesy: www.cliffnotes.com

Flow for Hypothesis Testing



Actual and Findings

- Found fact would be based on sample whereas actual fact would be based on population.
- We won't be able to directly find the actual fact due to the large size of population.
- Hence we analyze the sample and try to find the fact.
- It may happen that:
 - Found fact from the sample is same as actual fact
 - Found fact from the sample is exactly opposite of the actual fact.

Actual and Findings

- When found fact from sample is not same as the actual fact from the population then we would be wrong.
- We are interested in knowing the probability of us getting wrong.

Statistical Hypothesis

- Null Hypothesis: A **null hypothesis**, denoted by H_0 , is an assertive statement about one or more population parameters. This is the statement we hold to be true until we have sufficient statistical evidence to reject it.
- The **alternative hypothesis**, denoted by H_1 , is the assertive statement of all situations *not* covered by the null hypothesis.

Example

- H_0 : The mean sales are 1400.
- H_1 : The mean sales are not 1400.
- OR
- H_0 : The mean sales are less than or equal to 1400.
- H_1 : The mean sales are greater than 1400.

Example

- Consider the example:
 - H_0 : The mean sales are less than or equal to 1400.
 - H_1 : The mean sales are greater than 1400.
- Suppose we draw a random sample for testing the above hypotheses.
- We can be wrong in the either of the cases:
 - The mean sales for population are less than or equal to 1400 and our sample suggests it as greater than 1400.(Rejecting true H_0)
 - The mean sales for population are greater than 1400 and our sample suggests it as less than or equal to 1400.(Failing to reject the false H_0)

Errors

- Type I: Rejecting true H_0
- Type II: Failing to reject the false H_0
- $P(\text{Type I Error}) = \alpha$
- $P(\text{Type II Error}) = \beta$
- If we try to reduce α , then β increases and if we try to reduce β , then α increases.
- More serious is the error α . Hence a level for α is maintained. This level is called level of significance. It is denoted by α . Usually its is maintained as 0.1 or 0.05 or 0.01.
- More often maintained as 0.05.

Types of Errors

Decision	H0 True	H0 False
Reject H0	Type I Error (α) Producer's Risk	Correct
Fail to reject H0	Correct	Type II Error (β) Consumer's Risk

P - Value

- The ***p-value*** is the probability of getting a value of the test statistic as extreme as, or more extreme than, the actual value obtained, when the null hypothesis is true.
- The p-value is the smallest level of significance, α , at which the null hypothesis may be rejected using the obtained value of the test statistic.
- **Policy to be followed: When the *p*-value is less than α , reject H_0 , otherwise we do not reject H_0 .**

Power of a Test

- The **power** of a statistical hypothesis test is the probability of rejecting the null hypothesis when the null hypothesis is false.

$$\text{Power} = (1 - \beta)$$

- Power is the probability that your test will reject the null hypothesis when the null hypothesis is false, or the probability that you will detect a difference when a difference actually exists.

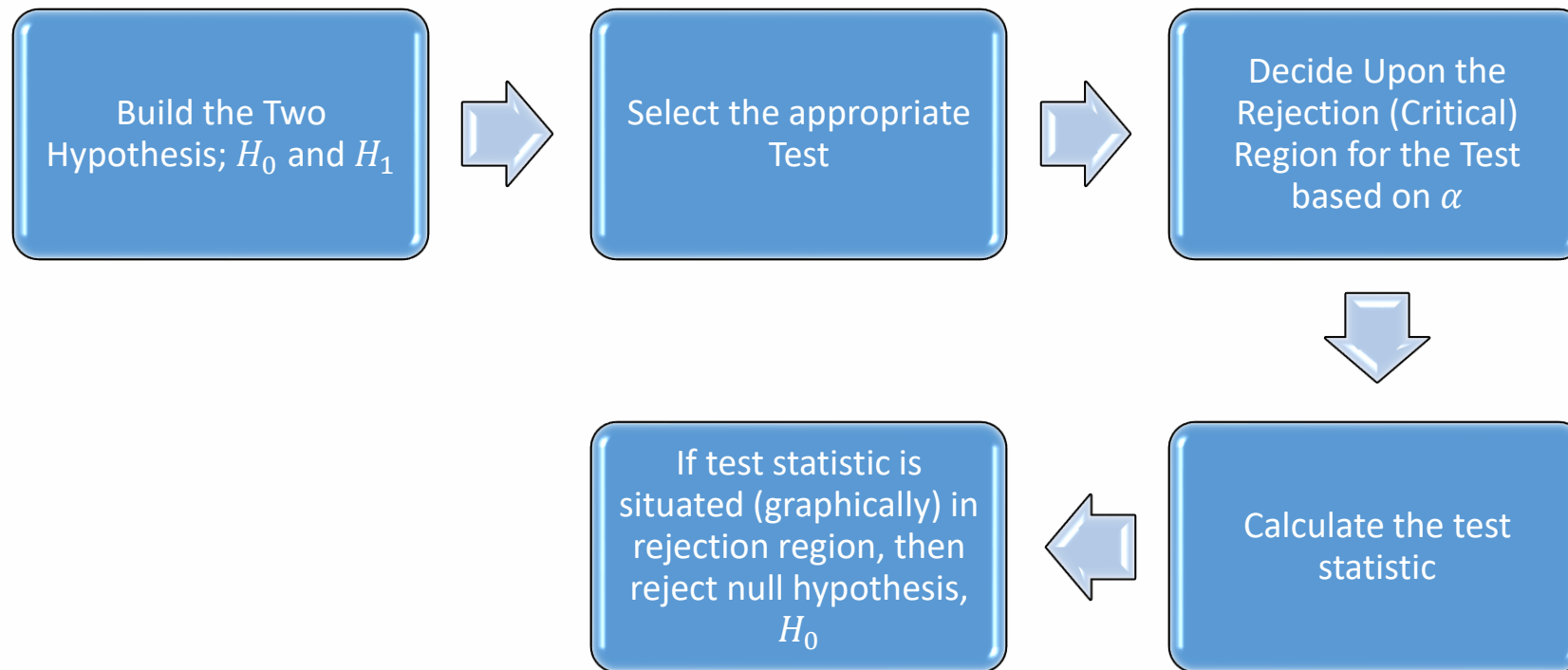
Tail of Test

- H0: The mean sales are 1400. i.e. $\mu = 1400$
- H1: The mean sales are not 1400. i.e. $\mu \neq 1400$
 - This is **two tailed** test as in H1, mean sales can be greater then or less than 1400.
- H0: The mean sales are less than or equal to 1400. i.e. $\mu \leq 1400$
- H1: The mean sales are greater than 1400. i.e. $\mu > 1400$
 - This is **right tailed** or **upper tailed** test as in H1, mean sales are greater than 1400.
- H0: The mean sales are greater than or equal to 1400. i.e. $\mu \geq 1400$
- H1: The mean sales are less than 1400. i.e. $\mu < 1400$
 - This is **left tailed** or **lower tailed** test as in H1, mean sales are less than 1400.

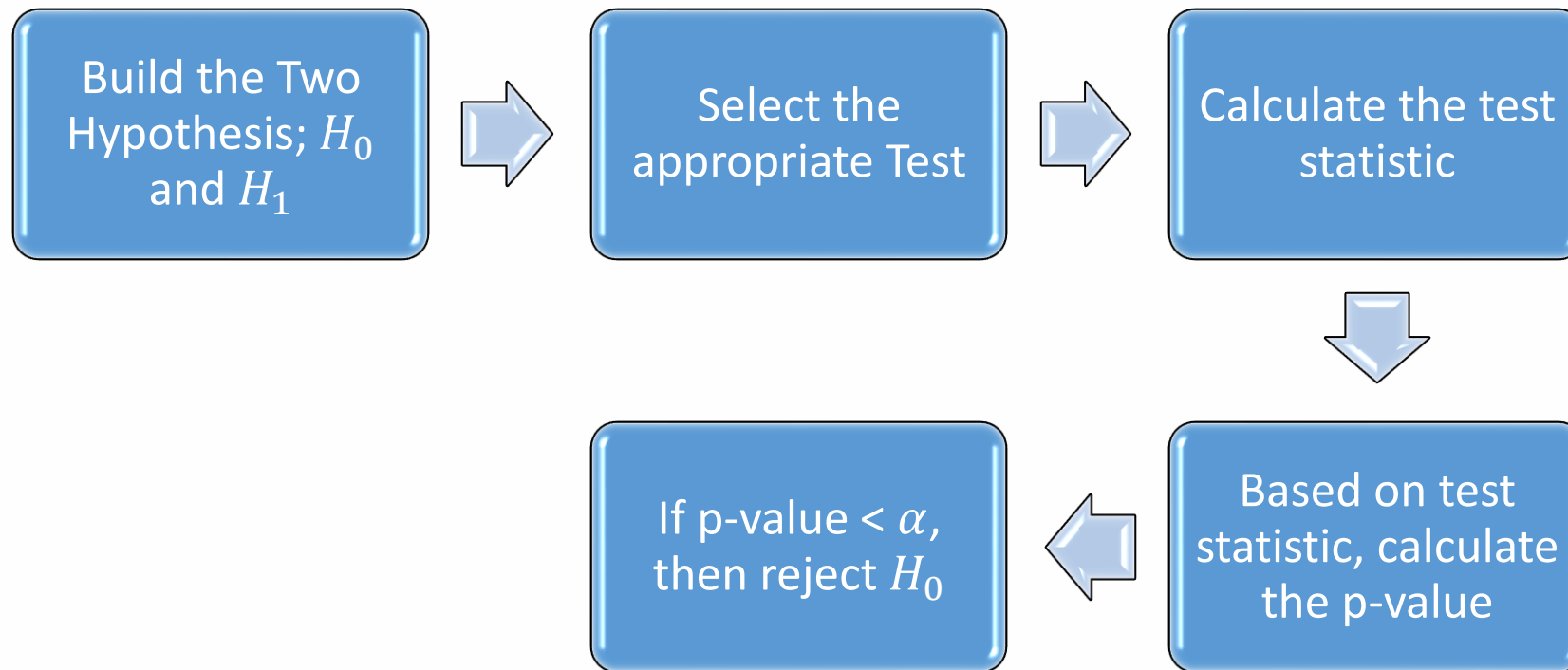
Testing of Hypothesis Approaches

- Critical Value Approach
- P-value Approach

Critical Value Approach



P-value Approach



Types of Hypothesis Tests

- Parametric: Tests which assume a particular distribution(Normal) of the population
 - t-test
 - F-test
 - Chi-square test for variance
- Non-Parametric: Tests which assume no particular distribution of the population
 - Median Test
 - Wilcoxon's Rank Sum Test
 - Mann-Whitney Test
 - Kruskal Wallis Test

Thank You