Inferential Statistics - Hypothesis

Acciojob/ Statistics-for-DataScience

Content

- Population & Sample
- Descriptive and Inferential Statistics
- Central Limit Theorem (CLT)
- Hypothesis
- Types of Hypothesis
- Steps in the process of Hypothesis Testing

Population and Sample

POPULATION

- A **Population** is a collection of all people, items, or events about which you want to make inferences or draw conclusion about. The size of the population can be finite or infinite

- **Not always convenient or possible to examine every member** of an entire population for analysis hence we have to deal with samples that can represent the population to draw conclusion. It is here we use sampling techniques to draw samples that best represent the population

- **Parameter** is the value that describe the characteristics of the population eg - ***Population Mean*** **(μ)** & ***Population S.D*** ***(σ)***

Population Parameters (μ,σ)

Sampling

Sample Statistics ***(x(bar),σ)***

Population and Sample

SAMPLE

- A **Sample** is a representative subset of people, items, or events from a larger population that you collect and analyze to make inferences. It is the group from which you collect the data from
- To represent the population well, a sample should be **randomly collected** and **adequately large.**

Eg - To understand the impact of AI on the productivity of different industries an analyst collected the sample of the productivity of employees on the scale of 1-10.

- **Statistic** is the value that describe the characteristics of the Sample eg - ***Sample Mean* (x(bar))** & ***Sample S.D* (s)** *are called Sample Statistics*

Population Parameters ($\boldsymbol{\mu}, \boldsymbol{\sigma}$*)*

Sampling

*Sample Statistics **(x(bar)**,s)*

Sampling Techniques

- Probability sampling techniques are in which every member of the population has a chance of being selected. They are mainly used in quantitative research and analysis. If you want to produce results that are representative of the whole population, **probability sampling techniques are the most valid choice.**

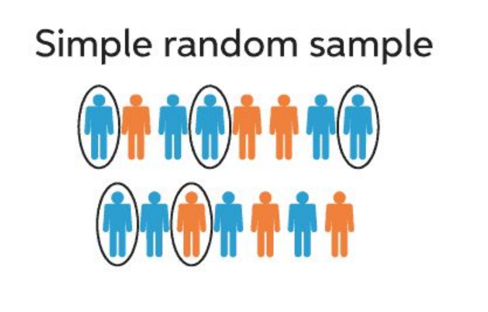
- **To draw valid conclusions from your results, you have to carefully decide how you will select a sample that is representative of the group as a whole. This is called a sampling method.**

Types of Sampling Techniques

Selection the most representative Sample from the population

- **Simple Random** - In a simple random sample, every member of the population has an equal chance of being selected. Your sampling frame should include the whole population. To conduct this type of sampling, you can use tools like random number generators or other techniques that are based entirely on chance.

  Example of Simple Random - You want to select a simple random sample of 1000 employees of a social media marketing company. You assign a number to every employee in the company database from 1 to 1000, and use a random number generator to select 100 numbers.

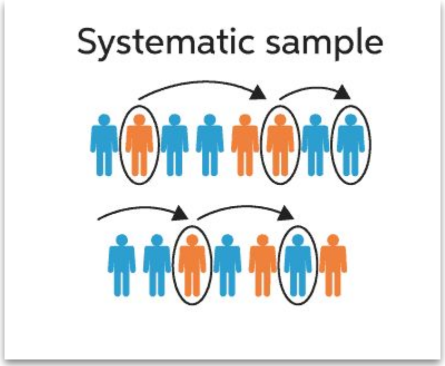


Types of Sampling Techniques

Selection the most representative Sample from the population

- **Systematic Sampling** - This is similar to simple random sampling, but it is usually slightly easier to conduct. Every member of the population is listed with a number, but instead of randomly generating numbers, individuals are chosen at regular intervals.

  Example of Systematic Sampling - All employees of the company are listed in alphabetical order. From the first 10 numbers, you randomly select a starting point: number 6. From number 6 onwards, every 10th person on the list is selected (6, 16, 26, 36, and so on), and you end up with a sample of 100 people.

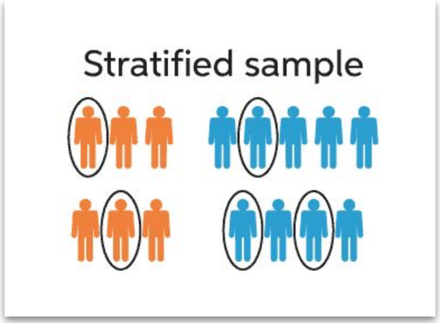


Types of Sampling Techniques

Selection the most representative Sample from the population

- **Stratified Sampling** - Stratified sampling involves dividing the population into subpopulations that may differ in important ways. It allows you draw more precise conclusions by ensuring that every subgroup is properly represented in the sample.

  Example of Stratified Sampling - The company has 1000 female employees and 800 male employees. You want to ensure that the sample reflects the gender balance of the company, so you sort the population into two strata based on gender. Then you use random sampling on each group, selecting 100 women and 80 men, which gives you a representative sample of 180 people (10%).

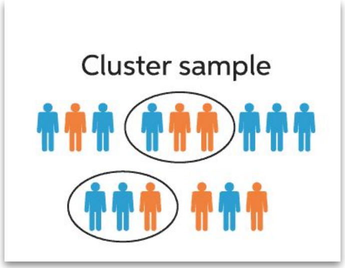


Types of Sampling Techniques

Selection the most representative Sample from the population

- **Cluster Sampling** - Cluster sampling also involves dividing the population into subgroups, but each subgroup should have similar characteristics to the whole sample. Instead of sampling individuals from each subgroup, you randomly select entire subgroups.

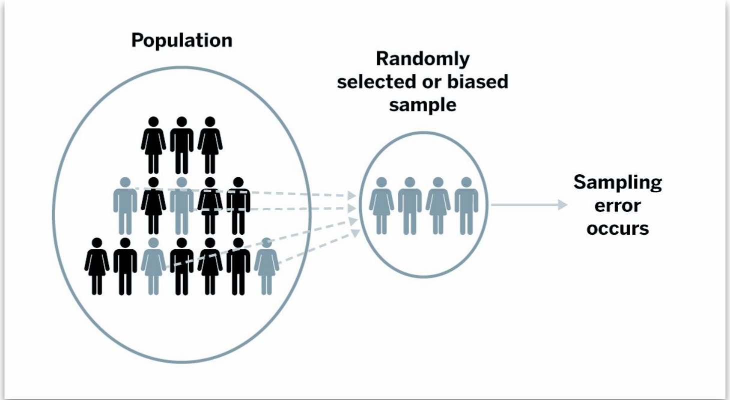
  Example - The company has offices in 10 cities across the country (all with roughly the same number of employees in similar roles). You don't have the capacity to travel to every office to collect your data, so you use random sampling to select 3 offices – these are your clusters.



Sampling Errors

Because the sample is not the whole population

- Statistical error that occurs when the analyst selects a sample that is not the representative of the population being studied.
- It is the difference between a sample statistic and the population parameter it estimates.
- There are tremendous benefits for working with samples. It is not only usually impossible to measure an entire population because they tend to be extremely large. Samples allow you to obtain a practical dataset with reasonable costs in a realistic timeframe.
- Randomness alone guarantees that your sample cannot be 100% representative of the population. Chance inevitably causes some error because the probability of obtaining just the right sample that perfectly matches the population value is practically zero. Additionally, samples can never provide a perfect depiction of the population with all its nuances because it is not the entire population.



Descriptive Statistics

Statistics

Descriptive

Inferential

Describes and summarizes the data

Measures of Central Tendency

- Mean,
- Median,
- Mode

Measures of Dispersion

- Variance
- Standard Deviation
- Range, Inter-Quartile Range
- Count, Maximum & Minimum

Shapes of Distributions

- Number of Peaks
- Symmetry
- Skewness &
- Kurtosis

Charts & Graphs

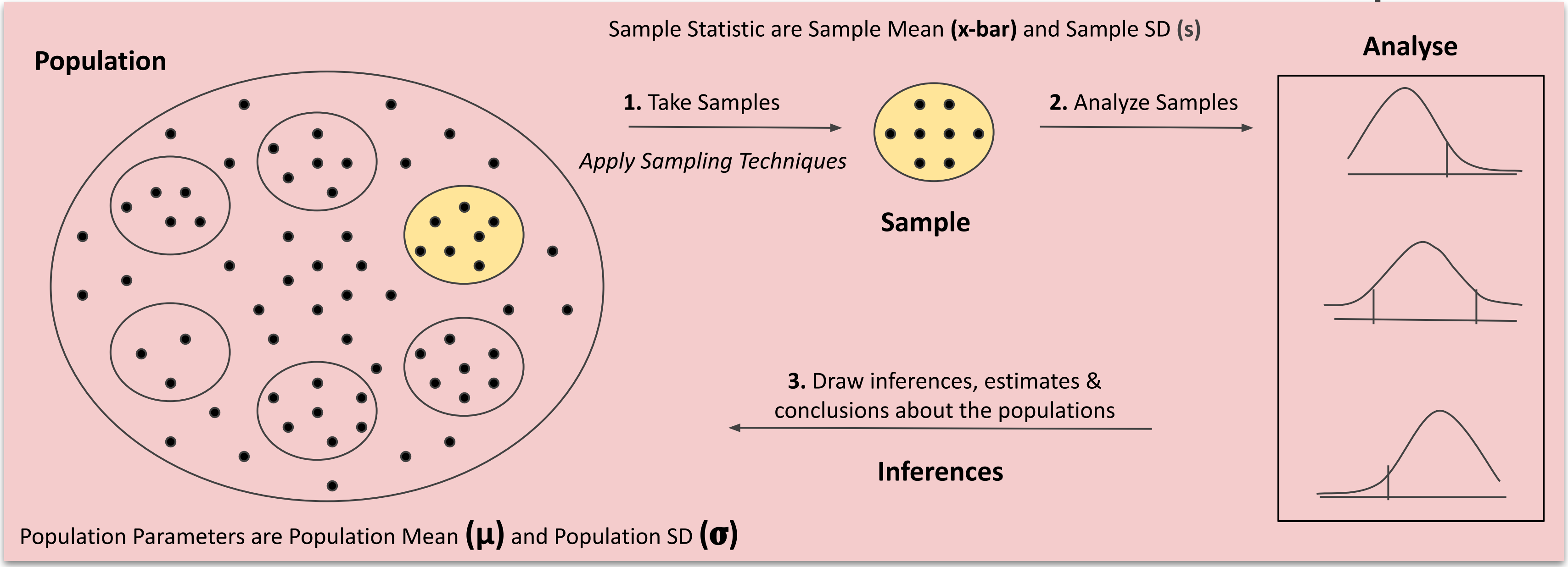
- Histograms
- Bar Charts
- Line Charts
- Pie Charts
- Box & Whiskers
- Scatter Plots
- Heat Maps

Inferential Statistics

Descriptive

Inferential

Use Sample statistics to make predictions about Population

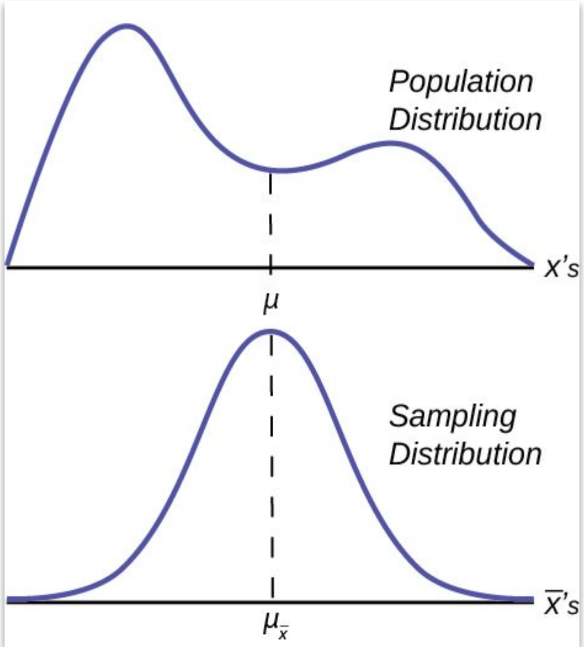


Central Limit Theorem (CLT)

Concept

The foundation of Inferential Statistics

- **CLT states that -** given a dataset with unknown distribution (be it uniform, binomial or random), **the Sampling Distribution of the Sample Means will approximate the Normal distribution.**
- **In short Average of all the Sample Means ($\overline{x}$) will be the Population mean (μ) and**
- **The distribution of Sample Means, calculated from repeated sampling, will tend to Normality as the size of your samples gets larger.**
- **Standard Deviation of the sampling distribution decreases as the size of the samples that were used to calculate the means for the sampling distribution increases.**



Assumptions of CLT

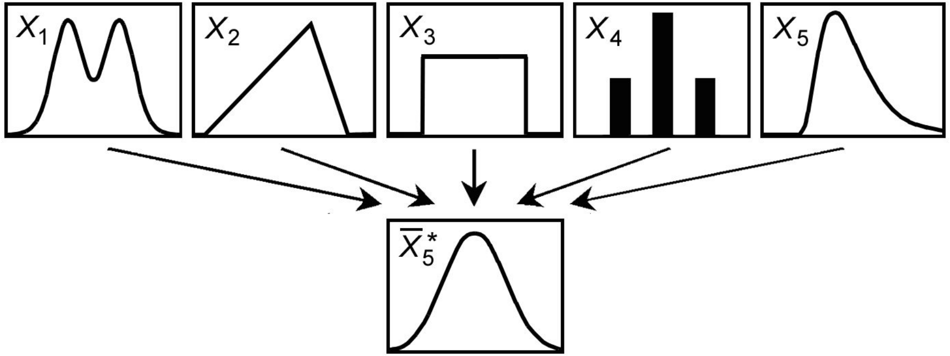
- *The data must be sampled randomly*
- *Samples should be independent of each other. One sample should not influence the other samples*
- *Sample size should be not more than 10% of the population when sampling is done without replacement*
- *The sample size should be sufficiently large.***

*** Requirements for accuracy and the Shape of the underlying population distribution. In practice, statisticians believe that a sample size of 30 is large enough when the population distribution is roughly normal. If the original population is not normal (e.g., is badly skewed, has multiple peaks, and/or has outliers), researchers like the sample size to be even larger.*

Simulation of Central Limit Theorem

Central Limit Theorem (CLT) - Example

Concept



Random Variable	Mean	Standard Deviation	Skewness	Kurtosis	Description
X_1	0.00	1.00	0.00	1.89	continuous
X_2	0.00	1.00	–0.41	2.41	continuous
X_3	0.00	1.00	0.00	1.80	continuous
X_4	0.00	1.00	0.00	2.00	discrete
X_5	0.00	1.00	1.62	7.89	continuous
$\overline{X}_5^*$	0.00	1.00	0.11	3.03	continuous

- The picture shows that taking sufficiently large samples from any type of population and plot the sample means, it will create a Normal distribution.
- This can be validated from the below tables where Skewness and Kurtosis of the 5th distribution tells you the the distribution is close to Normal.
- Further if you keep on increasing the samples the more close this distribution be to the Normal distribution.

Simulation of Central Limit Theorem

Central Limit Theorem (CLT)

Concept

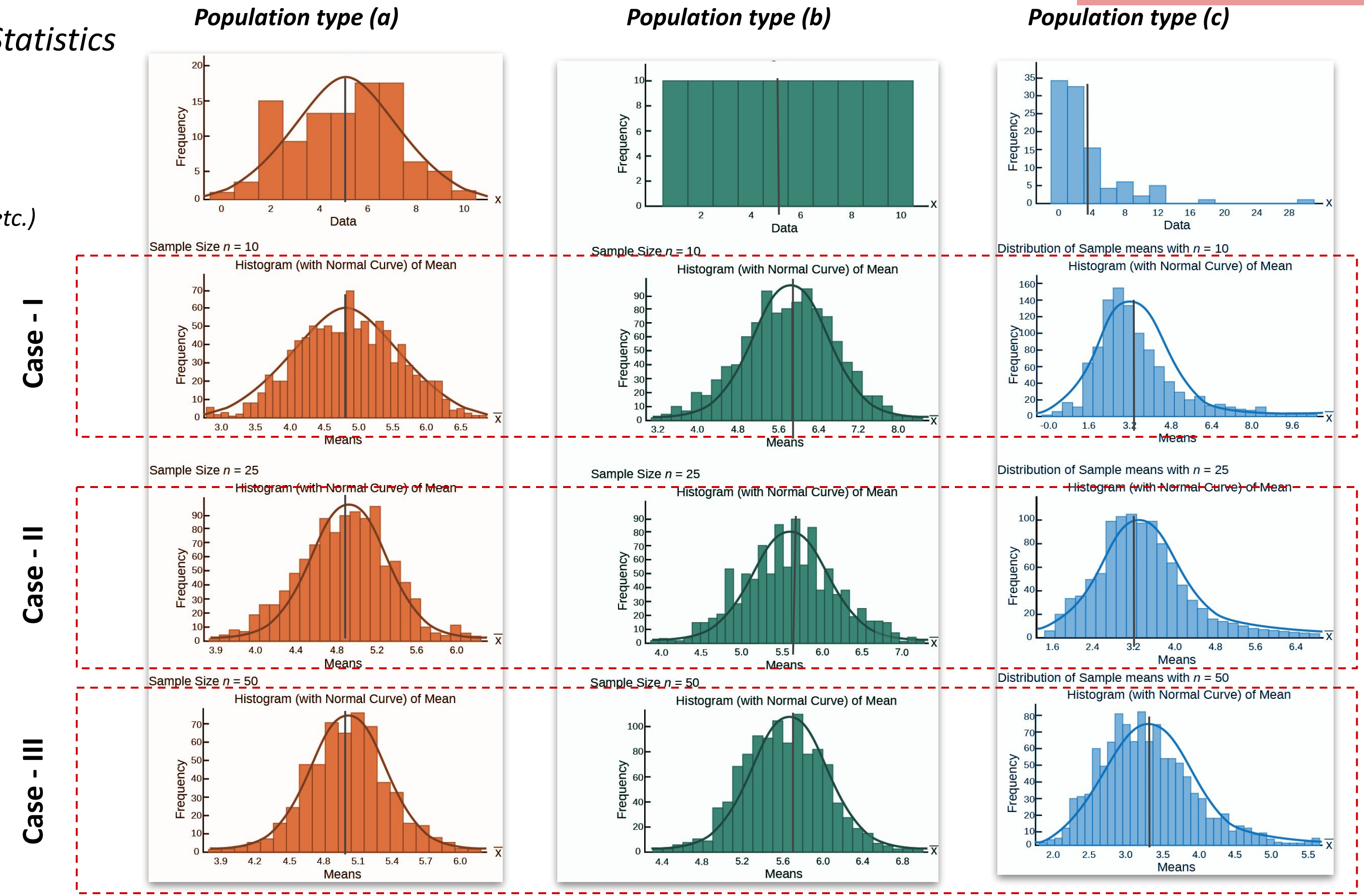
The foundation of Inferential Statistics

Population types are -

a) Normal
b) Uniform
c) Skewed (exponential, geometric etc.)

Random Samples drawn and plotted

- ***Case I*** *- 1,000 Random Samples of n= 10 drawn.*
- ***Case-II*** *- 1,000 Random Samples of n= 25 drawn.*
- ***Case-III*** *- 1,000 Random Samples of n= 50 drawn.*



Central Limit Theorem (CLT) - Applications

Concept

Uses and Applications

Figure - 1

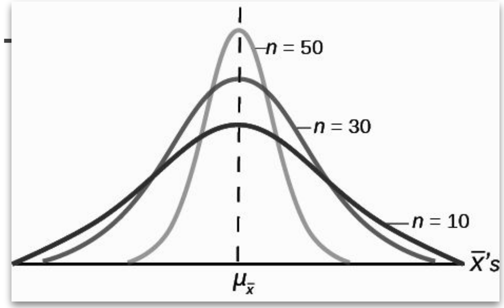


Figure - 1 *shows that* ***as the Sample size increases****, n goes from 10 to 30 to 50,* ***the Standard Deviations of the respective sampling distributions decrease*** *because the sample size is in the denominator of the Standard Deviations of the sampling distributions.*

Figure - 2

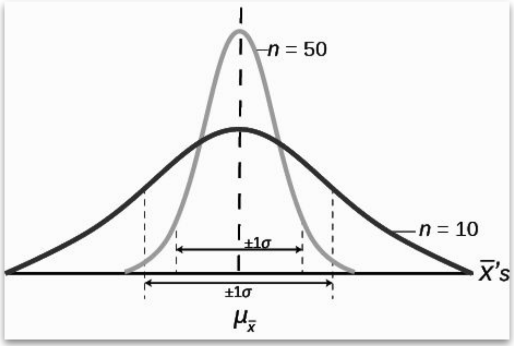


Figure - 2 *shows that for two sampling distributions from the same populations. All other things constant, the sampling distribution with sample size 50 has a smaller standard deviation that causes the graph to be higher and narrower. The* ***effect*** *of this is that for the same probability of one standard deviation from the mean, this distribution (with n = 50) covers much less of a range of possible values than the other distribution.*

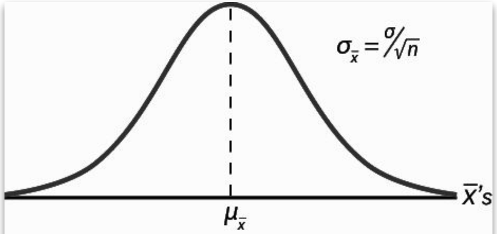


Figure - 3

Figure-3 *shows the Sampling distribution of Sample Means. Note that -*

- *Expected Value of average of Sample Means is the Population Mean*
- *The SD of the population mean is the given by the population SD divided by the Sample Size*

Hypothesis - Meaning

Hypothesis

- **Claim/ statement** or **assumption** about the population parameter
- The problem statement that leads you to some analysis is called Hypothesis. Problem statement means a statement that is *'significantly different'* from the as is situation.

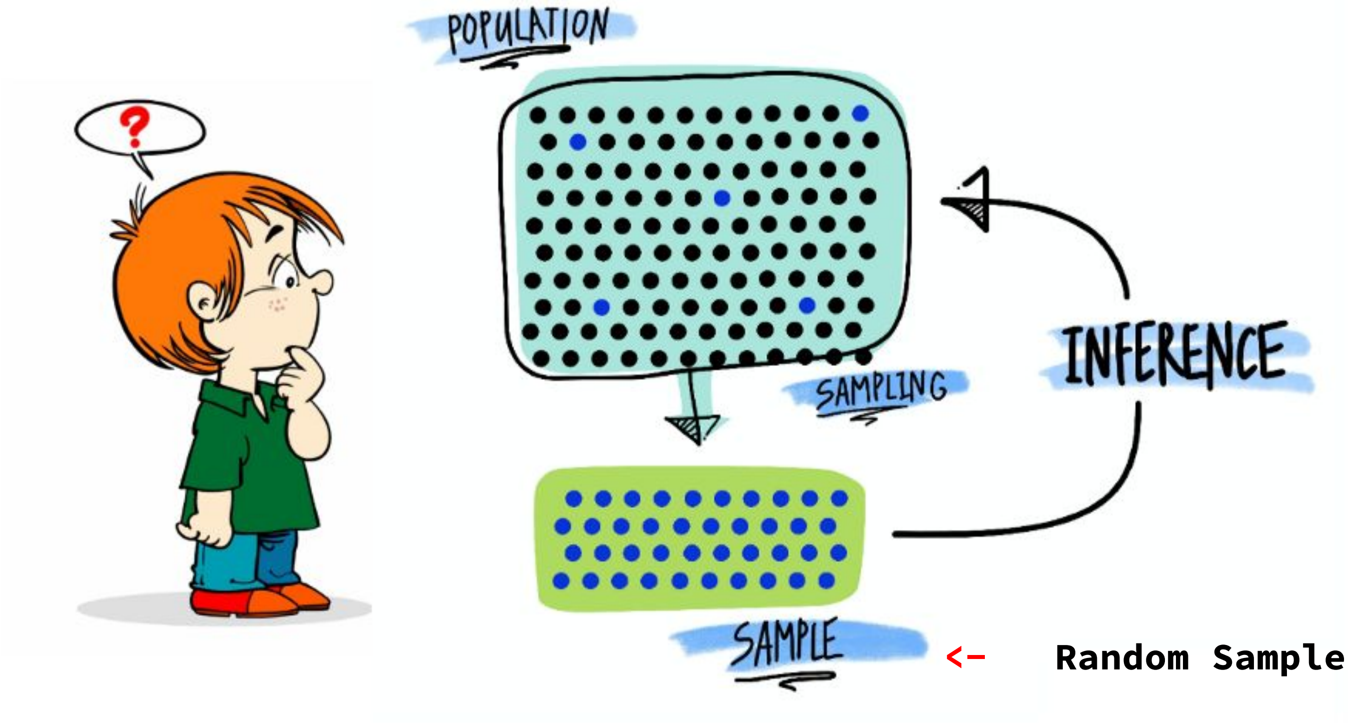
Hypothesis Testing

Technique of ***Inferential Statistics***, **where we test the claim** or statement of the analysts/ researcher about the population.

Examples of Claims Statements

- The average weight of a Male in a city is 80 kgs.
- Chewing gum makes the teeth stronger.
- Using social media more than 1 hour a day reduces the concentration levels.
- Next Year the Sales for the company is going to be INR 120 crores.
- Average age of the customers of Amazon is 40 years.
- Use analytics techniques improves the business performance by 20%.
- Average screen size of a mobile phone is 7 inches.
- Average time a iphone user is on phone is less than that of an android phone.
- Average speed of an Audi car driven by owner is 100 kms/ hour.
- The profitability of company is expected to reduce next year from 25% to 23%.
- The Social media has adversely impacted the concentration levels of the users.
- On an average a reader reads approximately 250 wpm.
- The new vaccine introduced in the market is more effective than the existing one.

Graphical Depiction of Hypothesis



<- Random Sample

Hypothesis - Foundation of the Research

- **First step** in the designing and conducting the research/ analysis

- Characteristics of Hypothesis -
    - **Clear**
    - **Specific**
    - **Testable (or researchable)**

- **Link between the Theory and Research** that leads to new discoveries

Hypothesis must satisfy the following requirements...

- Expressed in a **declarative statement**
- Postulates **relation** between variables
- **Reflect a theory** which will guide the research
- Be **brief** and **concise**
- Be **testable** and/ or provable

Types of Hypothesis

- Simple Hypothesis
- Complex Hypothesis
- Statistical Hypothesis
- Null Hypothesis
- Alternative Hypothesis

Simple Hypothesis

- Relationship between 2 variables

- One is called an Independent Variable (or Cause) and the other is called Dependent Variable (or Effect)

Example -

Older workers are more loyal to the companies.

Complex Hypothesis

- Relation between 2 or more Independent variables or 2 or more dependent variables

Example -

Global warming causes icebergs to melt which in turn causes major changes in weather patterns

Statistical Hypothesis

- To scientifically test the analysts hypothesis, a more formal hypothesis structure is required.

- This is done using the Statistical Hypothesis

- Two types - Null & Alternative Hypothesis

Types of Statistical Hypothesis

Null Hypothesis

*Is **an assumption** about the value of the Population parameter. It is usually the assumption with **'no change'.***

- ***'Null'*** condition exist
- **Nothing new** happening
- **Old theory** is still true
- **Old standard** is correct
- System is **in control**
- Assume **No Change**

H0

Types of Statistical Hypothesis

Null Hypothesis - H_0

- Must contain the condition of equality
- =, <=, >=
- Generally researcher don't believe Null
- Observed difference between Sample & Population arise because of **Random Chance Variation**
- We either -
  - **Reject** the H0
  - **'Fail to Reject'** the H0

H_0: $\mu = 5$

H_0: $\mu >= 25$

H_0: $\mu <= 150$

What is Random Chance Variation?

- It is the **inherent error** in any predictive statistical model. It is defined as the *difference between the predicted value of a variable (by the statistical model in question) and the actual value of the variable.*
- For a fairly large sample size, these errors are seen to be uniformly distributed above and below the mean and cancel each other out, resulting in Expected Value of zero.
- Examples -

  1) Time taken to cover the distance from point A to B by 100 persons in the same car

  2) Pulse rate of randomly selected 500 people in a locality

What is Random Chance Variation?

- The factors influencing the behaviour of a variable in question usually behave in a random way. So the outputs and the resulting chance errors also appear in a somewhat random fashion. Thus, **chance errors cannot be controlled,** however accurate the model be. **The presence of this error hence does not reduce the credibility of a model.**

Why - Fail to Reject?

- *"Fail to reject"* sounds like the **double negative**
- We **cannot conclusively Affirm** a Hypothesis but we can conclusively Negate it
- We assume that the Null hypothesis is correct **until we have enough evidence to suggest otherwise**
- Fact is that we can't prove a Negative

A **lack of evidence** only means that you haven't proven that something exists. It **does not prove that something doesn't exist. It might exist, but your study missed it.**

Case of Criminal Trials

- Start with an **assumption** that the **criminal is not guilty** unless proven guilty
- Prosecutor works to prove the criminal guilty if he is not able to prove the criminal guity, **the verdict says - 'not guilty'**
- The judgement does not say - the defendant is 'proven innocent'. Just that there wasn't enough evidence

Types of Statistical Hypothesis

Alternative Hypothesis

It is what is believed to be true if Null Hypothesis is False.

*It is a **Negation of Null**.*

- State ***'New Theory'*** is true
- **New standards** are there
- System is **out of control**
- Something **new** happening
- ***'Statistically Significant'*** change is observed

H1

Types of Statistical Hypothesis

Alternative Hypothesis - H_1 or H_A

- Must be True if H0 is False
- Does not contain the condition of equality i.e. $\neq$, <, >
- Negation of Null Hypothesis
- Usually what we are testing
- Claim we are testing or the initial claim that is to be tested

H1: $\mu \neq 5$

H1: $\mu < 25$

H1: $\mu > 150$

Examples of Null and Alternate Hypothesis

- **Claim: The average weight of a male has reduced from 80 kgs after installing open air gyms.**

  H0: Average weight of the males >= 80 kgs

  H1: Average age of the males < 80 kgs

- **Claim: The food experts believe that chewing gum makes the teeth stronger.**

  H0: Let us assume that chewing gum does not make the teeth stronger

  H1: Chewing gum makes the teeth stronger

- **Claim: The food experts believe that chewing gum makes the teeth stronger.**

  H0: Let us assume that chewing gum does not make the teeth stronger

  H1: Chewing gum makes the teeth stronger

Two tailed, right Tailed and left tailed tests

Two Tailed Test: Testing the claim on both the sides of the central tendency

Null Hypothesis: $\mathbf{H_0 : \mu = \mu_0}$

Alternate Hypothesis: $\mathbf{H_1 : \mu \neq \mu_0}$

Left Tailed Test: Testing the claim on the lower side of the central tendency

Null Hypothesis: **H0 : μ >= μ0**

Alternate Hypothesis: **H1: μ < μ0**

Two Tailed Test: Testing the claim on the upper side of the central tendency

Null Hypothesis: H0 : μ <= μ0

Alternate Hypothesis: H1: μ > μ0

Examples of Hypothesis

Example - 1

A company has held an 23% market share of product A.
There was increased marketing effort hence company believe that it’s market share is now greater than 23%.

H0: $p <= 23\%$

H1: $p > 23\%$

Examples of Hypothesis

Example - 2

A retail store wants to test if the average age of its customers is less than 40 years. It does a survey to gather the data of the age of customers.

H0: $\mu >= 40$

H1: $\mu < 40$

Examples of Hypothesis

Example - 3

The Hypothesis is conducted to test if the population mean is still 100. 25 Samples are collected and the Sample mean was 103.2 with a Sample Standard Deviation of 10.

H0: $\mu = 100$

H1: $\mu \neq 100$

Why perform Statistical Hypothesis Testing?

- Hypothesis test is form of inferential statistics with which **we draw conclusions about an entire population based on representative sample**
- There are lot of **benefits by working with a sample.** In most cases, it is simply impossible to observe the entire population to understand its properties. Alternative is to collect random sample and use statistics to analyze it. Also samples are much more practical and less expensive to work with
- There are **trade-off also to work with samples**. When you estimate the properties of a population from a sample, the sample statistics are unlikely to equal the actual population value precisely. The difference between the sample statistic and the population value is the *Sampling Error.* Difference observed in samples might be due to sample error rather than representing a true effect at the population level. Different samples will show different results. **Hypothesis testing incorporates estimates of the sampling error to help you make the correct decision.**

Substantive Hypothesis

- ***'Statistically significant'*** results are not always *'business significant'* outcomes.

- To **statisticians/ analysts** the word ***'significant'*** *means the result is* ***not merely due to chance*** *i.e. we 'Reject' the Null Hypothesis.*

- To **business/ decision makers** ***substantive result*** is when the outcome of a statistical study produces results that are *'important'* or *'large enough'* for them. They ***focus on ROI***.

Statistically Significant

- ***'Statistically significant'*** means that the results/ outcome of an event likely **did not happen by random chance. It means that the effect (changes) you observe in the sample also exist in the population.**
- Means test results are ***Significant*** or ***Meaningful.***
- When you draw a random sample from a population, there is always a chance that sampling error created the observed effect (read change). **How do we know whether the sample estimate reflects 'sampling error' or a true effect?**
- **It is 'Statistical significance' which tells us that the sample effect (the result shown by the sample that is different from population) is unlikely caused by sampling error. When we have statistically significant results, we conclude it is an actual effect existing in the population.**
- It indicates that the sample effect is unlikely caused by chance (ie Sampling error). In other words - we have evidence that the results we see in Sample also exist in population

Statistically Significant

- ***'Statistically significant'*** means that the results/ outcome of an event likely **did not happen by random chance. It means that the effect (changes) you observe in the sample also exist in the population.**
- Means test results are ***Significant*** or ***Meaningful***
- More specifically, weather your Statistic (from Sample) closely matches what value you would ***'expect'*** to find in entire population parameters
- Means that we have evidence that the results we see in Sample also exist in population
- *Eg - 120 people from the city were asked their sports preference. You would want the sample results to represent everyone's preference in the city. In other words you would want the report to have - 'Statistically significant' findings*

Errors in Hypothesis - Type - I (α) and Type - II (β)

	H_0 True	*H_0 False*
Reject H_0	**Type - I Error (α)** ***Reject a True H_0*** *False Positive (FP)*	**True Positive (1 - β)** **Power of Test/ 'Specificity'**
Accept H_0	**True Negative (1 - α)** **Confidence Level/ Confidence Coefficient/ Sensitivity**	**Type - II Error (β)** ***'Fail to Reject' a False H_0*** *False Negatives (FN)*

Note: *There is always a trade off between Type-I error and Type-II error i.e. if* α increases β *will decrease and vice versa.*

Errors in Hypothesis Comparison - Type - I (α) and Type - II (β)

	Type - I Error (α)	Type - II Error (β)
Meaning	Non-Acceptance of True H0	Acceptance of False H0
What is it	False Positive	False Negative
Represents	A False Hit	A Miss
Probability of Error	Level Of Significance (α)	Power of the Test (β)
Example	New drug is accepted but it is not effective	Rejecting a newly effective drug

Errors in Hypothesis - Example of Bank Customer

	H_0 True: Customer is Good	***H_0 False: Customer not Good***
Reject H_0	**Type - I Error (α)** ***Reject a True H_0*** ***Reject a Good customer.***	**True Positive (1 - β)** **Reject a bad Customer.** **Avoid Bad debts** **Power of Test**
Accept H_0	**True Negative (1 - α)** **'Do not Reject' a Good Customer.** **We keep good business**	**Type - II Error (β)** ***'Fail to Reject' a False H_0*** ***We accept a bad customer and thus we incur Bad debts***

H_0 : Customer is Good

H_a : Customer is not Good

Errors in Hypothesis - Example of Weather Forecast

	H_0 True: No Rain	***H_0 False: Rains***
Reject H_0 ***Take Umbrella***	**Type - I Error (α)** ***Reject a True H_0*** ***Take Umbrella when it does not Rain.***	**True Positive (1 - β)** **Take Umbrella when it Rains** **Power of Test**
Accept H_0 ***Do not take Umbrella***	**True Negative (1 - α)** **'Do not take Umbrella' when it does not rain**	**Type - II Error (β)** ***'Fail to Reject' a False H_0*** ***Do not take Umbrella and it Rains***

H_0 : No Rains

H_a : It will Rain

Hypothesis Testing - Steps

The heart of Inferential Statistics

1 Setting Null Hypothesis (H_0)

Assume No change.
The difference in Sample mean and Population mean is are simply because of random chance.

H_0: μ = Observed Mean ($\bar{x}$)

2 Setting Alternate Hypothesis (H_a)

Negation of Null Hypothesis.
A ***Statistically Significant*** *change is observed*
H_a *is usually what we are testing*

H_a: $\mu \neq$ Observed Mean ($\bar{x}$)

3 Set the Significance Levels (α)

One Tail test *- Significance Level will be on one side*
Two Tail test *- Significance level split on both sides*

Confidence Levels *is 1 - α. If α* = 0.05 then CI = 0.95

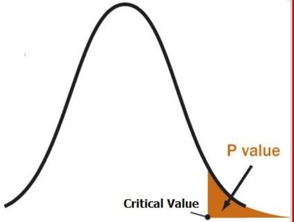
4 Select Distribution Test. Find Test Statistic to find Critical Values

Distribution Test		***Test Statistic***		
Z-Test		*Z-Score*		
t-Test	→	*t-Score*	→	*Critical Value (Threshold Value)*
F-Test		*F-Statistic*		
ANOVA		*ANOVA*		
Chi Square		*Chi Sq. Statistic*		

5 Calculating the P-Value

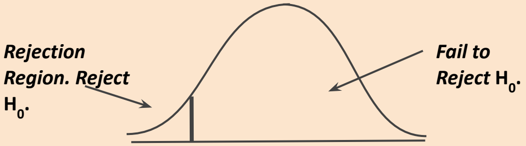
Calculated from the Critical Values you get from the Distribution Test

P-Value is the ***Area in the tail*** *of the Probability Distribution*



6 Take Decision. ACCEPT or REJECT.

Reject H_0: *If P-Value is less than α. Accept H_a.*
Fail to Reject H_0: *If P-Value is more than α.*



P-Value is the evidence against the Null (H_0)

Hypothesis Testing - Steps

Step - 1 & 2 Defining Null & Alternative Hypothesis

1

Setting Null Hypothesis (H_0)

Assume No change.

The difference in Sample mean and Population mean is are simply because of random chance.

H_0: μ = Observed Mean

2

Setting Alternative Hypothesis (H_a)

A Statistically Significant change is observed.

H_a is *usually what we are testing*

H_a: $\mu \neq$ Observed Mean

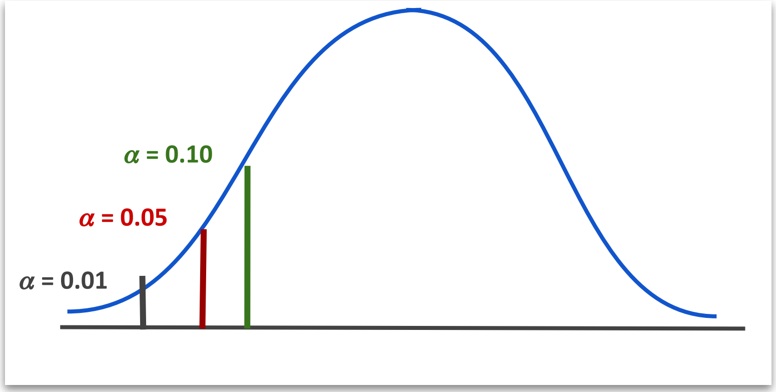
Hypothesis Testing - Steps

Step - 3 Setting Significance levels (α)

3

- **Significance level*** **(α)** is the maximum probability of accepting the **'Type-I Error'** i.e. Rejecting a True Null Hypothesis (and assume that a 'Significant Change' has occurred). Thus it **represents the probability that tests will produce *'statistically significant'* results when the null hypothesis is correct.**

- In Hypothesis the most common Significance Levels** are -
  - 1% (0.01)
  - 5% (0.05)
  - 10% (0.1)



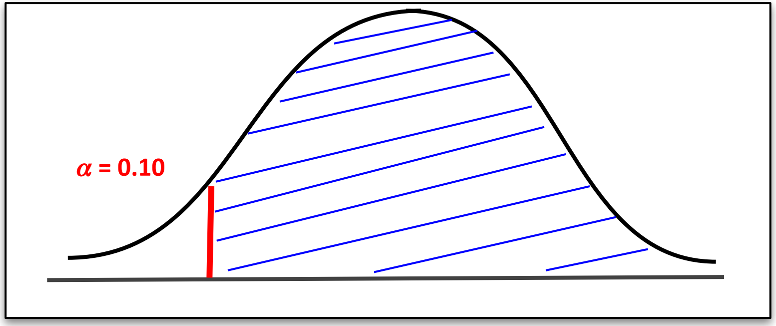
** **1 - Significance level (α) = Confidence Level (1-α)** is the Probability with which we accept the Null when True*

*** The selection of Significance level depends on business case of how stringent we want to test our assumption with*

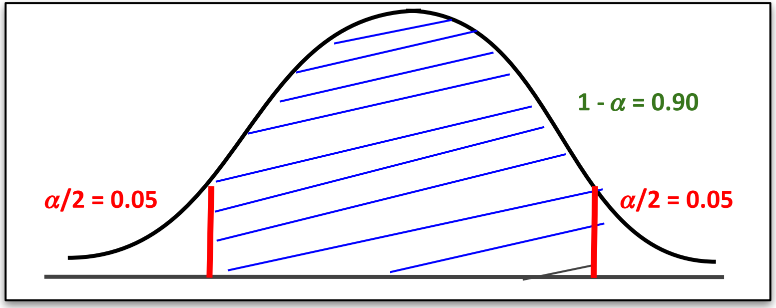
Hypothesis Testing - Steps

Step - 3 Setting Significance levels (α) for one tailed test and 2 tailed test

- For **'One Tailed Tests'** the level of Significance (α) lies on either side of the distribution - left or right



- For **'Two Tailed Tests'** the total level of Significance is split on 2 sides of the distribution and it will be $\alpha/2$ on either side



Hypothesis Testing - Steps

Selecting the Distribution Test/ Test statistic

4

- **Normal Distribution**
  - **z-test**
  - **T-test**
    - **Independent**
    - **Paired**
  - **ANOVA**

- **Non-Normal Distribution**
  - **Chi-Square (χ^2)**
  - *Mann-Whitney*
  - *Wilcoxon Sign Rank Test*
  - *Kruskal Wallis Test*

Test Statistic

- **Z-Score** *(from z-Test)*
- **t-Score** *(from t-test)*
- **F-Statistic** *(from ANOVA)*
- **Chi-Square (χ^2)** *(from Chi-sq test)*

Test Statistic with selected Significance Levels is used to calculate the **Critical Values (the thresholds).** *Earlier this was done through by looking at the tables but now we use excel (data analysis pack) and other tools to get these values.*

Critical Value

Critical Value is the value of test statistic (z-score, t-score, f-statistic or chi-sq) that divides non-Rejection region with that of the Rejection region.

Hypothesis Testing - Steps

What is Test statistic and its interpretation?

Test Statistic

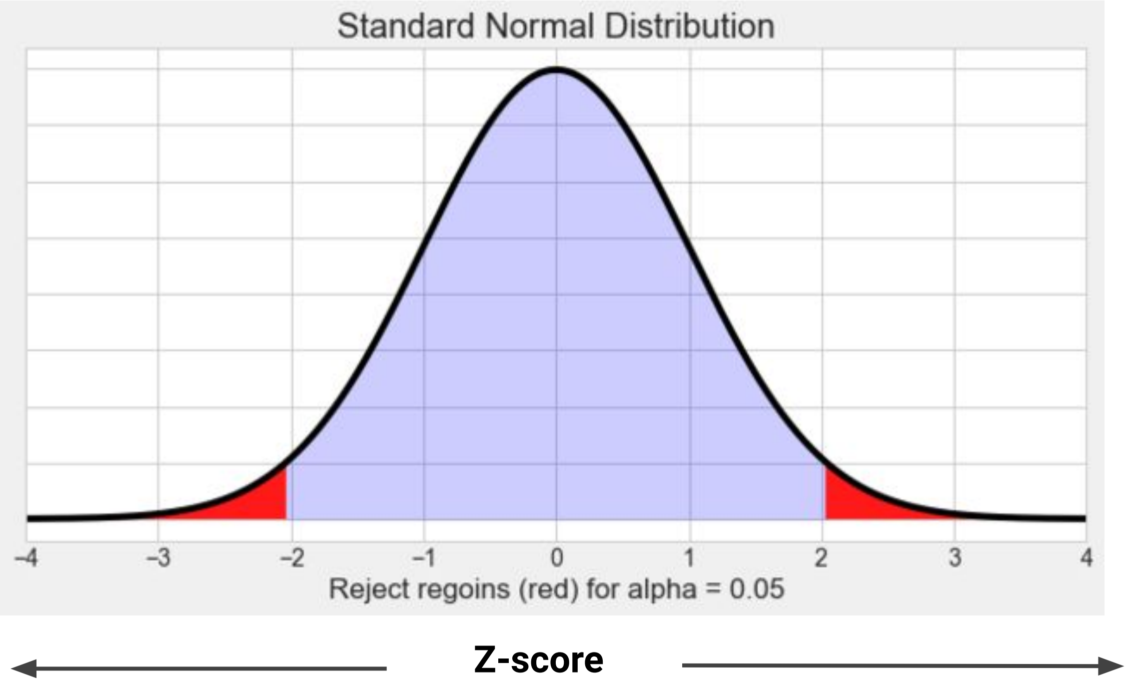
- A test statistic is a number calculated by a statistical test. Eg - z-score, t-score, F-score and chi-sq
- All these test statistic quantifies **how much sample diverges (or agrees) from the Null Hypothesis for different types of distributions.** Test statistic summarises your sample data (observed data) into a single number using central tendency, variation, sample size and number of predictor variables in the model.

Interpretation of test statistic

- It describes how far (or consistant) your observed data is from the null hypothesis of no relationship between variables.
- The **value of test statistic = 0 indicates that your sample data match the null hypothesis exactly**
- As a **test statistic value becomes more extreme, it indicates larger differences between your sample data and the null hypothesis**, thus we can reject the null and state that the results are 'statistically significant'. i.e. your data support belief - *'the sample effect exists in the population'*.
- **In nutshell - as sample mean moves away from the hypothesized mean in either the positive or negative direction, the test statistic moves away from zero in the same direction.**

Hypothesis Testing - Steps

Test statistic (z-score) for z-distribution (called Standard Normal Distribution)



Hypothesis Testing - Steps

What is a Critical Value?

Critical Value

- A critical value is derived from the **test statistic (used with significance levels) which defines the upper and lower bounds of a confidence interval**, or **which defines the threshold of statistical significance** in a statistical test. It describes how far from the mean of the distribution you have to go to cover a certain amount of the total variation in the data (i.e. 90%, 95%, 99%).
- These values also play an important role in both hypothesis tests. In hypothesis tests, critical values determine whether the results are statistically significant.
- Critical value is derived by looking at the z-table, t-table etc. It is in the terms of ‘test statistic’ only i.e. z-score (in z-test, t-score (in t-test) etc.

Use of Critical values

- **Test statistics that exceed a critical value have a low probability of occurring** if the null hypothesis is true. Therefore, when test statistics exceed these cutoffs, you can reject the null and conclude that the effect exists in the population. In other words, they **define the rejection regions for the null hypothesis.**

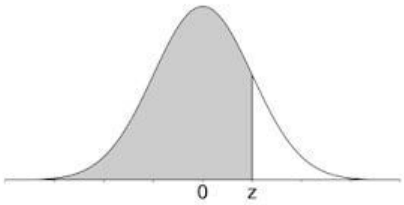
Hypothesis Testing - Steps

What does z Table show?

- Z-score is the test statistic which is the S.D. of the for a Standard Normal Curve.
- A z-table (also called the standard normal table) provides **the area under the curve to the left of a z-score**. This area represents the probability that z-values will fall within a region of the standard normal distribution.
- **Eg1** - if z = 1.25 the area to the left of this z is 0.8944 = 89.44%
- **Eg2** - if z = -2.23 then the area to the left of z is 0.01287 - 1.28%
- **Eg3** - Find the area of z-scores that are more extreme than ± 2.5 standard deviations from the mean.
  Since the SND is symmetric, you can find the area only on one side and then double it. Thus the area will be = 0.00621+0.00621 = 0.01242

Z	0.00	0.01	0.02
-2.7	0.00347	0.00336	0.00326
-2.6	0.00466	0.00453	0.00440
-2.5	0.00621	0.00604	0.00587
-2.4	0.00820	0.00798	0.00776

Table of Standard Normal Probabilities for Positive Z-Scores



z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441

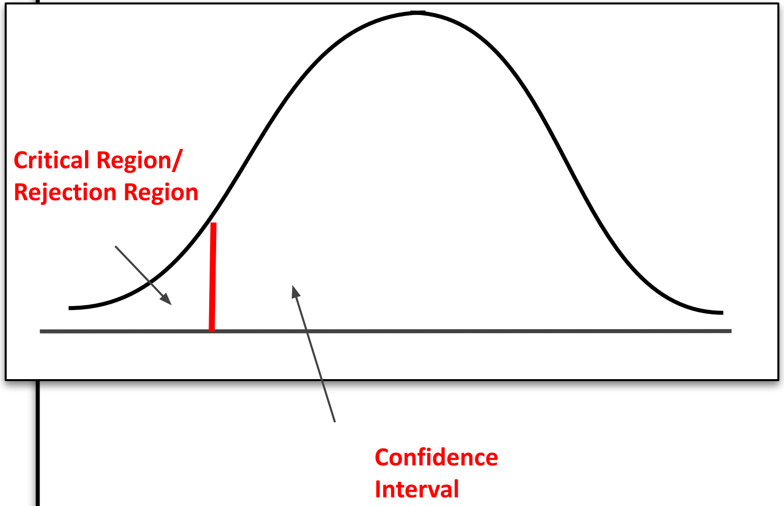
Hypothesis Testing - Steps

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
-3.9	0.00005	0.00005	0.00004	0.00004	0.00004	0.00004	0.00004	0.00004	0.00003	0.00003
-3.8	0.00007	0.00007	0.00007	0.00006	0.00006	0.00006	0.00006	0.00005	0.00005	0.00005
-3.7	0.00011	0.00010	0.00010	0.00010	0.00009	0.00009	0.00008	0.00008	0.00008	0.00008
-3.6	0.00016	0.00015	0.00015	0.00014	0.00014	0.00013	0.00013	0.00012	0.00012	0.00011
-3.5	0.00023	0.00022	0.00022	0.00021	0.00020	0.00019	0.00019	0.00018	0.00017	0.00017
-3.4	0.00034	0.00032	0.00031	0.00030	0.00029	0.00028	0.00027	0.00026	0.00025	0.00024
-3.3	0.00048	0.00047	0.00045	0.00043	0.00042	0.00040	0.00039	0.00038	0.00036	0.00035
-3.2	0.00069	0.00066	0.00064	0.00062	0.00060	0.00058	0.00056	0.00054	0.00052	0.00050
-3.1	0.00097	0.00094	0.00090	0.00087	0.00084	0.00082	0.00079	0.00076	0.00074	0.00071
-3.0	0.00135	0.00131	0.00126	0.00122	0.00118	0.00114	0.00111	0.00107	0.00104	0.00100
-2.9	0.00187	0.00181	0.00175	0.00169	0.00164	0.00159	0.00154	0.00149	0.00144	0.00139
-2.8	0.00256	0.00248	0.00240	0.00233	0.00226	0.00219	0.00212	0.00205	0.00199	0.00193
-2.7	0.00347	0.00336	0.00326	0.00317	0.00307	0.00298	0.00289	0.00280	0.00272	0.00264
-2.6	0.00466	0.00453	0.00440	0.00427	0.00415	0.00402	0.00391	0.00379	0.00368	0.00357
-2.5	0.00621	0.00604	0.00587	0.00570	0.00554	0.00539	0.00523	0.00508	0.00494	0.00480
-2.4	0.00820	0.00798	0.00776	0.00755	0.00734	0.00714	0.00695	0.00676	0.00657	0.00639
-2.3	0.01072	0.01044	0.01017	0.00990	0.00964	0.00939	0.00914	0.00889	0.00866	0.00842
-2.2	0.01390	0.01355	0.01321	0.01287	0.01255	0.01222	0.01191	0.01160	0.01130	0.01101
-2.1	0.01786	0.01743	0.01700	0.01659	0.01618	0.01578	0.01539	0.01500	0.01463	0.01426
-2.0	0.02275	0.02222	0.02169	0.02118	0.02068	0.02018	0.01970	0.01923	0.01876	0.01831
-1.9	0.02872	0.02807	0.02743	0.02680	0.02619	0.02559	0.02500	0.02442	0.02385	0.02330
-1.8	0.03593	0.03515	0.03438	0.03362	0.03288	0.03216	0.03144	0.03074	0.03005	0.02938
-1.7	0.04457	0.04363	0.04272	0.04182	0.04093	0.04006	0.03920	0.03836	0.03754	0.03673
-1.6	0.05480	0.05370	0.05262	0.05155	0.05050	0.04947	0.04846	0.04746	0.04648	0.04551
-1.5	0.06681	0.06552	0.06426	0.06301	0.06178	0.06057	0.05938	0.05821	0.05705	0.05592
-1.4	0.08076	0.07927	0.07780	0.07636	0.07493	0.07353	0.07215	0.07078	0.06944	0.06811
-1.3	0.09680	0.09510	0.09342	0.09176	0.09012	0.08851	0.08691	0.08534	0.08379	0.08226
-1.2	0.11507	0.11314	0.11123	0.10935	0.10749	0.10565	0.10383	0.10204	0.10027	0.09853
-1.1	0.13567	0.13350	0.13136	0.12924	0.12714	0.12507	0.12302	0.12100	0.11900	0.11702
-1.0	0.15866	0.15625	0.15386	0.15151	0.14917	0.14686	0.14457	0.14231	0.14007	0.13786
-0.9	0.18406	0.18141	0.17879	0.17619	0.17361	0.17106	0.16853	0.16602	0.16354	0.16109
-0.8	0.21186	0.20897	0.20611	0.20327	0.20045	0.19766	0.19489	0.19215	0.18943	0.18673
-0.7	0.24196	0.23885	0.23576	0.23270	0.22965	0.22663	0.22363	0.22065	0.21770	0.21476
-0.6	0.27425	0.27093	0.26763	0.26435	0.26109	0.25785	0.25463	0.25143	0.24825	0.24510
-0.5	0.30854	0.30503	0.30153	0.29806	0.29460	0.29116	0.28774	0.28434	0.28096	0.27760
-0.4	0.34458	0.34090	0.33724	0.33360	0.32997	0.32636	0.32276	0.31918	0.31561	0.31207
-0.3	0.38209	0.37828	0.37448	0.37070	0.36693	0.36317	0.35942	0.35569	0.35197	0.34827
-0.2	0.42074	0.41683	0.41294	0.40905	0.40517	0.40129	0.39743	0.39358	0.38974	0.38591
-0.1	0.46017	0.45620	0.45224	0.44828	0.44433	0.44038	0.43644	0.43251	0.42858	0.42465
0.0	0.50000	0.49601	0.49202	0.48803	0.48405	0.48006	0.47608	0.47210	0.46812	0.46414

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.50000	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.52790	0.53188	0.53586
0.1	0.53983	0.54380	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	0.61791	0.62172	0.62552	0.62930	0.63307	0.63683	0.64058	0.64431	0.64803	0.65173
0.4	0.65542	0.65910	0.66276	0.66640	0.67003	0.67364	0.67724	0.68082	0.68439	0.68793
0.5	0.69146	0.69497	0.69847	0.70194	0.70540	0.70884	0.71226	0.71566	0.71904	0.72240
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.75490
0.7	0.75804	0.76115	0.76424	0.76730	0.77035	0.77337	0.77637	0.77935	0.78230	0.78524
0.8	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234	0.80511	0.80785	0.81057	0.81327
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894	0.83147	0.83398	0.83646	0.83891
1.0	0.84134	0.84375	0.84614	0.84849	0.85083	0.85314	0.85543	0.85769	0.85993	0.86214
1.1	0.86433	0.86650	0.86864	0.87076	0.87286	0.87493	0.87698	0.87900	0.88100	0.88298
1.2	0.88493	0.88686	0.88877	0.89065	0.89251	0.89435	0.89617	0.89796	0.89973	0.90147
1.3	0.90320	0.90490	0.90658	0.90824	0.90988	0.91149	0.91309	0.91466	0.91621	0.91774
1.4	0.91924	0.92073	0.92220	0.92364	0.92507	0.92647	0.92785	0.92922	0.93056	0.93189
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	0.93943	0.94062	0.94179	0.94295	0.94408
1.6	0.94520	0.94630	0.94738	0.94845	0.94950	0.95053	0.95154	0.95254	0.95352	0.95449
1.7	0.95543	0.95637	0.95728	0.95818	0.95907	0.95994	0.96080	0.96164	0.96246	0.96327
1.8	0.96407	0.96485	0.96562	0.96638	0.96712	0.96784	0.96856	0.96926	0.96995	0.97062
1.9	0.97128	0.97193	0.97257	0.97320	0.97381	0.97441	0.97500	0.97558	0.97615	0.97670
2.0	0.97725	0.97778	0.97831	0.97882	0.97932	0.97982	0.98030	0.98077	0.98124	0.98169
2.1	0.98214	0.98257	0.98300	0.98341	0.98382	0.98422	0.98461	0.98500	0.98537	0.98574
2.2	0.98610	0.98645	0.98679	0.98713	0.98745	0.98778	0.98809	0.98840	0.98870	0.98899
2.3	0.98928	0.98956	0.98983	0.99010	0.99036	0.99061	0.99086	0.99111	0.99134	0.99158
2.4	0.99180	0.99202	0.99224	0.99245	0.99266	0.99286	0.99305	0.99324	0.99343	0.99361
2.5	0.99379	0.99396	0.99413	0.99430	0.99446	0.99461	0.99477	0.99492	0.99506	0.99520
2.6	0.99534	0.99547	0.99560	0.99573	0.99585	0.99598	0.99609	0.99621	0.99632	0.99643
2.7	0.99653	0.99664	0.99674	0.99683	0.99693	0.99702	0.99711	0.99720	0.99728	0.99736
2.8	0.99744	0.99752	0.99760	0.99767	0.99774	0.99781	0.99788	0.99795	0.99801	0.99807
2.9	0.99813	0.99819	0.99825	0.99831	0.99836	0.99841	0.99846	0.99851	0.99856	0.99861
3.0	0.99865	0.99869	0.99874	0.99878	0.99882	0.99886	0.99889	0.99893	0.99896	0.99900
3.1	0.99903	0.99906	0.99910	0.99913	0.99916	0.99918	0.99921	0.99924	0.99926	0.99929
3.2	0.99931	0.99934	0.99936	0.99938	0.99940	0.99942	0.99944	0.99946	0.99948	0.99950
3.3	0.99952	0.99953	0.99955	0.99957	0.99958	0.99960	0.99961	0.99962	0.99964	0.99965
3.4	0.99966	0.99968	0.99969	0.99970	0.99971	0.99972	0.99973	0.99974	0.99975	0.99976
3.5	0.99977	0.99978	0.99978	0.99979	0.99980	0.99981	0.99981	0.99982	0.99983	0.99983
3.6	0.99984	0.99985	0.99985	0.99986	0.99986	0.99987	0.99987	0.99988	0.99988	0.99989
3.7	0.99989	0.99990	0.99990	0.99990	0.99991	0.99991	0.99992	0.99992	0.99992	0.99992
3.8	0.99993	0.99993	0.99993	0.99994	0.99994	0.99994	0.99994	0.99995	0.99995	0.99995
3.9	0.99995	0.99995	0.99996	0.99996	0.99996	0.99996	0.99996	0.99996	0.99997	0.99997

Hypothesis Testing - Steps

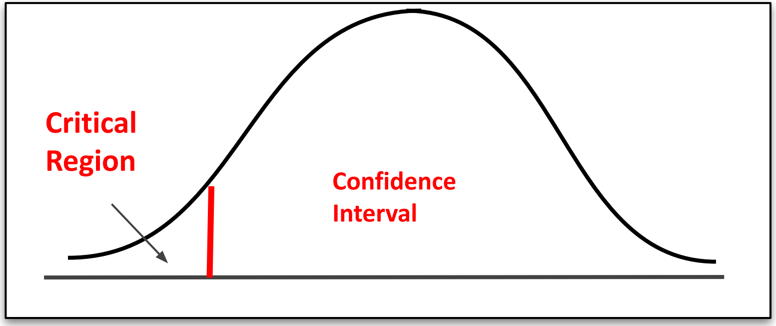
What is a Critical Region and Confidence Interval?



- Critical Region is also called the Rejection Region.
- Critical Region is the Set of all values of *test statistic* that would cause a *'Rejection of the Null Hypothesis'*
- If the observed test statistic is in the critical region then we 'Reject the Null' hypothesis and Accept the Alternative hypothesis.

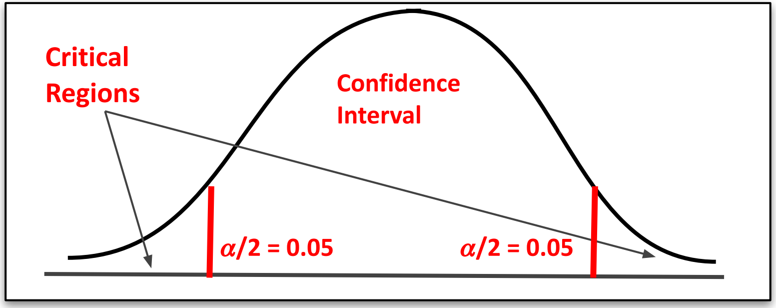
Hypothesis Testing - Steps

Critical Region and Confidence Interval



- Critical Region for 'One Tail Test'

- Critical Region for a 'Two Tailed Tests'



Hypothesis Testing - Steps

Step - 5 Calculating p-value (Observed test statistic)

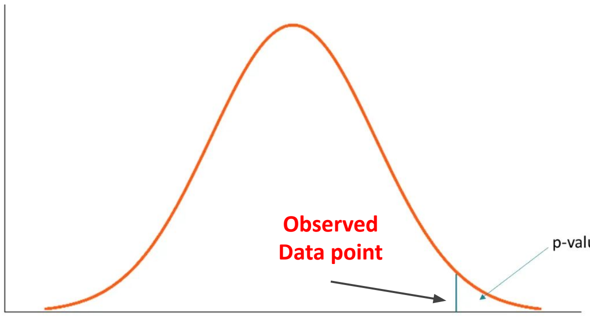
5

What is p-Value?

p-value measures the probability of obtaining the observed results, assuming that the null hypothesis is true. It is the likelihood of getting the observed results by chance alone. The lower the p-value, the greater the statistical significance of the observed difference.

- p-value is the 'actual risk' we carry in the model i.e. it measures the the probability of the effect (read change) observed in the sample, if the null hypothesis is True.

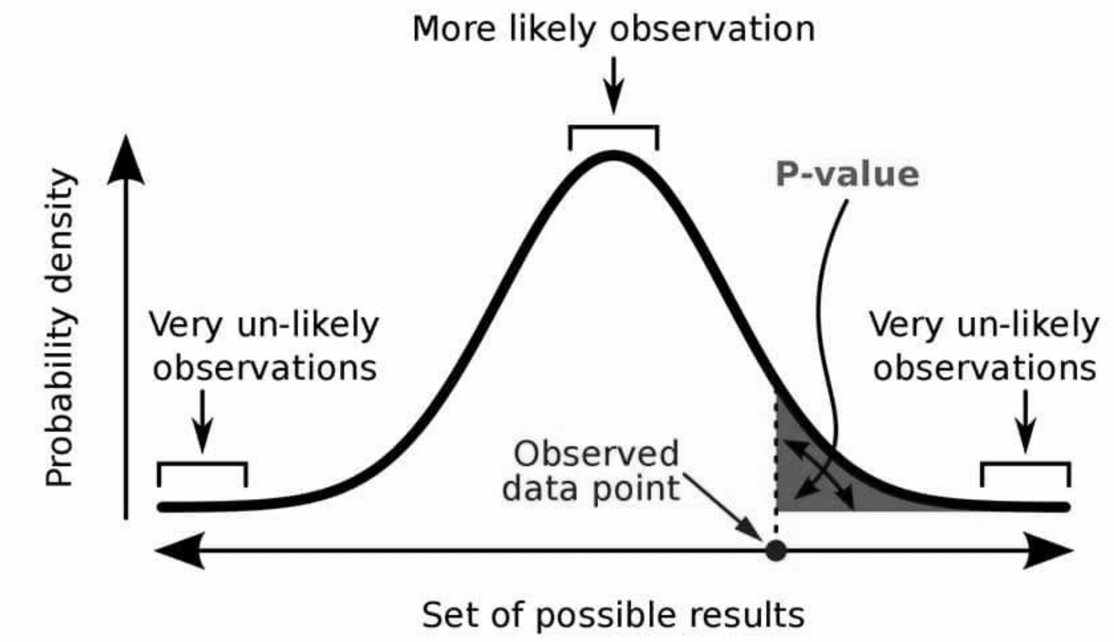
- It is the cumulative probability of the outcomes as extreme or more extreme than the observed value arising by chance. It is the Area under the tail of the probability distribution.



Hypothesis Testing - Steps

Step - 5 Calculating p-value (Observed test statistic)

5



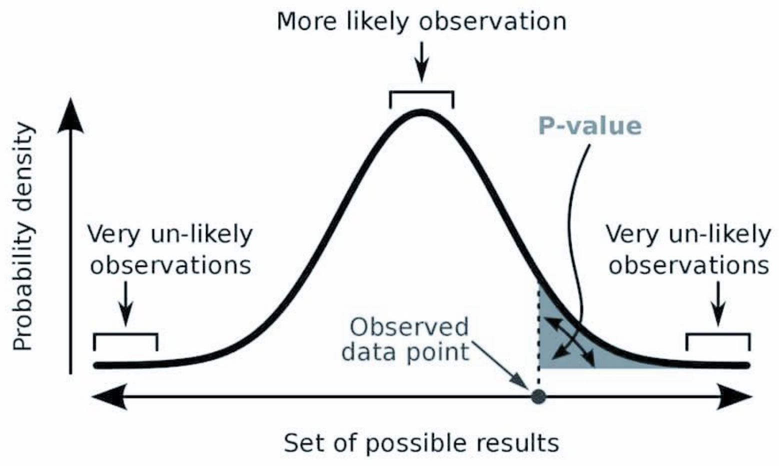
- p-value is the probability of the observed (or more extreme) result assuming Null hypothesis is True.
- A p-value, or probability value, is a number describing how likely it is that your data would have occurred by random chance (i.e. that the null hypothesis is true).

Hypothesis Testing - Steps

Step - 5 Calculating p-value (Observed test statistic)

5

- **P-value is the evidence against the Null Hypothesis**



- A **low p-value** (typically ≤ 0.05) **indicates the null hypothesis is unlikely to be true.**
- A **high p-value** (typically > 0.05) **suggests the observed results align with the null hypothesis.**

To Summarize:

- If the **p-value ≤ critical value, Reject the null hypothes**
- If the **p-value > critical value, 'Fail to reject' the null hy**

Smaller the p-Value, stronger the evidence that you should reject Null Hypothesis.

Hypothesis Testing - Steps

Step - 5 Calculating p-value (Observed test statistic)

5

Steps in calculating the p-Value

1. **Define the Null (H_0)and Alternative (H_1) Hypothesis -** The null hypothesis typically states that there is no effect or relationship between variables, while the alternative hypothesis suggests otherwise.
2. **Determine which distribution is to be applied and the test statistic accordingly -** The choice of test statistic depends on your data and hypotheses. Some common test statistics include the t-test, z-test, or chi-square test.
3. **Calculate the observed test statistic based on your sample data** - Using your sample data, compute the test statistic. This value quantifies the difference between your sample data and the null hypothesis.
4. **Find the probability of obtaining a test statistic at least as extreme as the observed value -** Find the probability of obtaining a test statistic at least as extreme as the observed value, under the assumption that the null hypothesis is true. **This probability is the *p-value*.**

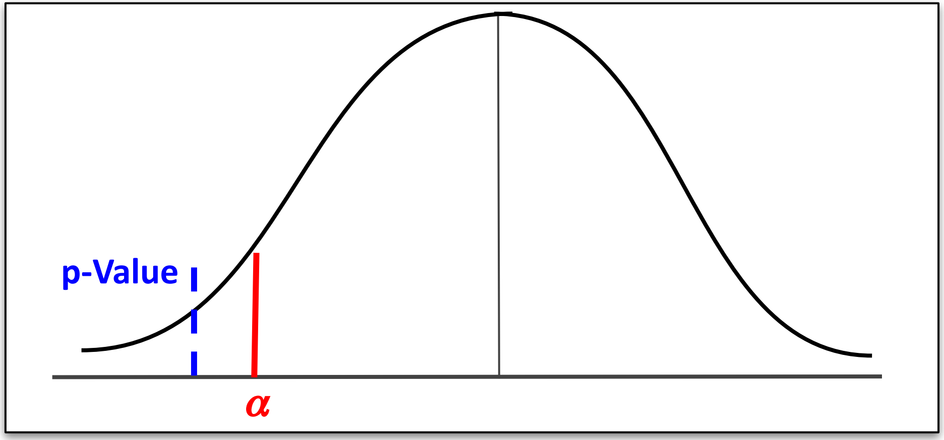
Hypothesis Testing - Steps

Step - 6 Take Decision. Reject or 'Fail to Reject'.

6

p-Value < Significance level (α)

REJECT THE NULL



Lower p-value means lower the likelihood that such a sample would have shown up. It means it is less likely that such a sample can be picked from the population. Means I am more confident that such a sample came from a different population. Hence we -

Reject the Null.

Hypothesis Testing - Steps

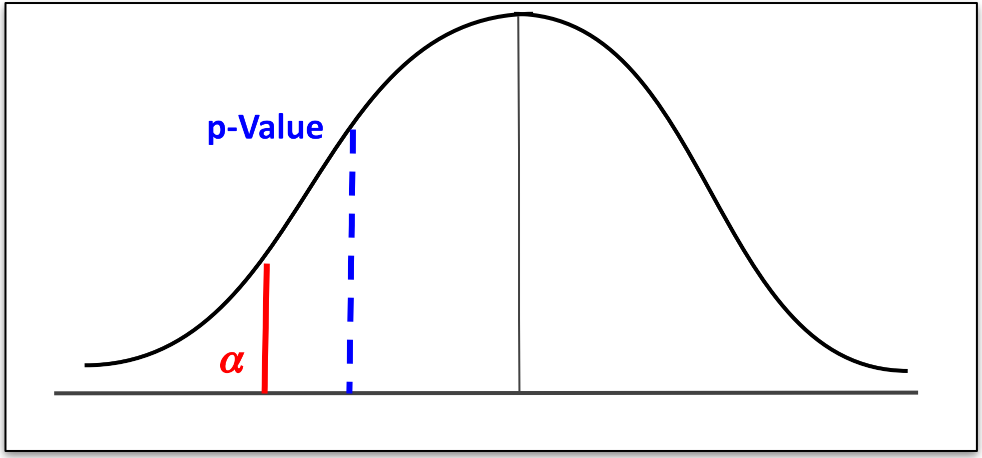
Step - 5 Take Decision. Reject or ‘Fail to Reject’.

p-Value > Significance level (α)

FAIL TO REJECT THE NULL

Higher P-Value means higher the likelihood that such a sample can appear. It means it is more likely that such a sample can be picked from the existing population. Means we are more confident that such a sample came from this population. Hence we -

‘Fail to Reject’ the Null.



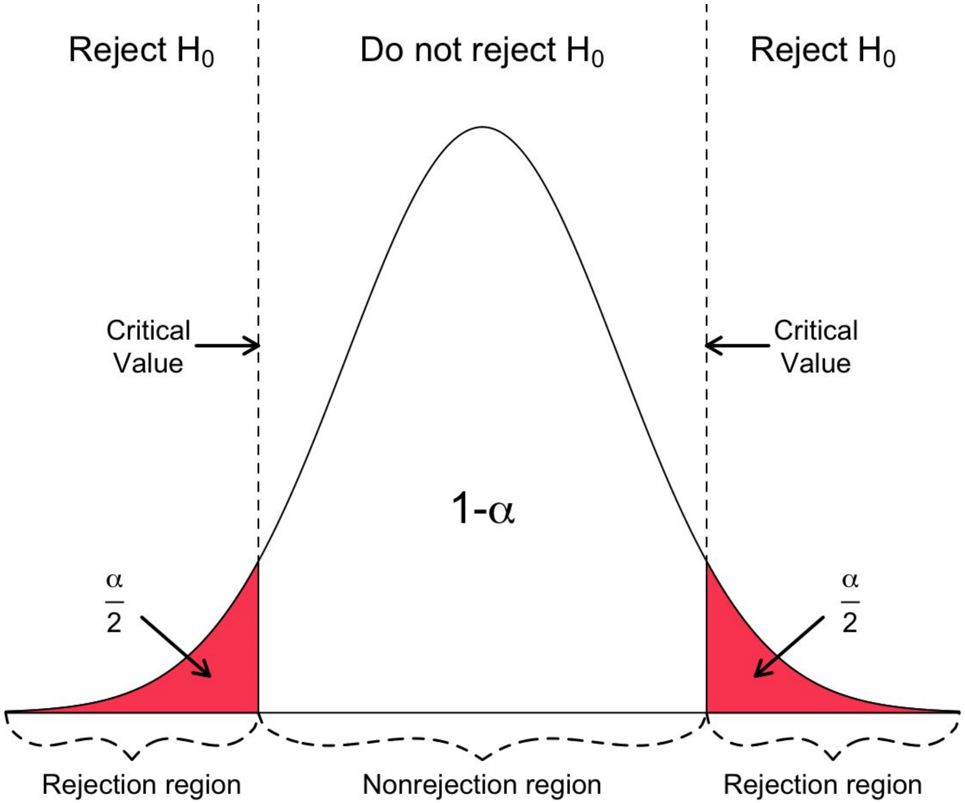
Understanding the Decision with the 'Risk' perspective

- **p-Value is the actual Risk we carry in the model i.e. when the event occurs.** Here Risk means the probability of Type-I error (α) i.e. we reject a True H_0. We can also call **p-Value as the 'Actual Significance level'** also. Significance level (α) is the desired probability of the Type-I error i.e. the acceptable level of risk.

- **When p-Value < α (Significance Level) -** It means **Actual Risk of the new model is less than the desired risk.** Thus we Reject the status quo (H_0) and we conclude that the change is Statistically Significant. **Reject H_0** (and accept H_a)

- **When p-Value > α (Significance Level) -** It means **Actual Risk (of moving to H_0) of the new model is greater than the desired risk.** Thus we Reject the status quo (H_0) and we conclude that the change is not Statistically Significant. **Fail to Reject H_0** (and reject H_a)

In NutShell



Reject H_0
Do not reject H_0
Reject H_0
Critical Value
Critical Value
1-α
$\frac{\alpha}{2}$
$\frac{\alpha}{2}$
Rejection region
Nonrejection region
Rejection region

Thank You!