

Six Sigma
Prof. Jitesh J Thakkar
Department of Industrial and Systems Engineering
Indian Institute of Technology, Kharagpur

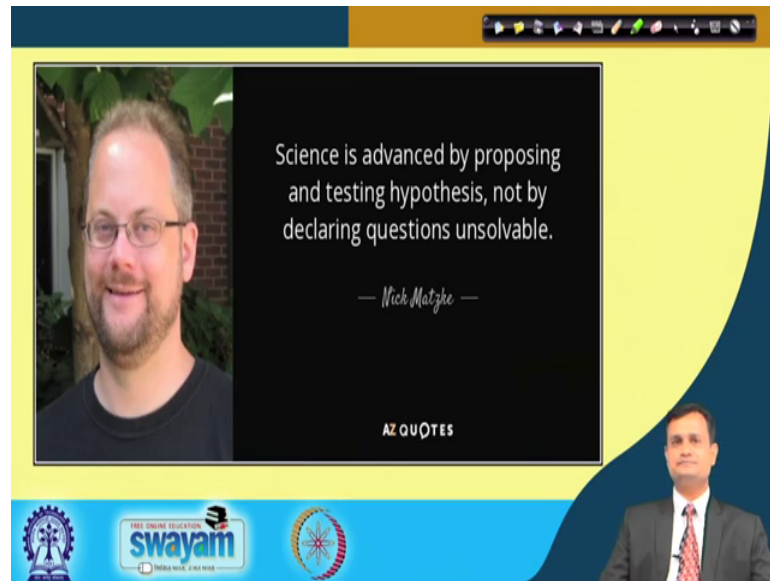
Lecture - 28
Hypothesis Testing: Fundamentals

Hello friends, welcoming you to the ongoing journey of Six Sigma. And we had a long discussion on process capability analysis almost we have spent 4 lectures discussing normal and non-normal process capability analysis with its application in minitab. And this lecture 28 we will focus on a very very important topic that is Hypothesis Testing Fundamentals.

So, if you recall when I talked about the basics of statistics. Then, we have divided the statistics into 2 parts; descriptive statistics and inferential statistics. We have seen that there are various ways and means like finding measures of central tendency, dispersion, shape, plotting the histogram, box and whisker plot many other ways to describe my data in more descriptive manner to enable the practicing people researchers for making the decision.

The another domain which typically deals with inferential statistic this domain we have not discussed so far. So, this is mainly dealing with hypothesis testing and we would like to say begin with our analyse phase in the DMAIC cycle with this lecture 28 that is hypothesis testing fundamentals.

(Refer Slide Time: 02:06)



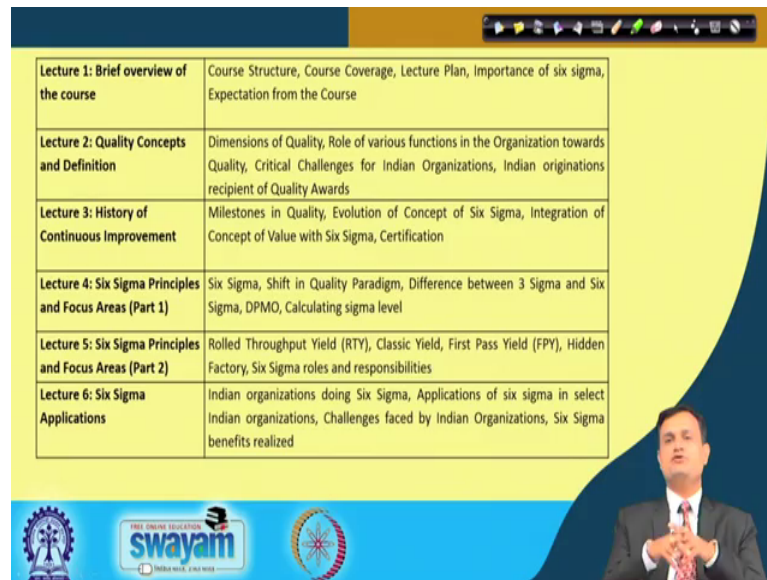
So, let us begin with a very interesting quote given by Nick Matzke and he says, that science is advanced by proposing and testing hypothesis not that by just declaring questions and solvable. You have to prove the theory, you have to test the theory. You have to check your assumptions. And in order to do so, you have to check your claim, you have to test your claim and your claim is basically they scientifically designed hypotheses which can really open a new direction for research, business or policy and practice.

(Refer Slide Time: 02:52)



So, before we really enter into the hypothesis testing fundamentals as we are beginning with a new phase in DMAIC cycle that is analyse, let me just give you a brief recap what we have done so far and where are we in our six sigma journey. So, quality fundamentals and key concepts we have discussed in week 1 and 2.

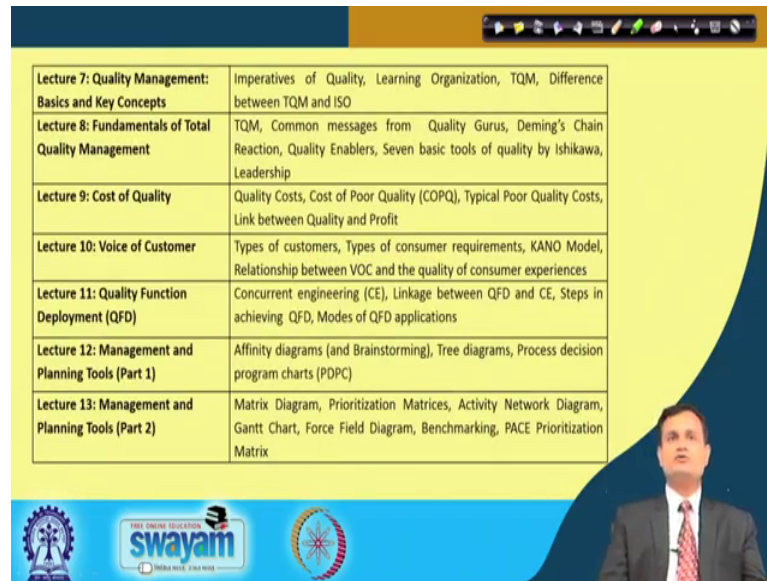
(Refer Slide Time: 03:18)



Lecture 1: Brief overview of the course	Course Structure, Course Coverage, Lecture Plan, Importance of six sigma, Expectation from the Course
Lecture 2: Quality Concepts and Definition	Dimensions of Quality, Role of various functions in the Organization towards Quality, Critical Challenges for Indian Organizations, Indian originations recipient of Quality Awards
Lecture 3: History of Continuous Improvement	Milestones in Quality, Evolution of Concept of Six Sigma, Integration of Concept of Value with Six Sigma, Certification
Lecture 4: Six Sigma Principles and Focus Areas (Part 1)	Six Sigma, Shift in Quality Paradigm, Difference between 3 Sigma and Six Sigma, DPMO, Calculating sigma level
Lecture 5: Six Sigma Principles and Focus Areas (Part 2)	Rolled Throughput Yield (RTY), Classic Yield, First Pass Yield (FPY), Hidden Factory, Six Sigma roles and responsibilities
Lecture 6: Six Sigma Applications	Indian organizations doing Six Sigma, Applications of six sigma in select Indian organizations, Challenges faced by Indian Organizations, Six Sigma benefits realized

Week 1; basically gave you the overview of the course quality concepts and definitions, history of continuous improvement, six sigma principles and focus areas. Lecture 4 and 5 part 1 and part 2 and then six sigma applications typically in some of the Indian organizations.

(Refer Slide Time: 03:39)

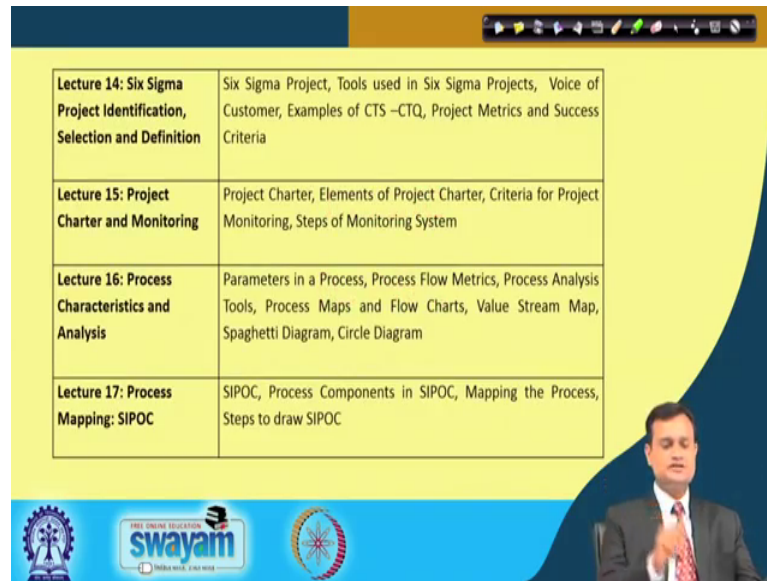


Lecture 7: Quality Management: Basics and Key Concepts	Imperatives of Quality, Learning Organization, TQM, Difference between TQM and ISO
Lecture 8: Fundamentals of Total Quality Management	TQM, Common messages from Quality Gurus, Deming's Chain Reaction, Quality Enablers, Seven basic tools of quality by Ishikawa, Leadership
Lecture 9: Cost of Quality	Quality Costs, Cost of Poor Quality (COPQ), Typical Poor Quality Costs, Link between Quality and Profit
Lecture 10: Voice of Customer	Types of customers, Types of consumer requirements, KANO Model, Relationship between VOC and the quality of consumer experiences
Lecture 11: Quality Function Deployment (QFD)	Concurrent engineering (CE), Linkage between QFD and CE, Steps in achieving QFD, Modes of QFD applications
Lecture 12: Management and Planning Tools (Part 1)	Affinity diagrams (and Brainstorming), Tree diagrams, Process decision program charts (PDPC)
Lecture 13: Management and Planning Tools (Part 2)	Matrix Diagram, Prioritization Matrices, Activity Network Diagram, Gantt Chart, Force Field Diagram, Benchmarking, PACE Prioritization Matrix

Week 2; we have talked about say quality management key concepts and fundamentals we have talked about fundamentals of TQM. Then lecture 8 cost of quality, lecture 9 cost of quality, lecture 10 voice of customer, lecture 11 QFD, 12 management and planning tools part 1, 13 management and planning tools part 2.

And we discussed many plan management and planning tools like benchmarking, affinity diagram, brainstorming metrics diagram, prioritization matrices, activity network diagram, force field diagram and so on. Then we entered in to define phase that is the first phase of DMAIC cycle and that was our week 3.

(Refer Slide Time: 04:33)

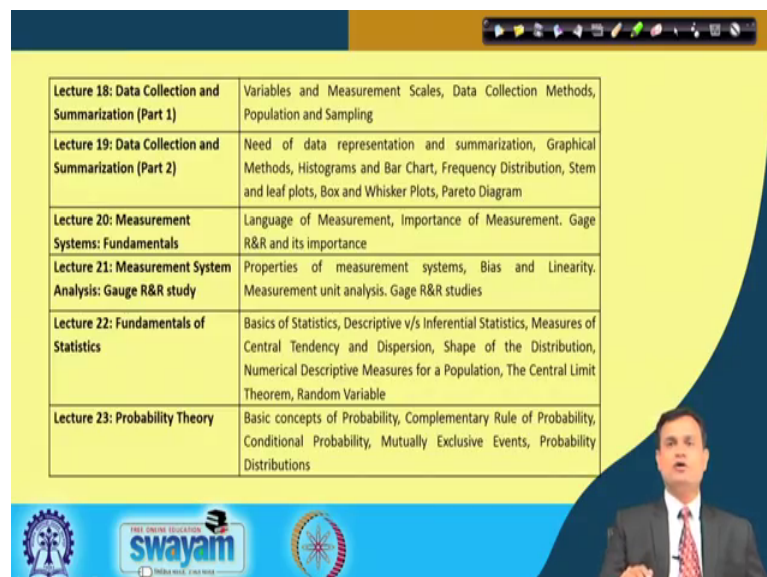


Lecture 14: Six Sigma Project Identification, Selection and Definition	Six Sigma Project, Tools used in Six Sigma Projects, Voice of Customer, Examples of CTS –CTQ, Project Metrics and Success Criteria
Lecture 15: Project Charter and Monitoring	Project Charter, Elements of Project Charter, Criteria for Project Monitoring, Steps of Monitoring System
Lecture 16: Process Characteristics and Analysis	Parameters in a Process, Process Flow Metrics, Process Analysis Tools, Process Maps and Flow Charts, Value Stream Map, Spaghetti Diagram, Circle Diagram
Lecture 17: Process Mapping: SIPOC	SIPOC, Process Components in SIPOC, Mapping the Process, Steps to draw SIPOC

The slide is part of a Swayam lecture series. It features a table with four rows of lecture topics. To the right of the table is a video feed of a male presenter in a suit. At the bottom, there are logos for Swayam and other educational institutions.

So in this phase, we talked about six sigma project identification, selection and definition. Lecture 15 project charter and monitoring that is the blueprint of the project, 16 process characteristic and analysis, 17 process mapping typically using a very important diagrammatic representation SIPOC. Then we have discussed in detail the measure phase we devoted week 4 and week 5.

(Refer Slide Time: 05:10)



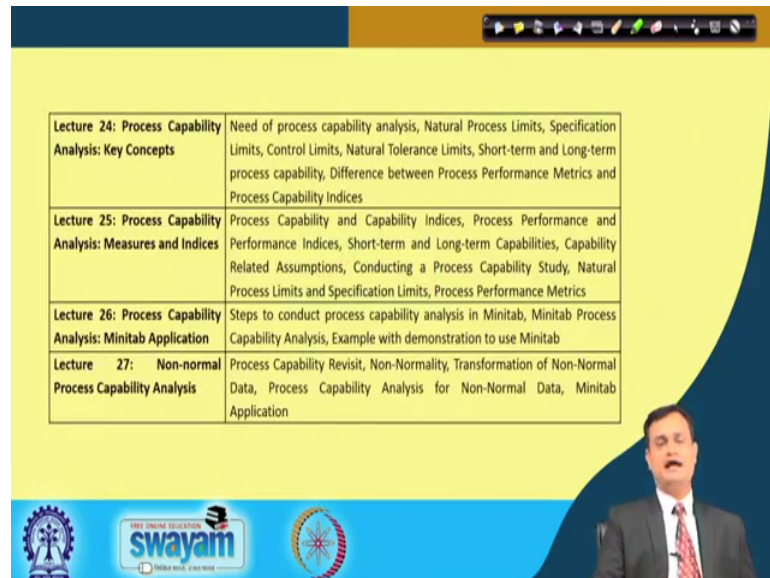
Lecture 18: Data Collection and Summarization (Part 1)	Variables and Measurement Scales, Data Collection Methods, Population and Sampling
Lecture 19: Data Collection and Summarization (Part 2)	Need of data representation and summarization, Graphical Methods, Histograms and Bar Chart, Frequency Distribution, Stem and leaf plots, Box and Whisker Plots, Pareto Diagram
Lecture 20: Measurement Systems: Fundamentals	Language of Measurement, Importance of Measurement. Gage R&R and its importance
Lecture 21: Measurement System Analysis: Gauge R&R study	Properties of measurement systems, Bias and Linearity. Measurement unit analysis. Gage R&R studies
Lecture 22: Fundamentals of Statistics	Basics of Statistics, Descriptive v/s Inferential Statistics, Measures of Central Tendency and Dispersion, Shape of the Distribution, Numerical Descriptive Measures for a Population, The Central Limit Theorem, Random Variable
Lecture 23: Probability Theory	Basic concepts of Probability, Complementary Rule of Probability, Conditional Probability, Mutually Exclusive Events, Probability Distributions

This slide continues the Swayam lecture series. It contains a table with six rows of lecture topics. A video of the presenter is visible on the right side. The Swayam logo and other institutional logos are at the bottom.

And measure phase week 4, we covered data collection and summarisation part 1 and part 2. Lecture 20 was on measurement system fundamentals, 21 measurement system

with gauge R and R study, lecture 22 fundamentals of statistics, 23 probability theory. Subsequently, we picked up few more topics in the measure phase in week 5.

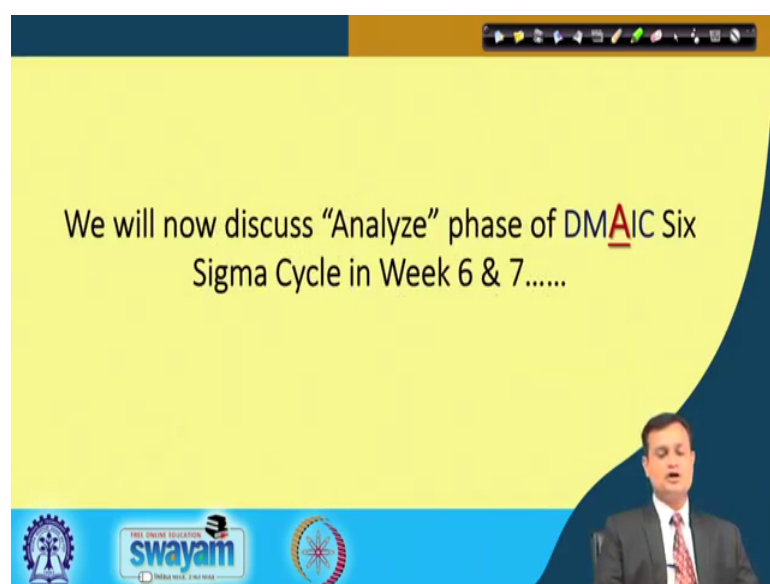
(Refer Slide Time: 05:35)



Lecture 24: Process Capability Analysis: Key Concepts	Need of process capability analysis, Natural Process Limits, Specification Limits, Control Limits, Natural Tolerance Limits, Short-term and Long-term process capability, Difference between Process Performance Metrics and Process Capability Indices
Lecture 25: Process Capability Analysis: Measures and Indices	Process Capability and Capability Indices, Process Performance and Performance Indices, Short-term and Long-term Capabilities, Capability Related Assumptions, Conducting a Process Capability Study, Natural Process Limits and Specification Limits, Process Performance Metrics
Lecture 26: Process Capability Analysis: Minitab Application	Steps to conduct process capability analysis in Minitab, Minitab Process Capability Analysis, Example with demonstration to use Minitab
Lecture 27: Non-normal Process Capability Analysis	Process Capability Revisit, Non-Normality, Transformation of Non-Normal Data, Process Capability Analysis for Non-Normal Data, Minitab Application

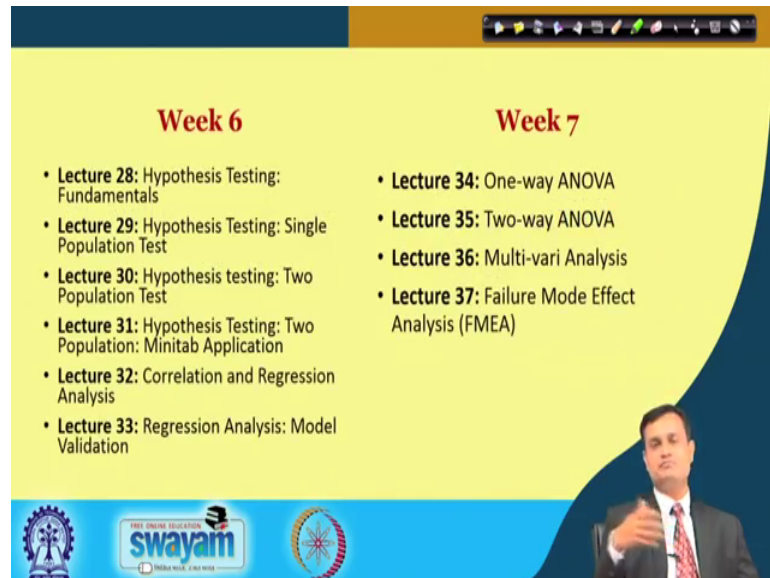
And this has included process capability analysis, some key concepts lecture 24, then lecture 25 process capability analysis measures and indices, lecture 26 process capability analysis with minitab application and lecture 27 non normal process capability analysis also we have included the minitab application.

(Refer Slide Time: 06:03)



So, with this we are now in the analysed phase we will now discuss the analysed phase of DMAIC six sigma cycle in week 6 and 7.

(Refer Slide Time: 06:18)



Week 6

- **Lecture 28:** Hypothesis Testing: Fundamentals
- **Lecture 29:** Hypothesis Testing: Single Population Test
- **Lecture 30:** Hypothesis testing: Two Population Test
- **Lecture 31:** Hypothesis Testing: Two Population: Minitab Application
- **Lecture 32:** Correlation and Regression Analysis
- **Lecture 33:** Regression Analysis: Model Validation

Week 7

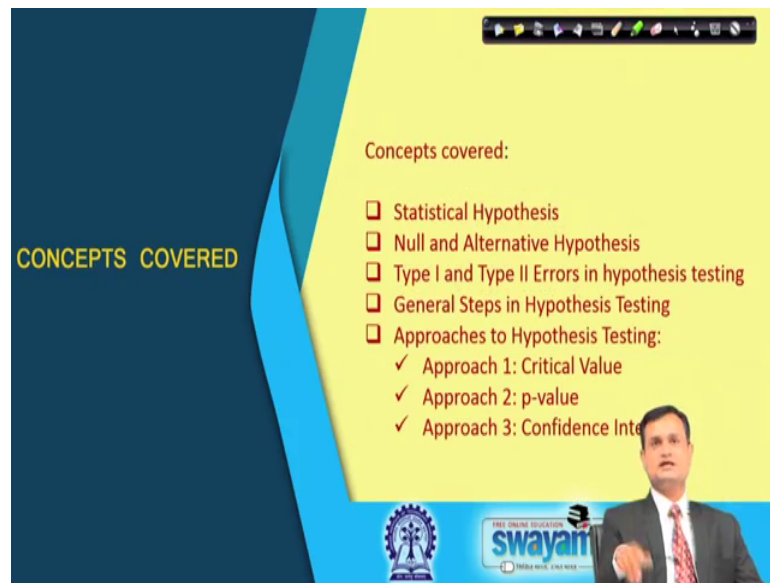
- **Lecture 34:** One-way ANOVA
- **Lecture 35:** Two-way ANOVA
- **Lecture 36:** Multi-vari Analysis
- **Lecture 37:** Failure Mode Effect Analysis (FMEA)

Logos at the bottom: IIT Bombay, swayam, and a circular logo.

Let me just give you the brief outline that we would be conducting, we would be executing in week 6 and week 7. So, week 6 this lecture 28 will help you to appreciate the concept of hypothesis testing. Lecture 29, hypothesis testing for single population test. Lecture 30, hypothesis testing for 2 population test. 31, hypothesis testing for 2 population test I would like to share the minitab application. Lecture 32, correlation and regression analysis and 33 regression analysis model validation.

Week 7, we will have 4 lectures. Lecture 31 one-way ANOVA, lecture 35, two-way ANOVA, lecture 36, multi-vari analysis and 37, failure mode effect analysis.

(Refer Slide Time: 07:15)



CONCEPTS COVERED

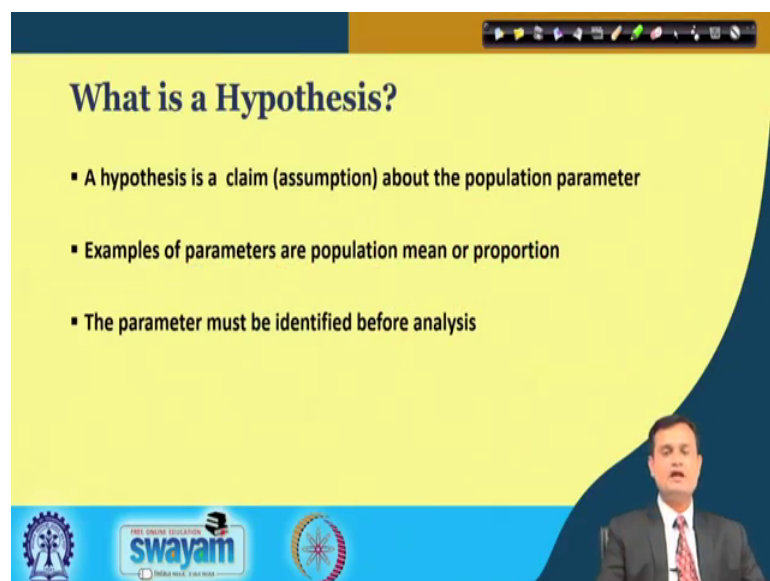
Concepts covered:

- ❑ Statistical Hypothesis
- ❑ Null and Alternative Hypothesis
- ❑ Type I and Type II Errors in hypothesis testing
- ❑ General Steps in Hypothesis Testing
- ❑ Approaches to Hypothesis Testing:
 - ✓ Approach 1: Critical Value
 - ✓ Approach 2: p-value
 - ✓ Approach 3: Confidence Interval

The slide features a dark blue background on the left with the title 'CONCEPTS COVERED' in yellow. The right side has a yellow background with a list of concepts. A small video inset of a man in a suit is in the bottom right corner. Logos for 'swayam' and 'INDIAN INSTITUTE OF TECHNOLOGY' are at the bottom.

So, with this recap let us try to focus on the important topics of this lecture 28 hypothesis testing fundamentals. So, we will try to see what is statistical hypothesis? What is null and alternative hypothesis? What is type 1 and type 2 error in hypothesis testing? What are the general steps in hypothesis testing? And typically there are three approaches for hypothesis testing approach number 1 critical value, approach number 2 p value, approach number 3 confidence interval.

(Refer Slide Time: 07:55)



What is a Hypothesis?

- A hypothesis is a claim (assumption) about the population parameter
- Examples of parameters are population mean or proportion
- The parameter must be identified before analysis

The slide has a yellow background with a dark blue curved border on the right. It contains a definition and examples of hypotheses. A small video inset of a man in a suit is in the bottom right corner. Logos for 'swayam' and 'INDIAN INSTITUTE OF TECHNOLOGY' are at the bottom.

So, what is hypothesis? In a very simple way I would like to say that its your claim about a particular problem or phenomena under investigation and you want to study this particular critical problem in a scientific manner so a hypothesis is basically a claim assumption.

I do not have the theory available, I want to test the theory either existing or propose a new theory and a scientific rigor is necessary to propose any new theory or test the existing theory then I would set a hypothesis about the population parameter. Just for example, I want to investigate the impact of worker morale on the productivity of the organization, its a very interesting phenomena. Now, there could be many factors that can impact the worker morale, it may be compensation, hygiene, incentives, recognition, many other things, transparent working environment.

You have various ways to measure the productivity. It may be labour productivity, it may be equipment productivity, it may be transactional productivity. Now, I want to test the theory for an organization or in general that worker morale or employee morale and the organizational productivity they are positively related. Its my claim, its not yet proven for a typical context fine. It might have been proven for US or UK or other context for some other organization.

I want to check it for my context for my organization then, this claim assumption is my hypothesis and I would like to test it through a systematic scientific procedure. So, examples of parameters are population mean and proportion and the parameters must be identified before we go for the hypothesis testing analysis.

(Refer Slide Time: 10:10)

Statistical Hypothesis

➤ **Definition**

- A statistical hypothesis is a statement about the parameter of one or more populations

Example

$$H_0 : \mu_{\text{female}} = \mu_{\text{male}}$$

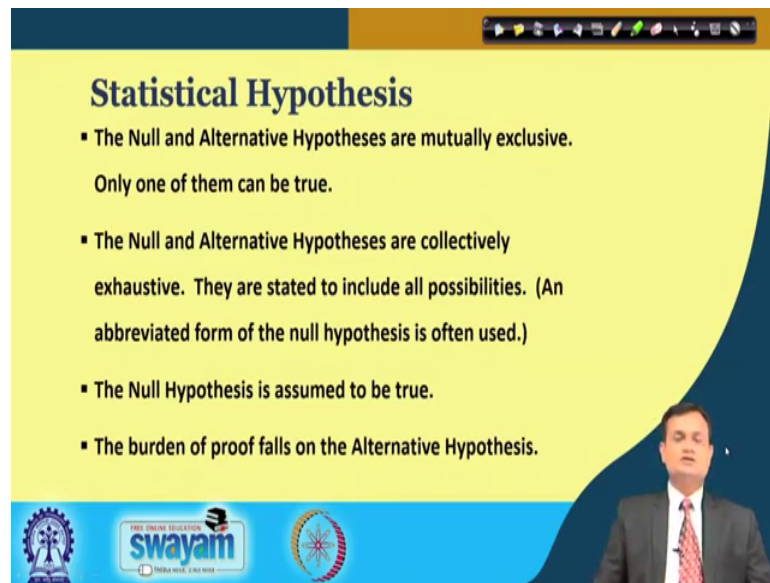
Where μ is the average IQ of the named gender

So, just see how we try to figure it out that the definition a statistical hypothesis is a claimed statement about the parameter of one or more population. Here I am citing just a simple, but interesting one that H_0 is typically called as null hypothesis 0 hypothesis null hypothesis. So, always null hypothesis assumes be conformance to the particular requirement and usually it is expressed in terms of equality and I believe that this is possible this can happen.

So, I will say μ of female and μ of male when I am trying to compare the μ for IQ their intelligence, I am trying to say that IQ of mean value of IQ of female and mean value of IQ of male it is equal. Obviously, I will take a sample and for that sample I would like to study because you cannot just take some heterogeneous sample and data and then try to study. Because qualification, upbringing, economic strata, food many things they have an impact on the individual IQ.

So, here for a particular group I want to study that IQ for female and male and I want to state it in a scientific manner null hypothesis that the IQ of female and IQ male mean value they are equal.

(Refer Slide Time: 11:59)



Statistical Hypothesis

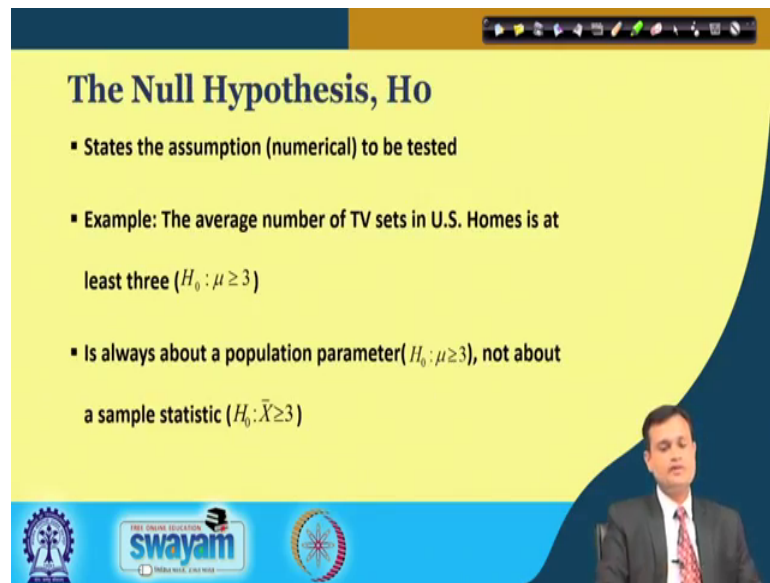
- The Null and Alternative Hypotheses are mutually exclusive.
Only one of them can be true.
- The Null and Alternative Hypotheses are collectively exhaustive. They are stated to include all possibilities. (An abbreviated form of the null hypothesis is often used.)
- The Null Hypothesis is assumed to be true.
- The burden of proof falls on the Alternative Hypothesis.

Logos at the bottom: Swamyam, and other institutional logos.

So, typically my null hypothesis and; obviously, I have null hypothesis when I do not agree with this the other hypothesis I can emerge that is called alternative hypothesis. And I believe that, null hypothesis and alternative hypothesis are mutually exclusive only one of them can occur. So, in this case either μ of that is the IQ mean IQ of female is equal to mean IQ of male or it is not equal. So, I have null hypothesis equal alternate hypothesis alternate claim not equal.

So, null hypothesis and alternative hypothesis are collectively exhaustive, they are stated to include all possibilities and abbreviation from null hypothesis is often used. So, null hypothesis is assumed to be true the burden of proof falls on the alternative hypothesis. So, here if I have null hypothesis true; obviously, alternate is false if I say I fail to accept my null hypothesis or in simple word it is rejected then I will say alternate hypothesis is true.

(Refer Slide Time: 13:20)



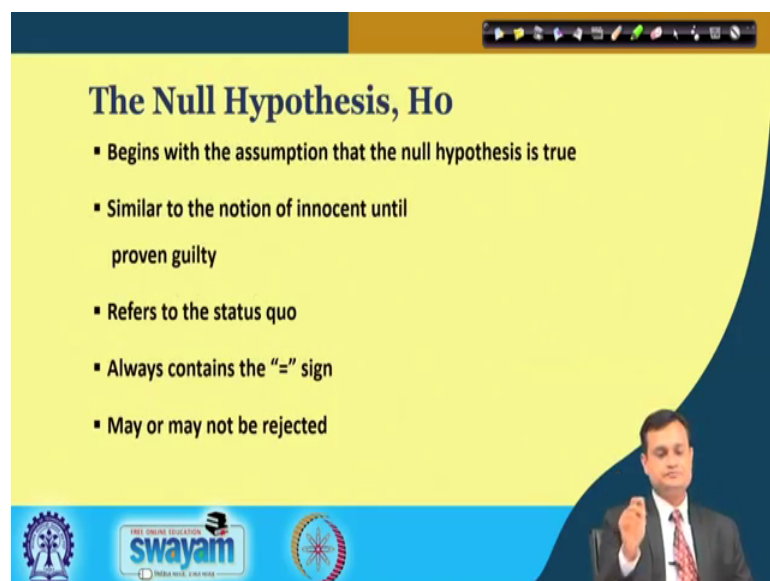
The Null Hypothesis, H_0

- States the assumption (numerical) to be tested
- Example: The average number of TV sets in U.S. Homes is at least three ($H_0: \mu \geq 3$)
- Is always about a population parameter ($H_0: \mu \geq 3$), not about a sample statistic ($H_0: \bar{X} \geq 3$)

The slide features a yellow background with a blue curved border on the right. At the bottom, there are logos for 'swayam' and other educational institutions, along with a small video inset of a man in a suit.

Now, just a little bit more idea on null hypothesis with example that, it is the assumption to be tested. And as I said usually, but not necessarily that it is expressed in terms of equality you can also have inequality just see the example the average number of TV sets in US homes is at least 3, this is my null hypothesis. It is 3 or greater than 3; it is always about a population parameter not about the sample statistic so $H_0: \mu \geq 3$ is my claim that I want to check, test scientifically.

(Refer Slide Time: 14:05)



The Null Hypothesis, H_0

- Begins with the assumption that the null hypothesis is true
- Similar to the notion of innocent until proven guilty
- Refers to the status quo
- Always contains the "=" sign
- May or may not be rejected

This slide is similar to the previous one, with a yellow background and blue border. It lists five characteristics of the null hypothesis. The bottom section includes the same logos and a video inset of the presenter.

So, my null hypothesis is something which begins with the assumption refers to the status quo that yes this will prevail or this is prevailing and usually as I said contains equal to sign, but not necessary and may or may not be rejected.

(Refer Slide Time: 14:25)

The Alternative Hypothesis, H_1

- Is the opposite of the null hypothesis
- e.g.: The average number of TV sets in U.S. homes is less than 3 ($H_1: \mu < 3$)
- Challenges the status quo
- Never contains the "=" sign
- May or may not be accepted
- Is generally the hypothesis that is believed (or needed to be proven) to be true by the researcher

The slide features a yellow background with a dark blue header and footer. A video feed of a man in a suit is visible in the bottom right corner. Logos for 'swayam' and other educational institutions are in the bottom left.

So, alternative hypothesis; obviously, when you say number of TVs in a home alternate hypothesis would be less than 3 if it is equal to or greater than 3, its null hypothesis. When I say alternate hypothesis, it is less than 3. So, alternate hypothesis typically is designated as H_1 .

So, H_0 is my null H_1 is my alternate and typically this hypothesis never contains equal to sign either it is in inequality form or it is non-equal to and this is the generally the hypothesis that is believed to be true by the researcher or needed to be proven.

(Refer Slide Time: 15:14)

Null and Alternative Hypotheses: Example

- A manufacturer is filling 40 Kg. packages with flour.
- The company wants the package contents to average 40 Kgs.

$$H_0 : \mu = 40 \text{ Kg}$$
$$H_a : \mu \neq 40 \text{ Kg}$$

Logos: Anna University, swamyam, Anna University

So, we can see the example of null and hypothesis together that, you have a manufacturing company and is filling 40 kg packages with flour now, company wants to package that contained the average of 40 kg so, I will say null hypothesis I want to check. I have collected the data, measured the package and now I want to prove that statistically really I am filling 40 kg. So, mu is equal to 40 kg alternate, so H 1 or H a both is ok, you can express it as H 1 or H a mu is not equal to 40 kg.

(Refer Slide Time: 15:57)

One-tailed and Two-tailed Tests

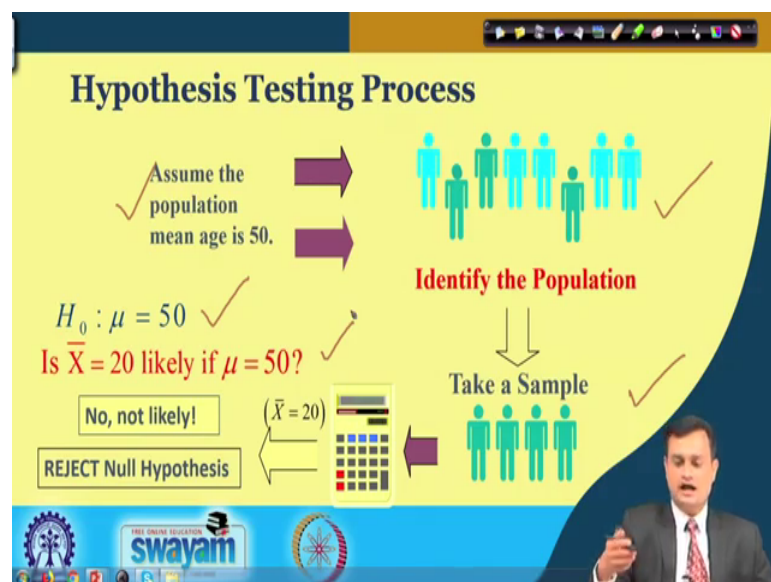
- One-tailed Tests
$$H_0 : \mu = 40 \quad H_a : P = 0.18$$
$$H_a : \mu < 40 \quad H_a : P > 0.18$$
- Two-tailed Test
$$H_0 : \mu = 12$$
$$H_a : \mu \neq 12$$

Logos: Anna University, swamyam, Anna University

I can have one-tailed test, I can have two-tailed test, you have seen the distribution, normal distribution typically and I can have one-tailed test two-tailed test. So, when I follow the one-tailed test it means, I am only bothered about either higher side or lower side just see here H_0 is μ is equal to 40 and H_a that is my alternative hypothesis that is μ is less than 40.

So, I would say that if it is less than 40 not acceptable, but if it is equal to 40 then I would prove that my claim about filling the bag with 40 kg is true. Two-tailed, I would say μ is equal to 12 μ is not equal to tail. So, I would like to explore both the sides of the distribution and I will conduct the two-tailed test.

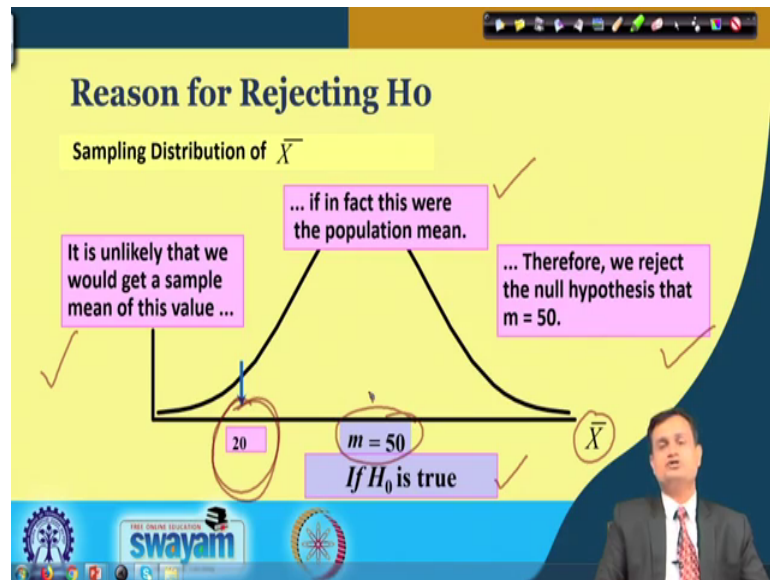
(Refer Slide Time: 16:56)



So, just see this to make your idea on hypothesis the word may sound little bit more difficult, but its a scientific and it is the simplest way to say its my claim about the population. So, you have the population on about which you want to test some of the claim may be average expectancy of the life, height, weight of the population, their socio-cultural values many a times. Now you take a sample and here I am assuming the population with mean age 50 and you can see that my H_0 says μ is equal to 50, so on an average life is 50 is \bar{X} bar 20 likely if μ is equal to 50, so this is what I want to check.

So, my hypothesis basically would help me to structure my problem and then to apply the scientific analysis to see that to what extent my claim about a population based on the sample is really true.

(Refer Slide Time: 18:13)



(Refer Slide Time: 18:19)

The slide is titled "Level of Significance". It contains a bulleted list of six points: "Defines unlikely values of sample statistic if null hypothesis is true", "Called rejection region of the sampling distribution", "Is designated by α , (level of significance)", "Typical values are .01, .05, .10", "Is selected by the researcher at the beginning", and "Provides the critical value(s) of the test". The slide includes a Swayam logo and a small video inset of a man in a suit.

So, just see this reasons for rejecting H_0 . And what you can see here that, what you can see here that sampling distribution is \bar{X} . So, I have taken a sample and I refer to the sampling distribution this is my \bar{X} sampling distribution I have taken maybe 100

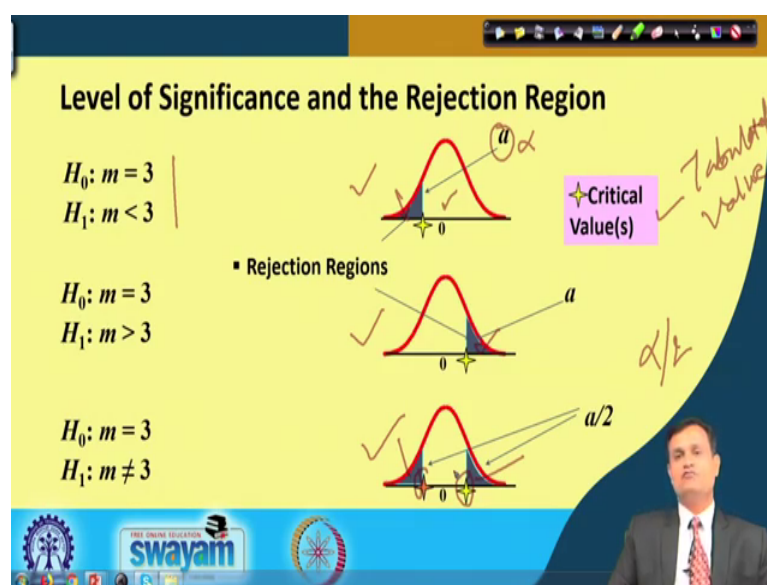
sample, each sample mean is taken and when I plot \bar{X} instead of let us individual value then, I am plotting the sampling distribution of \bar{X} .

So, it is unlikely that we would get a sample mean of this value that is 20 here and this is my 50 which says that H_0 is true and if in fact, this were the population mean therefore, we reject the null hypothesis that μ is equal to 50. I have not done any inferential analysis, just I am trying to show you and visually I am trying to say that you will say that yes this is too far and too far has to be proven statistically, but I would say that this is 20 and this where if this were the population mean then we reject the null hypothesis. But, this is not at all scientific I have to use the inferential statistics to prove that yes my claim about a particular population is true to what extent.

So, you have one term that is called level of significance. We always say that, when we deal with the probability theory and we draw our conclusions based on the sample. We always say that we cannot be 100 percent sure, we always try to check or test the hypothesis at particular level of significance. So, technically this defines unlikely value of sample statistic if null hypothesis is true.

So, this helps us to define the rejection region of my sampling distribution and typically it is denoted in all the textbook you will find; this is denoted as alpha level of significance. Typical values you can choose 0.01, 0.02, 0.05, 0.1 we will discuss what is the real interpretation of higher and lower value, but you can choose conveniently 0.01, 0.05. If this is selected by the researcher and at the beginning then he would be comfortable in checking the critical value finding from the table we had seen the various distributions and the table value. So, then he can prove that whether the null hypothesis is accepted or it is rejected.

(Refer Slide Time: 21:24)



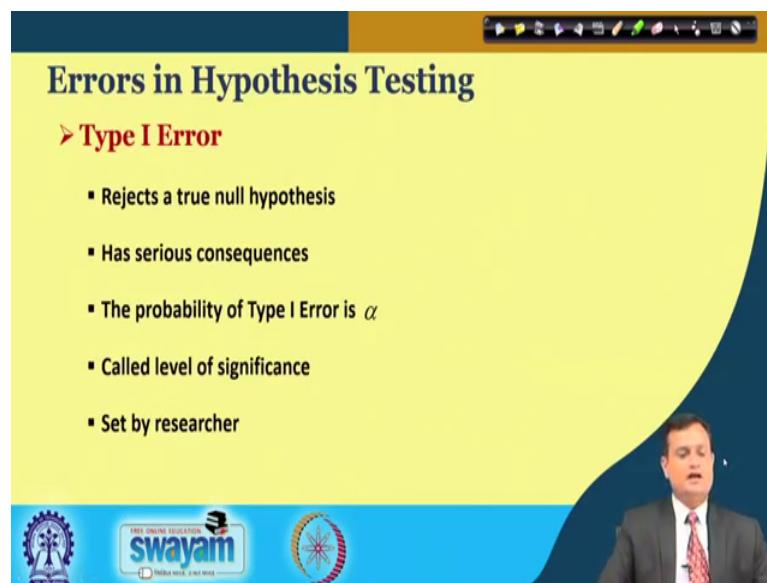
So, just see this will give you a very clear cut idea what I have here say I have null hypothesis m is equal to 3 and alternate hypothesis m less than 3. Now, I have the critical value α corresponding to my say particular claim and I want to this is my alpha rather you say not α , this is alpha I have this alpha and this is the level of significance.

So, this will give me a critical value. So, critical value is basically the tabulated value. So, from this statistical table for a given level of significance alpha you find the critical value and then you check the hypothesis, if it falls in this region particularly this shaded region then it is reject means null hypothesis my null hypothesis claim is reject and I will accept the alternate hypothesis. If it is fall, it falls in the non-shaded region then, I will accept the null hypothesis and alternate hypothesis is rejected.

Same way, you can have the rejection region on this side when you talk about only one side, one-tail test this is one-tail test, this is also one-tailed test and you have here two-tailed test. So, I will divide the alpha into 2 part alpha by 2 so suppose, you have selected 0.05 as the alpha then 0.025 and 0.025 will become alpha by 2; alpha by 2.

So, this region is the rejection region once again I am reminding you. And if, these regions are determined based on the critical values found from the statistical tables for a particular statistics of interest we will see it and if my calculated value falls in this region shaded region then it is reject if it falls in the non-shaded region my null hypothesis is accepted.

(Refer Slide Time: 23:51)



Errors in Hypothesis Testing

➤ Type I Error

- Rejects a true null hypothesis
- Has serious consequences
- The probability of Type I Error is α
- Called level of significance
- Set by researcher

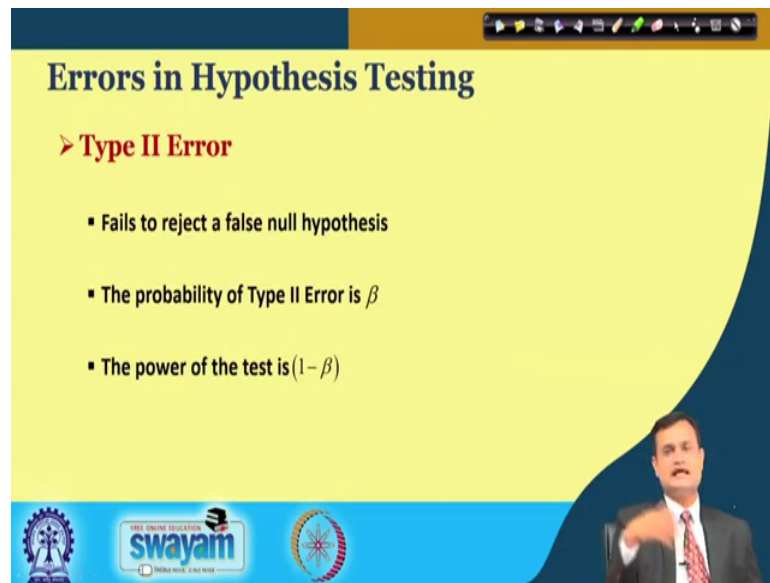
swayam

So, there are 2 errors when we deal with the probability and when we talk about the level of significance we should understand that there are 2 types of errors that a researcher, scientist, analyst has to accept and these are typically called as type I and type II error.

So, type I error means reject a true null hypothesis. It means, actually my null hypothesis is true, but for the chosen alpha given level of significance I say that my null hypothesis is rejected. So, if you recall, I gave you the example that suppose you have a person presented in the court against the jury before the jury and the decision is that he is guilty he should be punished, but actually the person is innocent.

Now in this case you are basically rejecting the null hypothesis that person is innocent and punishing him. So, this is your say type I error. So, there are certain serious consequences if you punish an innocent person and if you do not punish let us say, guilty there could not be that much of consequences because this fellow will be caught later on also.

(Refer Slide Time: 25:18)



Errors in Hypothesis Testing

➤ **Type II Error**

- Fails to reject a false null hypothesis
- The probability of Type II Error is β
- The power of the test is $(1 - \beta)$

So, this is alpha is said by the researcher. Type II error is to fail, to reject the false null hypothesis. It is just reverse, I say that null hypothesis is rejected, but here the null hypothesis is accepted, but actually it should be rejected. So, it is just reverse. So, in my way I will say that you are say declaring a guilty; declaring a guilty and innocent means you are not punishing the guilty you are accepting the say null hypothesis, but actually it is of reject.

So, this is something that basically leads to the error in any kind of statistical analysis, but we have to bear and we have to be judicious about deciding the type I and type II error. So, probability of type II error is typically beta and the power of test is 1 minus beta.

(Refer Slide Time: 26:23)

Result Probabilities

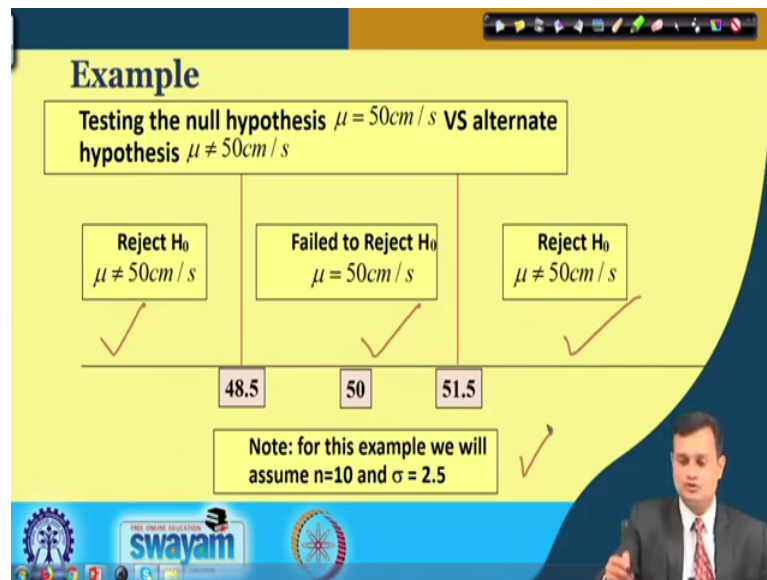
H_0 : Innocent

Jury Trial			Hypothesis	Test
	The Truth			The Truth
Verdict	Innocent	Guilty	Decision	H_0 True H_0 False
Innocent	Correct	Error	Do Not Reject H_0	$1 - \alpha$ Type II Error (β)
Guilty	Error	Correct	Reject H_0	Type I Error (α) Power ($1 - \beta$)

So, jury trial example I have given is well demonstrated here. Just see this to make it more clear you have innocent and I declare it as the innocent. So, there is a verdict you know and there is innocent correct. When I say, verdict is guilty and I declare it as innocent as a jury error when I say verdict is innocent and guilty error both is guilty this is correct. So, do not reject null hypothesis H_0 is true, I will say 1 minus alpha and if H_0 is false this is type 2 error. If H_0 is rejected, when H_0 is true then this is my type I error and when H_0 is rejected, when H_0 is really false, this is actually the power of my test.

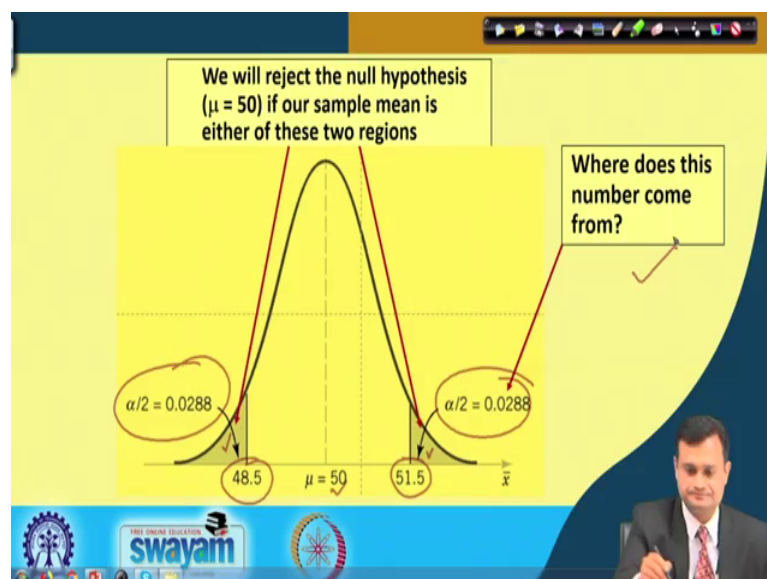
So, to what extent I am able to discriminate the truth from the lie or I am able to discriminate in quality context bad quality from the good quality this is where exactly the power of my statistical test lies.

(Refer Slide Time: 27:44)



So, we can see this through a small example, testing the null hypothesis μ is equal to 50 centimetre by per second versus alternate hypothesis that it is not this much of speed. So, you have these 3 options, 3 regions, reject H_0 μ is not equal to this, fail to reject that is μ is equal to 50 cm per second and reject H_0 , μ is not equal to 50 cm per s. So, you are basically trying to say check it with respect to 2 side 2 tail. So, for this example we will assume n is equal to 10 and σ is equal to let us say 2.5.

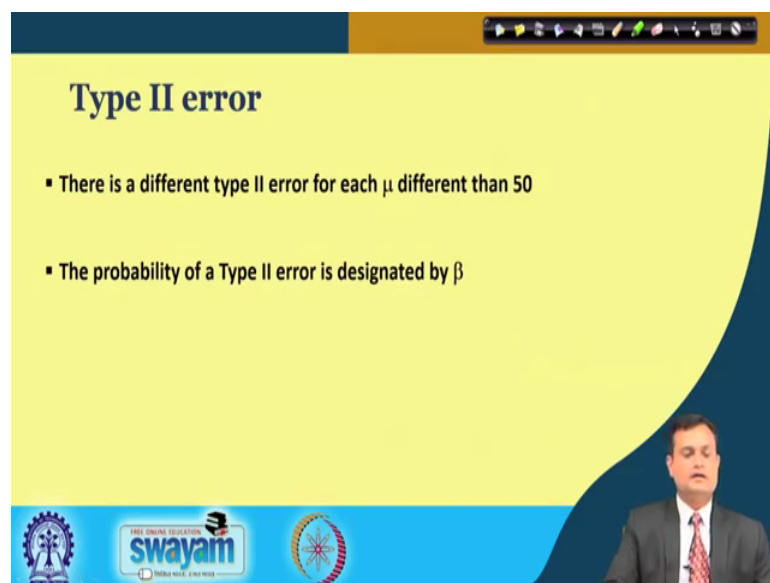
(Refer Slide Time: 28:32)



So, I would say that when I have put the alpha by 2. So, this is my alpha by 2, this is my alpha by 2 for the chosen value of my significance and I have the critical value here, I have the critical value here and I have mu is equal to 50.

So, we will basically reject the null hypothesis mu is equal to 50, if our sample mean is either of these regions. So, if it falls in this region, I will reject the null hypothesis this shaded region is my reject region. So, this value typically comes from the table your statistical table for the given level of alpha.

(Refer Slide Time: 29:24)



Type II error

- There is a different type II error for each μ different than 50
- The probability of a Type II error is designated by β

The slide features a yellow background with a blue border. At the bottom, there are logos for Swayam and other educational institutions, and a small video inset of a man in a suit.

Now, let us try to appreciate the importance of type II error. So, there is a different type II error for each mu different than 50 in this case and the probability of a type II error is designated by beta.

(Refer Slide Time: 29:42)

The Probability of Type II Error and Sample Size Decisions

- In most hypothesis testing situations, it is important to determine the probability of type II error associated with the test.
- Equivalently, we may elect to evaluate the power of the test.
- To illustrate how this may be done, we will find the probability of type II error associated with the test of
$$H_0: \mu = \mu_0$$
$$H_0: \mu \neq \mu_0$$
where the variance σ^2 is known.

So, typically if we see then little bit it is difficult complex, but I will make it simple. Then, in most of the hypothesis testing situation it is most important to determine type II error because it also decides the power of my test.

So, here the situation is like this H_0 is equal to μ is equal to μ_0 and H_0 is μ is not equal to μ_0 and the variance is sigma square which is known let us say.

(Refer Slide Time: 30:16)

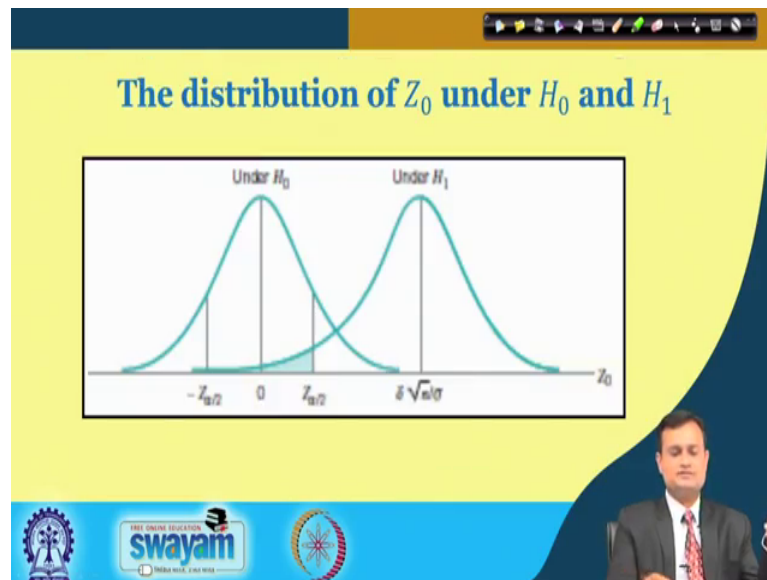
The Probability of Type II Error and Sample Size Decisions

The test statistic for this hypothesis is
$$Z_0 = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$
and under the null hypothesis the distribution of Z_0 is $N(0, 1)$.

- To find the probability of type II error, we must assume that the null hypothesis $H_0: \mu = \mu_0$ is false and then find the distribution of Z_0 .
- Suppose that the mean of the distribution is really $\mu_1 = \mu_0 + \delta$, where $\delta > 0$. Thus, the alternative hypothesis $H_0: \mu \neq \mu_0$ is true, and under this assumption the distribution of the test statistic Z_0 is
$$Z_0 \sim N\left(\frac{\delta\sqrt{n}}{\sigma}, 1\right)$$

Now, what exactly this μ and μ_0 means?

(Refer Slide Time: 30:22)



So, I am just trying to show you the figure first, I am saying that I have H_0 and I have H_1 . So, there could be a possibility that my process has shifted and now under this situation what is that region so you can very well see this shaded region this particular shaded region this will typically gives you the type II error. So, when there is a shift in the process what is that region that will prompt me to make the type II error, it means not rejecting the null hypothesis when it is really to be rejected ok or this is something that we can also see through the calculation, so just we can see.

So, I can find out the test static for this that is $Z_0 = \bar{x} - \mu_0$ divided by σ/\sqrt{n} and I have this statistics, that this Z is equal to $\bar{x} - \mu_0$ divided by σ/\sqrt{n} and I believe that this follows the normal distribution. So, here my μ_1 is little shift as I mentioned with δ in μ_0 and this Z_0 typically follows a normal distribution with $\mu = \delta/\sigma$ and $1/\sqrt{n}$ that is the shift divided by sigma and 1 has the standard deviation. So, for this suppose if I do the analysis for type II, then I can find this particular beta; I can find this particular beta.

(Refer Slide Time: 32:26)

- The distribution of the test statistic Z_0 under both hypotheses H_0 and H_1 is shown in Figure.
- We note that the probability of type II error is the probability that Z_0 will fall between $-Z_{\alpha/2}$ and $Z_{\alpha/2}$ given that the alternative hypothesis H_1 is true. To evaluate this probability, we must find $F(Z_{\alpha/2}) - F(-Z_{\alpha/2})$ where F denotes the cumulative distribution function of the $N(\delta\sqrt{n}/\sigma, 1)$ distribution. In terms of the standard normal cumulative distribution, we then have

$$\beta = \phi\left(Z_{\alpha/2} - \frac{\delta\sqrt{n}}{\sigma}\right) - \phi\left(-Z_{\alpha/2} - \frac{\delta\sqrt{n}}{\sigma}\right)$$

So, here when I say phi it is cumulative distribution your total probability over a horizon and this is $Z_{\alpha/2}$ minus δ square root n divided by σ minus cumulative probability minus $Z_{\alpha/2}$ minus δ square root n divided by σ . Its very simple, just see this figure you will have a clear idea.

So, here what I am trying to see that what is the region that I have between my minus $Z_{\alpha/2}$ and δ square root n and this and minus $Z_{\alpha/2}$ and this so you will end up with this particular region. So, to find the probability of this region basically, I have taken the subtraction of these two which is this one $Z_{\alpha/2}$ minus δ square root n divided by σ minus this smaller region.

(Refer Slide Time: 33:53)

Example

The mean contents of coffee cans filled on a particular production line are being studied. Standards specify that the mean contents must be 16.0 oz, and from past experience it is known that the standard deviation of the can contents is 0.1 oz. The hypotheses are

$$H_0: \mu = 16.0$$
$$H_0: \mu \neq 16.0$$

A random sample of nine cans is to be used, and the type I error probability is specified as $\alpha = 0.05$. Therefore, the test statistic is

$$Z_0 = \frac{\bar{x} - 16.0}{0.1/\sqrt{9}}$$

And H_0 is rejected if $|Z_0| > Z_{0.025} = 1.96$. Find the probability of type II error and the power of the test, if the true mean contents are $\mu_1 = 16.1$ oz.

The slide features a yellow background with a blue border. At the bottom, there is a blue banner with the 'swayam' logo and a small video feed of a man in a suit. The text is in black, and the formula is highlighted with a red circle and checkmarks.

So, you can find the type II error and just see the example small before we summarize this particular lecture. So, the mean contents of coffee cans filled on a particular production line are being studied and the standard specify that the mean contains must be 16 ozone and from the past experience it is known that the standard deviation of the can content is 0.1. And now I want to set the hypothesis mu is equal to 16 mu is not equal to 16, so I can find my Z statistic we have seen in this probability theory that Z_0 is \bar{x} minus mu x bar minus mu that is 16 divided by sigma divided by square root n because I am dealing with the sample.

So, as for the central limit theorem sigma divided by square root n my sigma is 0.1 and n is typically the sample size that is 9. So, here you will find that the value you get is Z_0 is basically if you see this Z_0 is basically greater than $Z_{0.025}$. So $Z_{0.025}$ is the critical value that you have obtained from the table and your value comes out to be greater than this which is 1.96. So you will say that my particular claim needs to be tested for this particular relationship and now find the probability of error and power of the test, if true mean contents our mu 1 16.1 ozone. So, I am now considering that delta part little shift and I want to compute the my type II error.


(Refer Slide Time: 35:54)

Solution

Since we are given that $\delta = \mu_1 - \mu_0 = 16.1 - 16.0 = 0.1$, we have

$$\begin{aligned}\beta &= \varphi\left(Z_{\alpha/2} - \frac{\delta\sqrt{n}}{\sigma}\right) - \varphi\left(-Z_{\alpha/2} - \frac{\delta\sqrt{n}}{\sigma}\right) \\ &= \varphi\left(1.96 - \frac{(0.1)\sqrt{9}}{0.1}\right) - \varphi\left(-1.96 - \frac{(0.1)\sqrt{9}}{0.1}\right) \\ &= \varphi(-1.04) - \varphi(-4.96) \\ \beta &= 0.1492\end{aligned}$$

- That is, the probability that we will incorrectly fail to reject H_0 if the true mean contents are 16.1 oz is 0.1492. Equivalently, we can say that the power of the test is $1 - \beta = 1 - 0.1492 = 0.8508$.


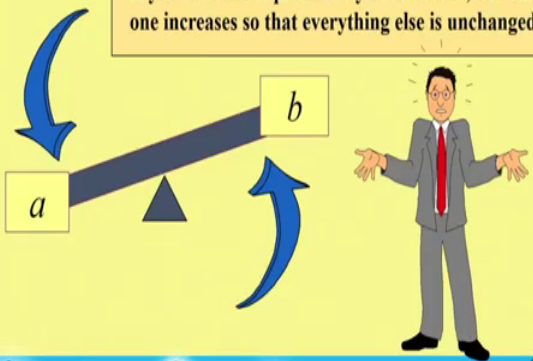


So, just see that how this can be done. So, now, I am using this expression which I have explained. This is typically the finding the shaded region for type II error I am just plugging in the values and what I find that beta is equal to 0.1492. So, 0.1492 is the type II error it means we will incorrectly fail to reject at 0, if the true mean contains are 16.1 is point 1 ozone is 0.1492 that is 14.92 percent or equivalently we can say that the power of test is 85 percent. So, this will help me to appreciate that to what extent I am critical and discriminative in applying my scientific analysis and I am able to make the decision.

(Refer Slide Time: 36:51)

Type I & II Errors Have an Inverse Relationship

If you reduce the probability of one error, the other one increases so that everything else is unchanged.



So, now the question comes what is the trade off between type I and type II error?

(Refer Slide Time: 38:59)

Factors Affecting Type II Error

- True value of population parameter
 - β Increases when the difference between hypothesized parameter and its true value decrease
- Significance level
 - β Increases when α decreases
- Population standard deviation
 - β Increases when σ increases
- Sample size
 - β Increases when n decreases

Diagram illustrating the relationship between β and other factors:

- β and α : β increases (up arrow) as α decreases (down arrow).
- β and σ : β increases (up arrow) as σ increases (up arrow).
- β and n : β increases (up arrow) as n decreases (down arrow).

Logos: IIT Bombay, swayam, and a circular logo.

So, just see this when beta increases typically the differences between hypothesis parameter and its true value increases. Significance level when you try to understand with respect to that, beta increases, alpha decreases. Population standard deviation beta increases when sigma increases, sample size beta increases when n decreases. So, there is lot of thing that even you can have in your control in order to set or in order to have a trade off between alpha and beta, but what is the right trade off.

(Refer Slide Time: 37:37)

How to Choose between Type I and Type II Errors

- **Choice depends on the cost of the errors**
- **Choose smaller Type I Error when the cost of rejecting the maintained hypothesis is high**
 - ✓ A criminal trial: convicting an innocent person
 - ✓ The Exxon Valdez: causing an oil tanker to sink

Logos: IIT Bombay, swayam, and a circular logo.

So, you must look at the particular situation and it is the situation problem which you are investigating will help you to decide what is the acceptable alpha level? So, as I mentioned in a criminal trial convicting an innocent person is not acceptable. So, in this case, you try to choose the lesser value of α for type I error level of significance.

(Refer Slide Time: 38:02)

How to Choose between Type I and Type II Errors

- **Choose larger Type I Error when you have an interest in changing the status quo**

- ✓ A decision in a startup company about a new piece of software
- ✓ A decision about unequal pay for a covered group

Logos at the bottom: Swamyam, Ministry of Education, Government of India, and a speaker in a suit.


So, choose larger type one error when you and say interest in changing the status queue. You want people to change, you want to have a new startup, you do not want to get discouraged and you want to prove that my business will bring more profit or it would be successful then higher alpha value is acceptable.

(Refer Slide Time: 38:26)

General Steps in Hypothesis Testing

▪ e.g.: Test the assumption that the true mean number of TV sets in U.S. homes is three (Known σ)

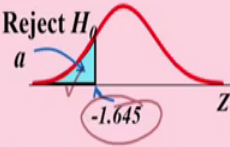
1. State the H_0	$H_0 : \mu \geq 3$
2. State the H_1	$H_1 : \mu < 3$
3. Choose α	$\alpha=0.05$
4. Choose n	$n=100$
5. Choose Test	Z test




So, general steps to summarize in hypothesis testing state H_0 , state H_1 , you can parallelly see the green box, choose alpha choose n , choose the test.

(Refer Slide Time: 38:39)

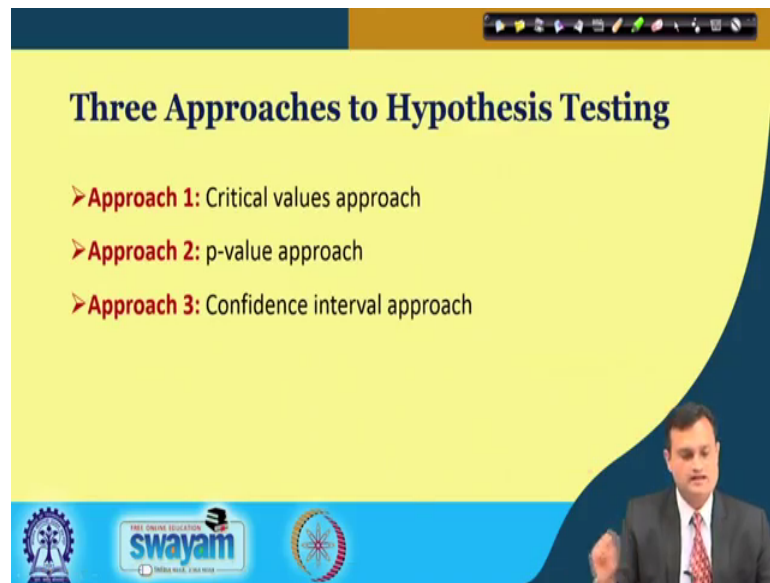
General Steps in Hypothesis Testing

6. Set up critical value(s)	 <p>100 households surveyed Computed test stat = -2, p-value = .0228 Reject null hypothesis The true mean number of TV sets is less than 3</p>
7. Collect data	
8. Compute test statistic and p-value	
9. Make statistical decision	
10. Express conclusion	



Here, let us say it is Z test decide the say particular region based on the critical value here the blue shaded region is the critical region and this is the critical value which I found from the table for a given. Say alpha setting and then compute these statistics and p value make the statistical decision whether my claim is really true or it is rejected.

(Refer Slide Time: 39:12)



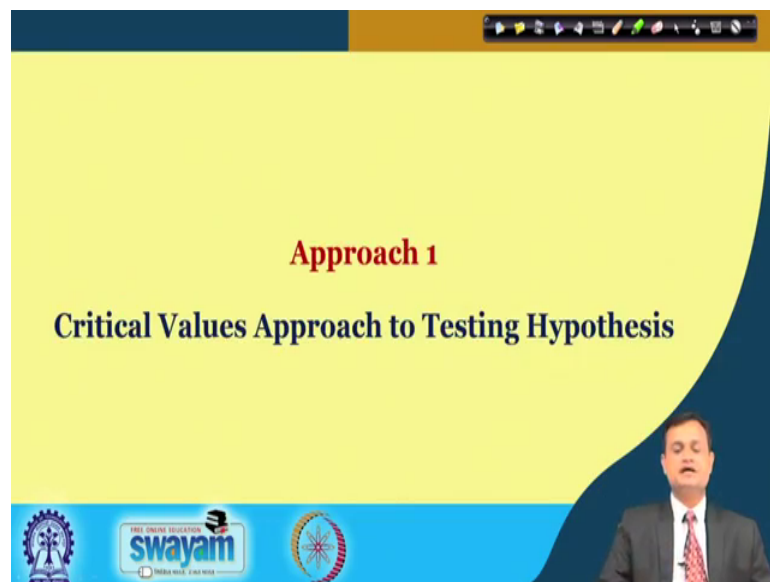
Three Approaches to Hypothesis Testing

- **Approach 1:** Critical values approach
- **Approach 2:** p-value approach
- **Approach 3:** Confidence interval approach

The slide features a yellow background with a dark blue curved border on the right. At the bottom, there is a blue banner containing logos for 'swayam' and other educational institutions. A small video inset of a man in a suit is visible in the bottom right corner.

So, we have three approaches to summarize critical value approach, p value approach, and confidence interval approach to test the hypothesis.

(Refer Slide Time: 39:21)

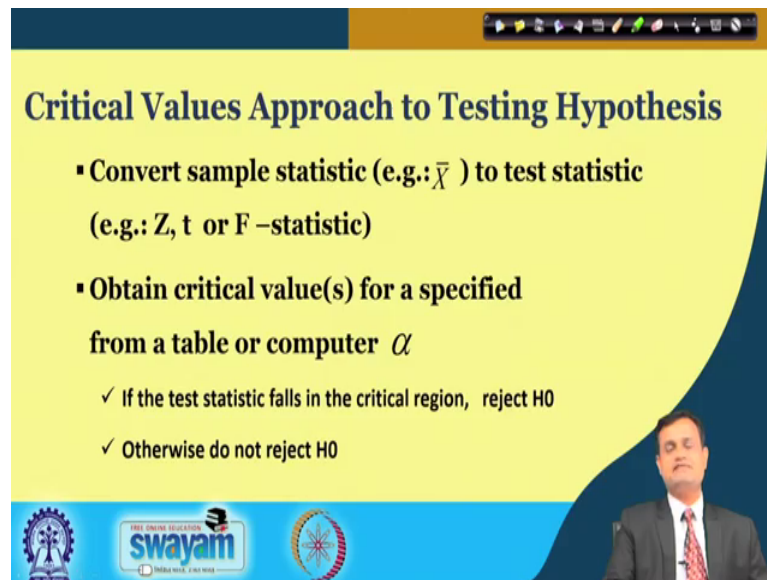


Approach 1

Critical Values Approach to Testing Hypothesis

This slide has a yellow background with a dark blue curved border on the right. The text is centered. At the bottom, there is a blue banner with logos for 'swayam' and other educational institutions. A small video inset of a man in a suit is visible in the bottom right corner.

(Refer Slide Time: 39:24)



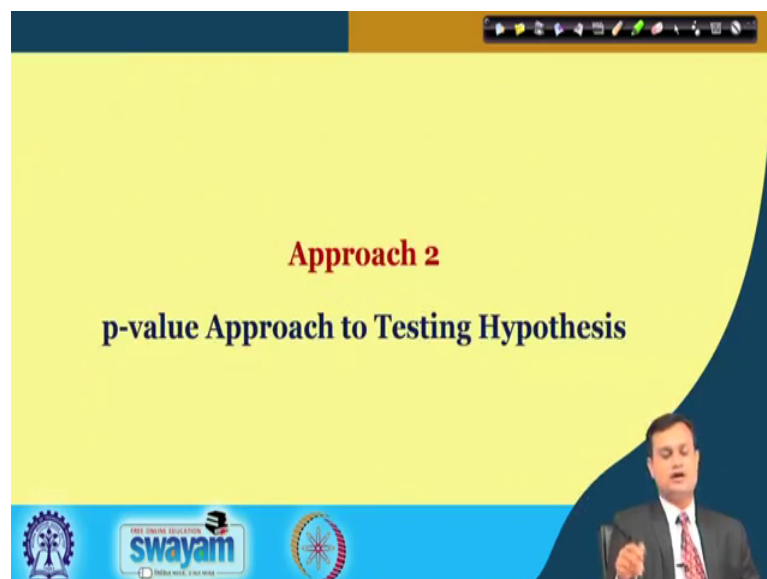
Critical Values Approach to Testing Hypothesis

- Convert sample statistic (e.g.: \bar{X}) to test statistic (e.g.: Z, t or F –statistic)
- Obtain critical value(s) for a specified α from a table or computer
 - ✓ If the test statistic falls in the critical region, reject H_0
 - ✓ Otherwise do not reject H_0

The slide features a yellow background with a dark blue curved border on the right. At the bottom, there are logos for 'swayam' and 'INDIA WISE, FINE WISE' along with a small inset video of a man in a suit.

Critical value approach, I have just discussed that I will find the critical value based on alpha and then I will try to check whether I should reject null hypothesis or I should accept.

(Refer Slide Time: 39:32)

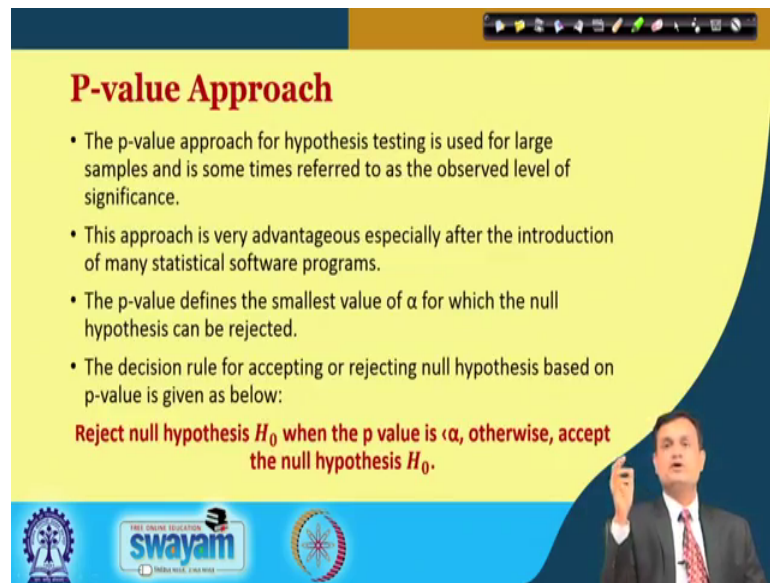


Approach 2

p-value Approach to Testing Hypothesis

The slide features a yellow background with a dark blue curved border on the right. At the bottom, there are logos for 'swayam' and 'INDIA WISE, FINE WISE' along with a small inset video of a man in a suit.

(Refer Slide Time: 39:35)



P-value Approach

- The p-value approach for hypothesis testing is used for large samples and is some times referred to as the observed level of significance.
- This approach is very advantageous especially after the introduction of many statistical software programs.
- The p-value defines the smallest value of α for which the null hypothesis can be rejected.
- The decision rule for accepting or rejecting null hypothesis based on p-value is given as below:

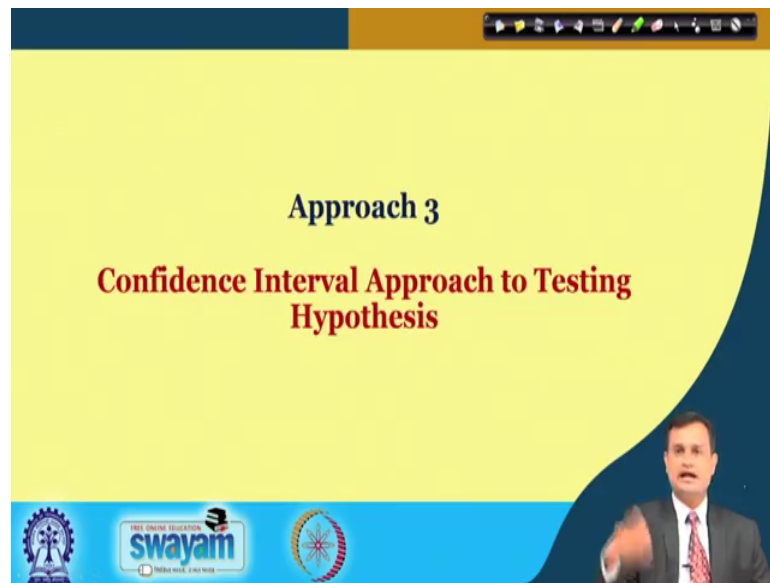
Reject null hypothesis H_0 when the p value is $\leq \alpha$, otherwise, accept the null hypothesis H_0 .

swayam

P value approach, I will try to find the probability value specific to critical value and then I will try to see that I have level of significance, I have observed level of significance that is the computed test statistics. Now you should feel comfortable with this technique or you should little bit revised. So, I have the observed level of significance, I have the critical level of significance, chosen level of significance and if my region is smaller it falls within the critical region. I will say my null hypothesis is reject.

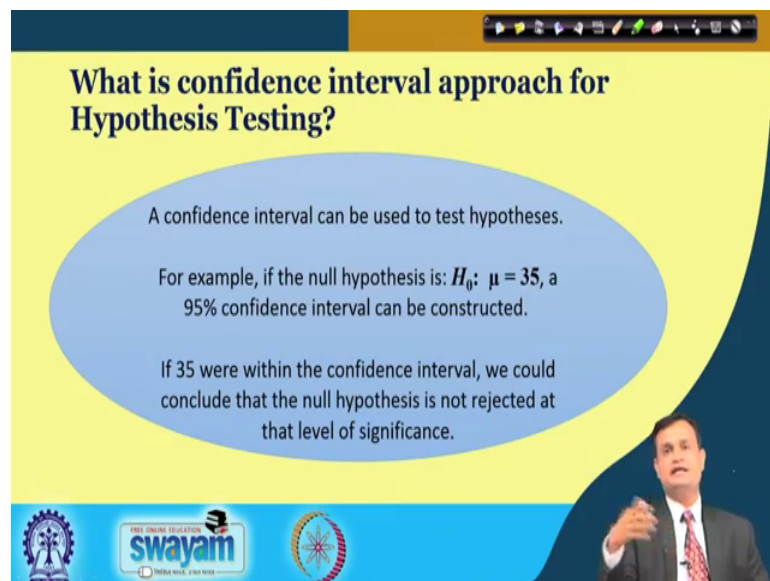
So, many a times softwares many softwares they use p value approach it is more intuitive to the managers because they know that what is the difference in probability at which they are rejecting null hypothesis. So, p value approach is more useful.

(Refer Slide Time: 40:28)



So, you have 3rd approach that is the confidence interval approach,

(Refer Slide Time: 40:32)




And here, you try to set the confidence interval that within what my value will fall. So, suppose my null hypothesis is 35, μ is equal to 35, I want to check that for a given confidence interval what is that interval within which majority of the time 95 percent of the time this value will fall.

(Refer Slide Time: 40:54)

Confidence Intervals for μ

- We can find confidence intervals for μ using the same reasoning as confidence intervals for p .
- The big idea is to cover the middle 95% (or whatever our confidence level is) of the area of the normal curve (sampling distribution of \bar{x}), we must go out a distance of z^* standard deviations on either side of the mean.
- The standard deviation of the sampling distribution of \bar{x} (σ/\sqrt{n}) depends on the sample size n and the population standard deviation σ .
- z^* is the critical value for our confidence level from Table
- The confidence interval will give us a reasonable range of values for our unknown population mean μ .




So, you have confidence interval construction using proportion mean.

(Refer Slide Time: 40:57)

Confidence Intervals and Tests of Significance

	C.I.	Test
Proportion	$\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$	$z = \frac{\hat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$
Mean	$\bar{x} \pm z^* \frac{\sigma}{\sqrt{n}}$	$z = \frac{\bar{x}-\mu_0}{\frac{\sigma}{\sqrt{n}}}$

Note: σ is generally unknown. Uses to estimate σ . When n is large, s is close to σ .



And you can use this μ plus or minus k sigma for proportion it would be \hat{p} plus or minus Z star square root of \hat{p} 1 minus \hat{p} hat divided by n for mean \bar{x} bar plus or minus Z sigma divided by square root n it will give you an interval and you need to check within what it falls.

(Refer Slide Time: 41:19)

Confidence Intervals and Tests of Significance

	C.I.	Test
σ known	$\bar{x} \pm z^* \frac{\sigma}{\sqrt{n}}$	$z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$
σ unknown	$\bar{x} \pm t^* \frac{s}{\sqrt{n}}$	$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$

Note: In this class, we will use z^* instead of t^* . If n is large, they should be fairly close. It will be an approximate level C confidence interval.

swayam

Same way, you have a situation sigma known, sigma unknown you can use the Z distribution or t distribution depending upon your sample size also.

(Refer Slide Time: 41:30)

General Procedure for confidence interval of μ

- Step 1. Define problem, pick level of confidence say 95%
- Step 2. Collect appropriate data, Random sample from population.
- Step 3. Compute \bar{x} as an estimate of μ
- Step 4. Check assumptions; if OK proceed further.
- Step 5. Compute interval
- Step 6. Interpret the intervals in the context.

swayam

So, there is a general procedure steps I have outline in the confidence interval. Define the problem, collect the data, compute \bar{x} and μ , check assumption if proceed further, compute interval then interpret the interval in the context and take the decision about accepting or rejecting null hypothesis.

(Refer Slide Time: 41:52)

1. Can we interchange the statements of null and alternate hypothesis?

2. Why null and alternate hypothesis are mutually exclusive?

3. Why level of significance is called the rejection region?

4. What are the assumptions of hypothesis testing?

swayam

Think it, can be interchange H_0 and H_1 null and alternate. Why null and alternate are mutually exclusive, why level of significance called rejection region and what are the assumption of hypothesis testing.

(Refer Slide Time: 42:06)

References

References:

- ❑ Aczel, A., Sounderpandian, J. and Saravanan, P., Complete Business Statistics, McGraw Hill Publication.
- ❑ David M. Levine, Timothy C. Krehbiel, Mark L. Berenson and P. K. Vishwanathan, Business Statistics, Pearson Publication.
- ❑ T. M. Kubiak, Donald W. Benbow, The Certified Six Sigma Black Belt Handbook, Pearson Publication.
- ❑ Forrest W. Breyfogle III, implementing Six Sigma, John Wiley & Sons, INC.
- ❑ Montgomery, D C. Statistical Quality Control: A modern introduction, Wiley.
- ❑ Mitra, Amitava. Fundamentals of Quality Control and Improvement, Wiley India Pvt Ltd.

swayam

Please go through couple of references to strengthen your idea. By the time, you must have developed confidence in understanding some of the vocabulary of hypothesis testing type I error, type II error, null hypothesis, alternate hypothesis, test statistics, confidence interval and I assume that this knowledge this terminology is clear to you. If

you have a doubt, please revise this lecture or see the relevant chapter in the reference book suggested so that, you can understand the remaining lectures with comfort.

(Refer Slide Time: 42:48)

Conclusion:

- ❖ Hypothesis testing is a well-defined procedure which helps us to decide objectively whether to accept or reject the hypothesis based on the information available from the sample.
- ❖ A Type I error is committed by rejecting null hypothesis when it is true. The probability of Type I error is called alpha (α) or level of significance.
- ❖ A Type II error is committed by accepting a null hypothesis when it is false. The probability of Type II error is beta (β).

So, we have a conclusion that hypothesis is a claim subject to type I and type II error and I want to test existing theory or prove new theory check new theory and hypothesis testing is required alpha is level of significance type 1 error beta is my type II error and $1 - \beta$ is the power of the test.

So, thank you very much for your patience in learning this particular topic. I would like to say, this is the most important lecture you must understand it thoroughly otherwise, the subsequent lectures you will have some difficulty in understanding when we will try to solve the example or execute the hypothesis testing for various situation.

So, please revise it, internalize it, if required listen this lecture 2 to 3 times or you refer the textbook. So, that your journey becomes comfortable in analysed phase. So, this is the first lecture of analysed phase we will keep discussing enjoy be with me.