Business Statistics
Prof. M. K. Barua
Department of Management Studies
Indian Institute of Technology – Roorkee

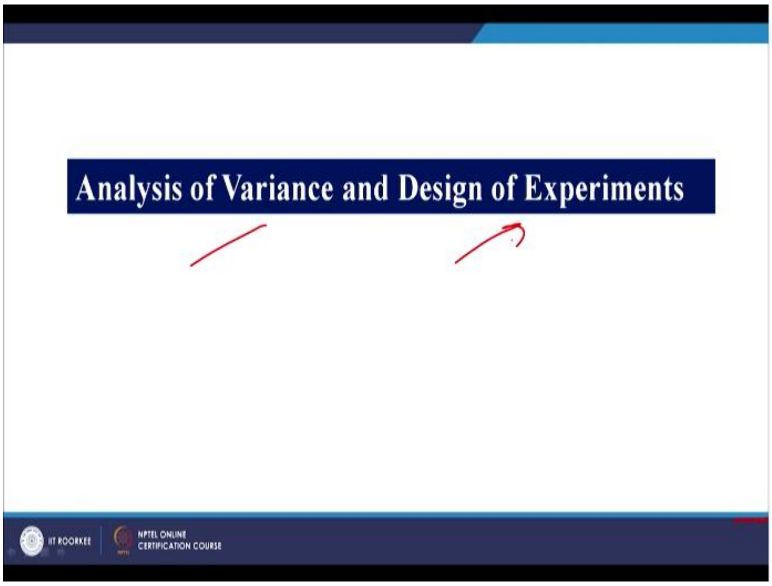
Lecture – 45
Design of Experiment (DOE)

Hello friends. I welcome you all in this session. As you are aware, in previous session we were discussing about hypothesis testing of two samples and we have done several examples on hypothesis testing of means, hypothesis testing of proportions and hypothesis testing of variance as well as standard deviation. In fact, we did solve the couple of examples on hypothesis testing of paired sample test where the samples were dependent on each other.

Now let us take a situation where you want to compare means of 3 different groups, then what to do or let us say if there are 4 groups and you want to compare means. So you can perform Z-test or T-test but you will have to perform several such tests and that would be quite an ineffective way of doing questions.

So there are some methods available which would be helpful in finding solution to problems where you have got 4 groups or 5 groups or 6 groups and so on right. So let us look at a new topic which is on analysis of variance and design of experiment.

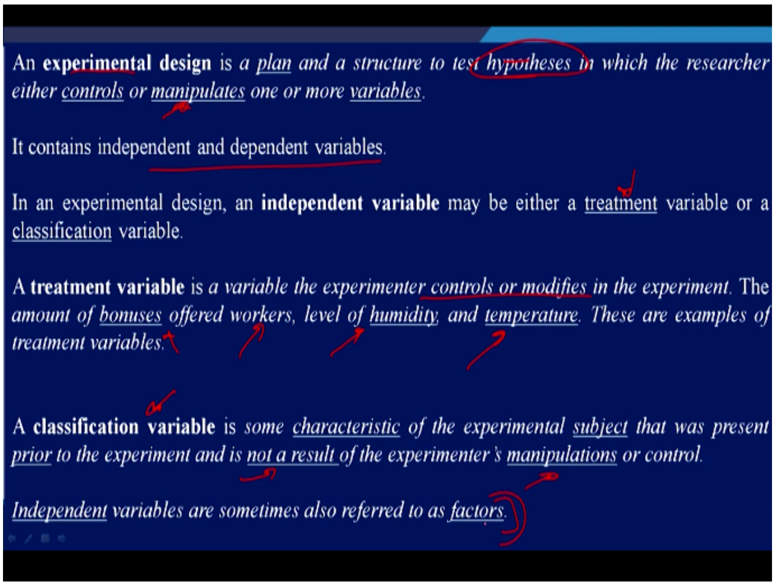
(Refer Slide Time: 01:53)



Analysis of Variance and Design of Experiments

So we will design certain experiments and there are some methods like analysis of variance to analyze outcome of experiments of the design.

(Refer Slide Time: 02:13)



An **experimental design** is *a plan and a structure to test hypotheses in which the researcher either controls or manipulates one or more variables*.

It contains independent and dependent variables.

In an experimental design, an **independent variable** may be either a treatment variable or a classification variable.

A **treatment variable** is *a variable the experimenter controls or modifies in the experiment*. *The amount of bonuses offered workers, level of humidity, and temperature. These are examples of treatment variables.*

A **classification variable** is *some characteristic of the experimental subject that was present prior to the experiment and is not a result of the experimenter's manipulations or control.*

Independent variables are sometimes also referred to as factors.

So experiment or say experimental design is nothing but the plan or the setting or the structure of testing hypothesis in which the researcher controls or manipulates or changes one or two variables. For example, let us say profit of an organization is a function of sales. So if we let us say increase sales, how the profit will increase or let us say in another example profit is a function of R&D investment.

So if we decrease R&D investment, how much profit will decrease? So you will always have some dependent variables and some independent variables. So whenever you make some changes in independent variable and see the effect of that change in dependent variable that is known as experimentation right. So you can either control or manipulate independent variable and you may have a situation let us say where there is a dependent variable.

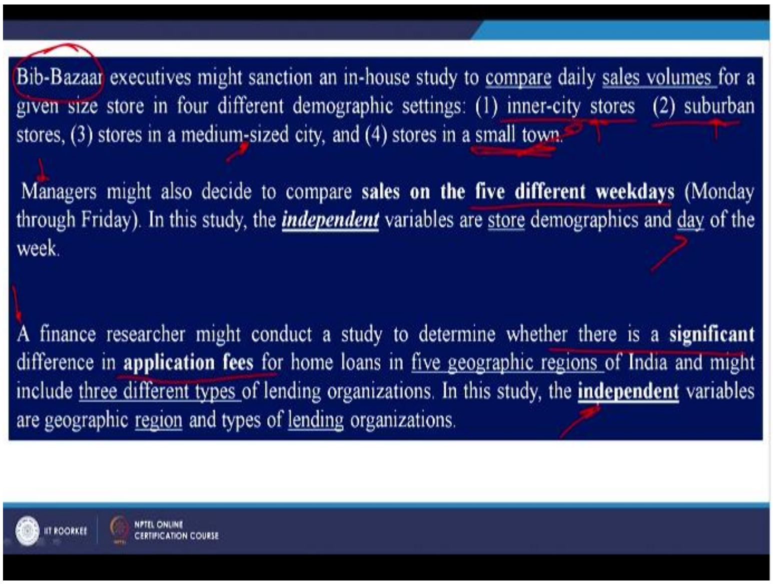
And you have got more than one independent variables, so you can in fact control or change more than one independent variable and see the effect of those changes on dependent variable. So as I said experiment contains dependent and independent variables. So independent variables maybe either treatment variable or classification variable, so let us look at what is the meaning of treatment variable.

In fact, treatment variable is a variable which an experimenter controls are modifies. Let us say the amount of bonuses offered to worker, then how the bonus would let us say improve its performance, improve productivity of the worker. So that bonus is nothing but a kind of treatment right. Let us say level of humidity in a workshop or in an assembly line and will see the effect of that humidity on performance of the worker.

Or the temperature performance on temperature, let us say temperature, how temperature is performing or affecting performance of a worker right. So these are examples of treatment variables. So whenever we change something, that variable is called treatment variable right. Classification variable is some characteristics of the experimental subject that was present prior to the experiment and is not a result of experimental manipulation or control.

We will look it in detail what classification variable is. So independent variables are also known as factors. Whenever we design any experiment, we can call independent variables as factors right. So let us look at couple of examples.

(Refer Slide Time: 05:24)



Bib-Bazaar executives might sanction an in-house study to compare daily sales volumes for a given size store in four different demographic settings: (1) inner-city stores (2) suburban stores, (3) stores in a medium-sized city, and (4) stores in a small town.

Managers might also decide to compare **sales on the five different weekdays** (Monday through Friday). In this study, the ***independent*** variables are store demographics and day of the week.

A finance researcher might conduct a study to determine whether there is a **significant** difference in **application fees** for home loans in five geographic regions of India and might include three different types of lending organizations. In this study, the **independent** variables are geographic region and types of lending organizations.

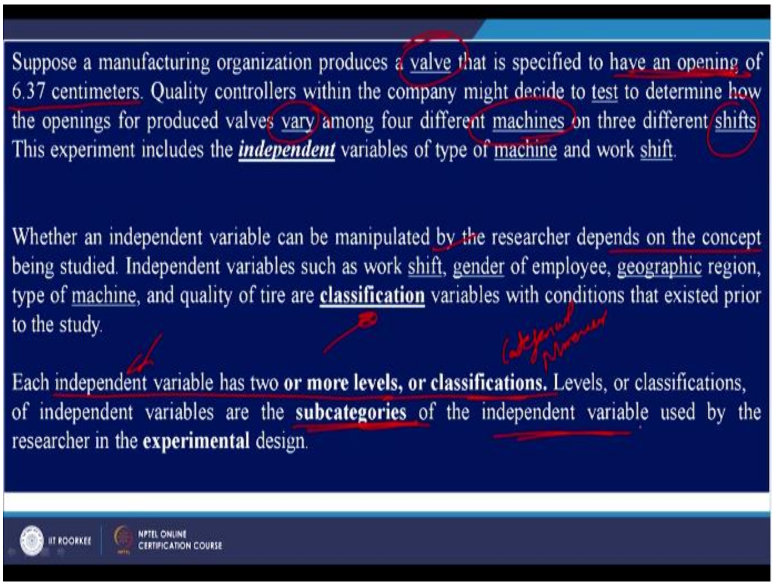
So let us say there is a very big retail chain of Big-Bazaar. So the executives of Big-Bazaar let us say want to study the amount of sales which is taking place in different stores. So let us say the executives have decided to have comparison of sales on these 4 stores right. Let us say one of them is in the city, the other one is in suburb area, the third one is in medium-sized city and fourth one is in small town right.

So there are 4 different stores of Big-Bazaar and the executives want to know how the sales figures are in this 4 stores right. So you can have one more example let us say the managers want to know how sales change over different weekdays. So one can compare sales figures of a particular store on 5 days right let us say Monday to Friday. So in this study, let us say so all the stores are nothing independent variables right.

Or in this example the week days are nothing but they are independent variables. Let us take one more example. A finance researcher might conduct a study to determine whether there is significant difference in application fees for home loans. So let us say there is a company let us for example let us take HDFC Bank right. Now it has got branches in 5 different areas and customers can apply for loan in those 5 areas right.

So the manager wants to know is there any significant difference between or is there any significant difference amongst fee which customers are paying in all those 5 areas right. So those 5 areas are nothing but kind of independent variable right and the fee would be the dependent variable okay.

(Refer Slide Time: 07:56)



Suppose a manufacturing organization produces a valve that is specified to have an opening of 6.37 centimeters. Quality controllers within the company might decide to test to determine how the openings for produced valves vary among four different machines on three different shifts. This experiment includes the ***independent*** variables of type of machine and work shift.

Whether an independent variable can be manipulated by the researcher depends on the concept being studied. Independent variables such as work shift, gender of employee, geographic region, type of machine, and quality of tire are **classification** variables with conditions that existed prior to the study.

Categorical Nominal

Each independent variable has two **or more levels, or classifications.** Levels, or classifications, of independent variables are the **subcategories** of the independent variable used by the researcher in the **experimental** design.

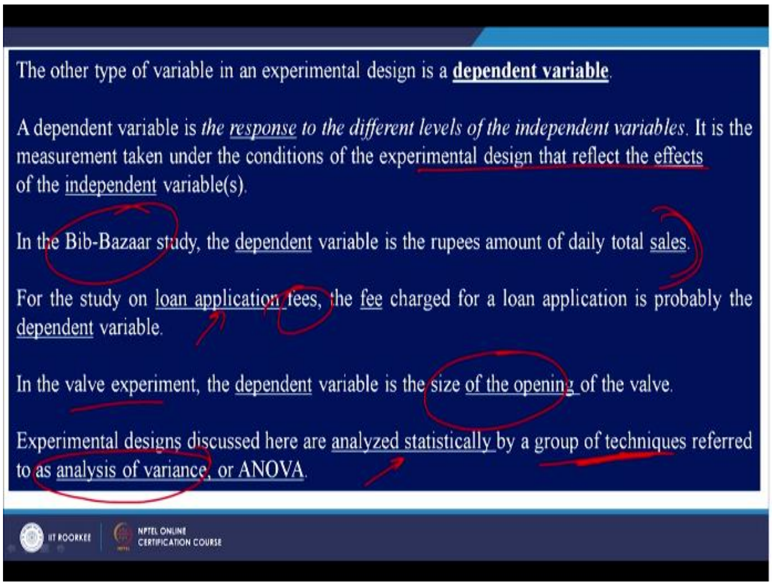
Let us take one more example. Suppose manufacturing organization produces valves and that has got different opening diameter. So let us say the valve that is specified to have an opening of 6.37 centimeters, quality controllers within the company might decide to test to determine how the openings for produced valves vary among 4 different machines on 3 different shifts. So let us say there is a manufacturing organization which is producing valves right.

And those valves can be produced by 4 different machines and can be in different shifts let us say morning shift and night shift right. So we want to know is there any variation in the diameter of the valve if we look at all these machines and these two shifts right. So will say that the 4 machines are nothing but they are independent variables and the shifts are again independent variables right.

So whether an independent variable can be manipulated by the researcher depends on the concept being studied. Independent variables such as shift or let us say number of machines or geographical areas or let us say the 4 stores of Big-Bazaar, all these are nothing but these are called classification variables right. So each independent variable can have two or more classification variables right.

So and these variables can be either categorical or numerical right. So let us say gender is one of the independent variable right. So we can have male and females right. These two are two levels of gender or let us say male, female and transgender. So for gender independent variable, there are 3 levels right. So these are levels are nothing but different subcategories right of independent variable which we use in experimental design.

(Refer Slide Time: 10:50)



The other type of variable in an experimental design is a **dependent variable**.

A dependent variable is *the response to the different levels of the independent variables*. It is the measurement taken under the conditions of the experimental design that reflect the effects of the independent variable(s).

In the Bib-Bazaar study, the dependent variable is the rupees amount of daily total sales.

For the study on loan application fees, the fee charged for a loan application is probably the dependent variable.

In the valve experiment, the dependent variable is the size of the opening of the valve.

Experimental designs discussed here are analyzed statistically by a group of techniques referred to as analysis of variance, or ANOVA.

So far we talked about independent variable. Let us talk about dependent variable. Dependent variable is nothing but the response. As I said, if we are testing let us say whether the application fee of loan is different in 3 different villages, so that application fees nothing but dependent variable that is response to some independent variable. It is the measurement taken under the conditions of experimental design that reflects the effect of the independent variable.

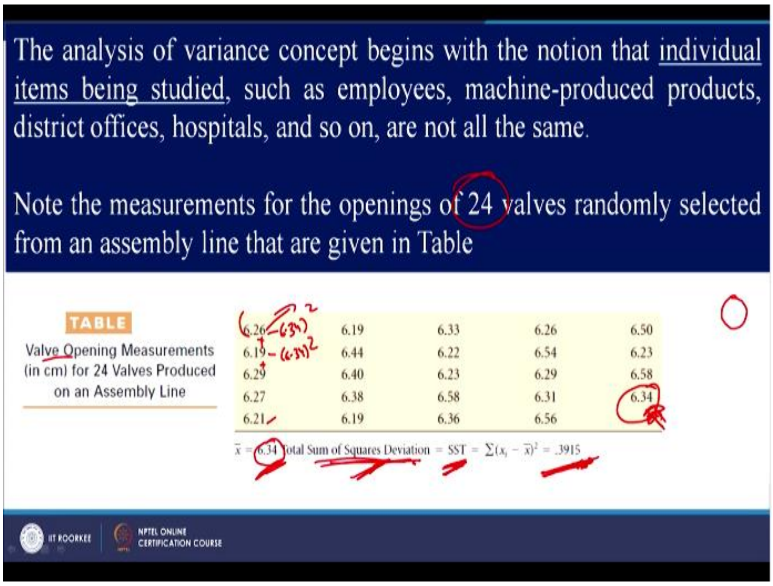
So let us take once again the same example. So you have got let us say profit is nothing but response variable or dependent variable and let us say R and D investment is independent variable or also known as factor right. So in Big-Bazaar example will say that dependent variable is nothing but sales, sales volume right. Now we can have let us say one more example over here.

So let us say in a store of Big-Bazaar or for any other store for that matter let us say customers are paying their bills either through cash or through debit card or through credit card or through some other means. So we can compare is there any significant difference amongst those modes of payments so far as sales in that particular store is concerned. So in loan application example, the fees application fees is nothing but the response variable or dependent variable.

In valve experiment example, the opening, the size of the valve is nothing but dependent variable. So experimental designs discussed here can be analyzed statistically by a group of techniques and so this group of techniques is referred as analysis of variance right. So what is analysis of variance? ANOVA in other words is how we are analyzing the outputs of different experiments.

And we use certain statistical techniques and those techniques are called ANOVA, analysis of variance. Let us look at this example. So you can have different examples and this is one example I have taken over here.

(Refer Slide Time: 13:53)



The analysis of variance concept begins with the notion that individual items being studied, such as employees, machine-produced products, district offices, hospitals, and so on, are not all the same.

Note the measurements for the openings of 24 valves randomly selected from an assembly line that are given in Table

TABLE

Valve Opening Measurements (in cm) for 24 Valves Produced on an Assembly Line

$(6.26-(6.34))^2$	6.19	6.33	6.26	6.50
$6.19-(6.34)^2$	6.44	6.22	6.54	6.23
6.29	6.40	6.23	6.29	6.58
6.27	6.38	6.58	6.31	6.34
6.21	6.19	6.36	6.56	

$\bar{x} = 6.34$ Total Sum of Squares Deviation $= SST = \sum(x_i - \bar{x})^2 = .3915$

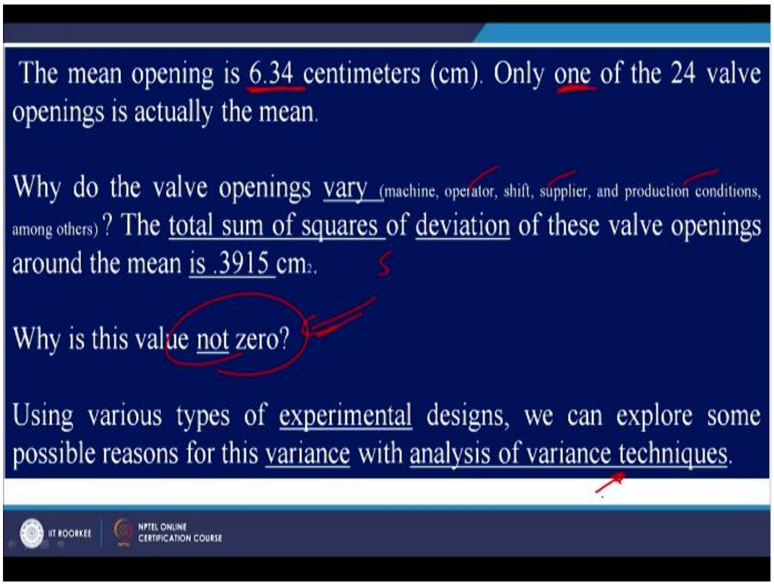
So let us say will take the valve opening example. So let us say there is something valve which has got the diameter like this right. So valves are being manufactured on 4 different machines right and let us say we have manufactured 24 valves and these are the diameters of these 24 valves. So let us say 6.26, 6.25 and so on right and the diameter of this valve is 6.34 right, so there are 24 valves okay.

When you take the mean of all these valves, it becomes 6.34. Now if you look at that there is only one valve which has got mean=6.34, it has happened just by chance, it is not necessary that you will have valves whose you know diameter is mean of all the valves right, it is not necessary. So for this example, this is the mean and the sum of squared deviation SST, sum of squared deviation is 0.390.

Actually, this sum of squared deviation should have been 0. Why this is not 0? That is the question okay. So sum of squared deviation, it is nothing but xi-x bar whole square summation of that right. So we know the xi is this right, so let us say this is xi-6.34 whole square plus this minus 6.34 whole square plus so on right. You just keep on adding all these, so that sum would be nothing.

Nothing but it would be.0.39 or approximately 0.40 right. So why this sum of squared deviation is not 0? That is the question.

(Refer Slide Time: 16:06)



The mean opening is 6.34 centimeters (cm). Only one of the 24 valve openings is actually the mean.

Why do the valve openings vary (machine, operator, shift, supplier, and production conditions, among others)? The total sum of squares of deviation of these valve openings around the mean is .3915 cm$_2$.

Why is this value not zero?

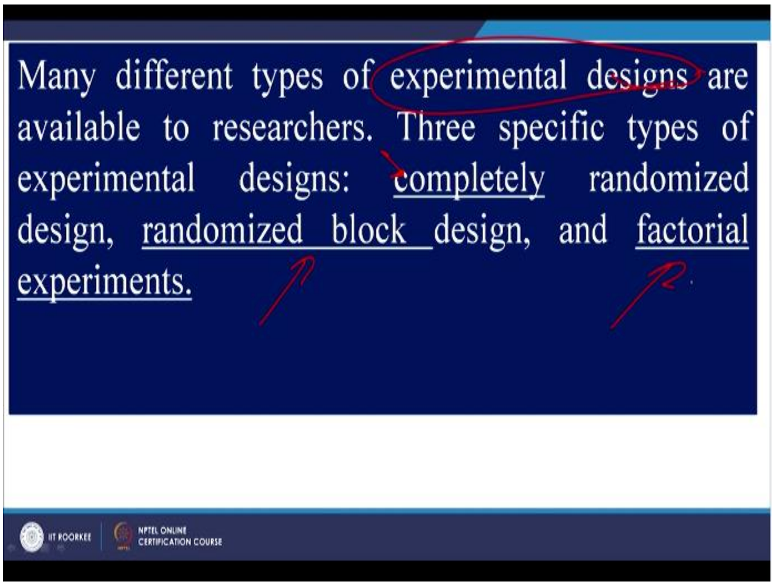
Using various types of experimental designs, we can explore some possible reasons for this variance with analysis of variance techniques.

So as I said there is only one valve whose diameter is 6.34 centimeter. So why in fact you can have multiple reasons why these diameters are not same when we produce these diameters, when we produce these valves. So there can be again multiple reasons, why these diameters are not same. This could be due to let us say different workers working on a machine, let us say in different shifts they are working.

And they are getting let us say raw material from different vendors. So there can be multiple reasons either let us say due to operator or supplier or production conditions and humidity and temperature and so on right. So these are different reasons why we are making or why valves of different diameters are being produced. So this is not 0, why this SST, sum of square of deviation is not 0 right.

So we will try to find out answers. We will try to find out why reasons, why this is not 0 right. So using various types of experimental designs we can explore some possible reasons for this variance with analysis of variance techniques okay.

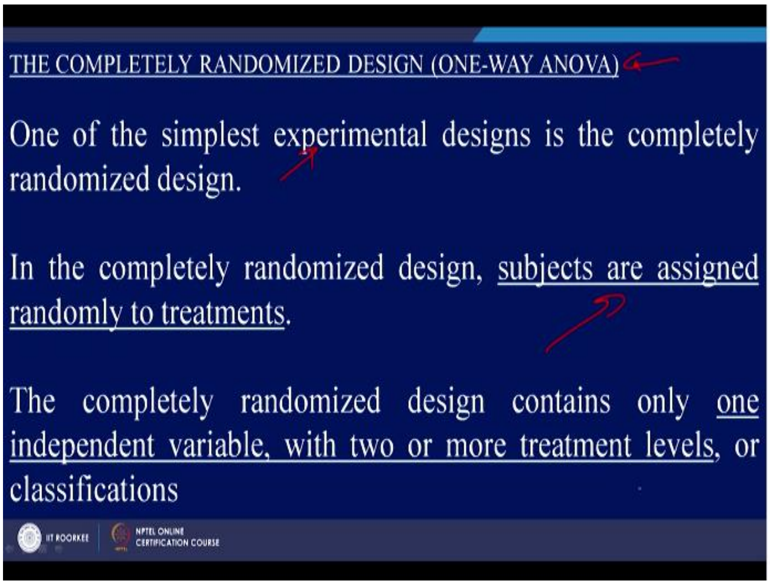
(Refer Slide Time: 17:45)



Many different types of experimental designs are available to researchers. Three specific types of experimental designs: completely randomized design, randomized block design, and factorial experiments.

So there are many different types of experimental designs are possible and there are basically 3 types of experimental designs available. The first one is completely randomized design, completely randomized, second is randomized block and third is factorial experiment. So let us look at what is completely randomized design.

(Refer Slide Time: 18:14)



THE COMPLETELY RANDOMIZED DESIGN (ONE-WAY ANOVA)

One of the simplest experimental designs is the completely randomized design.

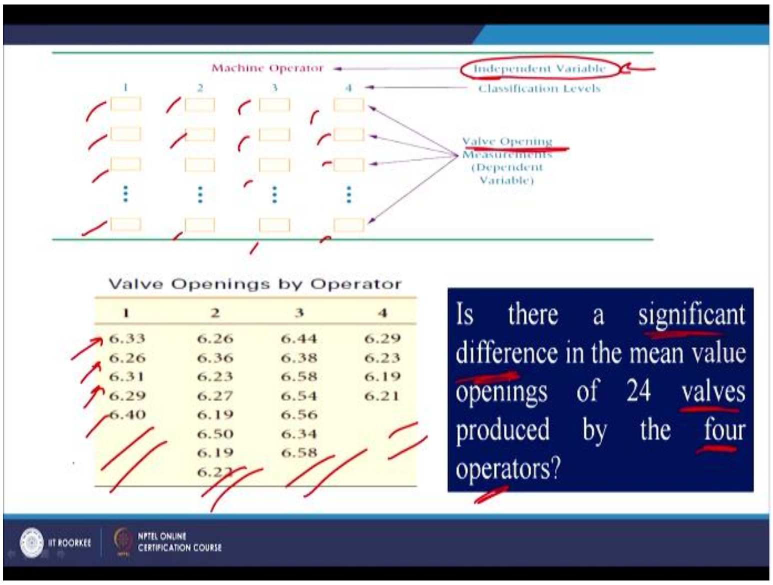
In the completely randomized design, subjects are assigned randomly to treatments.

The completely randomized design contains only one independent variable, with two or more treatment levels, or classifications

So the completely randomized design and we will use one-way ANOVA to find out the outcome of design of experiment. So this is one of the designs right completely randomized design. So we assign in fact this is the simplest amongst all as I said. So in the completely randomized design, subjects are assigned randomly to treatments. So that is why it is called completely randomized design.

So the completely randomized design contains one independent variable with two or more levels right or classification.

(Refer Slide Time: 19:08)



Valve Openings by Operator

1	2	3	4
6.33	6.26	6.44	6.29
6.26	6.36	6.38	6.23
6.31	6.23	6.58	6.19
6.29	6.27	6.54	6.21
6.40	6.19	6.56	
	6.50	6.34	
	6.19	6.58	
	6.22		

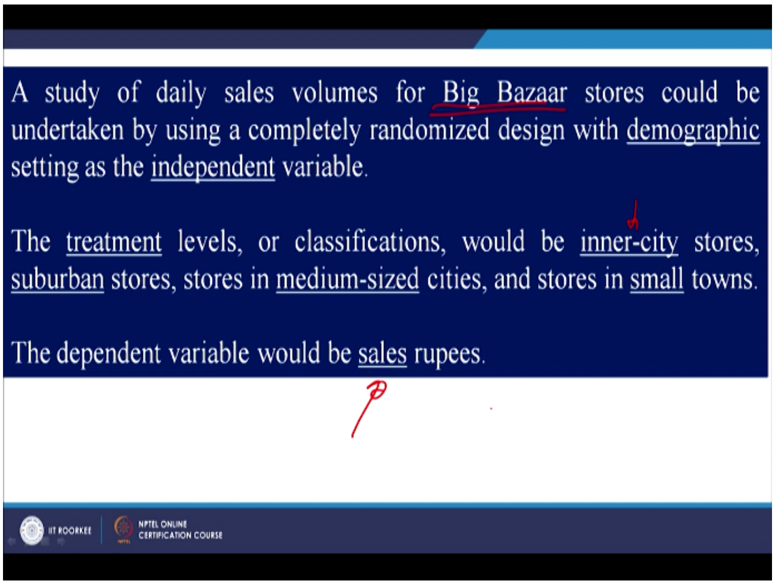
Is there a significant difference in the mean value openings of 24 valves produced by the four operators?

So let us look at this example. So will take the same valve manufacturing example. So let us say these valves are being manufactured on 4 different machines right. So machine 1 is producing these valves, machine 2 these valves, machine 3 these valves, machine 4 these valves right. So independent variable is you have got let us say there are 4 operators not 4 machines so machine operator is an independent variable.

There are 4 different operators and the classification level is let us say 1, 2, 3, 4 or A, B, C, D right. So the question is, is there any significant difference amongst means of all these valves produced by these 4 operators? So that is the question. Is there a significant difference in the mean value of diameter of these 24 valves produced by 4 operators? So this is an example right of independent variable.

And what would be the dependent variable by the way? So independent variables are 4 different operators right, dependent variable is nothing but these diameters right okay.

(Refer Slide Time: 20:35)



A study of daily sales volumes for Big Bazaar stores could be undertaken by using a completely randomized design with demographic setting as the independent variable.

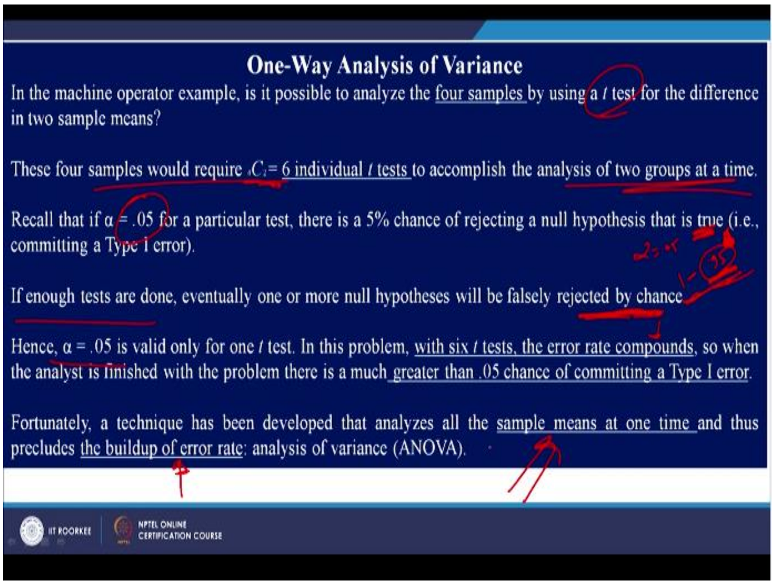
The treatment levels, or classifications, would be inner-city stores, suburban stores, stores in medium-sized cities, and stores in small towns.

The dependent variable would be sales rupees.

So let us look at one more slide. So what would be the dependent variable in case of Big-Bazaar? So in fact the demographical variables that 4 different stores let us say in urban, suburban, inside city and all those. So those 4 are independent variables and the levels or the treatments are those 4 cities right okay and the dependent variable is sales in terms of rupees or let us say in terms of quantity whatever it is okay.

So this is how you should know the difference between dependent variable and independent variable. So you will always have some subcategories of independent variables and those subcategories are called either level or classification variables. Let us look at what is one-way analysis of variance or one-way ANOVA.

(Refer Slide Time: 21:53)



One-Way Analysis of Variance

In the machine operator example, is it possible to analyze the four samples by using a *t* test for the difference in two sample means?

These four samples would require ${}_4C_2$ = 6 individual *t* tests to accomplish the analysis of two groups at a time.

Recall that if α = .05 for a particular test, there is a 5% chance of rejecting a null hypothesis that is true (i.e., committing a Type I error).

If enough tests are done, eventually one or more null hypotheses will be falsely rejected by chance.

Hence, α = .05 is valid only for one *t* test. In this problem, with six *t* tests, the error rate compounds, so when the analyst is finished with the problem there is a much greater than .05 chance of committing a Type I error.

Fortunately, a technique has been developed that analyzes all the sample means at one time and thus precludes the buildup of error rate: analysis of variance (ANOVA).

So if you look at the machine operator example, so we can find out the difference between let us say difference in means of these 4 workers using t-test because there are if you look at this there are 4 workers right. So we can perform t-test between first and second then first and third, first and fourth. So how many t-tests we will be performing? 3 right. Then, second and third, second and fourth, five and then second and six.

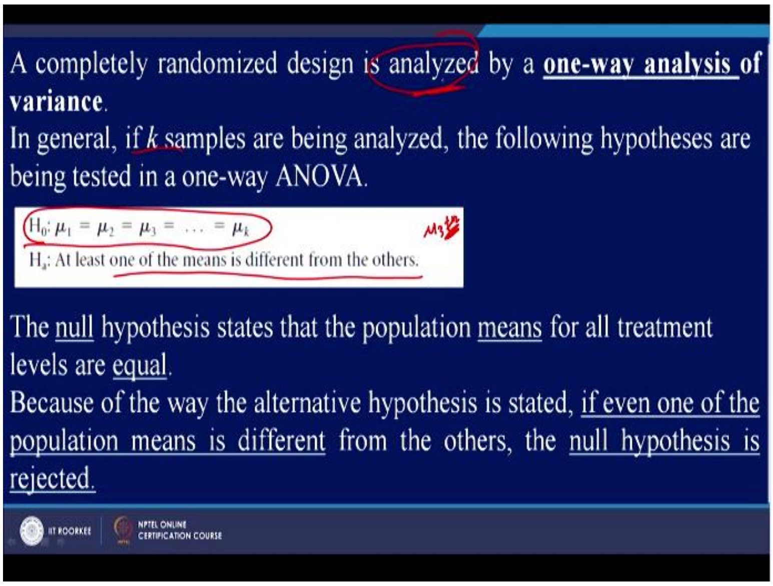
So to get the answer of this question we will have to perform 6 t-tests right. In fact, these 4 samples would be requiring 4C2 6 individual t-tests to accomplish the analysis of these two groups. So either you use t-test 6 times to solve this question or you use a technique called one-way ANOVA which in one stage itself will give the difference amongst 4 means right. We also know that when alpha is=0.5 so there is only just 5% chance of rejecting null hypothesis when it is true right.

So if you let us say perform 6 tests right, then the probability of error would increase right, is not it? So if enough tests are performed, eventually one or more null hypothesis will falsely rejected by chance right. So if this is valid then with 6 tests the error rate compounds, so there is a possibility of having more and more error right. Let me put it in this way, so probability of not making type 1 error is 0.95 right.

When I say alpha is=0.05, probability of not making type 1 error is 0.05. So when you t-tests 6 times so this will be the probability of not making type 1 error. So 1 minus this is the probability of making type 1 error which is very high, is not it? So that value 1-0.9 to the power 6 right is the probability of making error or probability of making type 1 error, so which is much greater than 0.05 right.

So a technique has been developed that analyzes all sample means at one time and thus precludes buildup of error rate. So you can minimize the errors which you build over performing 6 hypotheses testing right. So that is why ANOVA is used.

(Refer Slide Time: 25:14)



A completely randomized design is analyzed by a **one-way analysis of variance**.

In general, if *k* samples are being analyzed, the following hypotheses are being tested in a one-way ANOVA.

$H_0: \mu_1 = \mu_2 = \mu_3 = \ldots = \mu_k$ μ_3

H_a: At least one of the means is different from the others.

The null hypothesis states that the population means for all treatment levels are equal.

Because of the way the alternative hypothesis is stated, if even one of the population means is different from the others, the null hypothesis is rejected.

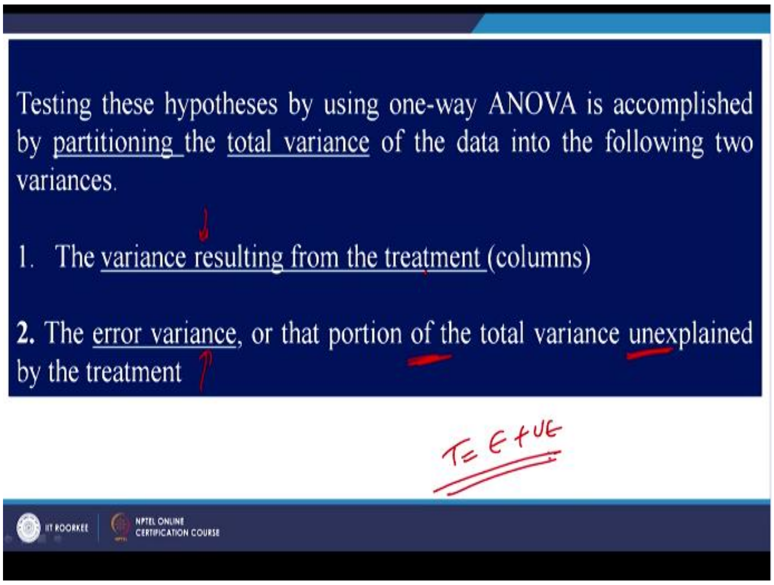
IIT ROORKEE | NPTEL ONLINE CERTIFICATION COURSE

A completely randomized design is analyzed by one-way ANOVA and what is this? So initially in this one-way ANOVA what we do, if there are let us say k samples, so you can have the following hypothesis as null hypothesis and alternative hypothesis. So we will say that all means are equal right if there is no significant difference amongst these means. So µ1=µ2=µk right.

So this is our null hypothesis and alternative hypothesis at least one of the means is different from the other mean. So let us say there are 6 means if mu 1 and mu 2 are not equal then alternative hypothesis would apply. So this is null hypothesis and alternative hypothesis. The null hypothesis states that the population means for all treatment levels are equal. So all hypothesis means are same right.

Because of the way alternative hypothesis stated even if one of the population mean is different from others, will reject the null hypothesis. So there is a possibility that mu 3 and mu 2 are not equal. Then, will reject the null hypothesis right. So a completely randomized design is analyzed by one-way ANOVA.

(Refer Slide Time: 27:01)



Testing these hypotheses by using one-way ANOVA is accomplished by partitioning the total variance of the data into the following two variances.

1. The variance resulting from the treatment (columns)

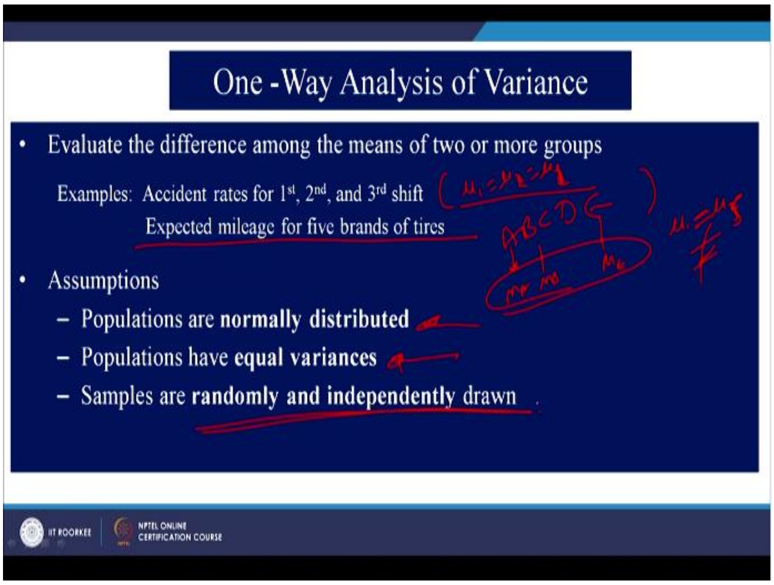
2. The error variance, or that portion of the total variance unexplained by the treatment

$T = E + UE$

So testing these hypotheses by using one-way ANOVA is accomplished by partitioning the total variance of the data into the following two variances. So we have said in the beginning in one of the slide that the sum of total variance was not 0 in the valve opening example. So why and what was the reason for that? So what we did, we in fact divided the total variance into two different variances called error variance and explained variance right or variance resulting from the independent variable or from the treatment.

And error variance or that portion of the total variance which is unexplained, so total variance is explained variance plus unexplained variance, is not it?

(Refer Slide Time: 28:00)



One -Way Analysis of Variance

- Evaluate the difference among the means of two or more groups

  Examples: Accident rates for 1st, 2nd, and 3rd shift ($\mu_1 = \mu_2 = \mu_3$)

  Expected mileage for five brands of tires

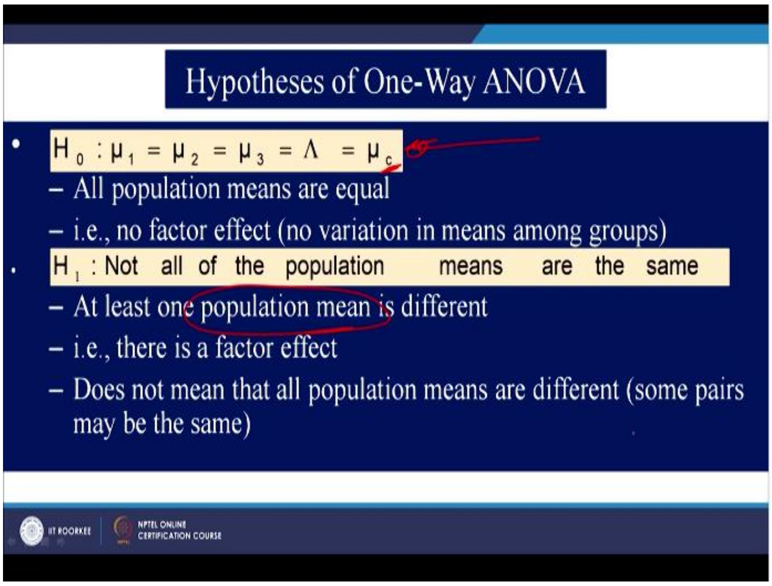
- Assumptions
  - Populations are **normally distributed**
  - Populations have **equal variances**
  - Samples are **randomly and independently** drawn

So we can evaluate difference among means of two or more groups. So you can have different examples, let us say the number of accidents taking place in 3 different shifts. So we can compare in which shift the accidents are more. So the null hypothesis can be framed like this. Let us say the number of accidents are same in all these 3 shifts, mu 1, mu 2, mu3 are same and alternative would be they are not same, is not it?

Let us look at this. Expected mileage of 5 brands of tires, so you have got let us say A, B, C, D and E these are 5 brands of tires. So when you measure the mileage of branded tires, you will get some mileage let us say mA, similarly mB and mC, mD right mE. So you are comparing all these 5 means. So again null hypothesis would be mu 1 to mu 5 all are same and not equal to its alternative hypothesis.

So whenever we go for hypothesis testing using one-way ANOVA, there are certain assumptions which are to be kept in mind. So we always say that the populations are normal distribution or normally distributed and the populations have equal variances, samples are randomly independently drawn. So if these conditions are, if these assumptions are fulfilled, then only we apply one-way ANOVA.

(Refer Slide Time: 30:05)

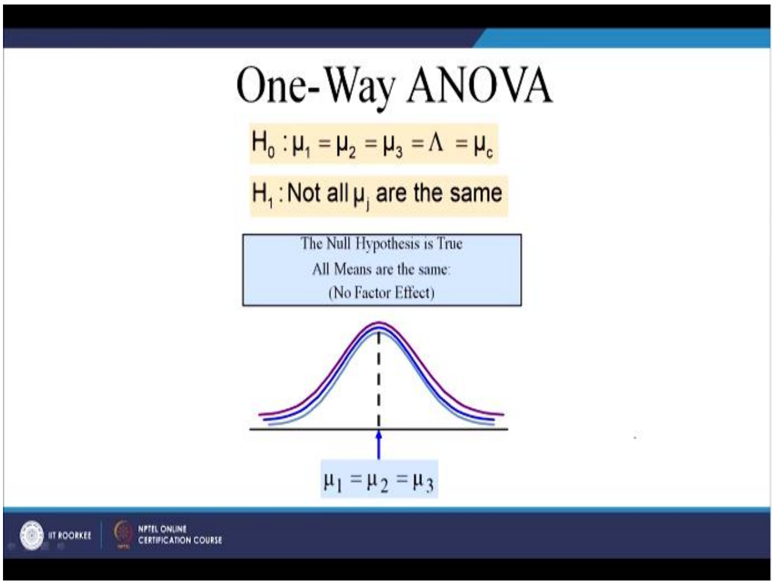


Hypotheses of One-Way ANOVA

- $H_0 : \mu_1 = \mu_2 = \mu_3 = \Lambda = \mu_c$
  - All population means are equal
  - i.e., no factor effect (no variation in means among groups)
- H_1 : Not all of the population means are the same
  - At least one population mean is different
  - i.e., there is a factor effect
  - Does not mean that all population means are different (some pairs may be the same)

We have already talked about this. So this is our null hypothesis, c is nothing but number of groups right. So mu 1 to mu 3 and so on up to mu c all are same right, all population means are equal and alternative hypothesis not all of the population means are same right. So at least one population mean has to be different to reject null hypothesis right.

(Refer Slide Time: 30:35)



One-Way ANOVA

$H_0 : \mu_1 = \mu_2 = \mu_3 = \Lambda = \mu_c$

$H_1 :$ Not all μ_j are the same

The Null Hypothesis is True
All Means are the same:
(No Factor Effect)

So before I start this slide let me summarize what we have done in today's session. In today's session, we have discussed in detail what is design of experiment, what is independent variable, what is the meaning of treatment variable, what is the meaning of classification variable, what is the meaning of dependent variable. Then, we have seen different examples of dependent and independent variables.

And when we compare means of more than 2 groups, then we use analysis of variance, not t-test. The reason is that when we have let us say 3 groups, then we will have to perform t-test more than one time. So let us say you will have to perform t-test in case of 3 groups, you will be having first second, first and third, second and third right, so 3 tests you will have to perform and when you do this, there is a possibility of having more and more error right.

There is a probability of rejecting null hypothesis. It would be let us say 1-0.95 to the power 3, that would be the probability right. So with this let me stop here. In next class, we will discuss some more points related to analysis of variance. Thank you.