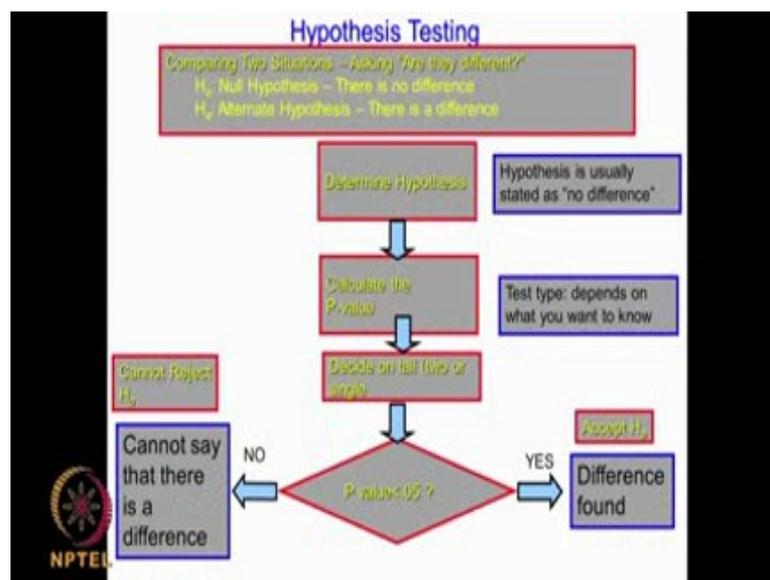


Biostatistics and Design of Experiments
Prof. Mukesh Doble
Department of Biotechnology
Indian Institute of Technology, Madras

Lecture - 08
Statistical tests

Hello again, welcome to the course on Biostatistics and Design of Experiments. In this class, we are going to briefly touch about various statistical test, that one should understand and how to go about using them. So, there are different types of statistical test, some are used for comparing mean of a data and some are used for comparing variance of data and some are used for comparing ratios of data and so on. So, we will talk about that.

(Refer Slide Time: 00:46)

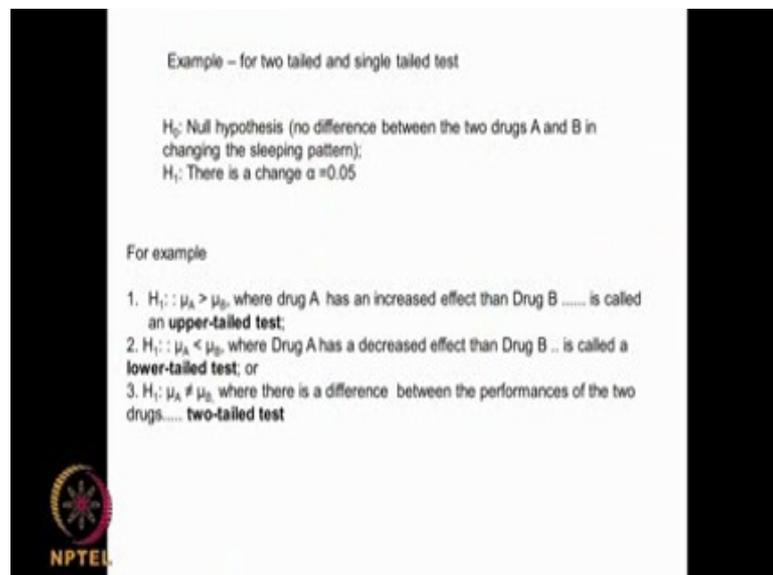


Before going into that, let me recall again how you go about doing this hypothesis, you have something called the null hypothesis, which is no difference or state as score. So, imagine I am comparing the IQ's of 2 different classes or class of students, then I would say the null hypothesis will be there is no difference in the IQ of students in class a and class b. Then we also have the alternate hypothesis, if I am going to say yes there is possible difference then I would say the alternate hypothesis will be the, IQ average of class a is different from IQ average of b. Then you can also have another situation where instead of just saying different, the IQ average of class a could be better than the IQ average of class b. That is we are comparing only the better part of it or the IQ average of

a class a is worst than that of that class b, then we are comparing only the worst part of it. So we have those 3 situations, no difference is the null hypothesis, there is a difference is the alternate hypothesis and that is called a two-tailed comparison that is a different, it could be greater or worst. Then we are comparing only 1 side of it greater side of it or worst side of it, then that is called one-tailed test.

So we decide on the hypothesis with the tail and then we also decide on the p value, that is am I going to test against 95 % confidence or am I going to test against 99 % confidence. So p of 0.05 indicates 95 % and p of 0.01 indicates 99 %. Once I decide on all these, that means I decide on the type of hypothesis, I decide on whether it is a single tailed or a two-tailed test, then I decide on the p I want to look at, then I calculate using different test something called t, and then from there I will calculate may be the p value and then I will say whether the p value which I have calculated is less than 0.05 or it is more. So if it is less than 0.05 obviously, there is a difference so obviously, I cannot accept null hypothesis so I have to reject the null hypothesis that means, I have to accept the alternate hypothesis. Now if the p value is greater than 0.05 for a 95 % confidence obviously, there is no reason for me to reject the null hypothesis.

(Refer Slide Time: 03:38)



Example – for two tailed and single tailed test

H_0 : Null hypothesis (no difference between the two drugs A and B in changing the sleeping pattern);
 H_1 : There is a change $\alpha = 0.05$

For example

1. $H_1: \mu_A > \mu_B$, where drug A has an increased effect than Drug B is called an **upper-tailed test**;
2. $H_1: \mu_A < \mu_B$, where Drug A has a decreased effect than Drug B .. is called a **lower-tailed test**; or
3. $H_1: \mu_A \neq \mu_B$, where there is a difference between the performances of the two drugs..... **two-tailed test**



So, imagine I am testing 2 drugs in the market which affects this sleeping pattern. So, the null hypothesis could be, there is no difference between the 2 drugs, the alternate there is a difference at a 95 % confidence interval. Now what type of hypothesis equations we

will put first, the two-tailed under a single tailed. For a two-tailed your alternate will be $\mu_A \neq \mu_B$, your null hypothesis will be $H_0: \mu_A = \mu_B$, that is the two-tail. For single tailed test null hypothesis could be μ_A is equal to μ_B , alternate hypothesis could be μ_A is greater than μ_B , that means drug a has an increased effect than drug b or μ_A could be less than μ_B , that mean drug a has decreased effect than drug b that is a lower tailed and upper tailed. Now there is something called error, there are 2 types of error.

(Refer Slide Time: 04:40)

Types of errors:

A type 1 error is rejecting the null hypothesis H_0 when H_0 is true.

A type 2 error is failing to reject the null hypothesis H_0 when the alternative hypothesis H_1 is true.

Hypothesis Test		
State Decision	Don't reject H_0	Reject H_0
H_0 True	Correct	Type 1/ α error
H_1 True	Type 2/ β error	Correct

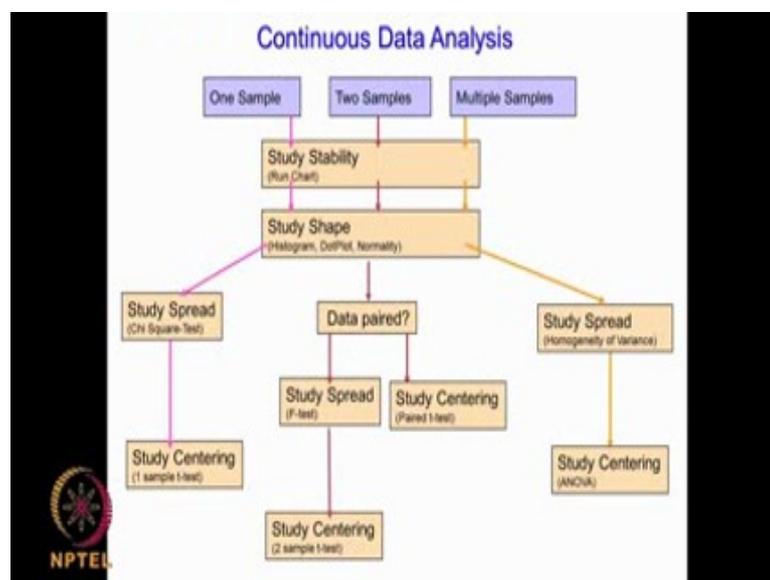
The maximum probability of a type 1 / α error, over the set of distributions specified by H_0 , is the *significance level* of the test or the size of the critical region.
(typically 0.1, 0.05, 0.01).



There is something called type 1 error and there is something called type 2 error. What is type 1 error? We are rejecting null hypothesis H_0 , when H_0 is true that means, we are rejecting the null hypothesis where as in reality the null hypothesis is true that is called type 1 error. Then you have the type 2 error, we fail to reject the null hypothesis when the alternate hypothesis is true. Whereas, instead of accepting alternate hypothesis we do not accept alternate, we accept null hypothesis. We have 2 situations, one is called the do not reject null hypothesis, null hypothesis is true null alternate hypothesis is true. So here this is the correct, null hypothesis is true you do not reject null hypothesis, similarly alternate hypothesis need to accept reject null hypothesis so this is the correct situation, whereas in some times although we need to, null hypothesis is true we end up rejecting the null hypothesis that is called the type 1 error or α error. Whereas in the other situation we do not reject the null hypothesis whereas, we have H_1 is true that is called the type 2 error and that is called the β error.

So type 1 error, is generally the probability 95 % or 99 % or 90 % which we make use of in our statistical calculation. That is where if you take a larger α then obviously you are very sure that you will not reject the null hypothesis, where when H_0 is true but then your significance level also gets affected by that. So this table is very important when H_0 is true, we reject the H_0 that is called the type 1 error. When H_1 is true, we fail to reject the null hypothesis that is called the type 2 error or β error. So type 1 error is called the α error and type 2 is called the β error. So we need to always, sort of balance between the α and the β error and generally we give more importance to the alpha error actually.

(Refer Slide Time: 07:31)



As you know we have a continuous data and alternatively we have the discrete data. Discrete data we use equations like Binomial or we use Poisson and so on. So when we have the continuous data there are many situations, we can have one sample that means I know the population details, I take a small sample and I am comparing with the population that is called a one sample. For example, I know the 12th standard average from a particular school is 95 %, so I take 10 students in that school and calculate their 12th standard average. I may get some \bar{X} , now I want to know whether this \bar{X} is related to the 95 % school average, which is more like a population or is it very far away. So here we are collecting only 1 set of samples and then comparing with the population that is called a one sample test or one sample t-test. Here we are comparing the mean of the sample with the population mean. Now what is Two samples, suppose I am comparing performance of drug A and performance of drug B and trying to tell there is no

statistically significant difference between their performance or there is significance difference that means, I am comparing the means of 2 samples that is why it is called the two sample t-test. Now again you can have Multiple samples, I may be comparing drug A, B, C, D so, I could be having lots of different means that is called a Multiple samples actually.

So there are many, many ways by which one could analyze these data one is called the Study Stable or Run Charts, then we can look at the Shape we can create a histogram and see whether it looks normally distributor, then we can look at the Data whether it is and so on actually. And then, there is something called Chi Squared Test and then there is something called t test, t-tests are generally meant for comparing means, chi square test are generally meant for comparing ratios. And there is something called F test, which is generally meant for comparing variances or spreads actually.

In the t test we have 3 types of t test, one sample t-test that means, I take only 1 sample and then compare it with the population or I can have a two sample t-test, where I am comparing 2 sets of samples or I may be comparing paired t-test that means, there is a relationship between the sample items with a and sample items in b. So when I am comparing variances there is something called F test or if I am having multiple samples and I am comparing variances then this is something called ANOVA analysis of variance test and if I am comparing the spread there is something called homogeneity of variance. There are different types of test that are possible and we are going to spend lot of time on each one of them actually. So we can have only one sample collected, comparing it with the population or we could be having two samples collected, comparing with the population or we could have multiple samples collected, comparing with the population. Then if I am comparing means then there is something called t test, the one sample t-test, the two sample t-test, paired t-test. If I am comparing variances there is something called F test, if I am comparing a variances of a large number of data sets or samples this analysis of variance. If I am looking at spread of the data I can use something called Homogeneity of variance and so on. So large number of tests are possible we will talk about each one of them in detail and we are going to spend with some examples also, so do not worry about it.

Refer Slide Time: 11:50)

Comparing Means

The t-test allow you to determine if there is a statistically significant difference in the means of two samples, or between the mean of a sample and an established value.

- T-tests are robust to non-normal data
- The test is more sensitive if the two groups being compared have similar standard deviations.

What tool should you use?

- 1-sample t-test : if you want to compare the mean of your data to a known value
- 2-sample t-test : if you want to compare the means of two samples
- Paired t-Test : if you are comparing the difference between two paired samples to zero.



So if you are comparing means, there is something called t test. We can get average of a sample one and then I get a mean of sample two and then I am going to find out whether both the means come from the same population or each of the mean come from 2 different population that is called a t-test, t-tests are quite robust even for non normal data. Generally we can say the standard deviations have to be similar, but there can be some difference in the standard deviation also but still, t test is good. In t-test we have 3 types 1 sample t-test, 2 sample t-test and paired t test.

So one sample t-test you are taking a sample out and then you are getting the mean and the variance of that sample and you are comparing it with the mean of the population like I gave you some examples actually like, I take 10 students from a class and then get their mean average, class average, remarks average then I compare it with the school average and try to tell whether these averages are far away from the school average or they fall into the same population, I can do that sort of study. I can collect IQ of 10 students in a university and then try to say whether the mean IQ falls within the university average IQ or it falls outside that, so that is one sample t-test here we are taking only 1 sample.

Two sample t-test if I am going to have 2 sets of samples, I am taking 10 students from 1 university 10 students from another university and getting their getting their IQ's and then comparing their IQ's and trying to say whether the IQ's are statistically different or there is no statistically significant difference between these 2 IQ's. So that is called two sample t-test because I am using 2 sets of samples.

Paired t-test, if you are pairing 2 sets of data then difference in result should be 0 for example, I take 10 cats and I test a drug A on the 10 cats and look at their outcome then on the same 10 cats I give drug B and look at the outcome. So the difference should be 0, if there is no difference between drug A and drug B. If there is a statistically significant difference away from 0 then I can say yes, drug A is different from drug B because, I have used the same volunteer cats and I am testing drug A seeing some performance change, then I am testing drug B seeing some performance change, if drug A and drug B have to be the same then the performance change we observed, the difference in performance change we observed should be equal to 0 that is called the paired t-test. So we are going to look at examples of each one of them. So you do not worry about that.

So, interestingly all these tests are looking at means that means averages. Average or mean of the samples which you are taking it out, whether it is one sample t-test, 2 sample. Now you may ask the question suppose instead of 2 samples if I have many more samples what will I do of course I can do, take 2 sets of samples 1 at a time and do a two sample t-test, but there is another approach which is much faster that is called ANOVA analysis of variance, it is called the 1 way ANOVA we will talk about that later in the course.

(Refer Slide Time: 15:52)

The slide is titled "Comparing Variances" and is divided into three main sections. The first section, "Homogeneity of Variance", describes it as a hypothesis test for comparing variances, valid for both normal and non-normal data. The second section, "Null Hypothesis", states that all variances are the same, with the mathematical expression $H_0: \sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2$ and the instruction to "Can not reject H_0 , when $p > 0.05$ ". The third section, "Alternate Hypothesis", states that at least one variance is different, with the mathematical expression $H_a: \sigma_A^2 \neq \sigma_B^2$ (for at least one pair) and the instruction to "Accept H_a when $p < 0.05$ ". An NPTEL logo is visible in the bottom left corner of the slide.

Now, if you are comparing variances that means, you are comparing these standard error. So here we are comparing variances, this is generally valid both for normal and non normal. So the H_0 will be

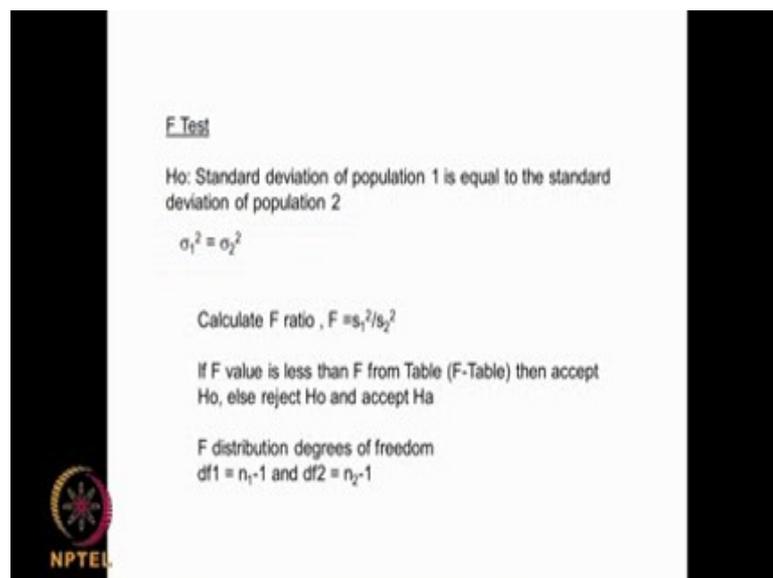
$$\sigma_1^2 = \sigma_2^2$$

and so on, we cannot reject H_0 when $p > 0.05$. The alternate could be

σ_A^2 is different from σ_B^2 , p is less than 0.05 we have to reject H_0 and accept H_a .

So you see the tests are there, for Comparing Variances. So in the t-test we are comparing means, here we are comparing variances.

(Refer Slide Time: 16:38)



The slide contains the following text:

F Test

H_0 : Standard deviation of population 1 is equal to the standard deviation of population 2

$$\sigma_1^2 = \sigma_2^2$$

Calculate F ratio, $F = s_1^2/s_2^2$

If F value is less than F from Table (F-Table) then accept H_0 , else reject H_0 and accept H_a

F distribution degrees of freedom
 $df1 = n_1 - 1$ and $df2 = n_2 - 1$

The slide also features the NPTEL logo in the bottom left corner.

One test that is there is called F test, that means I am comparing the variation from here sample 1 and a sample 2. So, the H_0 could be $\sigma_1^2 = \sigma_2^2$

σ_1^2 is equal to σ_2^2 that means variances are same or the alternate could be

σ_1^2 is different from σ_2^2 . So we calculate the f ratio which is given by $F = s_1^2/s_2^2$

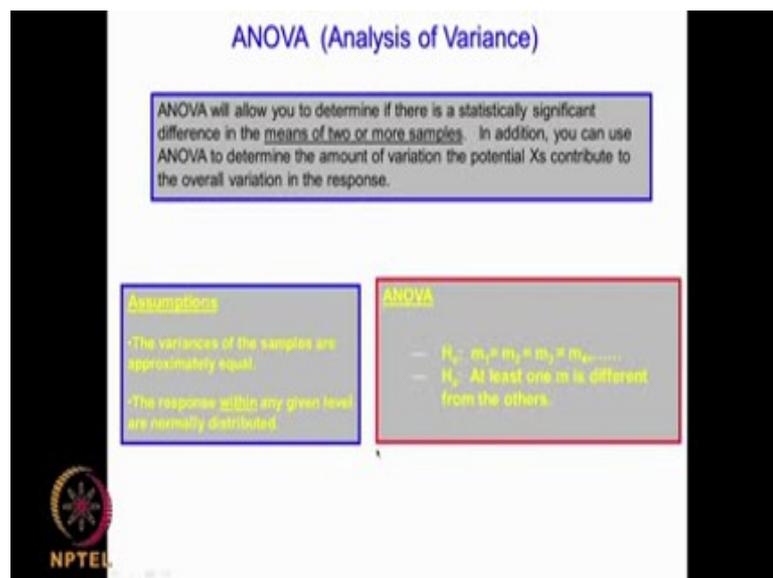
s_1 and the s_2 are the sample standard deviation, so s_1^2 square is the variance. There is a table called F table, for a 95 % or 99 % you will get F value and if the table F value is greater than the F you calculate, then you accept H_0 and if the table F value is less than the F you calculate, you reject H_0 and accept H_a . For different degrees of freedom, the degrees of freedom for data set 1

$$df_1 = n_1 - 1$$

if you have used n data sets and degrees of freedom for data set 2 is

$$df_2 = n_2 - 1$$
 if you have collected n_2 samples that is called the F test.

(Refer Slide Time: 17:52)



ANOVA (Analysis of Variance)

ANOVA will allow you to determine if there is a statistically significant difference in the means of two or more samples. In addition, you can use ANOVA to determine the amount of variation the potential Xs contribute to the overall variation in the response.

Assumptions

- The variances of the samples are approximately equal.
- The response within any given level are normally distributed.

ANOVA

- $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4, \dots$
- $H_a: \text{At least one } \mu \text{ is different from the others.}$

NPTEL

And then you also have ANOVA, when you are comparing a large number of data sets in the previous F test you have only 2 data sets, 1 and 2. So your saying

$\sigma_1^2 = \sigma_2^2$ alternative $\sigma_1^2 \neq \sigma_2^2$. So ANOVA is very, very powerful

because we can collect a large number of data sets I am comparing the IQ's of university A, university B, university C, university D and trying to find whether there is a statistical significant difference or not. I am comparing drug A, B, C, D in clinical trials, I want to perform analysis to find out whether there is a statistically significant difference, then I use ANOVA here.

The variances of the samples are approximately equal, the response within any given level are normally distributed these are the assumptions. So H_0 will be all these are same, whereas H_a is at least 1 variance is different then you get a p value less than 0.05 ok.

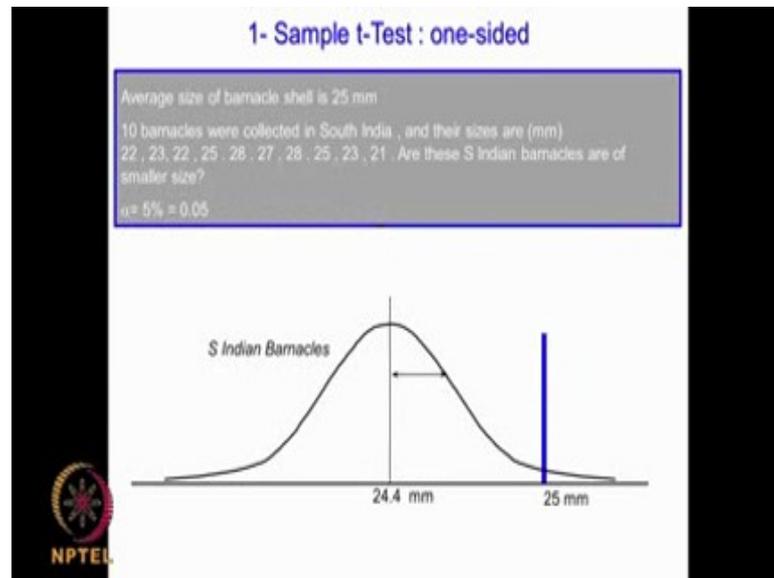
(Refer Slide Time: 19:00)



We have the t test, different types of t tests which are very powerful for comparing mean, then we have the test for comparing variances like F test and ANOVA and then for ratios we have something called chi squared test which we will talk about later. So if there is something called the power of the test or if H_1 is true, so that the distribution of X is specified by H_1 then the probability of rejecting the H_0 is the power of the test for distribution. H_1 is true, so you have 2 situations you do not reject H_0 that is beta error where as you reject H_0 . If H_1 is true and you reject H_0 that is called the power of the test for that distribution. So you have 2 types of error, alpha error where you should not reject H_0 but you end up rejecting H_0 . Whereas the beta error you have to reject H_0 but you do

not reject H_0 and that is also called the power of the test, the beta error is also called the power of the test. So these 2 terms are very, very important when you are deciding on the alpha error and the beta error.

(Refer Slide Time: 20:37)



Let us get into problems, the first problem is called the 1 sample t-test one-sided. The average size of barnacle shells is 25 mm, you know what is barnacle right? It is a marine organism, it is got a shell, it gets attached using glue to hard surfaces and then it feeds on it. So it is got a shell and average size is 25 mm. Now we have collected 10 barnacles in South India and we got their sizes, these are the sizes. Now are the South Indian barnacles of smaller size? That is the question or they are of greater size? So we take 95 % confidence. So this is the South Indian barnacles, the average comes out to be 24.4 but the population average is 25 mm, this is the statement. Now I want to know whether the this 24.4 is statistically significantly smaller or it comes from the same population at a p value of 0.05 or a 95 % confidence. So how do we that?

(Refer Slide Time: 21:51)

The slide is titled "1- Sample t-Test". It lists the null hypothesis as $H_0: \mu = \mu_0$ and the alternative hypothesis as $H_a: \mu < \mu_0$. The formula for the t-statistic is given as $t = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}}$. The degrees of freedom are noted as $df = n - 1$. Decision rules are provided: "If t < than the Table t accept Ho" and "If t is > than the Table t then reject Ho". The slide also indicates it is a "One tail test". An NPTEL logo is visible in the bottom left corner.

Simple,

$$H_0 : \mu = \mu_0$$

$$H_a: \mu < \mu_0$$

that is μ_0 is your original mean of the population, now they are same but alternate is μ is less than μ_0 that means, you want to know whether the average size of the barnacle shell from South India is less than this 25. So what do you do, you calculate t, if you remember this equation we had

$$\mu = \bar{X} \pm \frac{t_{df} S}{\sqrt{n}}$$

We rearrange that to get this t value, $\bar{X} - \mu_0$ divided by $s \sqrt{n}$. It is called df $n - 1$, if t which we calculate is less than the table t. So there is a table t for different degrees of freedom accept H_0 , if t is greater than t then the table t then reject H_0 this is called the, of course we are using 1 tail test.

(Refer Slide Time: 23:00)

size of barnacle shell

22
23
22
25
28
27
28
25
23
21

Avg 24.4
Std dev 2.590581
t calc -0.7377
df 9
t table 1.833
t calc < t table Accept H0

95% confidence limits for the mean::

$$\mu = \bar{X} \pm t_{df} \frac{S}{\sqrt{n}}$$

= 24.4 ± 2.26 * 2.59/√10
= 24.4 ± 1.85 = 22.24 to 26.25

So, what do we do? We get the average which is 24.4; we get the standard deviation which is 2.59. So the t calculated we can use here

$24.4 - 25 / s / \sqrt{n}$. So we get the t calculated and minus 0.73 for 9 degrees of freedom t table there is a t table here.

(Refer Slide Time: 23:29)

df	Probability in Two Tail (Confidence)					
	0.25	0.10	0.05	0.025	0.01	0.005
1	1.000	1.638	2.000	2.306	3.078	4.047
2	0.816	1.385	1.699	1.943	2.576	3.183
3	0.765	1.287	1.601	1.833	2.353	2.999
4	0.741	1.250	1.533	1.761	2.262	2.924
5	0.727	1.230	1.500	1.729	2.201	2.876
6	0.718	1.215	1.476	1.706	2.159	2.845
7	0.713	1.205	1.457	1.688	2.131	2.826
8	0.709	1.197	1.441	1.674	2.109	2.812
9	0.706	1.191	1.428	1.662	2.093	2.801
10	0.703	1.186	1.418	1.652	2.081	2.792
11	0.701	1.182	1.410	1.644	2.072	2.785
12	0.700	1.179	1.404	1.638	2.065	2.780
13	0.698	1.176	1.400	1.633	2.060	2.776
14	0.697	1.174	1.396	1.629	2.056	2.773
15	0.696	1.172	1.393	1.626	2.053	2.771
16	0.695	1.170	1.391	1.623	2.051	2.769
17	0.694	1.169	1.389	1.621	2.049	2.768
18	0.693	1.168	1.388	1.619	2.048	2.767
19	0.693	1.167	1.387	1.618	2.047	2.766
20	0.692	1.166	1.386	1.617	2.046	2.765
25	0.690	1.164	1.384	1.615	2.044	2.763
30	0.689	1.163	1.383	1.614	2.043	2.762
40	0.688	1.162	1.382	1.613	2.042	2.761
50	0.688	1.161	1.381	1.612	2.041	2.760
60	0.687	1.161	1.381	1.612	2.041	2.760
70	0.687	1.160	1.380	1.611	2.040	2.759
80	0.687	1.160	1.380	1.611	2.040	2.759
90	0.687	1.159	1.379	1.610	2.039	2.758
100	0.687	1.159	1.379	1.610	2.039	2.758
120	0.686	1.158	1.378	1.609	2.038	2.757
140	0.686	1.158	1.378	1.609	2.038	2.757
160	0.686	1.158	1.378	1.609	2.038	2.757
180	0.686	1.157	1.377	1.608	2.037	2.756
200	0.686	1.157	1.377	1.608	2.037	2.756
∞	0.685	1.156	1.377	1.608	2.037	2.756

I want you to look here for 95 % this is a two-tail test and the top 1 is the single tail test, in your problem we are talking about single tail test because we want to know whether this South Indian barnacles are of smaller size. So for 9 degrees of freedom go like this,

you go like this and read out 0.05, here you get for 9 degrees of freedom 1.833. So t table is 1.833, the t calculated is - 0.737. So t calculated is less than the t table so there is no reason for you to reject H0 at this condition. So what you can say is the South Indian barnacle, there is no statistical reason for saying, the South Indian barnacles are of smaller size. So South Indian barnacles come from the same population of 25 mm. In order to get your confidence limit on the mean which we calculated from the sample, as you know this equation

$$\mu = \bar{x} \pm \frac{t_{df} S}{\sqrt{n}}$$

mu is equal to x bar plus or minus t df s by square root of n the s by square root of n is called the standard error, right? I talked about this long time back.

Now you need to know that degree t value, now for these you have 10 data sets obviously the degrees of freedom is 9 you get t as 2.26 and the mean of this sample is 24.4. So

$$24.4 \pm 2.26 * 2.59/\sqrt{10} = 24.4 \pm 1.85$$

So the confidence limit 95 % confidence limit for the this mean is 22.24 to 26.25. One important point you need to remember is the t which you calculated will be for two-tailed both the sides because we are talking about \pm that is why we get 2.26. 4 mm as the mean of the sample with the standard deviation of 2.59 and you want to know whether the μ which you calculate = μ_0 or the alternate $\mu < \mu_0$, so for $\mu = \mu_0$. So what do you do, we calculate the t you know t from this equation if you adjust it is

$$t = \frac{\bar{x} - \mu_0}{\frac{S}{\sqrt{n}}}$$

$$\frac{-}{\sqrt{\quad}}$$

So you get minus 0.7377 for 9 degrees of freedom we have from the table, I showed you t table that top one is for single tailed, the bottom one is for two-tailed so for 5 and p of 0.05, 9 degrees of freedom you get 1.833. So obviously, the t which you have calculated is much less so there is no reason for you to reject the null hypothesis.

For 95 % confidence we say $24.4 \pm 1.96 s$ is equal to given here $2.59 / \sqrt{n}$. So you get the confidence limit for the mean a 22.82 to 26 so obviously your 25 falls within that that is why you are not able to reject the null hypothesis at this condition. So this table is very important as we can see this table gives you for different degrees of freedom the top 1 for single tail, the bottom 1 is for two-tail, So if I am interested in a single tail I will use 95 % this column, if I am interested in two-tail 95 % I use this column. As you go down, down, down as you can see for infinite degrees of freedom we get 1.96. So, as I said two-tailed 0.05 means a single tailed 0.025 because when you say 95 % two-tail the tails are divided half on both the sides that is why you get 0.05 by 2 which is 0.025 here. Do you understand? Do you understand the logic of this particular table, this table is very important when you are calculating t test, when you are calculating t based on means, whether it is 1 sample t-test the example which we saw or later on we are going to look at 2 sample t-test, paired t tests and so on. So this table is very important. So that you calculate t from the equation and then you compare with the t in the table for a different degrees of freedom, and then you say whether the t calculated is greater than t table, if it is greater than you have you can reject the null hypothesis, but if it is less we calculated less than the t table we cannot reject the null hypothesis.

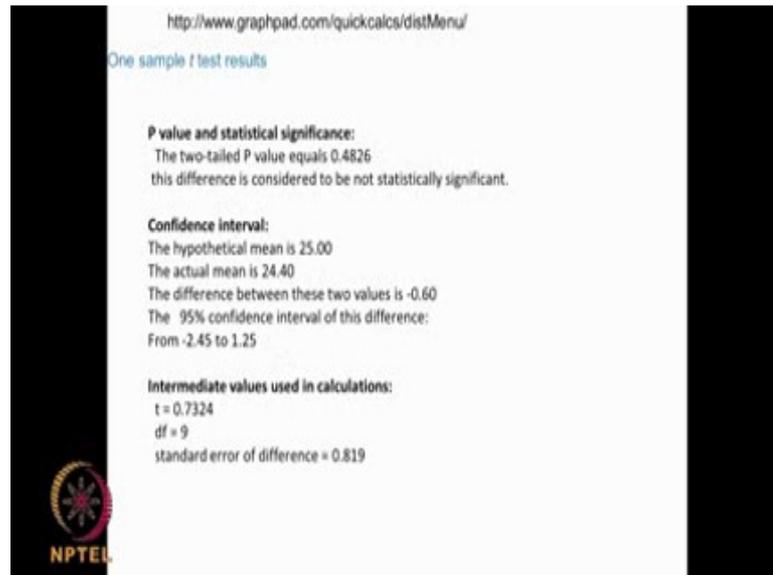
(Refer Slide Time: 28:40)

The screenshot shows a web-based calculator interface. At the top, the URL <http://www.graphpad.com/quickcalcs/distMenu/> is visible. Below the URL, the text "t value results" is displayed. Underneath, it shows "Probability = .05 df=9 Two tailed: t= 2.2622" and "5.000% of a t distribution has values greater than 2.262 or less than - 2.2622." Below that, it shows "One tailed: t= 1.8331" and "5.000% of a t distribution has values greater than 1.833" and "5.000% of a t distribution has values less than - 1.833". Further down, the text "P Value Results" is shown, followed by "t=-0.737 DF=9 The two-tailed P value equals 0.4799" and "this difference is considered to be not statistically significant." In the bottom left corner, there is a logo for NPTEL (National Programme on Technology Enhanced Learning) featuring a stylized gear and the text "NPTEL".

Now we can also use the particular software which I mentioned about the GraphPad software, which can also calculate the t value, given the probability value you can see for a 1 tail it is 1.833. So you compare with the t which you calculated, which is - 0.737. So

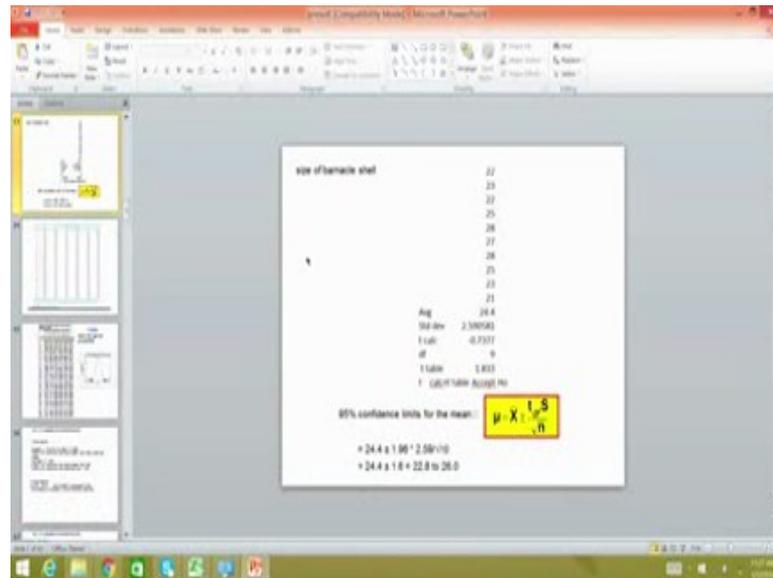
there is no reason for you to reject the null hypothesis, but if I take t value the - 0.737. And then the same software can be used to calculate the p value, the p comes out to be 0.4799. So obviously, you can say it is not statistically significant at all. It should have been 0.05 or less than only we can call it a statistically significant difference. So it is very useful for calculating this.

(Refer Slide Time: 29:39)



Or we can also use 1 sample t-test results, again the GraphPad software can do this and again the results as you can see it gives you a p value of this the difference is considered to be not statistically significant. The actual mean of the sample is 24.4, the hypothetical mean you want to reach is 25.0 and so on actually. So we will see how to do this, it is quite simple. We have this data set; I will show you how to do this.

(Refer Slide Time: 30:14)



So here we have the continuous data here we need to use this particular thing, here you have the one sample t-test as you can see here one sample t-test you can say continue. So we can even enter data like this or we can copy paste like this. So we can enter the data like that also. So I am comparing it with respect to 25, 25 is my population. Now I want to know the sample which is equal to 22 and so on. So I can copy this, copy this and I go to my I paste it here or if it does not get pasted. So obviously we can write, so we can write 22, 23, 22, 25, 28. So we say it 22, then we say 23, then 22, 25, 28, 25, 28, 27, then again we put 28, then again we put 25, then we put 23. So we have 10 data points we can put this here and then we can say you calculate now. So here we put 25 is the global average which you are interested in so we can say you calculate now. So by conventional the difference is considered to be not statistically significance, because the p value is coming out to be 0.4826. If we get p value less than, because the t value is 0.732 where as you want 1.833. So obviously it is not statistically significant different. So you can use the GraphPad software also to calculate, we can put in your data and you can use the GraphPad software also to perform this type of calculation.

It is quite useful software and this problem is quite simple. So you have the population mean and you have the sample. So from the sample you calculate the mean, from this sample you calculate the standard deviation and then you calculate the standard error which is given is by

S/\sqrt{n} , and then you know you can get the t value and then for 9 degrees of freedom you make use of this particular table. And for a single tailed test you use this, you go like this for 9 degrees of freedom you get 1.833, whereas t calculated is 0.7377. So obviously, we have no reason for rejecting the null hypothesis that is what it is. So I showed you how to calculate from the GraphPad software also actually. We will continue more on this one sample t-test as we go along.

Thank you very much for your time.