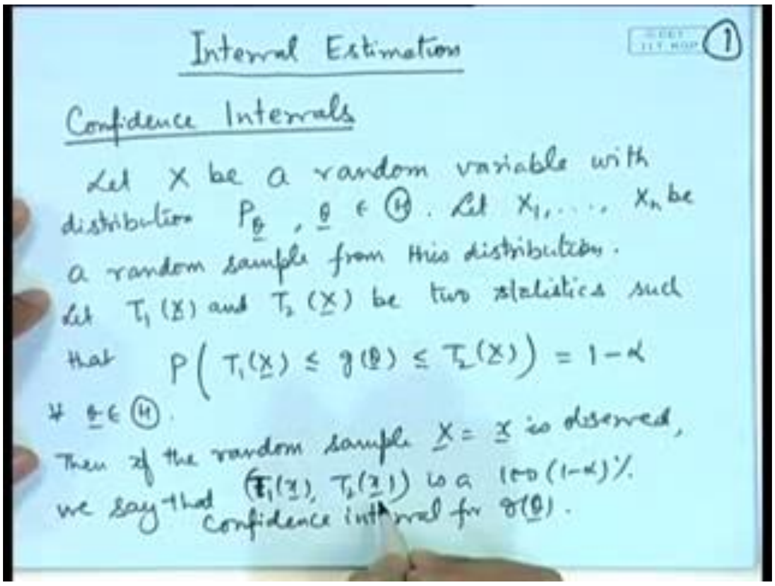
Probability and Statistics
Prof. Somesh Kumar
Department of Mathematics
Indian Institute of Technology, Kharagpur

Lecture - 61
Confidence Intervals – I

We will consider confidence interval estimation. Till now we have concentrated on the point estimation of a given parameter. So, in the point estimation we propose one value for the parameter to be estimated for example, if we are estimating say average income levels of persons of a particularly state, so we give a value say, we say the value is 2000 rupees per month. So, we are assigning a single value, but the problem with the point estimation is that that value is not necessarily the actual value because we do not know the true value; a more practical approach could be to given interval of the value rather than saying it is 2000 rupees per month, we may say the value is say 1900 rupees per month to 2100 rupees per month.

So, now since we are basing our decision on a random sample therefore, a certain probabilities associated with this statement, this is known as confidence level.

(Refer Slide Time: 01:29)



Interval Estimation

1

Confidence Intervals

Let X be a random variable with distribution $P_{\underline{\theta}}$, $\underline{\theta} \in \Theta$. Let $X_1, \ldots, X_n$ be a random sample from this distribution. Let $T_1(\underline{X})$ and $T_2(\underline{X})$ be two statistics such that

$$P\left(T_1(\underline{X}) \leq g(\underline{\theta}) \leq T_2(\underline{X}) \right) = 1-\alpha$$

$\forall\ \underline{\theta} \in \Theta$.

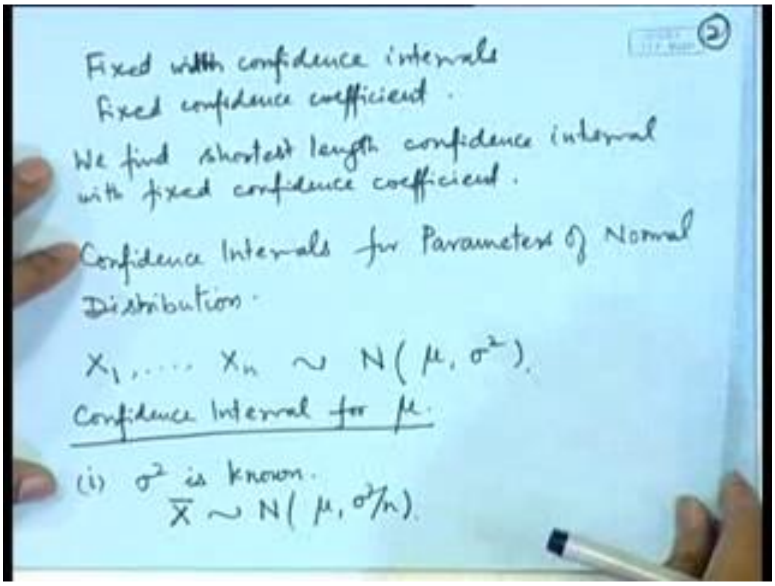
Then if the random sample $\underline{X} = \underline{x}$ is observed, we say that $(T_1(\underline{x}), T_2(\underline{x}))$ is a $100(1-\alpha)\%$ confidence interval for $g(\underline{\theta})$.

So, formally speaking we define a confidence interval as follows. So, we are discussing interval estimation and as I mentioned that since a probability statement is associated with that, we consider confidence intervals.

So let X be a random variable with distribution it is a P theta, theta belonging to the script theta. So, let X 1, X 2, X n be a random sample from this distribution. So, let say T 1 X and T 2 X be 2 statistics such that probability that T 1 X is less than or equal to say g theta, less than or equal to T 2 X is equal to 1 minus alpha for all theta, then if the random sample X is equal to small x is observed we say that T 1 x to T 2 x is a 100 1 minus alpha percent confidence interval for g theta. So, what is the interpretation of this? That if we take 100 samples and we calculate the value T 1 x and T 2 x then this interval is likely to include the g theta value 95 percent of the times, this is the meaning of the confidence interval because the probability of this statement the T 1 X less than or equal to g theta is less than or equal to T 2 X is 1 minus alpha therefore, 100 1 minus alpha percent of the times the interval T 1 X to T 2 X will include the true parameter value g theta. So, now, from here we understand that if I take alpha to be a small value say I take alpha is equal to 0.01; that means, 1 minus alpha is 0.99.

So that means we will have 99 percent confidence interval. So, the is smaller alpha value the larger is the confidence level, now how to find out the confidence intervals what you will called to be a desirable confidence interval? Naturally if we have higher confidence level then it is good because we can say that this is likely to include the true value; however, there is a certain contradiction here, because if we increase the probability then naturally the size of the interval will also increase. So, there is a conflicting goal that either we increase the length of the interval, either we increase the length of the coefficient that is magnitude of a coefficient or we decrease the length of the interval.

(Refer Slide Time: 06:00)



2

Fixed width confidence intervals
fixed confidence coefficient.
We find shortest length confidence interval
with fixed confidence coefficient.

Confidence Intervals for Parameters of Normal
Distribution.

$X_1, \ldots X_n \sim N(\mu, \sigma^2)$.

Confidence Interval for μ.

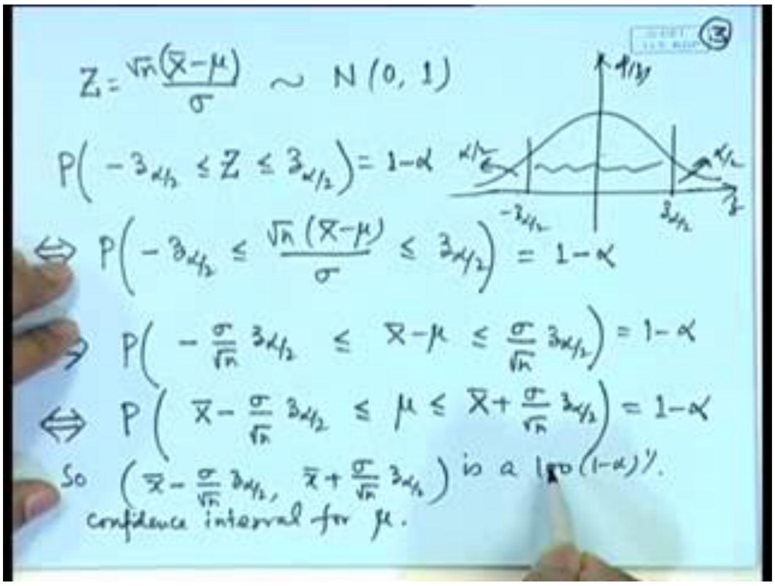
(i) σ^2 is known.
$\overline{X} \sim N(\mu, \sigma^2/n)$.

So, we have a theory of fixed width; fixed width confidence intervals or fixed confidence coefficient that means, we find shortest length confidence interval with fixed confidence coefficient.

This theory is closely correlated with the theory of finding out best tests of hypothesis, so we will not get into that right now. So, what is required here is that for a given parameter we should be able to find out 2 statistics such that this probability is independent of the parameter, because this is fixed value 1 minus alpha it is not dependent upon theta; that means, we should be able to consider certain quantity where g theta is involved, and I am able to write down the distribution of that which is free from the parameters. So, a well known technique here is called the pivoting technique and we make use of the well known sampling distributions that we have done; so let us considered confidence intervals for parameters of normal distribution.

So, we have the following set up that I have a random sample X 1, X 2, X n from a normal population with mean, mu and variance say sigma square. So, here the parameters are mean and variance and we are interested in the confidence intervals for mean and variance. So, let us considered say confidence interval for mu that is the mean; now there can be 2 cases one case is that sigma square is known. Now here we make use of the sampling distributions of X bar and s square. So, when we are considering mu let us considered X bar, we know that the distribution of X bar is normal with mean, mu and variance sigma square by n.

(Refer Slide Time: 09:06)



$$Z = \frac{\sqrt{n}(\bar{X}-\mu)}{\sigma} \sim N(0, 1)$$

$$P\left(-z_{\alpha/2} \le Z \le z_{\alpha/2}\right) = 1-\alpha$$

$$\Leftrightarrow P\left(-z_{\alpha/2} \le \frac{\sqrt{n}(\bar{X}-\mu)}{\sigma} \le z_{\alpha/2}\right) = 1-\alpha$$

$$\Rightarrow P\left(-\frac{\sigma}{\sqrt{n}} z_{\alpha/2} \le \bar{X}-\mu \le \frac{\sigma}{\sqrt{n}} z_{\alpha/2}\right) = 1-\alpha$$

$$\Leftrightarrow P\left(\bar{X}-\frac{\sigma}{\sqrt{n}} z_{\alpha/2} \le \mu \le \bar{X}+\frac{\sigma}{\sqrt{n}} z_{\alpha/2}\right) = 1-\alpha$$

So $\left(\bar{x}-\frac{\sigma}{\sqrt{n}} z_{\alpha/2}, \bar{x}+\frac{\sigma}{\sqrt{n}} z_{\alpha/2}\right)$ is a $100(1-\alpha)\%$ confidence interval for μ.

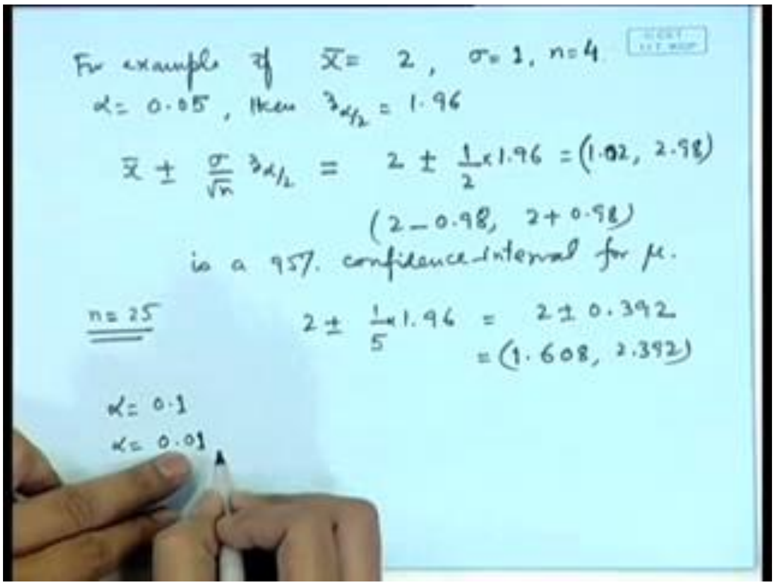
So, from here we can construct Z is equal to X bar minus mu divided by sigma by root n. So, this will have normal 0, 1 distribution; now you see this is important statement because here I have able to create a function of random variables and the parameter for which I need the confidence interval such that the distribution of this quantity is free from the parameters. So, now, if we make use of the normal distributions probability points, if this is the pdf of a standard normal distribution, then let us consider the point z alpha by 2 and minus z alpha by 2 that is this probabilities alpha by 2, this probabilities alpha by 2. Then I can write down the statement probability minus z alpha by 2 less than or equal to Z less than or equal to plus z alpha by 2 is equal to 1 minus alpha, that is the probability of this random variable line between this range is 1 minus alpha, we are showing f5 z here this is z.

Let us write down this statement elaborately. So, this is statement is equivalent to probability of minus z alpha by 2 less than or equal to root n, X bar minus mu by sigma less than or equal to z alpha by 2, is equal to 1 minus alpha.

We adjust this term we multiplied by sigma by root n and both the, so we get minus sigma by root n, z alpha by 2 less than or equal to X bar minus mu, less than or equal to sigma by root n z alpha by 2 that is equal to 1 minus alpha. Now this is equivalent to now if you look at this one this is equivalent to that mu is greater than or equal to X bar minus sigma by root n, z alpha by 2 and if I take this side then which is equivalent to mu is less than or equal to X bar plus sigma by root n, z alpha by 2. So, this statement is equivalent to X bar minus sigma by root n z alpha by 2, less than or equal to mu less than or equal to X bar plus sigma by root n z alpha by 2, this probability is equal to 1 minus alpha.

So, if we compare with the definition of the confidence interval, then this quantity is equal to T 1 X mu is g theta and X bar plus sigma by root n z alpha by 2 is T 2 x; that means, we are able to obtain 2 functions of random observables which include the given parameter with a given probability 1 minus alpha. So, we can say that x bar minus sigma by root n, z alpha by 2, to x bar plus sigma by root n, z alpha by 2 is a 100 1 minus alpha percent confidence interval for mu. That means, when I have the sample X 1, X 2, X n here sigma is known then easily I can calculate x bar sigma by root n and z alpha by 2 will depend upon the value of alpha; suppose I say alpha is equal to 0.05; that means, I want a 95 percent confidence interval. So, then z alpha by 2 means z of 0.025 from the tables of normal distribution 1 can see that this value is 1.96.

(Refer Slide Time: 13:16)



For example if $\bar{x} = 2$, $\sigma = 1$, $n = 4$

$\alpha = 0.05$, then $z_{\alpha/2} = 1.96$

$$\bar{x} \pm \frac{\sigma}{\sqrt{n}} z_{\alpha/2} = 2 \pm \frac{1}{2} \times 1.96 = (1.02, 2.98)$$

$(2 - 0.98, 2 + 0.98)$

is a 95% confidence interval for μ.

$\underline{\underline{n = 25}}$

$$2 \pm \frac{1}{5} \times 1.96 = 2 \pm 0.392$$
$$= (1.608, 2.392)$$

$\alpha = 0.1$

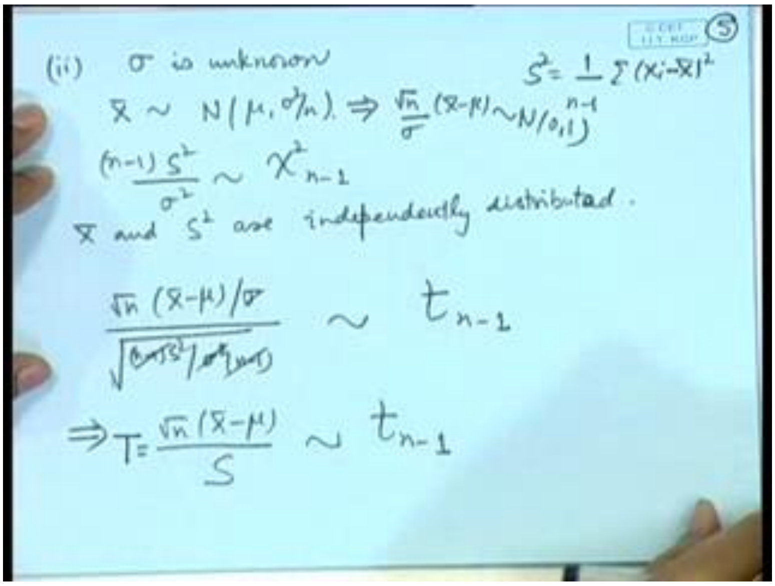
$\alpha = 0.01$

So, for example if say X bar is equal to some value say 2, say sigma is equal to 1, suppose n is equal to say 4 and alpha is equal to say 0.05; then you have z alpha by 2 is equal to 1.96. So, x bar plus minus sigma by root n z alpha by 2, what this value will be equal to 2 plus minus 1 by 2 1.96 that is 2.98 as the upper value and 1.09 1.02 as the lower value; basically we are getting 2 minus 0.98 to 2 point plus 0.98. So, this is the 95 percent confidence interval.

For mu if we had a larger sample size we can easily see that as the n will increase the sample size is more than sigma by root n will become a smaller value and the length of the confidence interval will shrink. That means, accuracy will be more for example, suppose here n is equal to say 25, then this value will become 2 plus minus 1 by 5, 1.96. So, the value is actually 0.392. So, that is equal to 1.8062, 2.392. So, you can see that the interval is sharper and therefore, you have more confidence in a smaller interval rather than a large interval and this will give you more precise information.

We can see another thing suppose I have alpha is equal to 0.1 or suppose I have alpha is equal to 0.01, then you can see that if I have alpha is equal to 0.1 that means, I need only 90 percent confidence here that is 1 minus alpha will become 0.96. Naturally the interval will become is smaller, if I have alpha is equal to 0.01; that means, it say 99 percent confidence interval, so the interval length will become more. So, depending upon your compromise situation that how much confidence you need for the parameter, you can appropriately adjust your interval.

(Refer Slide Time: 16:46)



(ii) σ is unknown $\quad S^2 = \frac{1}{n-1}\sum (X_i-\bar{X})^2$

$\bar{X} \sim N(\mu, \sigma^2/n) \Rightarrow \frac{\sqrt{n}}{\sigma}(\bar{X}-\mu) \sim N(0,1)$

$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$

$\bar{X}$ and S^2 are independently distributed.

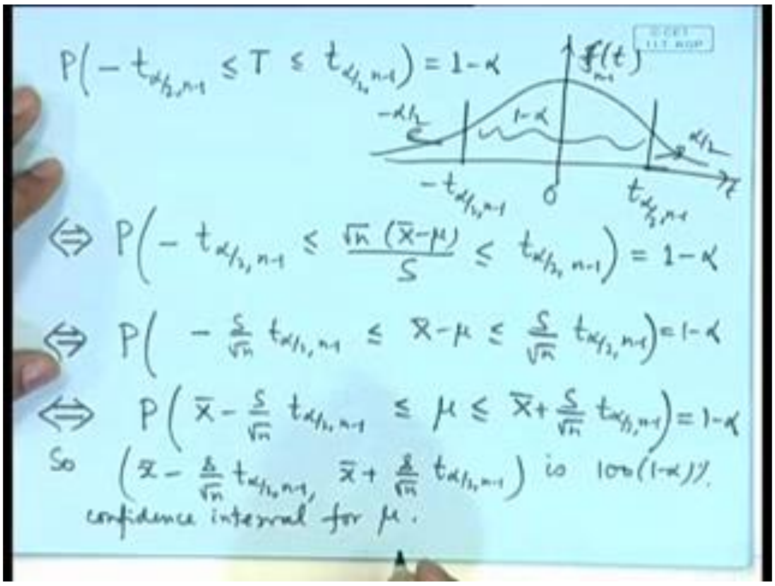
$$\frac{\sqrt{n}(\bar{X}-\mu)/\sigma}{\sqrt{(n-1)S^2/\sigma^2(n-1)}} \sim t_{n-1}$$

$$\Rightarrow T = \frac{\sqrt{n}(\bar{X}-\mu)}{S} \sim t_{n-1}$$

Now, let us take another case here; because in general sigma may be unknown, then naturally you can see here that this confidence interval cannot be utilized, because this requires knowledge of sigma. So, here we make use of the sampling distribution of S also. So, we have X bar follows normal mu sigma square by n, then if we consider n minus 1 S square by sigma square, where S square was defined as 1 by n minus 1 sigma x i minus X bar whole square that is the sample variance then this follows chi square distribution on n minus 1 degrees of freedom.

We also know that X bar and S square are independently distributed, this we had done in the theory of sampling distributions. So, if I have a normal variable I can convert it to a standard normal that is root n by sigma X bar minus mu, this follows normal 0, 1. So, root n X bar minus mu by sigma divided by, root n minus 1 S square by sigma square into n minus 1, this follows T distribution on n minus 1 degrees of freedom. So, after adjustment of this coefficient you can see that this is equivalent to root n X bar minus mu by S, this follows t distribution on n minus 1 degrees of freedom. Once again you observed this statement this involves the random observables the parameter for which I need the confidence intervals that is mu, and the distribution of this function of observation and the parameter is a distribution which is known; that means, it does not depend upon the unknown parameters mu and sigma square therefore, once again we can make use of let me call this quantity as T.

(Refer Slide Time: 19:01)



$$P\left(-t_{\alpha/2,n-1} \le T \le t_{\alpha/2,n-1}\right) = 1-\alpha$$

$f_{n-1}(t)$
$-\alpha/2$
$1-\alpha$
$\alpha/2$
$-t_{\alpha/2,n-1}$
0
$t_{\alpha/2,n-1}$
t

$$\Leftrightarrow P\left(-t_{\alpha/2,n-1} \le \frac{\sqrt{n}(\bar{X}-\mu)}{S} \le t_{\alpha/2,n-1}\right) = 1-\alpha$$

$$\Leftrightarrow P\left(-\frac{S}{\sqrt{n}} t_{\alpha/2,n-1} \le \bar{X}-\mu \le \frac{S}{\sqrt{n}} t_{\alpha/2,n-1}\right) = 1-\alpha$$

$$\Leftrightarrow P\left(\bar{X}-\frac{S}{\sqrt{n}} t_{\alpha/2,n-1} \le \mu \le \bar{X}+\frac{S}{\sqrt{n}} t_{\alpha/2,n-1}\right) = 1-\alpha$$

So $\left(\bar{x}-\frac{s}{\sqrt{n}} t_{\alpha/2,n-1}, \bar{x}+\frac{s}{\sqrt{n}} t_{\alpha/2,n-1}\right)$ is $100(1-\alpha)\%$ confidence interval for μ.

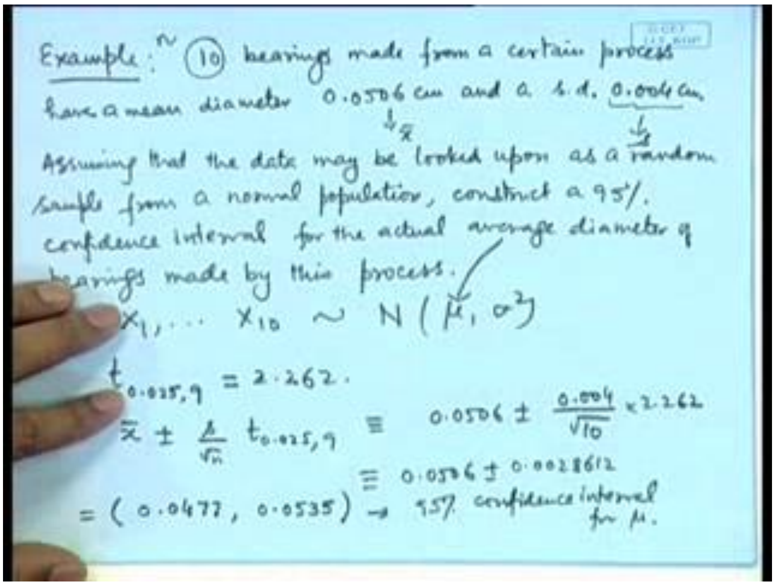
Then the distribution of T as we know it is a metric about the origin. So, if we have t alpha by 2 n minus 1, and minus t alpha by 2 n minus 1, this is the density of a t variable on n minus 1 degrees of freedom; then this probability alpha by 2, this probability is minus alpha by 2. So, these inter major probabilities 1 minus alpha. So, we can write down by statement probability minus t alpha by 2, n minus 1 less than or equal to T, less than or equal to t alpha by 2 n minus 1 is equal to 1 minus alpha.

So, from manipulating this statement we will be able to derive the confidence interval for mu. So, let us see this, this is equivalent to probability of minus T alpha by 2 n minus 1 less than or equal to root n X bar minus mu by S less than or equal to t alpha by 2, n minus 1. So, this probability is equal to 1 minus alpha. So, once again if we multiply by S by root n we get minus s by root n t alpha by 2 n minus 1, less than or equal to X bar minus mu, less than or equal to S by root n, t alpha by 2 n minus 1 that is equal to 1 minus alpha.

So, once again after simplification this condition is equivalent to mu is greater than or equal to X bar minus S by root n t. alpha by 2 n minus 1 and mu is greater than or equal to X bar plus s mu is less than or equal to X bar plus s by root n T alpha by 2 n minus 1.

So X bar minus S by root n, t alpha by 2, n minus 1, less than or equal to mu less than or equal to X bar plus S by root n, t alpha by 2 n minus 1. So, this probability is equal to 1 minus alpha. So, X bar minus s by root n, t alpha by 2 n minus 1, 2 x bar plus s by root n T alpha by 2 n minus 1 this is a 100 1 minus alpha percent confidence interval for mu. So, how to obtained the confidence interval in a practical situation, we observe the sample, we calculate the mean and the sample variance and then we look at the confidence level that we want; suppose alpha is equal to 0.1 then we look at t at 0.05 now suppose there are ten observations when we look at t 0.059 from the tables have the t distribution and evaluate this value. So, that will be a 90 percent confidence interval for the parameter mu.

(Refer Slide Time: 22:45)



Example: (10) bearings made from a certain process have a mean diameter $\underset{\bar{x}}{\underset{\downarrow}{0.0506}}$ cm and a s.d. $\underset{s}{\underset{\downarrow}{\underbrace{0.004}}}$ cm.

Assuming that the data may be looked upon as a random sample from a normal population, construct a 95% confidence interval for the actual average diameter of bearings made by this process.

$$X_1, \dots X_{10} \sim N(\mu, \sigma^2)$$

$$t_{0.025,9} = 2.262.$$

$$\bar{x} \pm \frac{s}{\sqrt{n}}\, t_{0.025,9} \equiv 0.0506 \pm \frac{0.004}{\sqrt{10}} \times 2.262$$

$$\equiv 0.0506 \pm 0.0028612$$

$$= (0.0477,\ 0.0535) \rightarrow 95\% \text{ confidence interval for } \mu.$$

Let me give one example here. So, 10 bearings made from a certain process have a mean diameter 0.0506 centimeter and a standard deviation 0.004 centimeter; that means, the sample has been collected on 10 bearings and its mean that is x bar has been calculated and the standard deviation. That means, s has been calculated and n is 10 here, assuming that the data may be looked upon as a random sample from a normal population construct a 95 percent confidence interval for the actual average diameter of bearings made by this process. So, if you are considering X 1, X 2, X10 as the 10 observations on this bearings so that means, these are denoting the diameters then this is following normal mu sigma square.

So, this actual average diameter is mu here; now we do not know the values of mu and sigma square so from the sample x bar has been calculated and s has been calculated here, n is equal to 10 here. So, if we want a 95 percent confidence interval; that means, we need the value of t 0.025, 10 minus 1 that is 9. So, this value from the tables of T distribution can be found and that is 2.262. So, we calculate here x bar plus minus s by root n, t 0.0259 that is this value is 0.0506 plus minus 0.004 divided by root 10 into 2.262 which is equivalent to 0.0506 plus minus 0.0028612. So, if we evaluate the lower and upper limits it is 0.0477 and 0.0535. So, this is a 95 percent confidence interval for mu.