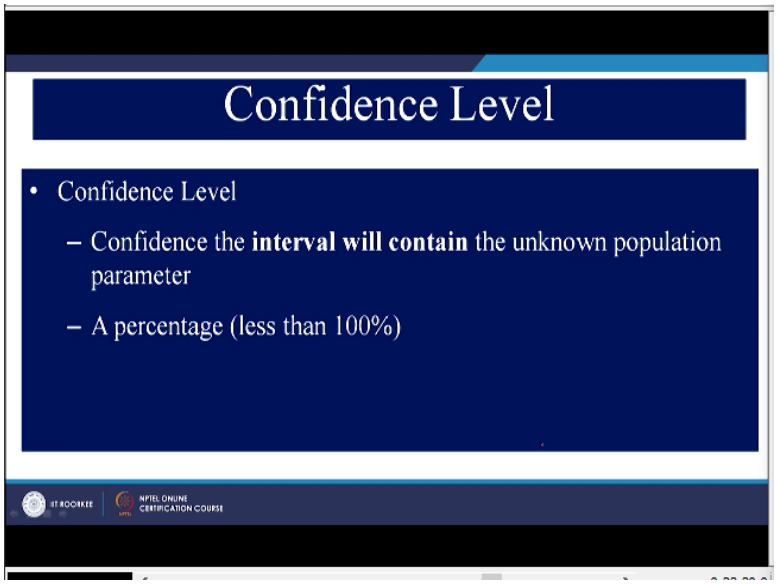
Business Statistics
Prof. M. K. Barua
Department of Management Studies
Indian Institute of Technology-Roorkee

Lecture-26
Method of Estimation

Good afternoon everyone I welcome you all in this session as you are aware in previous session we discussed about sampling and sampling distribution. In fact in one of the very few lectures I talked about types of statistics very nice and that there are 2 types of statistics descriptive statistics and inferential statistics. So for whatever we have seen in this course we have seen everything about descriptive statistics.

Now onwards will look at inferential statistics and I said there are 2 ways of inferential statistics, there are 2 types of inferential statistics or 2 methods of inferential statistics the first one is estimation and the second one is hypothesis testing. So we will look at first one which is estimation or will look at first the confidence interval estimation.

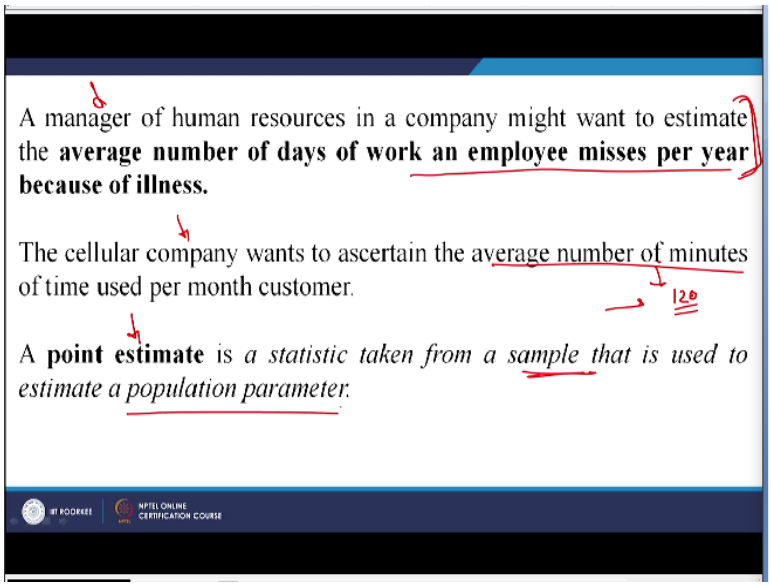
(Refer Slide Time: 01:21)



Generally what happens whenever we try to do something we try to let say in how much time it will take from Roorkee to Delhi right by bus. Then you always estimate the travel time whether it would be 4 hours or 4 hours 30 minutes or let say if you are if you want to cross road then you always estimate the speed of vehicles coming from either sides right. And then only you take

decision whether you should cross the road or not. Similarly in business world as well manager have to make several estimations let's say manager and HR manager wants to know what is the on an average how many employees miss the office in a year right.

(Refer Slide Time: 02:21)



So, one method is that he should note down how many employees did not come to the office on all 365 days or for whatever number of days that office was remained open or what the manager can do. You can take average number of days work and employee misses per year, so he will take few employees and he will find out that on an average 10 employees did not come for let say 10 days.

So on the basis of that the manager can estimate about the population information. In this case what the average number of days all the employees of the organization miss the office right. Let us look at one more example let say cellular phone company wants to know the average talk time of the mobile phone users. Now the mobile company has got let say bills of all the customers who are using that company's mobile phone.

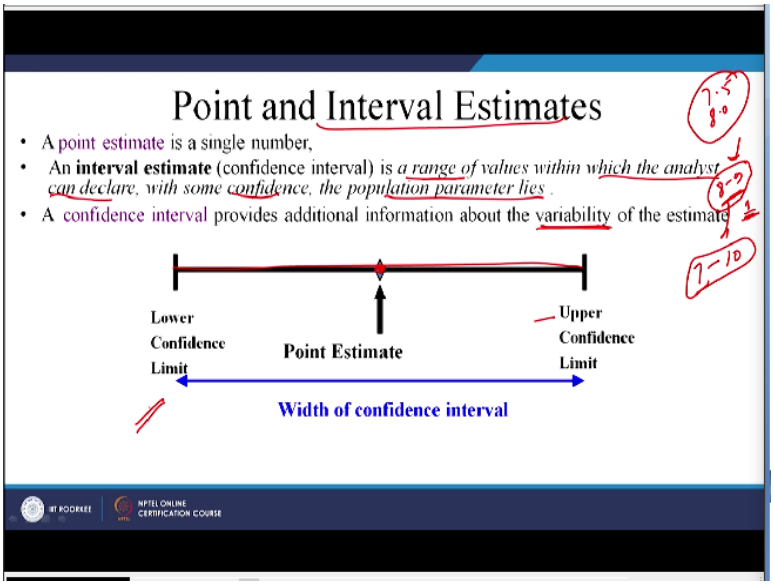
So, rather than looking at all those bills what company can do is, it can select few bills and can estimate about the average time of all the customers of that mobile phone company. So, this is known as and the in fact there are 2 types of estimates you got point estimate and the other one is

interval estimate. So, whatever we have seen over here in this example and in this example **we** we just calculated point estimation.

In point estimation is something wherein in you will have just one number right. For example cellular company wants to ascertain the average number of minutes. So, let say for simplicity answer is 120 minutes right, so this is just one number. So, this is known as point estimate right, so of course you say statistics taken from sample which we used to estimate about population parameter right.

So, apart from point estimation you always have one more estimation is called interval estimation right to give you one more example. Let say what would be the India's GDP in next financial year.

(Refer Slide Time: 04:58)

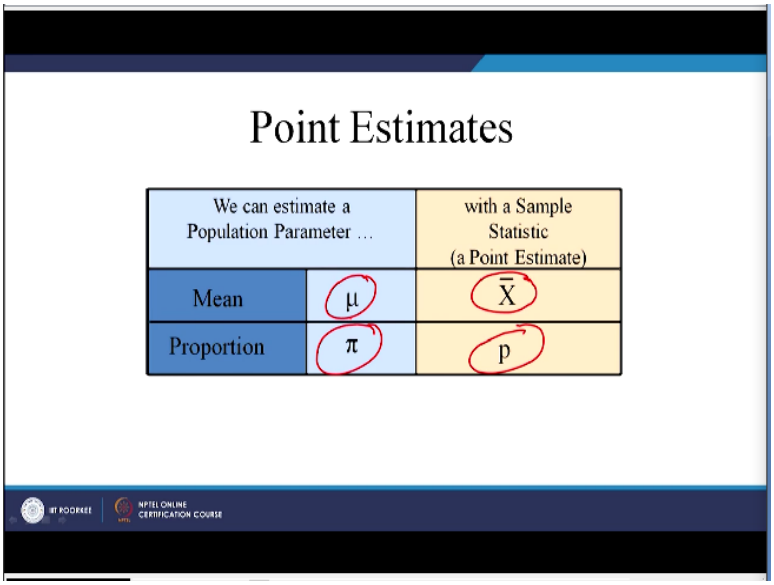


So, let say some of you will say 7.5% or let say the growth rate would be 8% so on. So, this is nothing but point estimation right. So, you can always have something call interval estimation. Interval estimation is a range, so let say the India's GDP growth rate in next year would be 8 to 9%. So, this is interval estimation right, so the range is just 1, right. So an interval estimate is a range of values within which the analyst can declare with some confidence the population parameter.

So, will say that the India's GDP would lie between 8 to 9% right, so it provides additional information about variability of the estimate right how much variability is there in estimation. So, let say India GDP is the interval is 7 to 10, so here variability is more, here variability is less. So, this is let say your point estimate and this is your interval estimate, so you will have upper limit and lower limit right.

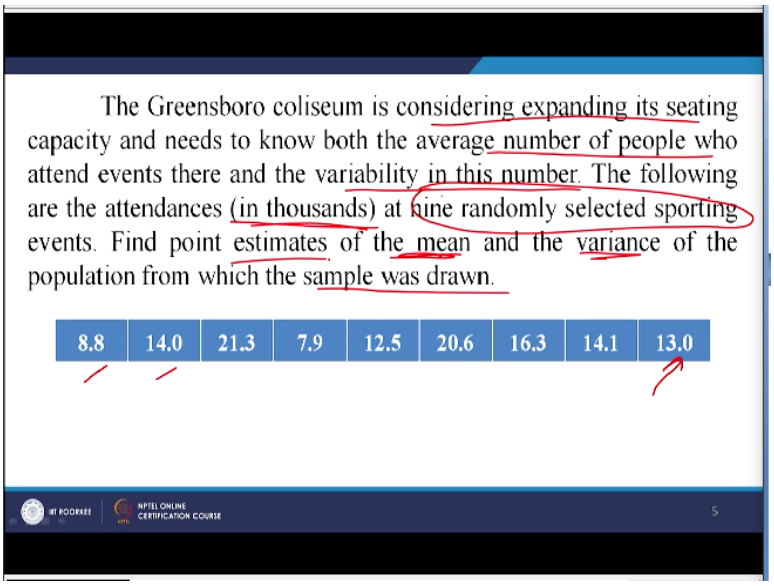
So, what is estimation, estimation is too easy simple statistics which will help you in knowing about population parameter. And you got either Point estimation or interval estimation.

(Refer Slide Time: 06:37)



So, we will say that this is sample mean is estimator of population mean simple proportion is estimator of population proportion right.

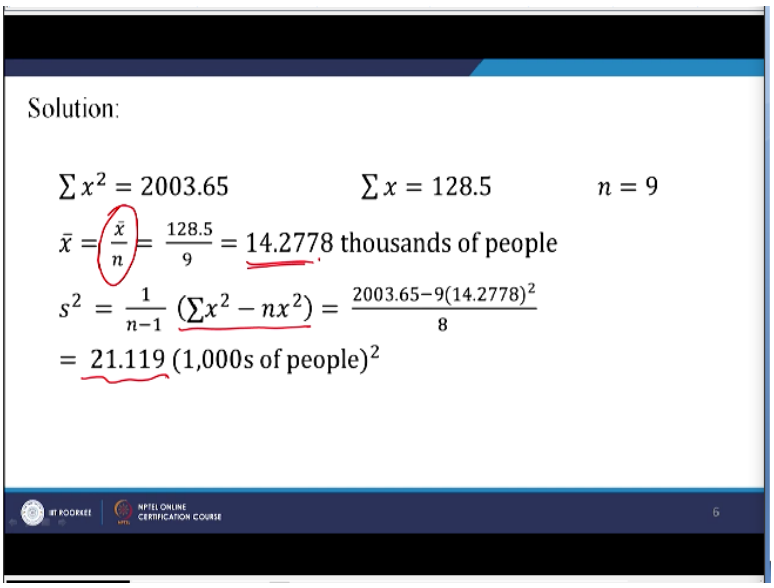
(Refer Slide Time: 06:50)



So, let us look at this example, the Greensboro Coliseum is considering expanding it's seating capacity and needs to know both the average number of people who attend events there and the variability in this number. The following are the attendances in thousands 9 randomly selected sporting events, find the point estimate of the mean and the variance of the population from which sample was drawn.

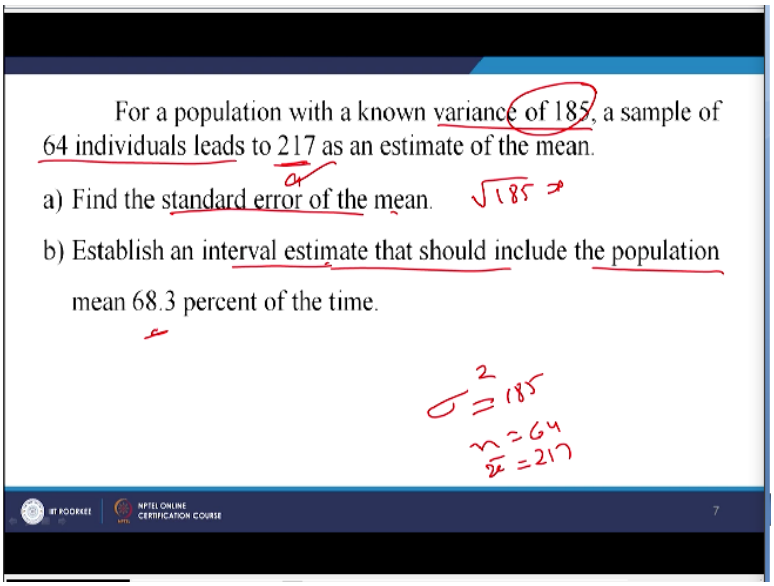
So, these are total 9 random you know 9 samples which we have selected randomly. So, would you do it was the question we have to estimate the pollution parameter right or the mean first of all. So, calculate mean and variance, so it is very simple just summation of all these divided by 9.

(Refer Slide Time: 07:56)



So, this is what simple mean, it is 14.27 and for calculating variance sample variance is $(\sum x^2 - n\,x^2)/n-1$. So, variance is 21.119 right, so this was just an example on point estimation.

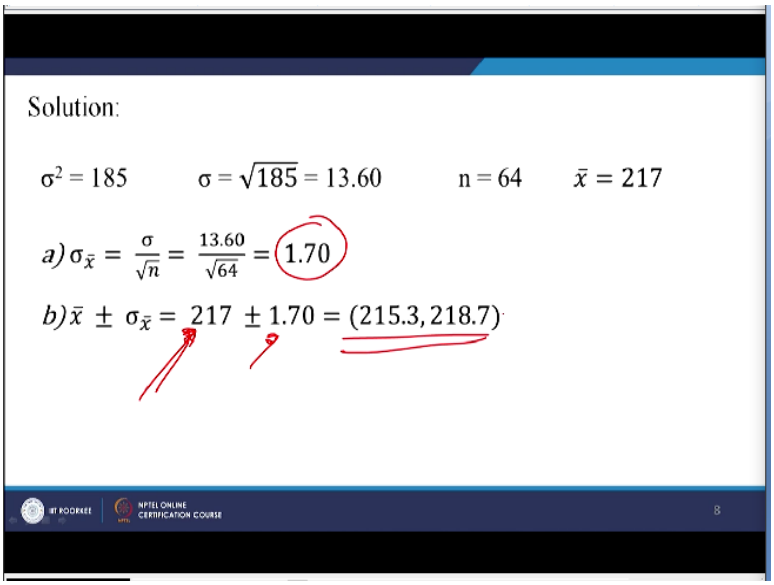
(Refer Slide Time: 08:26)



Let us look at this question for a population with known variance now variance is known here 185 and a sample of 64 individual, so n is 64 leads to this much mean. So X bar is 217, now we have to find out standard error of the mean, in fact we have seen this type of question in sampling and sampling distribution chapter as well. Establish an interval estimate, so the first part is find a standard error establish.

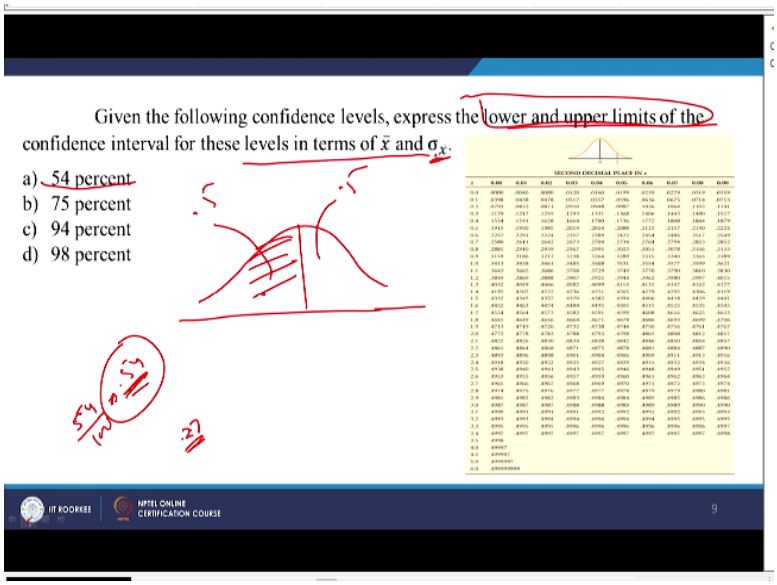
So, let us find out standard error since variance is there you just take under root of this right. So, this would be the standard error, so one standard error is known we can calculate interval estimate that should include population mean this much percent of the time ok.

(Refer Slide Time: 09:30)



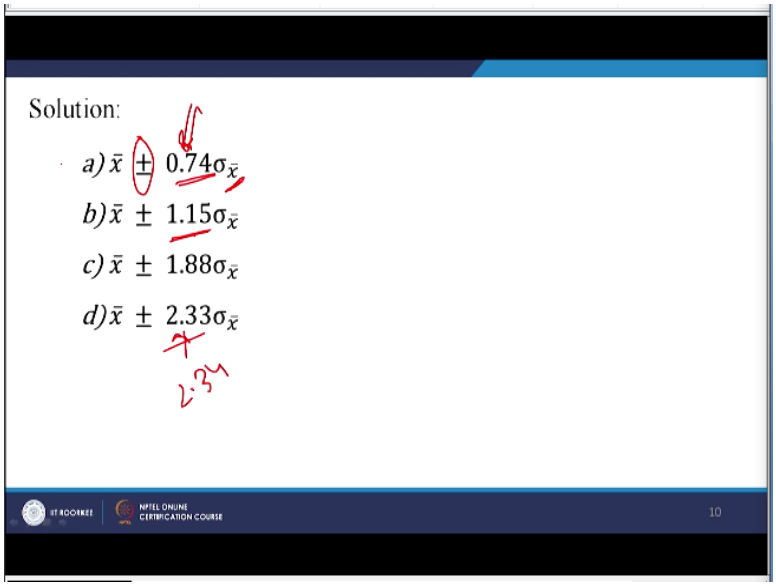
So this is your standard error 1.70, now since this is known you know the population mean sorry you know the sample mean right is 217. So 217 ± this 1.70, so this is the interval, so what is our answer, establish an interval estimate that should include the population mean this much percentage of time. So what is that interval, interval is 215.3218.7 right.

(Refer Slide Time: 10:13)



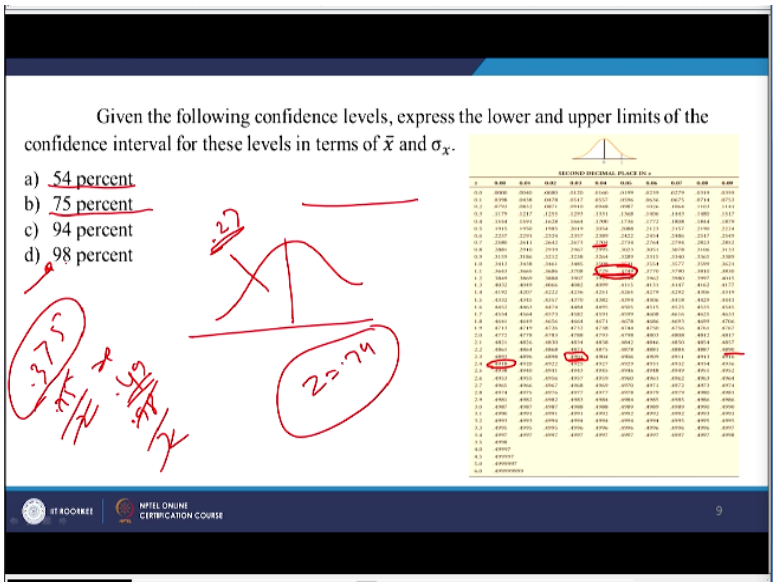
Let us look at one more question this is quite an interesting one, given the following confidence level express the lower and upper limit is of the confidence interval for these syllables in terms of X bar and standard error right. So we have to find out upper and lower limits, so how to do that let us look at this is the answer to this question.

(Refer Slide Time: 10:46)



So x bar ±0.74 standard deviation or standard error right, so 0.74 from where did we get 0.74 we know that this area is 50% this area is 50%. So, when we say 54%, so what we have to do it 54/100 it is In fact this becomes 0.54. So, this side is how much would be this side if you divide 0.54 into 2 parts then it would be 0.27 one side it is 0.27. So, let us call it ok let me erase this first.

(Refer Slide Time: 11:48)



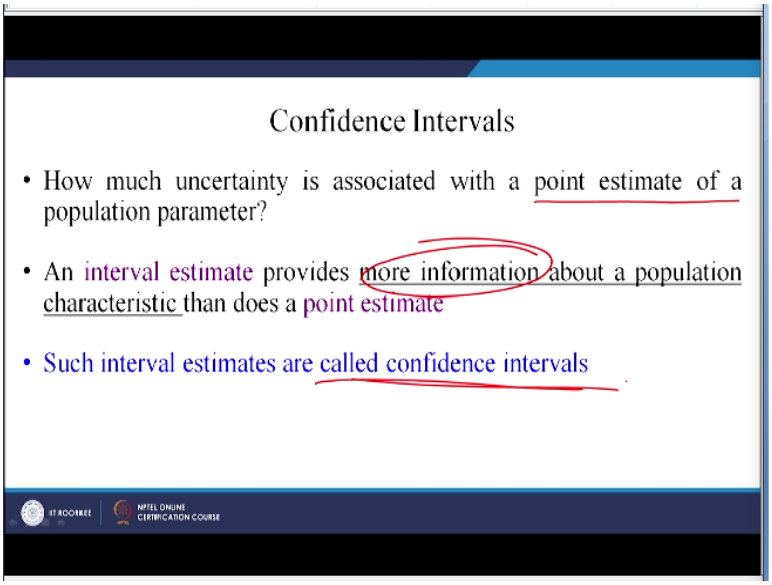
So when this is 54%, so each side is 0.27 right, so this 0.27, so, now let us look at this table where probabilities 0.27. So, this point no sorry when the area is where this probabilities 0.27 let us look at this where it is 0.27 is somewhere here right this is here. So, this is Z values 0.74, so

this what we have written here 0.74. And the next one is 75% how to find out this probabilities this 75%, 1/2 of this is what?.

So, just 0.75/2, so this is 0.3 then 7 right approximately 0.375 ok. 0.375 is where is 0.37. So, 0.3536 this is somewhere here 0.375 and somewhere here right 0.374. So, somewhere here right, so this is 1.14 or 15 is it so. This is exactly 1.15 ok, so for remaining one how would you calculate what it would be for this let's say 0.98/2. So, it would be 0.49, 0.49 is where is 0.4890.49 let us look at this.

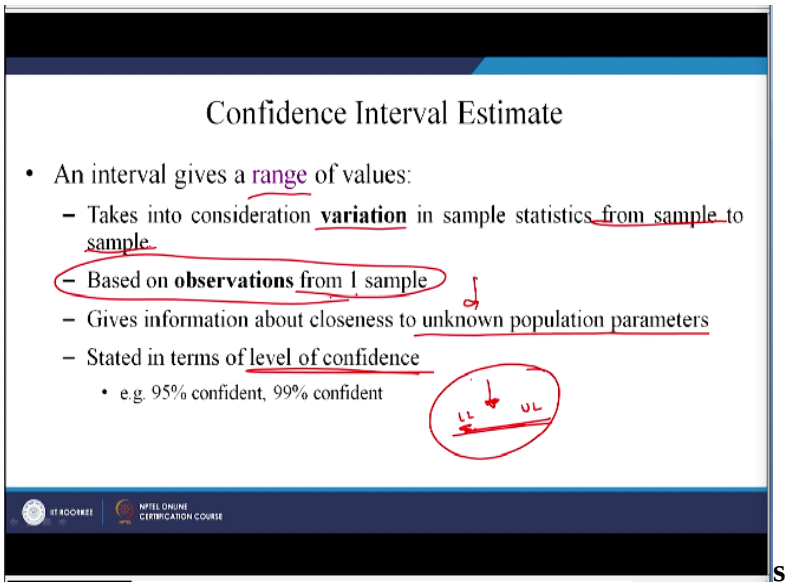
So, this is let us call it 2.40 or let say exactly it is 2.33 let us look at 2.33 2.3s, 0.49 then this is 0.4901 this what we are looking for it. So, approximate value you have to write over here, so even if you are right let say 2.34 then also the answer would be correct right. So, this how you should be calculating this Z values if you are given these probabilities right in terms of percentage.

(Refer Slide Time: 14:49)



Let us look at confidence interval once again, so it tells you how much uncertainty is associated with point estimate of a population parameter. So, an interval estimate provides more information rather than point estimation right such interval estimates are called confidence interval right. So, the interval should be called confidence interval, so because we will ensure that we are 95% confidence, we will say that the 90% confidence and so on right.

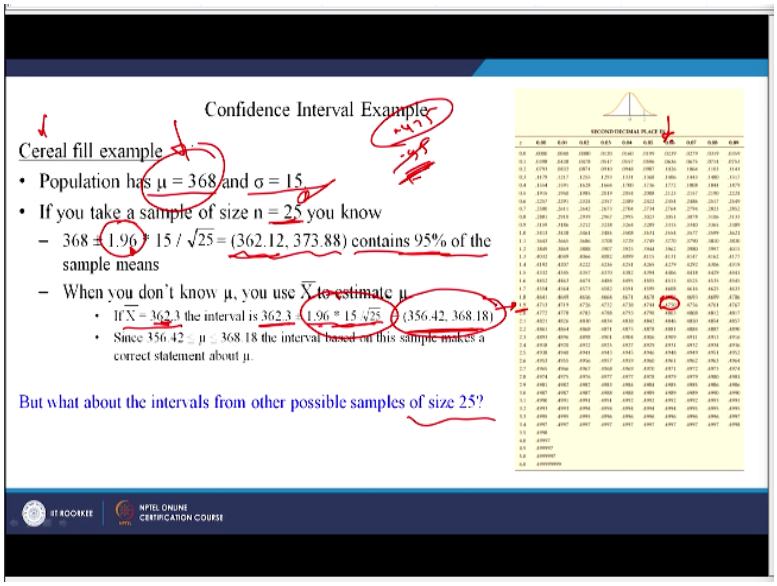
(Refer Slide Time: 15:30)



So it gives you a range I have already talked about this. It takes into account variation in sample statistics from sample to sample right based on observations. Generally we calculate interval estimate on the basis of just one sample, there is no need of taking several samples to know about population parameter, gives information about closeness to unknown population parameter we do not know the population parameter, stated in terms of level of confidence.

So, there are 2 things, there is something on confidence level and there is something called confidence interval. So the confidence interval is the range within which population parameter will fall right, so you have got upper limit and you have got lower limit right, so this is in this population parameter will fall. So, this confident interval right, confidence level is the probability that the population parameter will fall in this range. So, one is probability and the other one is range, let us take there is an assembly line and serial is being filled in plastic.

(Refer Slide Time: 16:49)

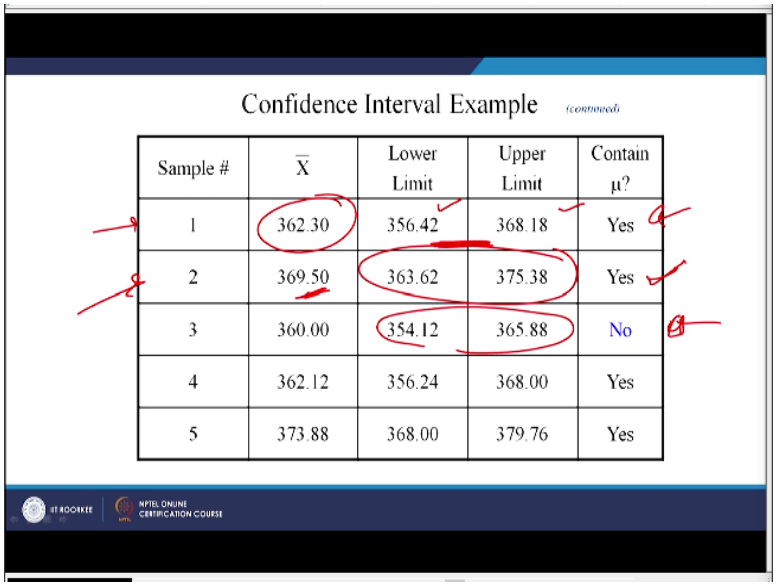


Let say pouch or in a plastic container and we know that the weight of the container from past data the population mean is this and standard deviation of weight is 15. Now if you take a sample of 25 unit is or 25 bags and when you come up with 95% confidence interval then you will say that this is the confidence interval. So, we say that when we take a sample of 25 packets the mean weight the interval of the mean weight would be 362.1 2, 373.88.

And how to get this is something which you need to keep in mind, so this mean 368 ± this is what you have to calculate from Z table right. So, when you say 95% means what 0.95 right, so 0.95, so 50% of this would be point let say 0. 475. So, 0.475 is where in this table 0.46, 0.475 is here right this is the object value 0.475. So, Z values 1.96 this 1.9 and this is 6 let us take an example where population mean is not there.

And it is not given then what you should do you should take sample mean sample mean instead of population mean. So, let us put sample mean over here 362.33 ± same value right. So, this is your interval estimate, so 356.42 to 368.18, so what we have said that if we take a sample of size 25 and sample mean is this. Then the confidence interval would be this much, now if you take another sample of 25 sample size you will have this mean indifferent. This value will remain constant but your interval will change.

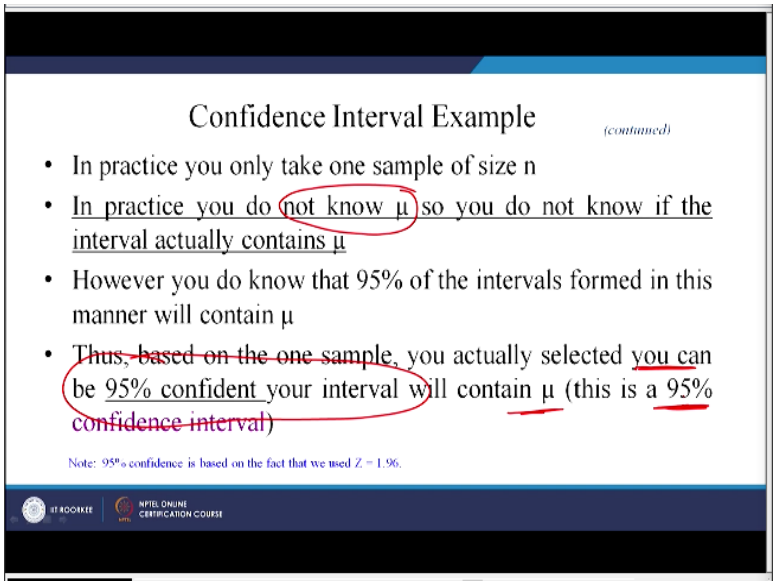
(Refer Slide Time: 19:51)



So, what we are saying what we have just done is this our sample mean was this when we to took 25 samples lower limit was this, upper limit was this. And whether this interval contains population mean, what was population mean 368, yes it does contain. Let us take one more sample of 25 unit is it is sample mean is this, does it contain population mean, does this range contain population mean yes here it does not.

In this case it does not contain, so when here you taken 5 samples each time sample size was 25 and you calculated confidence interval. And in one of them the population mean was not there.

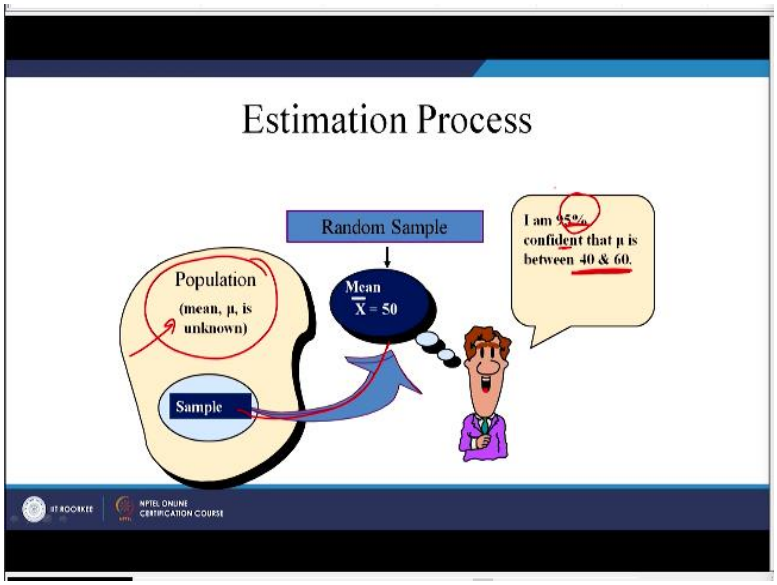
(Refer Slide Time: 20:50)



So, what is the meaning of 95%, generally in real life as I said will take just one sample and will develop confidence interval. In real life we do not know population mean because if we know population mean then in fact you would also be knowing something called standard deviation. So, that is the different issue altogether, so let us focus on this way, in real life we really do not know population mean.

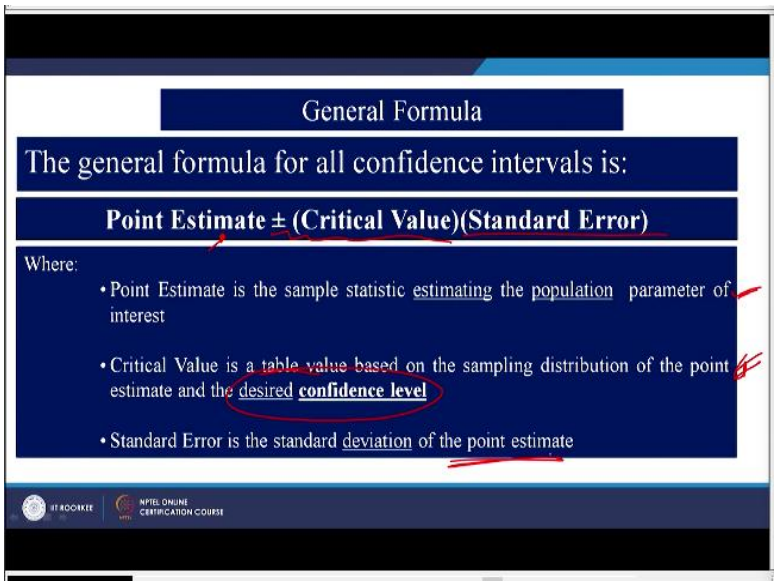
So what we are saying based on one sample which you have selected 95% of the confident you can be 95% confident your interval will contain mu and this is known as 95% confidence interval. So, what is the meaning of 95% confidence interval, the based on the sample which you select you can be 95% confident that your interval will contain population mean and this the meaning of 95% confidence interval.

(Refer Slide Time: 22:06)



So, what we have done today is we have estimated population proportion from sample right. So, this population, so mean is unknown initially will take a sample let say it is mean is 50 and will say I am 95% confident that the population mean would be between 40 and 60 right. So, you can have different confidence level or confidence interval right, this is known as estimation process.

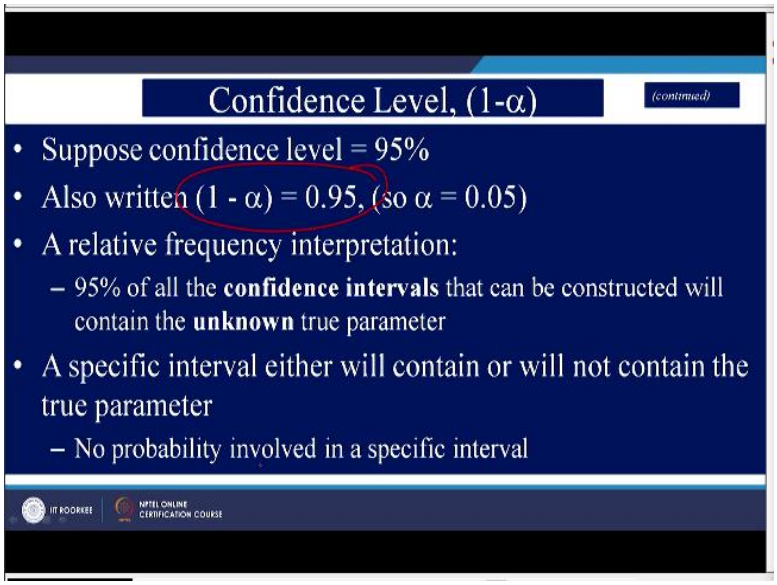
(Refer Slide Time: 22:40)



So, this is the general formula though we have already seen this and we are solved an example as well, so point estimate is equal to ± critical value multiplied by standard error. So, this is the simple statistic estimating the population parameter right this point estimate critical value of, of course it will come from Z table right. And this table value would be depending upon desired confidence level, for 90% it would be different, for 91% it would be different and so on.

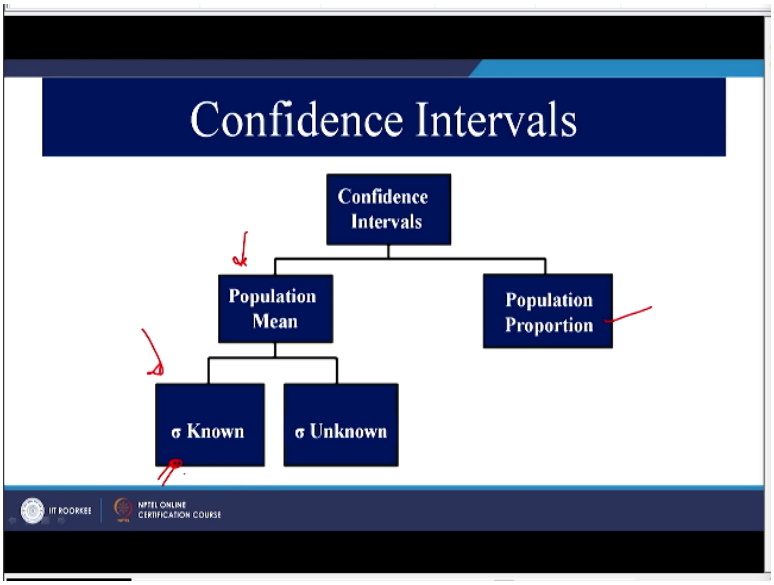
And standard error is the standard deviation of the point estimate right. Confidence level as I have already said confidence level is the probability right.

(Refer Slide Time: 23:30)



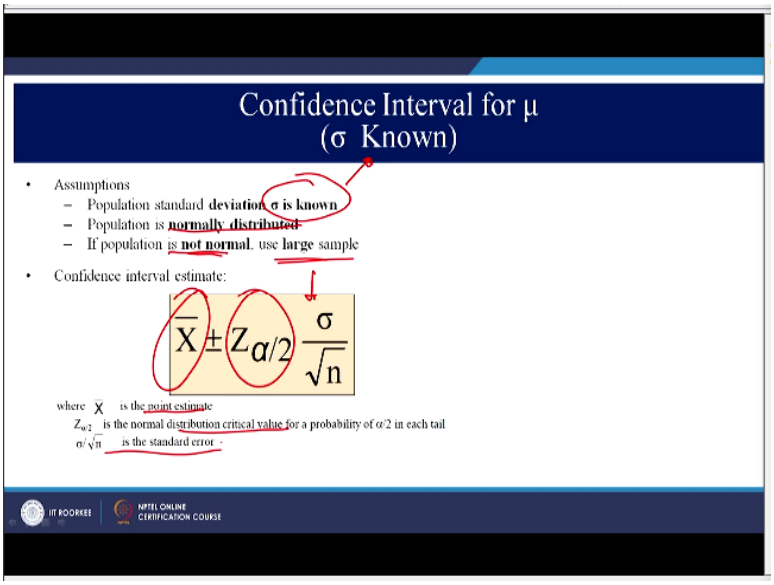
So generally we write confidence level is 1- alpha ok, the remaining things as we have already discussed.

(Refer Slide Time: 23:43)



So, let us find out confidence interval for population mean. Now we can calculate confidence interval for population proportion as well population variance as well right. So, we can different types of population parameters right, so first will look at population mean confidence interval for population mean. And in this case first will look at when standard deviation of population is known right.

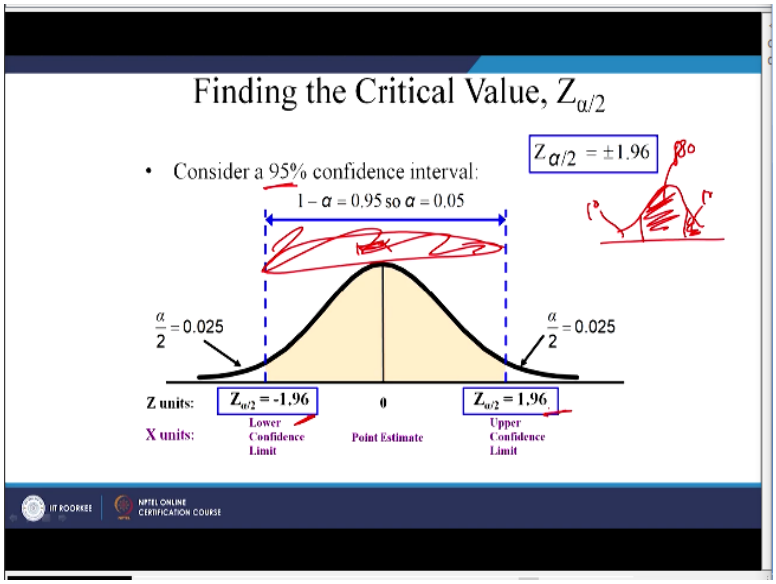
(Refer Slide Time: 24:16)



So, we will take this situation where standard deviation is known ok, so there are couple of assumptions you have to make. The first is the population standard deviation is known, second is population is normally distributed right it is not for non-normal population right, it is for normal distribution if population is not normal then will have to take a large sample right. So, the confidence interval estimate is this.

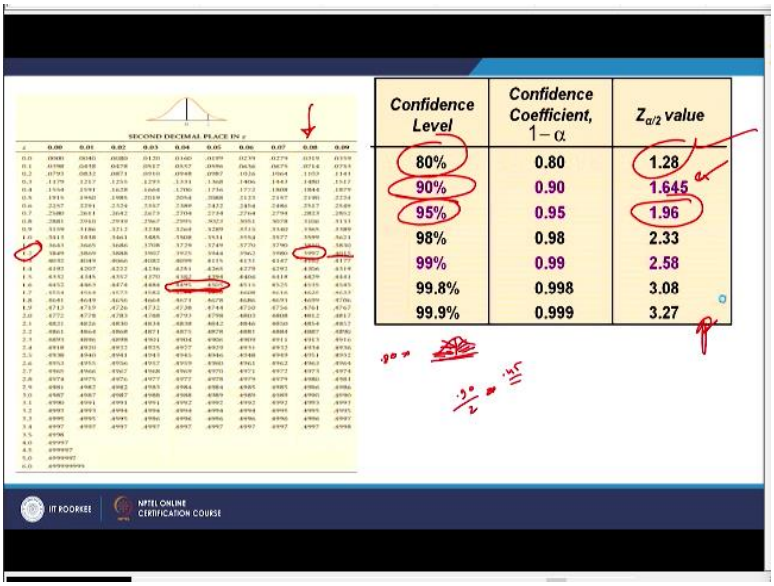
The sample mean critical value and the standard error, this what we have seen in the formula. So, this is point estimate or sample mean, this is normal distribution critical value and this is standard error ok.

(Refer Slide Time: 25:15)



So, let us look at 1 question, before this as I said this is your confidence level, this is generally we represented by 1-alpha. So when I say confidence level is let say 90%, so this area is 90% and the remaining area is 10% this side 10% this side sorry if this is 90% then this area is 5% and this remaining areas 5%. So, if let say confidence level is 80% right, so this is 80%, this is 10% and this is 10%, so this total becomes 100%. And for 95% confidence interval we have already seen Z values 1.96.

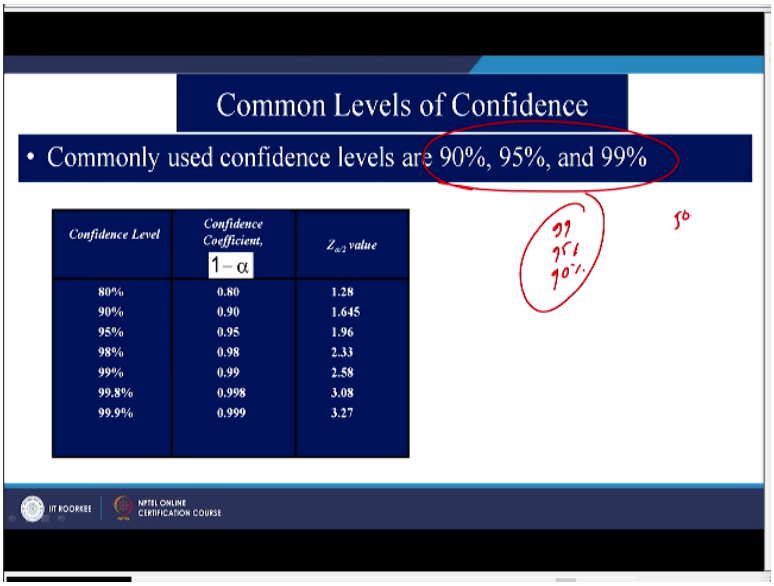
(Refer Slide Time: 26:14)



Let us look at other values in Z table, so when confidence level is 95% Z value is 1.96 but what if it is 80% how would you calculate this, what is the Z value when confidence level is 80%. So 80% means what 0.80 it means both sides you will have 0.4, this side and this side 0.4. So let us look at where is 0.4 in this table. So, 0.4 this is 0.4 approximately in fact this is even closer to that type 0.3997 right approaching 0.4.

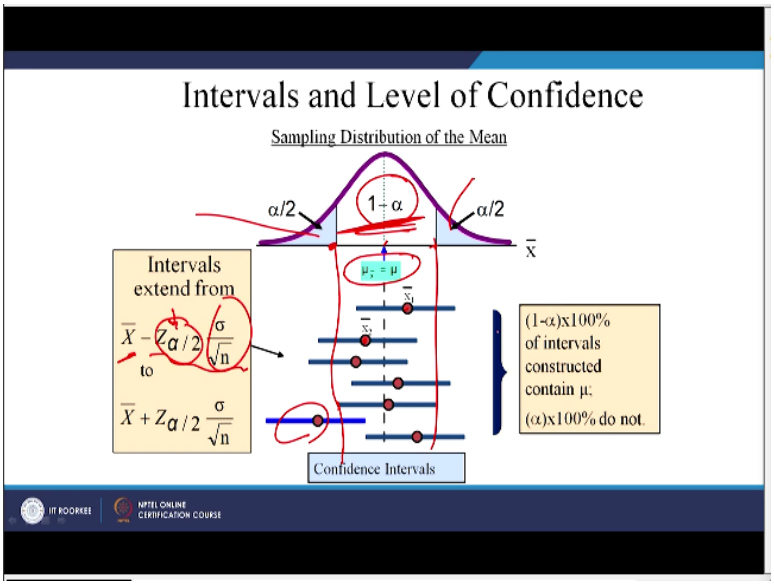
So, will take 1.28, so this 1.28 it this 1.2and this is 0.08 right, so just add these to right. what about let us say 90% what is the Z value when confidence level is 90%, so this 0.9/2 right, so it is 0.45. So, 0.45 is where this is somewhere, so will say this 1.64 right we say you can this 1.64 and if you want to take average you just take average you will get exact 0.45. So, this would be 1.645, so this how you should be calculating Z value for a given value of confidence level.

(Refer Slide Time: 28:09)



This is a similar table generally you should keep in mind that we always estimate our point estimation and interval estimation on these 3 confidence level 99%, 95% and 90% right. Generally we do not test our hypothesis and other things let say 50% or 40%, so just keep in mind that it is always 90%, 95% and 99%.

(Refer Slide Time: 28:49)



So, this interval and level of confidence, so this is your confidence level let say this is 90% and this is 90%, this is 5%, this is 5%.. So total is 100%, so this is your population parameter mu here and these are your sample mean let say first sample mean is this right, for some other sample, sample mean is this and so on right. So, all these are in this range right except this one which is outside the limit ok, so you can easily calculate this lower limit and upper limit.

So, X bar-critical value, this is critical value and this is standard right and similarly this side as well ok. Let me summarize what we did in today's session, we have looked at estimation and estimation is nothing but is a kind of inferential statistics wherein on the best of mean of the sample we try to estimate population mean. And we have seen something like **con** point estimation and interval estimation.

In point estimation you have just one value, in internal estimation you got range. We also seen how to look at Z value from table for a given confidence level. So, whatever is the confidence level if let say 80, so 0.4, 0.4 both sides of distribution from mean and look at the value of Z where probabilities 0.40. So, with this let me complete today's session will have a couple of examples on estimation in next session, thank you.