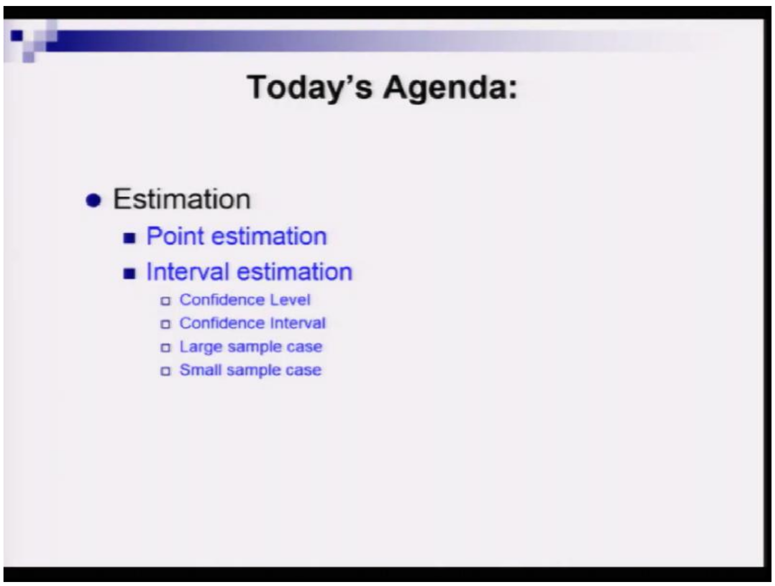
Applied Statistics and Econometrics
Professor Deep Mukherjee
Department of Economic Sciences
Indian Institute of Technology Kanpur
Lecture 9
Estimation (Part I)

Hello, friends. Welcome back to the lecture series on Applied Statistics and Econometrics. So today, we are going to start our discussion on estimation theory. So, we are going to cover a lot of topics. So, let us have a look at the agenda items.

(Refer Slide Time: 00:31)

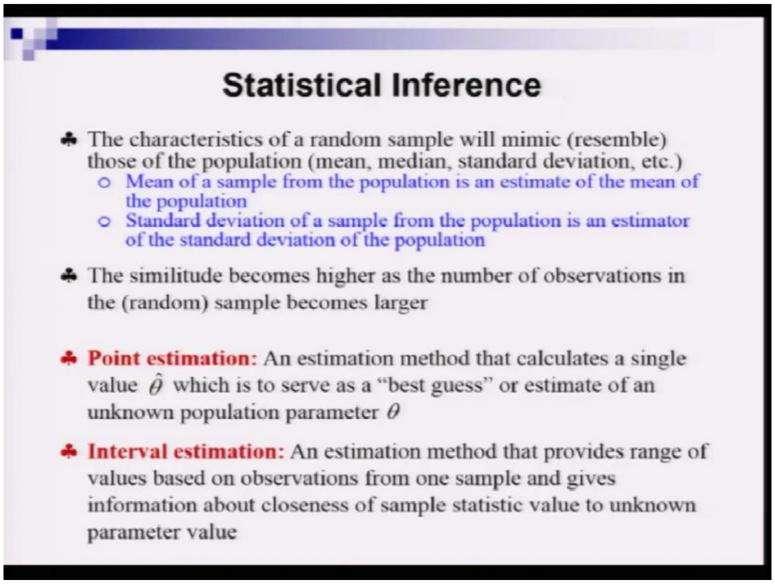


So, under the statistical estimation theory, we will discuss the case of point estimation, very briefly. But we are going to then discuss the case of interval estimation at length and under this interval estimation category we are going to introduce some new concepts to you, namely confidence level, confidence interval, and then we are going to discuss the large sample and small sample cases.

So, before we start the discussion on statistical estimation theory, let me briefly remind you about the problem at hand. So, we are actually in the first step of statistical inference, so, we actually have a large population or universe, but we do not have resources to reach out to the each element of this universe.

So, we actually have drawn a small sample from the population and we are trying to infer about population characteristics by looking at some sample characteristics. So, that is basically the background and the entire field of statistical inference helps us to make accurate guesses about population characteristics from sample characteristics.

(Refer Slide Time: 01:47)

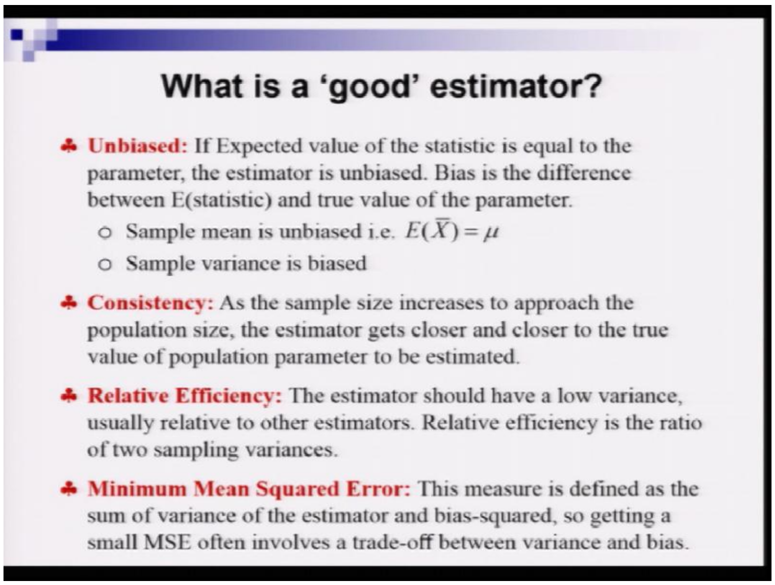


So, the characteristics of a random sample will basically mimic or resemble those of the population characteristics namely population mean, population median, and population variance, etc. Now, note that, mean of a sample from the population is an estimate of the mean of the population and standard deviation of a sample from the population is an estimator of the standard deviation of the population.

Now, this resemblance or similitude or likeliness becomes much more higher as the number of observations in the random sample becomes larger. So, in the last class only we have discussed the case of point estimation and interval estimation. So, I have given you the definitions there, I have reproduced the definitions again, so that you can remind yourself. But basically in a nutshell, point estimation is basically giving you a particular value of the sample statistic and that is your best guess about the unknown parameter value theta.

And in the case of interval estimation, you get a range of values as the best guess about the unknown population parameter theta. So, at this moment, let us look at some desirable properties of estimators. We are going to study a lot of estimators in this course. Now, how do you rank them? Because there are many estimators available when you are trying to draw some statistical inference from a sample about the population characteristics.

(Refer Slide Time: 03:26)



So, we are going to talk about 4 desirable properties or features of estimators and if these all 4 satisfied by 1 particular estimator, you can call that as a good estimator. Now, the most important property of a good estimator is the property of unbiasedness. So, what is unbiasedness? So, if expected value of the statistic is equal to the parameter value, then the estimator is unbiased. Note that, I have been talking about this feature of sample statistic from the very beginning, that sample statistic actually is a random variable associated with a probability distribution.

So, of course, there will be an expected value of the sample statistic as you keep on drawing repeated samples from the same population. So, if you look at the mean, if that mean is equal to the true value of the unknown population parameter, then you can call that estimator as unbiased. So, the related concept is bias and that is defined as the difference between the expected value of the statistic and the true value of the population parameter to be estimated.

So, now, we are going to talk about 2 sample statistic that we have seen several times in this course. So, first is sample mean and that is used to get a proxy measure for the population mean and that is an unbiased estimator, because we can prove, in fact, you can also prove that expectation of x bar is equal to mu. But note that, sample variance is a biased estimator and we have spoken about it.

So, the second property that is in the list is the property of consistency. What do we mean by consistency? Well, we have spoken about this at length in the previous class, but it is not a bad idea to have a very brief recap of a minute or so. So, how do I define formally in simple words without involving any symbols, so, I define it in the following way; so, as the sample size small n increases to approach the population size, so n tends to infinity in mathematical terms, then the estimator gets closer and closer to the true value of the population parameter that we need to estimate.

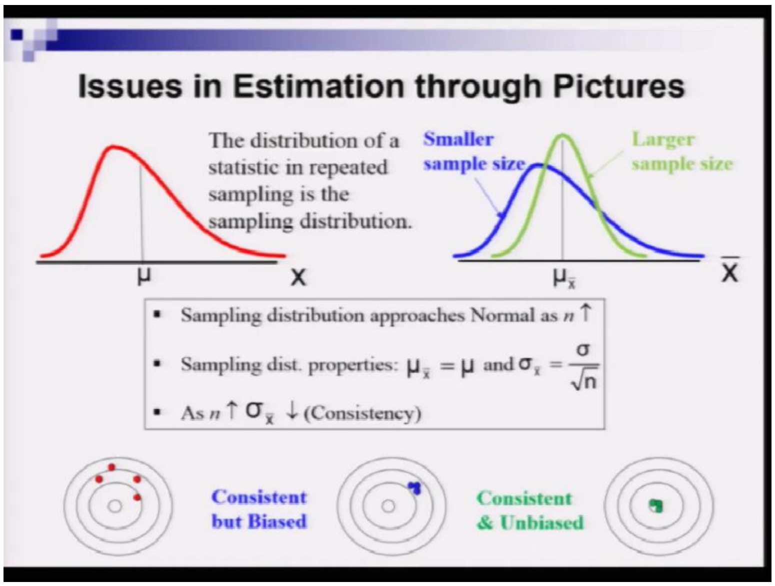
So, that is the feature of consistency. Now, the third in the least is the feature called relative efficiency. Now, efficiency is a very interesting and important property that we are going to visit again, when we will be dealing regression. But as of now, I just want to give you a flavor of that property just want to define it formally.

So, the estimator should have a low variance amongst all the available or potential estimators and that is usually relative to the other estimators. So, relative efficiency can be computed and that is the ratio of 2 sampling variances. So, this is basically the sampling variance of the sample statistic, I mean.

Now, the last but not the least in this list of criteria for good estimator is called minimum of mean squared error or that sometimes abbreviated as MSE. Now, this measure is defined as the sum of variance of the estimator and bias-squared. So, getting a small MSE often means that there is a trade off between the variance and the bias.

Now, we are going to look at a pictorial depiction of sampling estimation theory. And I believe that through this diagrammatic slide, many concepts will be much more clearer to you, even from the previous lecture. Because in the previous lecture, we have spoken about certain theoretical concepts, which may look at it, in words or in symbols, so maybe diagram will be a better representation of that concept and not only the concepts from the previous lecture, but also what we have been talking in this lecture also, I have tried to give you a visual representation of these concepts.

(Refer Slide Time: 07:58)



So, let us start with population and there is a continuous random variable associated with that population and let that random variable be denoted by capital X. We are interested to know the unknown value of the population mean mu and the population actually has a distribution of course, because of this random variable x and that is denoted by these red colored curve. Note that, I have not assumed symmetry, because we do not have to. So, this is an asymmetric distribution.

Now, what we would do, we have to draw a sample and if possible repeated samples from this population of same size. And then, we are going to proxy the population mean by the sample mean. So, sample mean is basically my statistic. So, if I now draw our distribution of the sample statistics, then we get some other distributions and let us have a look at these distributions.

Now, note that, here at this moment, I would like to mention 2 different theorems that we have spoken about in the previous lectures and they are law of large numbers and central limit theorem. So, I just want to give you a brief recap of both of these major theorems, so that you can understand why the graphs are behaving in certain manner.

So, I will first start with law of large numbers. So, in a very simple language, it says that as sample size n increases, the sample statistics get closer and closer to the population characteristics. And what the central limit theorem says? Central limit theorem says that, sample statistics computed from the means are approximately normally distributed regardless of the parent distribution.

So, now, let us look at the next diagram and here, I am going to plot the values of sample mean along the x axis, our horizontal axis and you see that I am showing you 2 distributions or 2 sampling distributions, one is for the smaller sample size and one for the larger sample size.

So, as you can see, as we increase the number of n by following the law of large numbers, the distribution actually moves much more closer to the true value of the unknown population parameter. And by the virtue of central limit theorem, the derived distribution of the sample statistic is going to converge towards an approximately normal distribution.

So, it can be approximated by a normal distribution and that is what I am showing here. So, if you look at that blue curve or the blue shaped curve distribution, that is basically corresponding to a smaller sample size and the lemon green or the light green colored curve is showing you the sampling distribution for sample mean for a larger sample size.

And as I am dealing with a larger sample size, these theorems; law of large numbers and the central limit theorem are both in place. And hence, we see that we are going to get more or less like symmetric bell shaped curve, which can be useful for normal approximation and I get the mean of the sampling distribution at mu of x bar, which is very close to the unknown population parameter mu.

So, we now look at 2 other interesting properties of the sampling distribution. I have already spoken about it for the mean, but let me complete it. So, if we are interested in the properties of a statistical distribution, the 2 quantities are of importance and they are mean and variance or standard deviation, if you take the square root of variance.

So here, for the sampling distribution, we can say that mu of x bar is equal to mu. So, the mean of the sampling distribution is very close and in limit it is exactly equal to the true value of the population parameter. And the standard deviation of the sample mean sigma x bar is given by sigma over root n. Now, the second formula of the standard deviation leads me to the property of good estimator and that is consistency. Hopefully, you have understood that that mu of x bar equal to mu that talks about the unbiasedness property because that is very simple.

So now, let me talk about the consistency property for the sigma x bar formula. So, here you can very well see that as number of observations or sample size increases, then sigma of x bar actually falls. And that is basically the consistency property talks about. So, as you are dealing with large and large sample, then the concentration of the mass for the sampling distribution actually comes closer towards the true value of the unknown population parameter. So, the variance actually declines and that is what consistency means.

Now, I am going to explain a little bit more about these concepts of consistency and bias with help of 3 concentric circles diagram. Now, you focus on the bottom part of the slide. Here, I am going to show you 3 concentric circles to explain certain things. So, let us look at the first set of concentric circles, which are at the right hand side of the slide, when we draw a sample and suppose, we draw 3, 4 more samples then from every sample I get a value of the sample statistics, a sample mean.

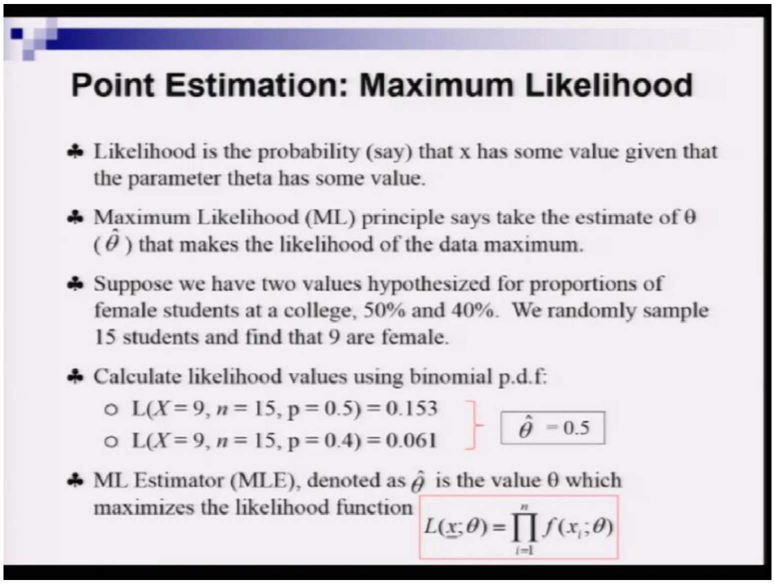
And then this gives me some approximate proxy for the true value of the population parameter. But, if they are very different from the true value of the population parameter, then there is a bias. And note that, all these 4 red dots actually are giving me the biased estimates for the population parameter.

However, note that, it is not only the case of bias in this diagram, we can also see the issue of inconsistency or imprecision. Why? Because what we see are these 4 dots are actually lying in different locations of this bull’s eye or the concentric circles diagram and they are way apart. So it is like very random shot by an inexperienced archer or shooter, and that is why they are pretty much off the target.

Now, the second or in the middle diagram, I am going to explain you the concept of a consistent but a biased estimator. So here you see, you see again, 3 dots, and these 2 dots are in blue. And you see that although they are off from the target, they are not very close to the target. But at least these 3 shots at the bull’s eye are coming very close of each other. So, you can say that they are consistent, but they are biased.

Now finally, we are going to look at the best possible scenario and here in the concentric circles diagram, in the left hand side of the slide, I am going to show you again 3 bullet points or 3 green circles and they are actually hitting the bull’s eye of this concentric circles. And not only they are hitting the bull’s eye, but they are also pretty close to each other. So you can say that the estimator that I am talking about in this particular case is consistent and is unbiased also, because the difference between the sample statistic value and the population parameter value, true unknown population parameter value is minimal.

(Refer Slide Time: 16:42)



Now, we are going to discuss the case of point estimation in greater detail. So, for point estimation, we have several methods available like method of moments, method of maximum likelihood and the method of least squares. But we are going to look at the method of maximum likelihood in this lecture and we are going to look at the least squares method in the regression context in the later part of the course.

So what is likelihood? Likelihood is the probability say that random variable x has some value, given that the parameter theta also has some value. Now, maximum likelihood sometimes is abbreviated as capital ML. So it is a principle, which says that take the estimate of theta, which is theta hat that makes the likelihood of the data maximum.

And this definition it is looking very bookish, it is going to be much clear to you, if I give you this particular example. So this is a very simplified example, but I hope that this will make the concept of maximum likelihood a bit more clear. Suppose we have 2 hypothesized value for the proportion of female students at a college and they are 50 percent and 40 percent. Now we want to know what is the true value of the population parameter? Is it 50 percent or is it 40 percent?

So from the population of the college students, we randomly draw a sample of 15 students and we find that in that sample 9 are female. Now, from the sample observation, can I make any inference that which of our hypothesized values of proportion of female students is correct? Yes, we can. So note that this is basically the case of a binomial distribution, we can actually apply the binomial distribution here. So if we now plug the values or parameters in the binomial PDF, we can get the probability values.

So let us assume that the true population parameter value is 50 percent. So the p the probability of success in the binomial distribution case is 0.5. So if that is the case, then how can I compute the probability? Because likelihood is basically nothing but the probability. So if we are looking for likelihood, we have to calculate the binomial probability.

So here, I am showing you the first case where capital L is basically the likelihood, it is nothing but the other way to express probability. So X is capital 9 that is the value of the binomial random variable and small n is basically one parameter that is 15. And small p is also a parameter and we have hypothetically assumed its value to be 0.5. So it gives me a probability figure of 0.153.

Now, I move to the other case, where the values of the X and n are same, but the value of p has changed from 0.5 to 0.4 because that is the second hypothesized value for the true population parameter. And if we assume that this is indeed the case, then that leads to a probability number 0.061. Now, note that, out of these two probabilities, definitely first one is higher. So we can draw an inference that our assumption of population parameter being 0.5 actually has higher probability from the data.

So we can make this decision that will my data or my sample that I have, it actually has lend more support to the assumption or hypotheses that population, true population parameter value is indeed 0.5. So we then infer that, okay, out of these 2 choices, for population parameter values, 0.5, and 0.4, I will go with 0.5. So that is basically the simple example.

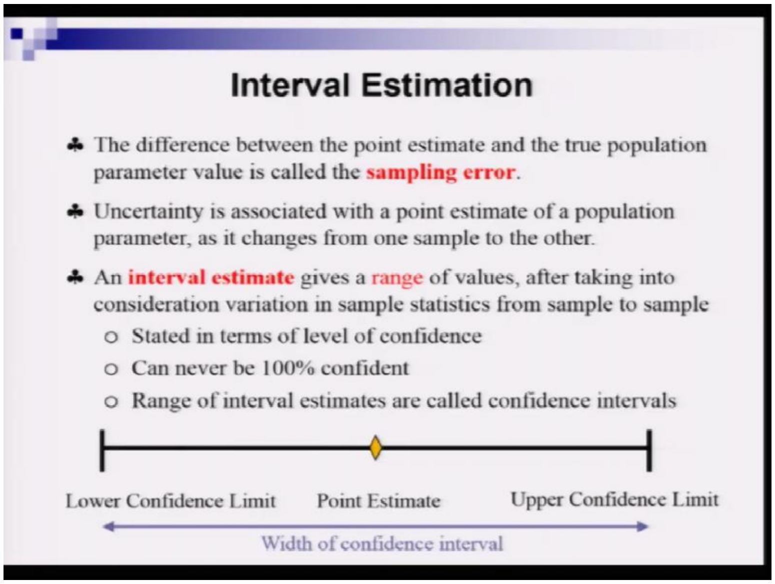
Now, we want to actually give you the mathematical or statistical expression for the maximum likelihood estimator, because finally, maximum likelihood is a principle. But it leads to an estimator, which is used to get the estimate for an unknown population parameter. So we have to now see how maximum likelihood estimator actually works. So MLE, is denoted by a theta hat and that is the value of theta which maximizes the likelihood function. So how do I write the likelihood function?

Note that that is basically a function of the unknown population parameters where the data is given. So likelihood function is denoted here by capital L of x given theta. And that is basically a product of n number of probabilities. Why? Because, if we have a sample of size n, then each of these observations in the sample has a probability of being chosen from a particular population. And then, these probabilities are independent, because I assume that they are drawn in a simple random sample way.

So we need to multiply all these individual probabilities. And that is how we get the maximum likelihood function. Now, the likelihood function, we need to maximize that one and it is a formidable shape, because it is product of many probability numbers. So it can be extremely complicated functional form. So sometimes, to make it simple, we have to take log on both sides and then make it linear, and then we can actually maximize this function.

But anyway, that is not our focal point. So, we actually stop right here on the maximum likelihood estimator case. Later on, in the course, when we are going to cover the case of discrete choice models, there I will show you how this MLE is very useful, and you will see that how MLE estimators are found by maximizing the likelihood function or log likelihood function.

(Refer Slide Time: 23:24)



So, next, we move on to the case of interval estimation. We have already defined what is interval estimation. Now, let me focus on the difference between the interval estimation and the point estimation. So, note that, when we get a particular sample and from the sample, we calculate a sample statistic value that works as an estimate for the unknown population parameter value.

So basically, the sample statistic formula is basically the estimator formula and the actual value that we get from that estimator formula, actually is the estimate for the unknown population parameter say theta. So, there is a difference between the point estimate and the true population parameter and that is called the sampling error. So now, let me explain what is sampling error in detail.

So when you get one particular sample and you get one sample statistic value, then that sample statistic value will change from one sample to the other. If you get to 2, 3 or 10 more samples, you are bound to see different numbers for the sample statistic value. And of course, the difference between this true value to population parameter theta and the sample statistic theta hat that you are obtaining from different samples are going to be different and this is called the sampling error, because this actually depends on the variations in sample.

So, of course, there is uncertainty associated with the point estimate, because we cannot be sure that if we get one particular value of sample statistic as an estimate for the unknown population parameter, whether that is indeed the true parameter value or not, whether it is very close to the population parameter or not, we have no clue. So, it is very uncertain to use or make use of a point estimate.

So, that is why statisticians prefer to work with the interval estimates where you actually take care of this sampling variation into account and then you try to get more information from the sample that you have in your hand. And that is why interval estimation, we are going to focus more in this lecture and we are going to study it in more detail.

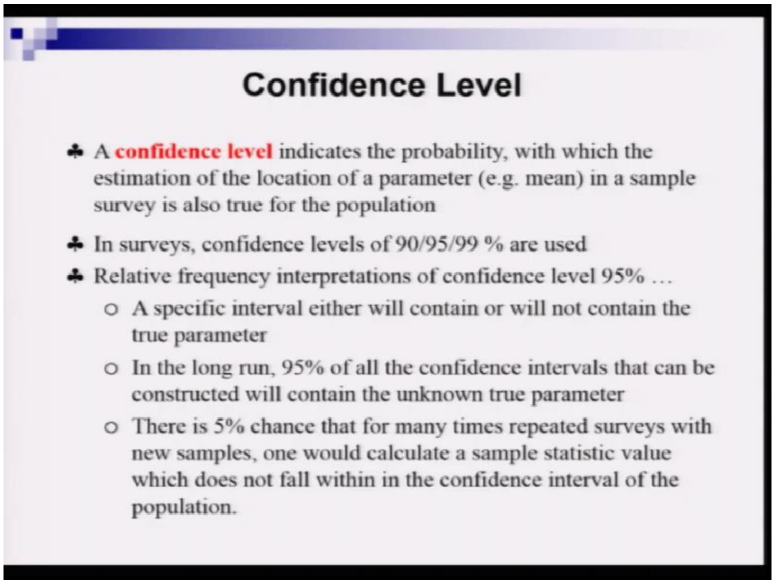
So, an interval estimate gives a range of values after taking into consideration of the variation in sample statistics from sample to sample. So, there are 3 interesting features that I would like to mention about interval estimate. So, these interval estimates are stated in terms of level of confidence, and then what is level of confidence that we are going to define and describe and explain in detail in the next slide.

And as there is probability associated, we cannot be 100 percent sure, we cannot be 100 percent confident about what we get. So, that is the second feature of interval estimates. And the third feature is that range of interval estimates are called confidence intervals. So, the diagram that I am showing you at the bottom of the slide will probably make things a bit clear. So, this diamond sign, the orange colored diamond sign in the middle of this thick black line, that is basically my point estimate.

Now, if I am going by point estimation method, then this is one single value that I have. But this number is bound to change from one sample to the other, as I keep on drawing samples from the same population of same size. So, lower confidence limit and upper confidence limit values are basically defining the range. And we are saying that well, within this range any particular value can happen and there is a probability associated with that. And this gap between this upper confidence limit value and the lower confidence limit value, the difference between these two values is called the width of the confidence interval.

So, if you remember, when we were discussing in the last class about sample size determination formula, I talked about width and I had sign w. And then, I said that, it is basically twice the margin of error, so I am talking about that w here in this diagram. Now, if you divided this by 2, then actually you get the difference between the lower confidence limit and the diamond in the middle, which is a point estimate, that is basically your margin of error.

(Refer Slide Time: 28:23)

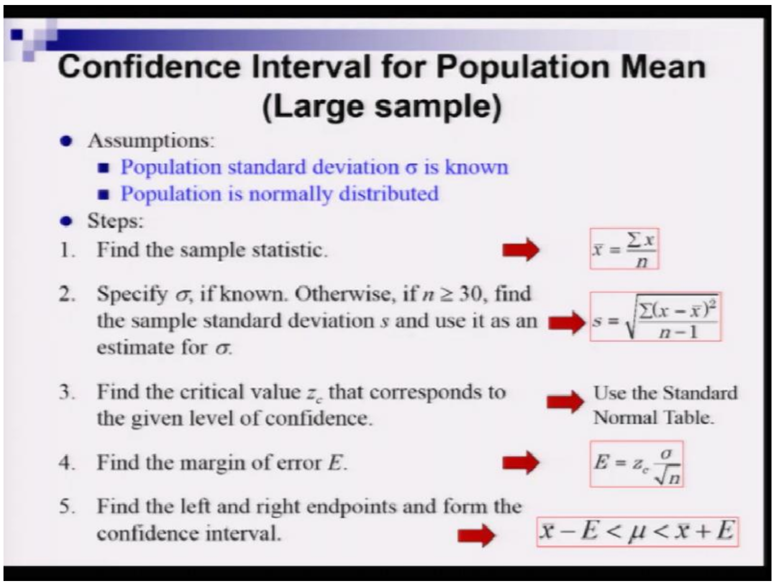


So it is time to define confidence level formally. So, a confidence level indicates that probability with which the estimation of the location of a parameter, may be mean, in a sample survey is also true for the population. Well, so to explain the idea of confidence level, we assume the value of 95 percent because in the applied statistics work and applied econometric work, you will find that more people are using this case of 95 percent confidence level. So, we have picked 95 percent to explain.

So, the first point about the features of this confidence level is that that a specific interval either will contain or will not contain the true parameter. So I am saying that when you have drawn samples repeatedly of same size from the same population, then you construct the interval for each of your samples. Now, a specific interval may not contain the true parameter and while the others are containing the true parameter. So this feature is going to be clear in 1 of the later slides where I am going to show you a diagram.

The second feature, which is interesting, that tells you that in the long run 95 percent of all the confidence intervals that can be constructed will contain the unknown true parameter. So, we can derive another conclusion from the second bullet point or interpretation and that is the last one here. So, there is 5 percent chance that for many times repeated surveys with new samples, one could calculate a sample statistic value, which does not fall within the confidence interval of the population.

(Refer Slide Time: 30:36)



So next, we are going to talk about how to generate the confidence interval for the population mean in the large sample case. So it implies that I am considering a case where n is greater than 30. So I will start with 2 assumptions. And first assumption says that population standard deviation sigma is known, although this is a very hard assumption or strict assumption to make, because in reality, many, many times sigma is unknown to us.

And the second one is a fairly straightforward assumption that we assume that population is normally distributed because we are dealing with a large sample. So law of large number tells us that normality can be assumed. So now I am going to show you 5 steps, following which you can generate the confidence interval for the population mean. So let us start with calculating the sample statistic value for the location parameter, which is the sample, from sample we get to know the sample mean.

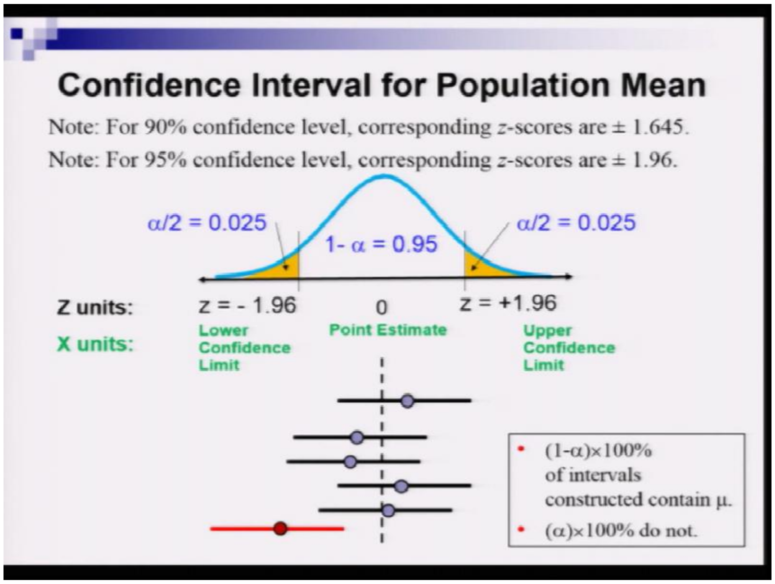
So that is x bar equal to summation x divided by n, that is quite straightforward. Then if sigma is known, then it is fine. But if it is non-known, if it is not known, then we assume that it is a large sample, so n is greater than 30. So we find the sample standard deviation s and we can use it as an estimate for sigma. So I am showing you the formula for s, the sample standard deviation. So note that, we are dividing by the number n minus 1 for the degrees of freedom issue that we discussed earlier.

Now, next step is to find the critical value from the standard normal table and that critical value is denoted by a zc. So that should correspond to a given level of confidence. So we have to assume a priori what level of confidence we are happy with. So if we are happy with the 90 percent confidence level, then you assume that or if you say, no, I am happy with 95 percent, you can go with that as well. But it is important to make that assumption in step 3, that what is the level of confidence that you are aiming for.

So after that, you find your corresponding critical value. Then you are in, step 4, where you have to find the margin of error. We talked about margin of error, hopefully, you remember that. So now I am showing you the formula how to compute it. So margin of error is given by capital E, and that is basically zc times sigma over root n. And once you get a measure for your margin of error, you need to subtract that number from the sample mean, and then you need to add that number to the sample mean.

So, the first one will give you the lower confidence limit and the second one will give you the upper confidence limit. So now you can say that I have generated confidence interval which will contain the unknown value or population parameter mu. And if you have assumed 95 percent confidence level, then you can say I am 95 percent confident that my unknown population parameter value mu will lie between x bar minus e and x bar plus e. So, this thing is again going to be talked about in terms of a diagram in the next slide.

(Refer Slide Time: 34:29)



So, now, I am going to show you a diagrammatic representation of what we have discussed. This diagram is also going to help you to understand some features of confidence level, level that we have discussed. So before we start talking about the diagram, I would like to start with 2 notes. It is important to remember two very vital z scores. Now, I have shown you the standard normal table, I have also taught you how to refer to the standard normal table and pick the z values whenever it is required.

But sometimes when you are in a hurry, you may not have time to consult the table, you may not have the table in hand. So you have to remember these two magic numbers, these are very useful magic numbers for the purpose of applied statistics and econometrics. So remember that 90 (percent), if you assume confidence level of 90 percent, then the corresponding z scores are plus minus 1.645. And if you assume the 95 percent confidence level, then the corresponding z scores are plus minus 1.96. So, if you remember these 2 numbers, you do not have to even consult the statistical table.

Now, let us see, what do we have in the diagram? So the light blue color bell shaped curve is the standard normal curve and along the horizontal axis, I am showing you first the z units. And as you know that, there is mapping between z axis and the x axis because of the standard normal transformation concept. So of course, for each and every value of z, that I am showing you here, you can actually find the corresponding x values. So you can also say that the x units can also be derived and they are shown in green color here.

So now, note that, a standard normal curve will have a mean at 0 and that is basically at the center of the distribution. And I am telling you that I am assuming here the 95 percent confidence level, because that is the most common in applied statistics and applied econometrics work.

So the corresponding critical values are plus and minus 1.96. So on the z axis, I am showing you these two values plus 1.96 and minus 1.96. So the minus 1.96 actually, if I transform that number into x units, then I actually get a lower confidence limit. And if I transform this 1.96 z value into x units, then I get the upper confidence limits. So if I now, transform the z value equal to 0 into x, then I get the value of the point estimate.

So now, note that, I am going to talk about the area under the normal curve and they give certain probabilities. So let us start with the mass which is there in the middle of the curve. So this area in white, which is below the bell shaped normal curve and bounded by two vertical lines placed on minus 1.96 and plus 1.96, that is basically giving me the value of 1 minus alpha and that is 0.95. So that is basically my 95 percent confidence level.

So it will be bounded by the lower confidence limit, which is x bar minus z times sigma divided by root n and x bar plus z times sigma divided by root n. So of course, the residual area is 0.05 and you see there are 2 equal parts for the residual probability and alpha divided by 2 gives you 0.025. So it is symmetric curve that is why you see the equal mass at two different tails of the distribution.

And we will again come back to this diagram, when we will be starting hypothesis testing, because this is very important diagram and it is very useful. So we have to get this diagram in our head and in our mind, so that even we know, if we fold our eyes we can visualize this and can apply this two; the confidence interval and hypothesis testing problems.

So, now, the explanation for the normal bell shaped curve and the confidence interval is over. But let me talk about something more in terms of the diagram that will help you to understand about the confidence interval and confidence limit in a better manner. So, we can generalize this diagram and we can have 2 bullet points to summarize the main results in generalized form. And, these are two red bullets at the corner of the slide. So, first one tells you 1 minus alpha multiplied by 100 percent of intervals constructed contain mu and alpha times 100 percent do not.

So what do I mean by these summary points? So these I would try to explain in terms of the bottom part of the diagram. So here, you see corresponding to that point estimate in X unit or 0 value in terms of the Z units, I have a broken vertical line. And now, I keep on drawings samples from the same population, but the samples are of same size small n. Now, of course, as there is sampling variation, the sample statistic value, which is basically the point estimate obtained from the individual sample is not going to be the same, it is going to vary from sample to sample and that is what is shown in the parallel straight lines that I have shown here.

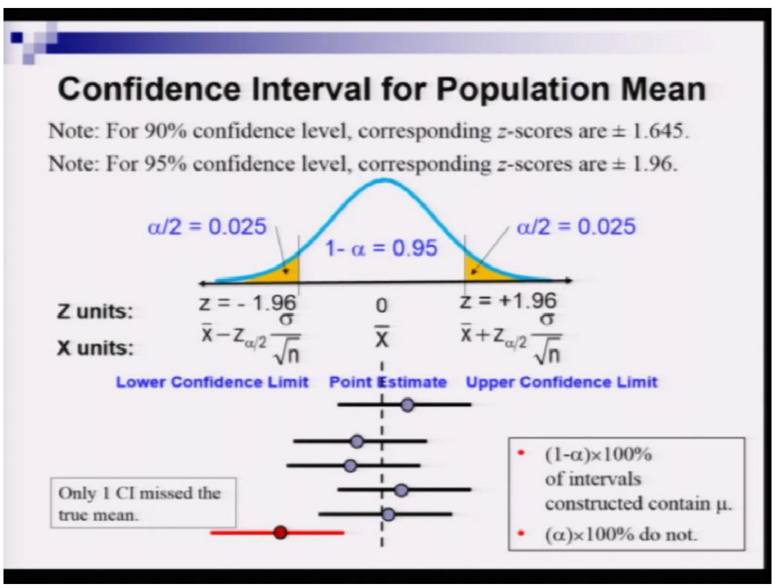
So, all these parallel straight lines are basically giving you the confidence interval that you obtain from one particular sample, given that there is a sample mean X bar. So if we look at the first line, you see that circle there of purple color in the middle, so that is basically X bar 1. So that is basically the sample mean that you calculate from your sample 1.

So, the second line you come down similarly, you see that purple color circle on that line. So, that line gives you the confidence interval that you can generate from the obtained data in the sample number 2, and if you calculate the sample statistic value X bar 2 say, so, this purple circle actually indicates its value. So, you note that, there is a difference between X bar 1 and X bar 2. So, now I keep on drawing samples again and again, again and again from the population, but all of them are of same size.

So, now, you see different sample means are generated X 1 bar, X 2 bar, X 3 bar, X 4 bar, X 5 bar, now you come to X 6 bar. So, you see that x 6 bar actually is far away from the point estimate or the Z value equal to 0, and this is kind of not falling in the interval. And we can say that well, this particular confidence interval that we generated from sample number 6 may not contain the true population parameter value, because it is in very much away from the center of the distribution. So, next, we are going to talk about the other cases of confidence interval generation.

So, we are done with the discussion on how to find confidence interval for the population mean in the case of large sample. But, the diagram that I have shown you in the last slide, the entire slide, actually can be made a bit better, so, that it is even more informative for you and it will be easier for you to consult as a note. So, I have made certain changes to the previous slide that I have shown and let us have a look at it and hopefully, we like it better, because it contains more information after the changes I made.

(Refer Slide Time: 43:18)



So, here, in the slide, you can see that I have replaced what was there as X unit; and here, I have replaced the lower confidence limit and the upper confidence limits, those were in green by the statistical symbol. So actually this formula, actually gives you the value of the lower confidence limit in this particular case and that is why it is very easy for you to remember because now, you can associate the formula with the diagram itself.

Similarly, for the upper confidence limit case, I am again showing you the formula X bar plus Z of alpha divided by 2 times sigma over root n. So, this particular formula gives you the value of the upper confidence limit. So, now, I think with this revised diagram, it will be even more clear to you how to find the confidence interval. I will end the discussion by saying that, please do not be under the impression that when we talk about this 95 percent confidence interval or 90 percent confidence interval, we are actually trying to have any probability interpretation from there.

We know the probability concept is associated with confidence interval finding, but that actually is embedded in the Z value or the critical value that we have to consult statistical table and then find the value. But this 90 percent and 95 percent confidence level that I spoke about, there actually it is better if you look at this number only in terms of percentage. So, we are done with today's lecture, we could not cover the case of small sample case in today's lecture. So, in the next lecture we will begin the discussion by looking at the small sample case. Thank you.