Multivariate Procedures with R

Prof. Shalabh

Department of Mathematics and Statistics

IIT Kanpur

Week – 06

Lecture – 27

Point and Interval Estimation

Hello friend, welcome to the course Multivariate Procedures with R. So, after completing our discussion on different types of probability distributions, now we are entering into another area that is about statistical inference and this is a very important component of statistics. If you try to recall, what do we expect in statistics or what do we expect from statistical tool, we are trying to observe a small sample but we want to have the conclusion about the whole population. For example, a sample of medicine is given to a small number of patients and in case if the medicine works well then it is given to everybody in the world. And in case if the medicine is not working on that a smaller number of people then it is concluded that the medicine is not good and it is not used and it is not given to any human being. So, if you try to understand what are we trying to do, we are just observing a small sample and based on that we are trying to take the conclusion for the whole population.

Now the question is this, how can you validate that whatever observations you have taken on the basis of a small sample or whatever conclusions you have taken on the basis of a small sample or whatever statistical inferences you have taken on the basis of a small sample they will be valid for the entire population. So, this is the broad area which we are going to discuss in the next couple of lectures. In the lecture today we are going to talk about the estimation of parameters. What is this? Suppose if I ask you that how much time do you take in going from your home to your college? Then you have two options.

You will say okay it takes 20 minutes or it takes between 15 and 25 minutes. Now if you try to understand this whole phenomena that how you have reached to these values 20 or between 15 and 25. Inside your mind you have some data from the past. Every day you go to your college and you have some idea that every day how much time I am going to take and you have simply taken the arithmetic mean of those values and based on that

you have given me two answers. One answer which is value at a point that is 20 minutes and another answer is, it is in the form of two intervals 15 minutes and 25 minutes.

Now there are many questions that how do you got these values? How you have got the value at a point? How you have got the value in the form of an interval? And how can you convince others that these values are good which you have obtained on the basis of arithmetic mean? Why not geometric mean? Why not harmonic mean? Why not median or something else? Why not you have used the sample variance to compute this answer? So all these things they come under the purview of estimation of parameters. So, as we have discussed that there are two types of estimations when obtaining the value at a point which is called point estimation and when we are trying to obtain the value in the form of an interval then it is called as interval estimation. Well this statistical inference this is actually a whole semester course in when you try to do MSc statistics program but anyway my objective here is to give you that much information which is required to understand the lectures in the further course. So, I will be here brief, short and I will be using here the minimum mathematics because I want to bring all my participant at the same platform somebody may have done MSc status somebody may not have done it. So my objective is that to bring you on the same platform and to inform you that how much knowledge do I need and if needed you can do yourself, you can read yourself, you can study yourself from the books.

So, with this objective let us begin this lecture and try to understand this concept. In this lecture most of the things are not going to be mathematical it is basically for you to understand this concept. So let us begin our lecture. So in this lecture we are going to talk about point estimation and interval estimation. So first question comes here what is the need for drawing the statistical inferences.

So we know that statistics deals with drawing conclusion from the observed data on the basis of a given sample and whatever conclusions you have drawn on the basis of a sample they have to be remain valid for the entire population how to ensure this thing. So how to ensure that the value based on a smaller sample will be will remain valid over a big population. So for that the first question comes over here that you have to take a sample suitably from the population and then based on those observation you have to use a proper statistical tool which can give you the correct answer. So the first thing is this that we try to choose a suitable sampling procedure that means you want to collect some observation from the given population and then we try to analyze the sample item that mean the observations and then we expect that we will be able to draw some conclusion about the population as a whole. And we assume that whatever is my sample that is giving me the relevant information.

So how to get it done? So whatever is the whatever be the characteristic which we want to study in a given population that is represented by a random variable. For example in the population of some human being there can be different characteristics like their height, weight, age etc. So whatever we want to study we try to represent them by a random variable and we try to characterize its property by a related probability density or say probability mass function depending on whether the random variable is assumed to be continuous or discrete. And based on that whatever be the suitable statistical tool for this type of random variables we employ them over this sample to draw the statistical inferences. And for this some statistical assumptions are required for collecting the observation as well as the statistical tool and that is the job of the statistician or the mathematical statistics so that some proper assumptions are made so that we get a correct answer.

Now the next question here is what is the sample? So in case if I say that if I have random variable X1, X2,… Xn which are independent and identically, this is briefly called as iid that is a popular name that means independent and identically. Independent means all these X1, X2,… Xn they are independent and when I say identically that means each of this Xi is coming from the same probability distribution. So under such a case these random variables constitute a random sample of size n from the common distribution capital. And this distribution f represent the distribution of random variables in the population not in the sample. So that is why a sample is drawn from this population and it is expected that the sample observation will also possesses the same characteristic which are present in the population f and that is why we always assume without even writing that the sample is always representative.

From the sample is representative that means all the properties of the population they are also available inside the sample. And then the question is once you have got the sample then what you have to do? Then there is a need for statistical tool for drawing the statistical inferences and for that various statistical method are used to make decisions and to draw conclusions about the population on the basis of the given sample of data and this aspect of statistics is generally called as statistical inference. And there are several methodologies, techniques which are utilized on the basis of the given sample of information to draw the valid statistical conclusion. And when you are trying to draw sample which is random and once you are handling the real set of data there is going to be some inherent randomness. So, the advantage of using the statistical framework is that that the framework of statistical inference allow us to infer from the sample data about the population of interest at a pre-specified uncertainty level and knowledge about the random process which is generating the data.

So, that means you are trying to first make decision that what is the level of uncertainty in your statistical inferences which you are trying to draw from the given sample for the whole population. For example, we consider the problem of estimation of parameters in a given context that some new medicine has been devised, has been developed and we want to test that how effective is the medicine in controlling the fever of the people. So, now what do we do? This medicine is administered to a group of patients and its effect is recorded by measuring the time of control of fever. That means medicine is giving and it is observed that how much time it can control the body temperature. Now, you will see that once the medicine is giving to different sets of people they are not going to report the exactly the same time but there is going to be some difference among the values.

Why this difference is coming? So, the effect is essentially coming which is a very natural variation and it is coming because of the difference in the characteristics of the patient like age, body structure, body weight etcetera and we have no control over it. We cannot say that we want patients only of the exactly the same age, same body structure, same body weight etcetera. So, now this difference is coming in the values due to different reasons. But anyway we want to make a statistical conclusion such that such random variations are as minimum as possible and we want to estimate the mean time to control the body temperature based on a sample of data to compute a number that in some sense is a reasonable value of the true population mean. That means yes if the medicine is good that will have some value by which it can control the value for the time in which the medicine can control the body temperature, but that value is not known to us because this is the population value.

So, what are we going to do? We are going to get this data and based on that we have to use some statistical tools so that we can get a single value. Why single value? If you say ok this person the medicine can control the body temperature in this person for 5 hours, in this person for 6 hours, in this person for 8 hours etcetera, then it is very difficult to understand this type of language and to make a valid conclusion which is useful for the patients. For example, on the other hand if I can say ok on an average this medicine is going to control the body temperature for say 7 hours, then it is easily understandable and it can be used for various jobs, various purposes. So, this type of number is called as point estimate. So, we want to have procedures for developing the point estimates of parameter that have good statistical properties.

Good statistical properties means they can give us the good and reliable values. So, the primary goal in statistical inference is to find a good estimate of the population parameters. It can be the say for mean, for variance, for skewness, for kurtosis or any

other function of the parameters. And these parameters are associated with the probability distribution which is believe to characterize the population. For example, if I say that the parent population is normal mu sigma square, then the values of mu and sigma square are the parameters which are giving us the entire information about the population.

So, in case if these parameters are known, then one can characterize the entire population. But in practice the trouble is that these parameters are not known to us, they are they are unknown. So, the objective is that how to know them or how to estimate them right. So, one can estimate them or find out the value of the parameter on the basis of the given sample of data as a function of sample values. For example, if I want to find out the mean, then for example, this arithmetic mean of the sample observation is a function of the sample values.

And the values of these parameters can be obtained at a point as well as in the form of an interval. For example, in the beginning of the lecture, I told you that if you want to measure that how much time are you going to take in going from your home to your college right. So, that can be 20 minutes or between 15 and 25 minutes. For this 20 minutes is a point estimate whereas, 15 and 20 they are the bounds of the interval estimate. So, when the values of parameters are obtained at a point, the estimation procedure is called as point estimation.

And when the values of parameters are obtained in the form in an interval, the estimation procedure is called as interval estimation right. So, now we are not going to deal with the point estimation and interval estimation in detail, but definitely I would like to give you that much of information which is enough for you to understand the way we are going to develop various types of tools in the further lectures right. So, whenever you are trying to get these values, either point estimate or this interval estimate, there have to be some criteria that how you can say that whether they are good or bad. For example, if I say a student is said to be good, if the student is attending all the classes, they are doing the assignment, getting good marks in the examination etcetera. Similarly, somebody has different types of characteristics and based on that we always try to decide whether the person is good or bad.

Similar is the story with the statistical inference also. We have different types of criteria on which we try to judge whether the values which we have obtained on the basis of a sample for the entire population are they good or bad right. So, we assume that suppose this X1, X2,… Xn all in lower case alphabets, suppose we denote them by here X, they are the observations of a random sample from a population of interest. For example, if

you want to see that whether the medicine is effective in controlling the body temperature, then you are going to collect some people who are going to have, who are already having fever or the higher body temperature and then you will try to give the medicine to them and you will try to see that how much time the medicine can control the body temperature to normal. So, in case if you try to give it to the first patient, the recording will be indicated by small x1.

If you try to give it to the second patient, then the recording will be obtained and indicated by here is small x2 etcetera right. So, this random sample represent the realized value of a random variable capital X and it can be said that this X1, X2,… Xn are the n observation collected on the random variable X. So, you will see that whenever we are trying to develop any procedure, we always said that ok let X1, X2,… Xn be a random sample obtained from some population on the random variable X right. And then we try to consider a statistics T x which is used to estimate the population parameter suppose theta. The theta can be a scalar value or theta can be a vector value right.

You can recall that we had discussed the definition of a statistic which is a function of random variable. And in such a cases, in cases we always say that this T x what we have obtained on the basis of given sample of data, it is an estimator of the theta. And in order to indicate that we are trying to estimate the the parameter theta using T x, we indicated like this that we will write the theta and we will put here a cap or a hat and we write the theta hat is equal to T x. So, now if you try to see this capital X is a random variable. So, this T x is also a random variable.

So, this theta hat will also a random variable and so it may also and it will also have the probability distribution right just like a any other random variable right. And when we are trying to calculate the value of the T on the basis of given sample of data, then we try to indicate it by here T x where now this x here is say lower case like this one T x right. So, you can see here that the T x is a random variable when where x is capital X and whereas, this T x where x is lower case like this here, then it is an observed value which is dependent on the sample values which have been obtained. So, for example, in case if you try to write this arithmetic mean of capital X1 capital X2 … capital Xn by 1 upon n summation i goes from 1 to n xi, then this is an estimator and this is a function of random variables and so this is a statistics which is indicated by here T x. So, now you obtain the value of x1, x2,… xn some numerical values and then you try to find out the arithmetic mean of those values and it is indicated by here 1 upon n, i goes from 1 to n say lower case xi and it is indicated by here T x where now this x is lower case, these are the sample values.

$$T(x) = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Then we say that this T x with lower case x is the estimated value of capital T x which is obtained on the basis of sample values small x1, small x2, … small xn. So, this is how we try to interpret and in a common language we always say this T x is an estimator of mean. Now in case if you try to recall whenever we are trying to mean with the real data there is always randomness and because of which whenever you will try to draw a sample, the sample values are going to be different. For example, in case if I say suppose you take a number from here 1 to 100 and suppose you try to draw here 10 numbers randomly right. Do you think that every time you will get the same number? No.

So, similarly the sample values which are actually the realization of a random variable they also lead to different values in different sample and each sample leads to a different value of the estimate of the population parameter. For example, if I say suppose there are suppose 20 students in the class and if you want to choose 5 students and suppose all are identical and then they can be chosen with the equal probability then there are 20 c 5 ways or 20 choose 5 ways in which the 5 students can be chosen. And yeah if you try to find out their average height or average weight they are going to vary from one sample to another sample right. So, this is what I mean that when I say that that means, sample leads to a different value of the estimate of the population parameter. Now we assume that this population parameter is fixed value at least as of now we are assuming it to be a fixed, but I can inform you that the parameter can also be random, but definitely unless and until I inform you specifically we are going to assume that our parameters are fixed in the entire course right.

Now the next question comes once you have obtained the value of the parameter on the basis of given sample of data which you are saying that different samples will lead to different values how are you going to judge whether the values which you have obtained they are good or bad. So, in order to understand it we have certain properties of this point estimators well these are very detailed properties, but we are not going into that much detail, but I simply want to give this idea. So, that whenever I am using this terminology in future possibly you can understand what I am trying to say. So, in the case of point estimators a good estimator need to have some desirable statistical properties and they are unbiased estimator, efficient estimator, consistent estimator, sufficient estimator and completeness. Well, these properties are trying to indicate different characteristic of the estimator and they are not written in any order right and having one property does not entirely implies that that other properties are also there and absolutely there is no ordering the way I have written here.

I have to write them in certain order, but then there is no order you can write them in any order right. So, now let me try to give you a very quick idea about these properties. So, first we try to consider the unbiased estimator that is unbiased estimator right. So, an estimator should be close in some sense to the true value of the unknown parameter that is what we expect that when I say that how much time are you going to take in going from your home to your college then do you have telling me that it is going to take 20 minutes. Now suppose if I really start going from your home to your college what do you expect will it take exactly 20 minutes possibly there will be certain variations it may be 21 minutes, it may be 24 minutes, it may be 18 minutes, but then you think that if it is taking 18 minutes and you are telling me 20 minutes is it a good value? So, that is what I mean when I say that an estimator should be close in some sense to the true value of the unknown parameter because remember one thing the parameter the value of the parameter that we want to know it is completely unknown to us right.

So, a point estimator theta hat is to be an unbiased estimator of the population parameter theta if expected value of theta hat is equal to theta right. So, in case if the estimator is not unbiased then we can it is called as biased estimator of theta and then we try to compute its bias that the bias of theta hat is given by here expected value of theta hat minus theta and this is a finite sample property that even if you have a small sample this will hold true and we expect from this property that all the values of theta hat to be equally spread around the two sides of the theta like if theta is here we expect that all those values are equally spread on the left hand and right hand side of the value theta. For example, if you are telling me that you take 20 minutes to recall it on your home then it can be 18 minutes or it can be 22 minutes also it can be 19 minutes or it can be 21 minutes also right. Similarly, we have another property that is efficiency and this efficiency is trying to measure the variability of the estimators right. You can recall that we have considered different measures of the variability and we are going to use here the concept of variance right.

So, the variance of an estimator theta hat of the parameter theta is defined here as say expected value of theta hat minus expected value of theta hat whole square indicated by here VAR inside the parenthesis theta hat. So, suppose the parametric space of this theta is suppose capital theta. So, theta will belong to capital theta. So, suppose there are two unbiased estimator of theta, theta can be estimated by theta 1 hat as well as theta 2 hat and suppose both of them are unbiased. Now, how to choose that which estimator I have to use? Then we try to impose the criteria of efficiency and we try to compute the variance of both the estimator and then we try to choose the estimator which has got a lower variance.

$$Var(\hat{\theta}) = E[\hat{\theta} - E(\hat{\theta})]^2$$

So, out of these two parameter theta 1 hat and theta 2 hat, theta 1 hat is said to be more efficient than theta 2 hat under the criteria of variance for estimating theta when variance of theta 1 hat is less than or equal to variance with the 2 hat for all values of theta belonging to theta and variance of theta 1 hat is less than the variance of theta 2 hat for at least one value of theta belonging to capital theta. So, in simple words I would say that you are going to choose an estimator which has got the lower variance. Now, the third property of estimator is about the consistency. So, we always expect in a good estimator that as the sample size increases the values of the estimator should get closer to the parameter being estimated right. It is something like this suppose somebody asked you that ok well you have told me that you are going to take 20 minutes when you try to go from your home to your college, but now there are two people one person says that ok my this value is based only on two observation that I have gone to the college only twice and so I am giving you this value 20 and say another person says ok I am going to the college for the last one year every day and so I am giving you the value 20.

$Var(\hat{\theta}_1) \leq Var(\hat{\theta}_2)$ **for all** $\theta \in \Theta$.

$Var(\hat{\theta}_1) < Var(\hat{\theta}_2)$ **for at least one** $\theta \in \Theta$.

So, definitely whose values you are going to believe more. So, obviously we will try to believe on the value which is based on more number of observations. So, in some sense that we are trying to say that as the sample size is increasing the value of the estimator is getting closer to the true value that is unknown to us and this property of estimator is referred to as consistency and the estimator are called as consistent estimator and contrary to unbiasedness this is a large sample property whereas, unbiasedness is a small sample property right. So, this property of consistency inform us that as the sample size and increases the probability that theta hat n is getting closer to theta is approaching 1 right. So, that means in simple words you can say that the value of the theta hat n is getting closer to the parameter theta as the sample size is getting larger right.

So, for example, if I try to take a sample from the normal distribution say normal mu sigma square and we try to compute its arithmetic mean. Suppose if I say that the arithmetic mean based on here say sample size is small n it is given here as say 1 upon n summation i goes from here 1 to n say here xi right. So, in this case if I try to find out the expected value of x bar n this will come out to be here mu and variance of x bar n will come out to be like sigma square upon here n right. So, and in that case this x bar n will

follow a normal distribution with mu and variance sigma square upon n. So, you can see here that variance here is sigma square upon n which is going to 0 as n going to infinity that means the variance is becoming smaller and smaller as the sample size is becoming larger.

$$Var(\bar{X}_n) = \frac{\sigma^2}{n}.$$

And therefore, I can say that x bar n that our sample mean based on the sample of size n is a consistent estimator of here mu right. So, definitely I am not going into that much detail there are very strong mathematical definition about the consistency, but my idea here is to give you this idea that when I say that the estimator is consistent what you have to understand. Similarly, we have two more properties say sufficiency and completeness of an estimator right. I am not going to give you here more details, but I will give you this specific idea. The idea behind the concept of sufficiency is that we want to reduce the dimension of the estimator and it is a basically dimensionality restriction property.

And in simple words I can say that if estimator is sufficient then it contains all the information about the parameter which is contained in the sample. For example, if I say suppose you want to know about the population mean right. And suppose if I say here that the sample mean based on the values here x1, x2,… xn is the sufficient estimator of the population mean. What does this mean? If you try to see sample mean is a single value say here x bar n, where there this x1, x2,… xn it is a dimension here is n these are n values. So, now if you try to see whatever information is contained inside this x1, x2,… xn that is also available inside the x bar n.

So, instead of looking into this n value you are going to look only into one value and that is the idea about the sufficiency of an estimator right. Similarly, when we have more than one estimator of a parameter and if we want to restrict some more conditions so that I can see the unique estimator then we use the concept of completeness. So, the idea behind the concept of completeness is to get a unique estimator. But anyway, well I am not going into those details because they require some good mathematical skills which is beyond the scope of this course right ok. So, now after this we come to another aspect of interval estimation.

So, as I told you in the beginning of the lecture that we can estimate a parameter in the at a point which is called as point estimator and if I am trying to estimate the parameter in the form of an interval then it is called as interval estimator of the parameter. For example, you take 20 minutes to reach to your school is a point estimator when you say that I take 15 to 25 minutes then it is my interval estimator right. So, in order to yeah the

same example I have written here to understand the concept of this interval estimation that suppose there is a student who wants to know the time taken to travel from his home to the college and suppose the student make 30 trips and note down the time taken in every trip. So, that is 20 is going to have say x1 up to x30, 30 values. So, now to get an estimate of the expected time and can find out here the arithmetic mean suppose I come talk to be 30 minutes then this is here the point estimate of the expected travelling time and but the problem is that it may not be 100 percent correct because every time the student is travelling or the student travel after this unit of information it may not take exactly 30 minutes right.

So, that may in many situation it may not be appropriate to say that the student will always take exactly 30 minutes to reach the college right. And so there can be a variation of some seconds, some minutes in every value in every trip right. So, to incorporate this feature the time can be estimated in the form of an interval. So, if I try to make here a statement like that the time varies mostly between 25 to 35 minutes for it gives us more information it is more informative. So, the advantage here is that in the interval estimation you can say that as soon as I say that the time varies that means I have taken care of the mean as well as variation of the data.

And this interval that 25 minutes comma 35 minutes this provide a range in which most of the values of this travel time are expected to live and this is essentially the concept of interval estimation right. So, now I can comprehend all this information that an interval estimate of a population parameter is called as confidence interval. And the length of the interval reflect the uncertainty about the parameter say mu for example, right. And if we try to make the interval wider then the uncertainty also increases about the location of the parameter mu right. For example, if I say somebody says that ok you will take suppose 25 to 35 minutes and somebody says ok you will say take 10 to 100 minutes.

Then the width of the interval in that first case will be 35 minus 25 which is equal to 10 minutes whereas, in the case of this 10 to 100 the width of the interval is going to be 100 minus 10 which is equal to here 90 minutes right. So, if you make the interval wider then the uncertainty about the parameter also increases. So, the information about the precision of estimation is conveyed by the length of the interval and a short interval implies precise estimation right. Suppose if I say that if a parameter has been estimated by two different procedures then which of the estimator is better then we would like to compute its confidence interval and whose confidence interval is shorter that will be said to be more precise in terms of the width of the confidence interval because it takes care of the mean and variation both right. And but we cannot say with 100 percent guarantee that any interval will always contain the true value.

So, we cannot be certain that the interval contain the true, but unknown parameter, unknown population parameter. So, but the thing is this we have to minimize the uncertainty and that is measured by the level of confidence right. So, the confidence interval is constructed so that we have a high confidence that it does contain the unknown population parameter right. So, in this sense the location of the interval will give us some idea about where the true, but unknown parameter mu lies right. Well, I have taken here the parameter to be here mu because you will see in the further lectures usually I will be constructing the interval and I will be using the interval for the population mu.

So, and then to just to give you a clear information that in the point estimation I have used the symbol theta and now I am using the symbol mu. So, now the next question come I can give you a very quick idea that what do I really mean by whatever I have said right. So, that how are you going to construct the intervals. So, there is a there is a popular way by which I can define the interval estimates of a parameter. So, suppose I have a parameter here theta right and the confidence interval of a parameter theta is a random interval which has got two bounds one is lower bound and say another is upper bound.

Lower bound is indicated by theta hat l, l will mean lower and upper bound is indicated by theta hat u is upper and these two upper bounds are such that the unknown parameter theta is covered by a pre specified probability of at least 1 minus alpha. So, this I am trying to write in this particular way this is here theta, theta hat is lying between theta hat l and theta hat u and both theta hat l and theta hat u they are the function of random variable x right and this probability is greater than or equal to 1 minus alpha where this x which I am writing here is a x underscore this is a function of x1, x2, xn. And this value here 1 minus alpha this is the probability and this is also called as confidence level or confidence coefficient right. And this value theta hat l x this is the lower confidence bound or the lower confidence limit and theta hat u x this is the upper confidence bound or say upper confidence limits right. So, now in case if you want to construct such a confidence interval for any unknown parameter theta then it is here like this.

$$P_\theta\left[\hat{\theta}_L(\underline{X}) \leq \theta \leq \hat{\theta}_U(\underline{X})\right] \geq 1 - \alpha$$

where $\underline{X} = (X_1, X_2, \ldots, X_n)$.

Suppose this x1, x2,… xn be a random sample of an observation. Now after that we are going to find a statistic g which is a function of x1, x2,… xn and the parameter theta such that this g x1, x2,… xn theta depends on both the sample and theta. But the probability

distribution of the g does not depend on the parameter theta or any other unknown parameter. And this type of function g x1, x2,… xn is called as vital quantity. So, you will see that whenever we want to construct any such a vital quantity we try to create here a statistics whose distribution does not depend on the parameter theta. For example, I can give you here one example we have done is suppose xi follows normal mu sigma square.

So, now xi probability distribution is dependent on the two parameter mu and sigma. But if I try to take the distribution of xi minus mu upon sigma whole square i goes from here 1 to n then this will follow a chi square which is equal to n degrees of freedom right. And similarly for the t distribution also that is square root of n x bar minus mu upon s this follows a t distribution and that is also true for the f distribution also. So, you can see here that these statistics they are the function of x1, x2,… xn mu that is equal to theta, but their probability distribution is independent of mu. For example, chi square does not depend on mu, t does not depend on mu, but they depend only on the degrees of that is all right.

So, this is what I meant and now let us come to an end to this lecture right. So, now you can see here that in this lecture we have covered a lot right. So, the length of this lecture is much much shorter than the information which has been conveyed to you and as I said in the beginning that I am not my interest is not to to make you understand about the whole the entire theory of statistical inference, but I want to use the point estimation, I want to use the confidence interval estimation. And then later on whatever are my estimators based on that I want to conduct the test of hypothesis, I want to construct different types of tool for discrimination, principal component etcetera etcetera. So, in order to fulfill those objective for the multivariate procedure, we are trying to clear here that background. We are trying to understand here this basic the basic concept and my objective is that that when we are going to those tools then all my participant should have a same level of knowledge at least with respect to this course.

Now the question comes if you do not know about these topics my request is that you please try to open a statistics book and try to read about these concepts. I am not asking that you should go into that much depth, but if you go and if you try to understand the mathematically that is definitely going to be more useful. The advantage will be that we are going to understand here some basic standard procedures, but in practice if something else is happening some assumption is getting violated then in that case how are you going to modify your tool that answer will come only when you understand this basic concepts. So, while dealing with the topics of this multivariate procedures my objective is to learn from the basic concepts and once your basic concepts are clear then I am 100 percent confident I am fully sure that if you have to develop some more tools or you want to

understand those tools which are not being covered in this course you can understand them very easily and if you want to modify any of the tool which we are going to discuss in the lecture it would not be difficult for you. So, you try to have a look into different topics in different books try to understand them and I will see you in the next lecture till then goodbye. Thank you.