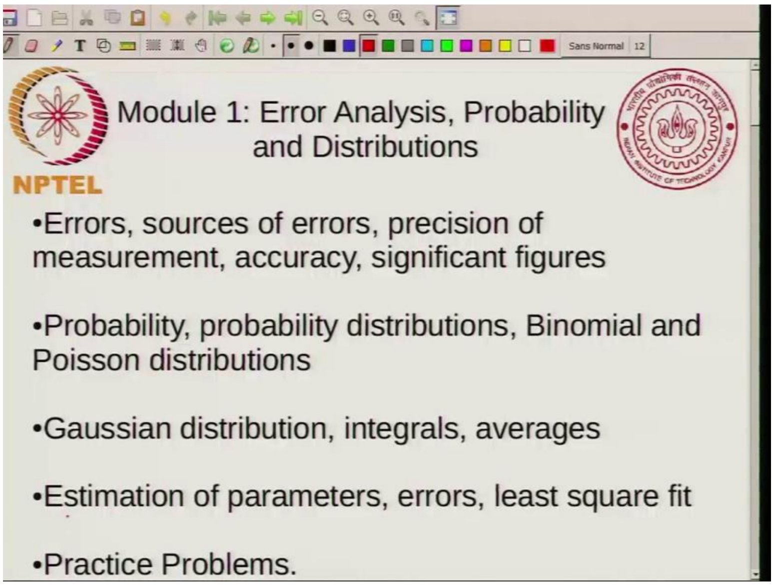
Mathematics for Chemistry
Prof. Madhav Ranganathan
Department of Chemistry
Indian Institute of Technology, Kanpur

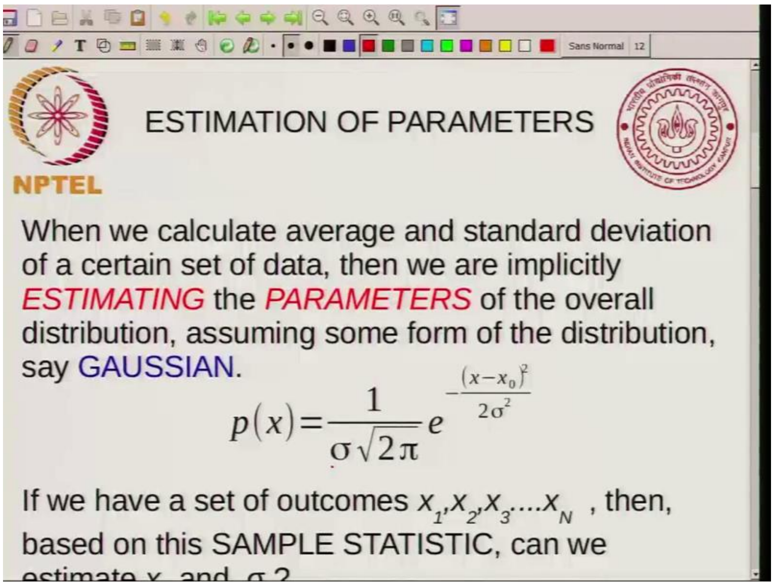
Module - 01
Lecture - 04
Estimation of parameters, errors, least square fit

(Refer Slide Time: 00:15)



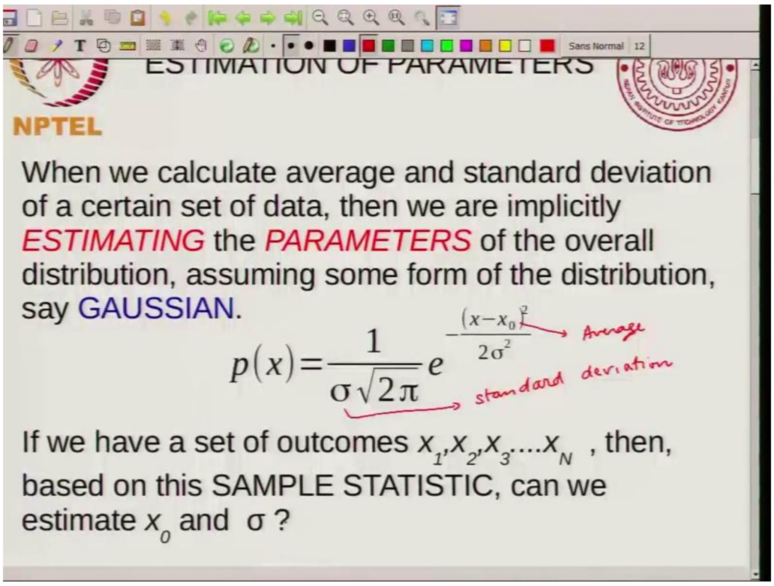
So, today's topic is going to be the topic of estimation of parameters, errors and least square fit and this will be the last topic in module one. The next lecture I will be doing practice problems.

(Refer Slide Time: 00:27)



So, what do we mean by estimation of parameters? I mean this is a very general terminology that is used in a lot of different context. Now, in fact, what I have said is that when we calculate things like average and standard deviation of a certain set of data. Then, this is something that we are implicitly doing we are estimating the parameters of the overall distribution. Assuming some form of distribution for example, you might be assuming a Gaussian distribution. So, the Gaussian, so, if you assuming that your data is distributed as a Gaussian then the Gaussian distribution have the parameters this is the average and this is the standard deviation.

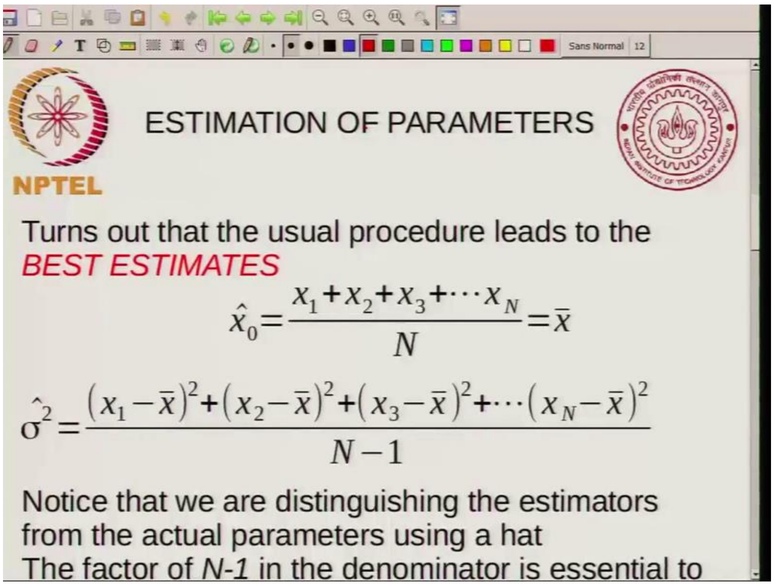
(Refer Slide Time: 01:11)



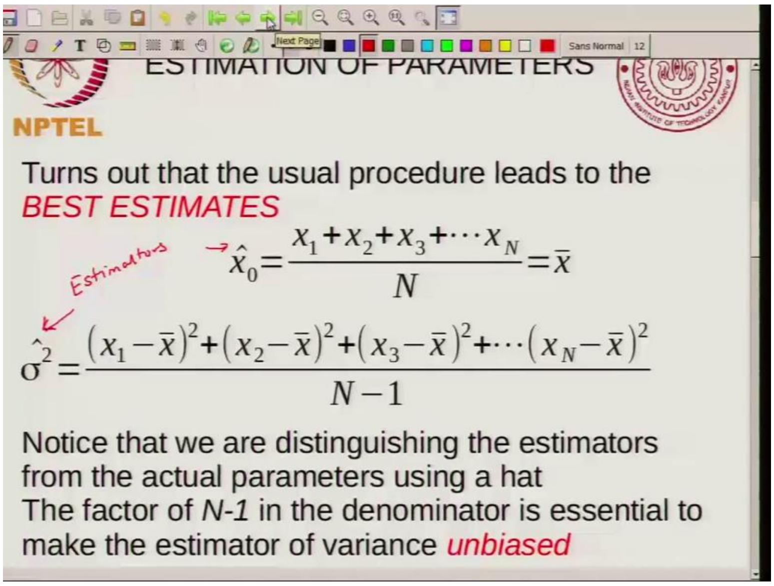
So, these are parameters of the distribution. So, the distribution p of x has these 2 parameters x 0 which is the average and sigma which is a standard deviation. So, when we are calculating averages of certain data, you can implicitly assume that what you are doing is you are trying to estimate x 0. You are trying to get an estimator of x 0. So, suppose you have a set of outcomes let us say x 1 x 2 x 3 up to x n you just have a finite number of n experiments. Then based on these n experiments you are trying to estimate the true average, which is what appears in the distribution and you are trying to estimate the few true value of sigma.

So, this is one way of stating the problem. So, that the idea of estimation of parameters becomes clear.

(Refer Slide Time: 02:18)



(Refer Slide Time: 03:43)



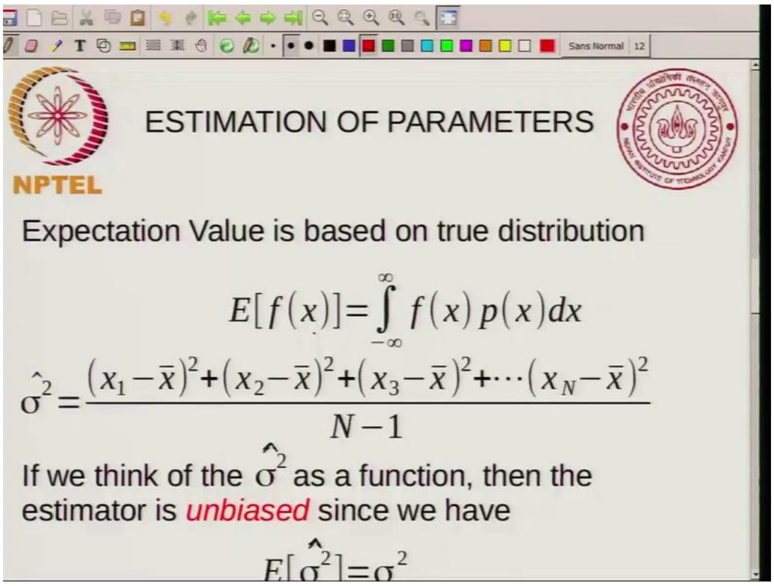
So, the, it turns out that the usual procedure of estimating, turns out to be the best estimate. So, here I put x 0 hat to distinguish from x 0 just to remind you x 0 is something that appears in the actual distribution. So, x 0 and sigma are the ones that appear in the actual distribution and what we are trying to say is, we are taking a finite number of points and we are trying to estimate x 0 and we are trying to estimate sigma. Based on this finite number of points and it turns out that through very formal methods you can show that the best estimate, best estimate of x 0. So, the estimator of x 0 the best estimate is the average of these quantities and the best estimator of sigma square which is the square of the standard deviation of the variance. So, the best estimator of that turns out to be this form. So, you take x 1 minus average of x, calculated in this form and you square it and go all the way and you divide by n minus 1. So, you take x 1 minus x bar square plus x 2 minus x bar square and so on you do for all the N data points you take the difference from the calculated average and you divide by N minus 1.

Now, it might always be a little confusing as to why, this N minus from 1 factor appears when you are calculating the standard deviation and this actually turns out to be a very the factor of N minus 1 and then in the denominator it is actually essential to make this what is called an unbiased estimator. So, if you take n then formally it becomes a biased estimator, but. So, formally you should take n minus 1 of course, if your n is very large and then N and N minus 1 are almost the same. So, it does not matter, but formally this should be n minus 1 in order to make this an unbiased estimator.

So, again let me emphasize that we are differentiating, we are putting a hat on top of x 0 or all these hats that are there on top of x 0 and 0 and sigma square these emphasize that these are estimators. So, x 0 hat refers to an estimator of x 0 similarly sigma square hat refers to an estimator of sigma square and this is not the same as x 0, this is not the same as sigma square. So, sigma square is something that appears in the actual distribution. So, that is x 0 something that appears in the actual distribution and they should be distinguished from the estimators of these quantities.

So, now let us come to this. So, we said that the factor of n minus 1 and the denominator is essential to make sigma square an unbiased estimator and we want we will see that we will see why that is so in a in a few minutes.

(Refer Slide Time: 05:30)



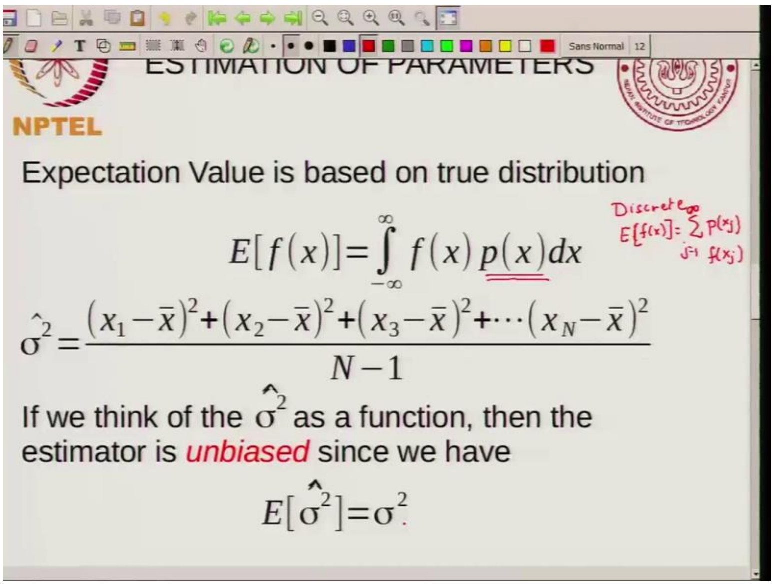
But, before that, let us define this quantity called the expectation value. So, the expectation value of any quantity or a function is based on the true distribution.

So, suppose f of x is some function of x and x the quantity the variable x the random variable x is distributed according to p of x. So, if p of x is the distribution of x and f of x is some function of x then what is the expectation value of f of x and that is given by this integral. So, you integrate f of x p of x d x from minus infinity to infinity. Here, I have taken the specific case where the range of x is minus infinity to infinity and have also considered the specific case where x is a continuous variable. You can do this for a discrete variable also where you would not have an integral, but you will have a sum, in the case of discrete variables.

So, suppose you had discrete variables, then what you would have is this estimator of f of x, would look like would look like sum over j equal to 1 to all the allowed values. If you might go all the way to infinity and what you will have is p of x j times f of x j. So, it would look like that. So, in the case of discrete you will have instead of having an integral you have a sum and everything will work out will work out exactly the same way.

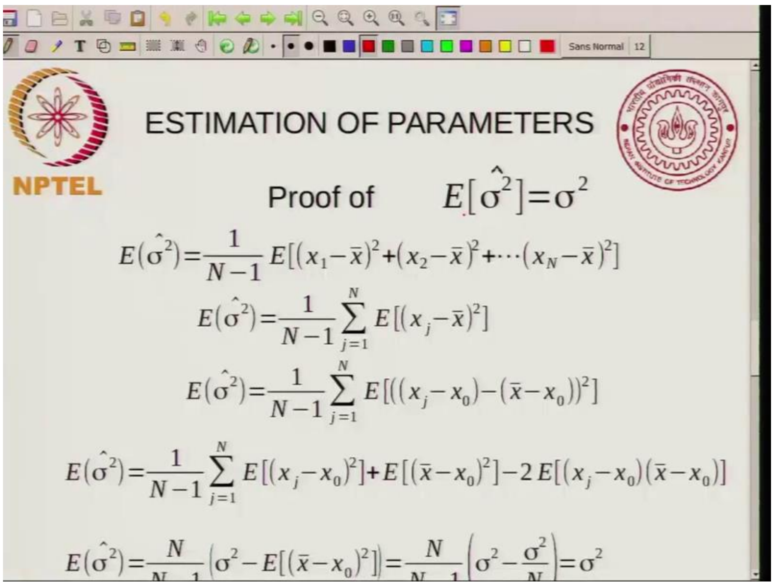
Similarly, now let us look at the estimator of sigma square, we said is x 1 minus x 1 bar square, x 2 minus x 1 bar square, x bar square and so on. Notice again I have used x bar and not x 0 x bar is the actual average that is calculated from the n readings and we said that this should be divided by n minus 1 and what you can say is that, what does it mean I said a few minutes ago that sigma square this is an unbiased estimator.

(Refer Slide Time: 07:40)



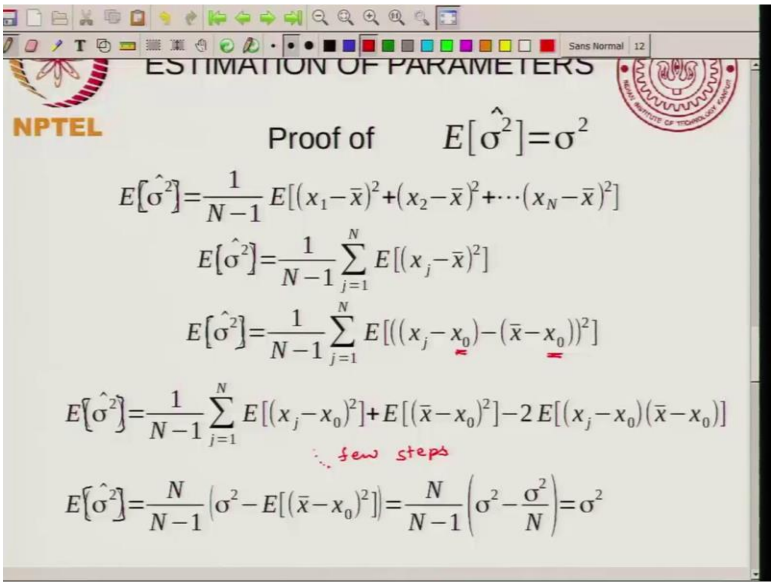
So, what does it mean to say that the estimator is unbiased? The estimator is unbiased if the expectation value of sigma square estimator, if the expectation value of the estimator is equal to the sigma square. So, the expectation value of the estimator of sigma square should be equal to sigma square, only then the estimator is said to be unbiased.

(Refer Slide Time: 08:09)



So, now we can go ahead and we can calculate the expectation value of sigma square. So, you can just go ahead and you can. So, what we will try to prove is that the estimate the expectation value of the estimator of sigma square is equal to sigma square and that will make it an unbiased estimator.

(Refer Slide Time: 08:24)



So, the proof of that is a not very difficult and so for example, you can there should be is square bracket the estimator of. So, the estimator of sigma square, can I using square brackets for the estimators. So, we will just be consistent. So, the estimator of the expectation value of the estimator of sigma square that is given by 1 by n minus 1, the expectation value of this sum. So, so again just to remind we had this estimator of sigma. So, what we are asking is, what is the expectation value of this quantity? That is same as expectation value of this quantity. Now, if you have a constant in the estimator you can take it outside the estimator. So, what we are left with is exactly this quantity. So, the 1 minus n I am to have taken outside and what I am left with is estimator of or the expectation value of all these things.

So, now I can write that in short form as the sum over j equals to 1 to n the expectation value of x j minus x bar square and again there is a property that the estimator of some of quantities is equal to the sum of estimator. So, I took the estimator inside these sums. So, I wrote it as a estimator of this plus estimator of this plus estimator of this and so on.

So, I can write the estimator of this sum of x j minus x bar square as the sum of the estimators and what I do is, I do a little trick I write this as x j minus x 0 minus x bar minus x 0. So, I am adding and subtracting x 0. So, if you add and subtract x 0 from this expression then you can write it in this form.

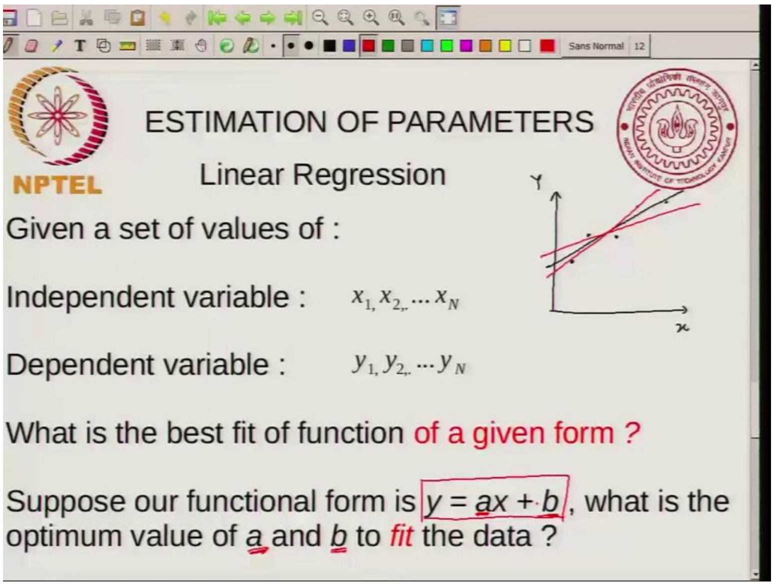
And now you can go ahead and you need to, you can expand this square. So, if you expand the square you will get 3 terms and you can again take the x the expectation value of this of the sum as the sum of expectation values and after this, there are a few steps in this derivation, but basically you can show that this whole thing works out to n by n minus 1 times sigma square minus expectation value of x minus x 0 square x bar minus x 0 square x bar is the calculated average then this is the true average.

And this quantity you can show is just sigma square by n. So, what you get is n over n minus 1 times sigma square minus sigma square by n which comes out to be equal to sigma square. So, actually this factor of n minus 1 is very important to make sure that, expectation value of sigma square is equal to expectation value of the estimator of sigma square is equal to sigma square. There are a few steps in this derivation. So, I will just write a note here that there are a few steps between these 2, but you can try to work it out and you can look up standard reference books and this is worked out there.

The point is I want to make from this exercise is that when you write your expression for sigma square there should be n minus 1 in the denominator and that is explained by this. Secondly, the idea that bite when you have a finite set of data just by the process of calculating average and standard deviation you are actually estimating the properties of the true distribution.

So, this idea of estimation of parameters is something that is in the background of what you do as a very normal course of action.

(Refer Slide Time: 12:20)

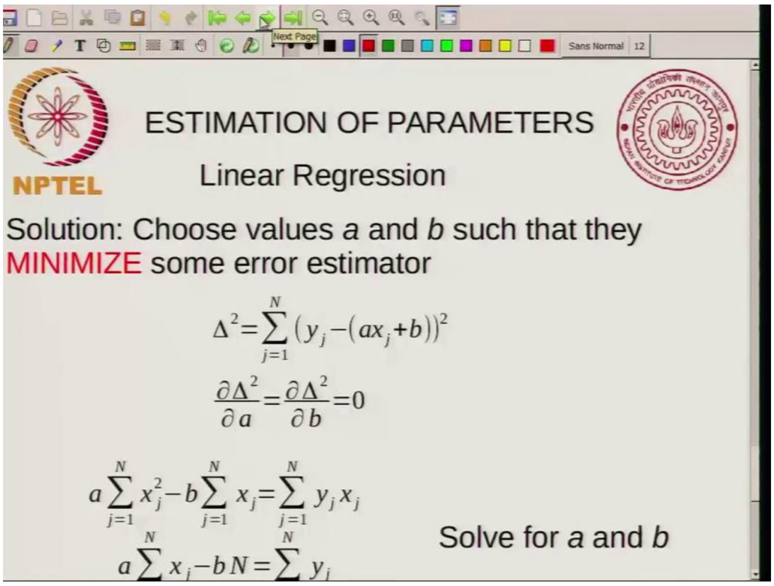


Now, the next example, where you are actually estimating parameters and again this is something that is done as a very routine course of action this is linear regression and you often do this in your especially during a laboratory experiment. So, what you might have is you might have a given set of values, you might have a set of values of some variable, which we call the independent variables which is x 1 x 2 up to x n and then you have a values of some dependent variable which we call y 1 y 2 up to y n. So, you have a table of data of x and y and you want to know; what is the best fit of certain of a function of a given form? So, suppose you have a data, you want to fit a straight line through it that is a good example that is a very common example. So, suppose your functional form is y equal to a x plus b. So, what is the best straight line fit through that data?

So, then you want to ask what is the optimum value of a and b to fit this data. So, this a and b are parameters of the straight line. So, what is the best value of a and b that fits the certain data and this is something all of you are used to doing. In fact, in fact you are used to drawing graphs for example; you might have done something like this, where you have various data that you plot. So, this is x and y you plot and then you sort of draw a straight line through that data and so the question is how do you choose how to draw the straight line? Should I draw it this way or should you draw it or should you draw it this way or should you draw it some other way how do you decide? So, the best choice of the straight line fit through the data is what this question asked.

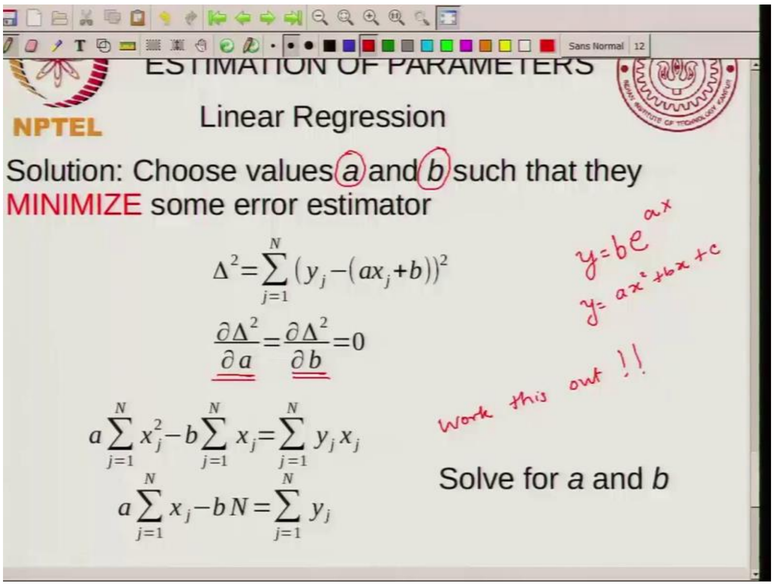
But what is important to say here is that, if you are just given data then you can fit it to a straight line, you can fit it to some parabola, you can fit it to an exponential, you can fit it to various functional forms and this process of estimation of parameters will say that, suppose I fit it to a certain form what is the optimum value of parameters. So, given that you fit this data to the form y equal to a x plus b which is the straight line. So, given that you fit it to this form what is the best choice of a and b. So, this is what is given by this idea of regression analysis and as I said a x plus b is just 1 functional form I can take many different functional forms and I can fit the same data to that. The procedure will remain the same; I will just illustrate this procedure by taking this particular example.

(Refer Slide Time: 15:03)



So, how do you implement this procedure? So, according to this procedure what you are supposed to do is to you choose values of a and b.

(Refer Slide Time: 15:11)



So, you choose your a and b such that they minimize some error estimator and I will not go into the details, but the error estimator that is minimized is given by delta square which is a square error. So, what you do is you take y j that is y at a certain data point and you subtract it from a times x j, x j's the value of x at that point plus b and you square this.

So, if all the points, who are exactly on the straight line then this this would be exactly o, but if the points are not on the straight line then each of these would be nonzero you square it. So, you get a positive quantity. So, you sum all these you will get some positive quantity and what you try to do is, you try to minimize this square error that is the procedure.

Now, what can you vary to minimize you minimize it with respect to a and b. So, you choose your value of a and b. So, that delta square is actually minimized. So, that is the formal solution and as I said, you can do this I am illustrating this for a linear fit, where y is where I am fitting the data to a straight line I could also do this for some other functional form and whatever the parameters of that functional that is what will enter into this.

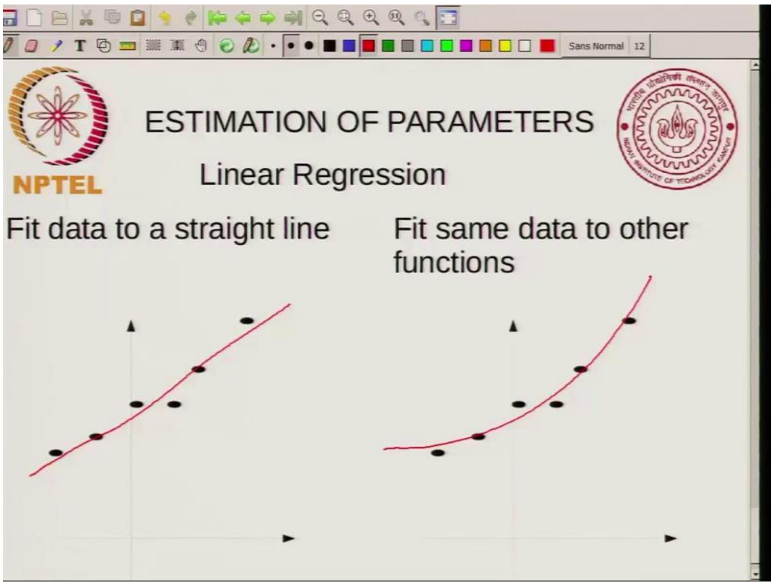
So, if you want to minimize this quantity with respect to a and b then you should have that first derivative with respect to a and first derivative with respect to b both should be 0 and you can calculate the first derivatives by expanding these quantities and so on I will not go into the details, but what these 2 will give you is these will give you 2 equations. So, this will give you 1 equation this will give you another equation. So, each of these derivative conditions will give you an equation. So, the first equation I am writing in this form I just rearrange terms and I can write this as a times the sum over j equal to one of x j square. So, you sum the squares of all the values of x and then you have a minus b times sum over j equal to 1 to n of x j and this should be equal to sum over j equal to 1 to n of y j x j.

Similarly, if you take a times sum over j equal to 1 to n of x j minus b times n then you should get sum over j equal to 1 to n of y j. So, it turns out you can work this out I mean based on these conditions you can easily work out both of these and I encourage again I encourage all of you to work these steps. So, work to work this out. So, what you get from this is that you get 2 equations and if you know all the x j's. So, you know all the y js. So, you know all these sums all these sums are known and so you just have 2 simultaneous equations for a and b and you can work out the values of a and b. So, you can solve these and you can get the values of a and b and this once you know the values of a and b you know how to draw your line.

So, this is the process of linear regression and as I stated, I can take other functional forms for example, I can take another functional form for example, y equal to e to the a x. I can take a functional form like this and if I want I can put a b b e to the a x. So, I can I can do the same thing, I just replace instead of a x j plus b I will have b e to the a x j and I can just go through the same exercise and I can get.

So, you can apply this procedure to any function and you need not have only 2 functions you can have 3 functions. For example, you can have y equal to equal to a x square plus b x plus c. So, now, you have 3 parameters and you can fit you can find the best estimator of a, b and c for this given this data.

(Refer Slide Time: 19:25)



So, this is a very powerful technique and what you get is you often have this data and once you calculate a and b you know that a is the intercept and a is the slope and b intercept. So, you know your line you know your line and you will get some straight line fit through the data. You can also fit the data you can take the same data and fit it to other functions for example, I could fit this to a parabola and I might get something like this.

So, what the estimation of parameters or linear regression will tell you is, what is the best choice of parameters for a given functional form? So, suppose you want to fit it to a x plus b then what is the best choice of a and b suppose you want to fit it to a parabola what is the best choice of the parameters of the parabola.

So, again I emphasize that you know this same data can be fit to different functions and in each case the linear regression analysis will give you the best fitting function. So, I will end this module here. So, in these 4 lectures we have covered we have looked at several different topics starting from errors, sources of error precision measurement accuracy significant figures we looked at probability, probability distributions, we looked at the binomial and the Poisson distribution then we looked at the Gaussian distribution and we calculated integrals and averages and then and then we saw the process the procedure of estimator of parameters where we talked about errors and we talked about the least square fit method.

So, I hope through all these steps I mean there is an underlying theme that is there in all these is that you imagine that your experimental results or your data is some sort of random variable that is distributed according to some distribution and once you have that underlying picture then all these procedures are they follow very naturally. So, in the next lecture what I am going to do is I am going to do a few practice problems from these topics. So, that will be for the next lecture.

Thank you.