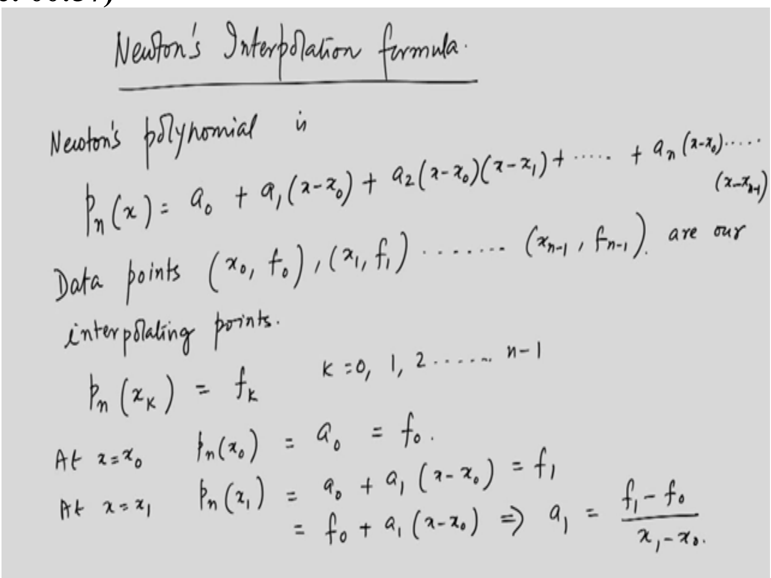
Numerical Methods and Simulation
Techniques for Scientists And Engineers
Saurabh Basu
Department of Physics
Indian Institute of Technology- Guwahati

Lecture 06
Newton's interpolation formula, statistical

As I said earlier that in the Lagrange's interpolation formalism we could not improve the accuracy by including one more data point or extending our set by one more data point the Lagrange's basis polynomials which we have described would have to be computed once again all over again from the start. In fact these Newton's interpolation formula that we are going to describe now takes care of that difficulty okay.

(Refer Slide Time: 00:57)



So, we can take into account one more data point in our polynomial and can improve its accuracy. So, Newton's polynomial is so n stands for the order of the polynomial as I said earlier this is the linear the index independent term then it is a linear term and this is the quadratic term and so on. And we have all the way up to a n x - x 0 and x - x n - 1 so we have n data points starting from x 0 to x n - 1 and we are going to construct a polynomial nth order polynomial according to this where a 0 a 1 a 2 our coefficients which have to be determined.

So, the data points that you obtain as I told you that these are the data points obtained from either a computer experiment or a real experiment suppose you are you know tabulating data for a very large class this is from the perspectives of a teacher that when he or she has gotten the marks for a particular course for a large class. So, there is a roll number versus the marks that one obtains and then one makes a table out of it and then one can actually find out the performance of the class from the data by doing certain analysis.

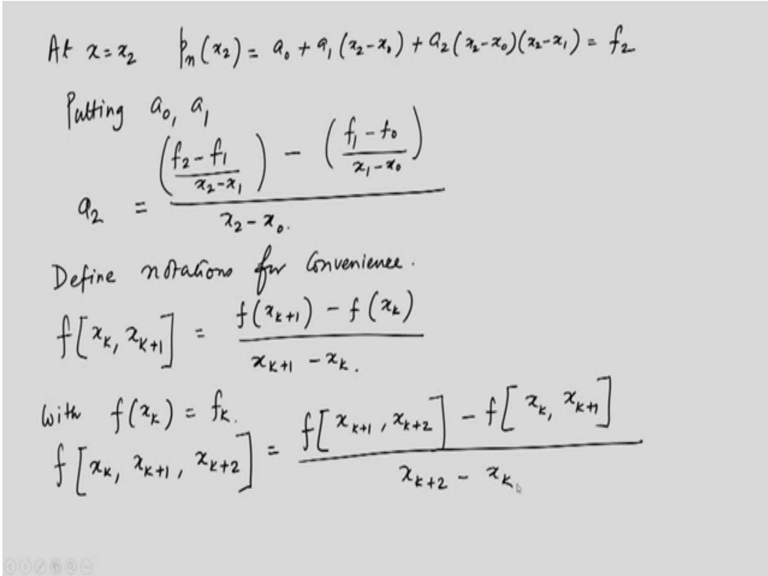
So, whether we have really a normal distribution I am going to describe that or what is the full width at half maximum or the variance of a distribution and so on. Here we are simply talking about these are mostly monotonic functions and even if they are non monotonic it does not matter I mean there are nonlinear functions that we are anyway talking about. But even there could be points of inflection where the there is a non monotonicity.

But in any case we are talking about this polynomial and a set of data points. So, the data points are given by x 0 f 0 we call it f 0 which is a value of the function at x equal to x 0 and then you have a x 1 f 1 similar meaning goes that is the value of the function at x equal to x 1 is called f 1 and similarly you have all the way up to x n - 1 f n – 1. So, these are the interpolating points. So, which says that p n at x k x k some value of x is called as f k where k equal to 0, 1, 2 all the way up to n – 1.

So, at x equal to x 0 p n of x 0 = 0 which is nothing but equal to f 0 so at x = x1 PN of x 1 = a 0 + a1 x – x 0 that is a linear polynomial which we are going to call as f 1. And since your f 0 is a 0 we can simply write it as f 0 + a 1 x – x 0 and this gives us a 1 equal to f 1 - f 0 divided by x 1 – x 0 which is typically the slope. So, far we have talked about the 0th order polynomial which is a constant term and we have talked about the linear polynomial that is of the first order.

So, we have at similarly then proceeding at x equal to x 2 make I am doing this because you need to actually get these all these coefficients in terms of this functions and I guess showing up to a 2 would be enough you can build up this other functions a 3 the coefficients a 3 a 4 etcetera from the values of the functions.

(Refer Slide Time: 06:26)



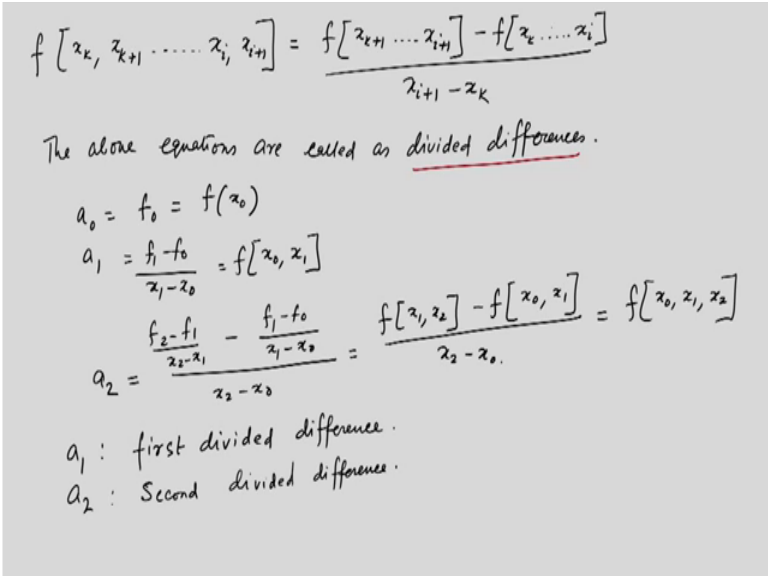
p n of x 2 it is equal to a 0 + a 1 x 2 - x0 + a 2 x 2 – x 0 x 2 – x 1 equal to f 2 so that is the value of the function at x equal to x 2. So, if you put the values of a 0 and a 1 of course a 0 is equal to f 0 and a 1 is what we have just obtained in terms of this basically the slope which is f 1 – f 0 divided by x 1 – x 0, so if you put both then the a 2 comes out to be f 2 – f 1 divided by x 2 – x 1 – f 1 – f 0 divided by x 1 – x 0 whole divided by so let us just write the left hand side a little carefully, so this is a 2 and this is x2 - x0 okay.

So, we define slightly cryptic may not be easier but for our convenience we use these notations we call it f x we write it in square bracket x k x k + one because you see always we are talking about the difference in that functions values of the functions at two points, so this we can call it as f x k + 1 - f of x k divided by x k + 1 – x k so with f of x k denoted by f k we can write down another function which is of a larger you know just like your a 2 which involves 3 functions the values of the functions at 3 points namely f 1 f 2 and f 0.

We are we can define these functions as x k x k + 1 just like x 0 x 1 and x 2 so it is x k x k + 1 and x k + 2 is simply like your now this has to be written in square bracket because we are writing these functions in I mean these values x k + 1 which are data points in the single bracket but whenever we are defining these the left-hand side which is a function of two points then we are writing it in square bracket.

So, if that is the case then we have a x k + 1 and the x k + 2 and f x k and x k + 1 and this is like x k + 2 - x k okay. So, these are your notations that we have defined this for our convenience.

(Refer Slide Time: 10:22)



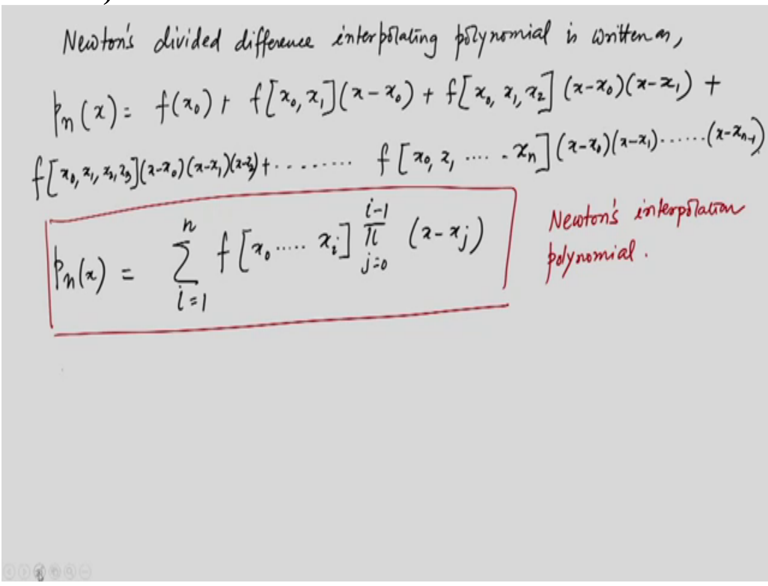
Let us go with these notation and define the most general one so that is f x k x k + 1 going up to x i x i + 1 that is f x k + 1 going up to x i + 1 - of x k and going up to x i and I have a difference in x i + 1 - x k that is the two extreme points the difference between the two extreme points. So, these are called as divided differences okay. So, these above equations okay, so, this is a new term that you have so they are called divided differences these f functions. So, what we have in terms of our known quantities that is these coefficients.

So, our coefficients are like a 0 which is equal to f 0 which is nothing but f x 0 then you have a a 1 which is an f 1 - f 0 divided by x 1 - x 0 which is nothing but a function of two things which is x 0 and x 1 similarly a 2 is f 2 – f 1 divided by x 2 – x 1 – f 1 – f 0 divided by x 1 – x 0 this is like x 2 – x 0, so let us again align to remove any confusion that you may have while working out so that is your a 2 which is nothing but equal to f of x 1 x 2 - f of this has to be a square bracket x 1 and divided by x 2 – x 0 okay.

So I believe this is clear that we are trying to again talk about another polynomial just like yesterday or rather the last time that we talked about the Lagrange's polynomial. So, this is that a Newton's interpolation polynomial which uses this divided differences and you could figure it out that it is not a problem to include one more data point and improve the accuracy because it just requires you to calculate these functions once more.

So, we have this can be written as so this is like a f and x 0 x 1 and x 2, so if I wanted x 3 also to be included I can I can write down f x 0 x 1 x 2 x 3 by simply incorporating the divided difference in this particular fashion. So, it is the name is that so a 1 is called as the a 0 is of course the 0 with order divided difference a one is called as a first divided difference and a 2 is likewise the second divided difference.

(Refer Slide Time: 14:47)

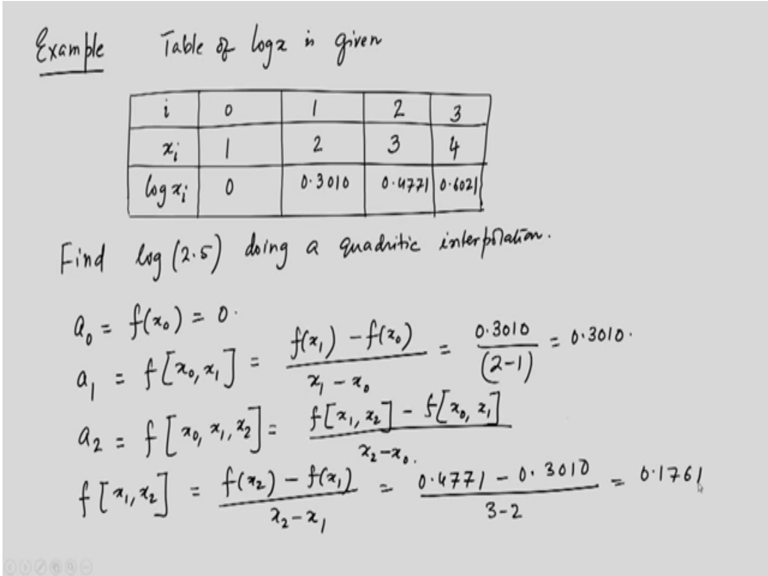


So, with this we can write down the Newton's just use this word once again to make it very obvious that there is the difference divided difference method. Divided difference interpolating polynomial is written as p n x that is f of x 0 + f of x 0 x 1 x - x 0 that is a linear term and then we have a quadratic term and this x - x 0 into x - x1 and then we have we can continue we can have a cubic term which would simply have a if you want me to write this so there is a cubic term which would be simply you know x 0 x 1 x 2 x 3 and then you have a x – x 0 x – x 1 into x – x 2 and so on okay.

So this plus if you carry on like this then we will have a x 0 x 1 and so on and x n and x – x 0 x – x 1 all the way up to x - x n - 1 okay. So, this is the procedure that we should follow now you see that if you have suppose you need to do a quadratic interpolation. And in which case you may want to actually include 3 which is the minimum number of points that you should include in order to do a quadratic interpolation but suppose you want to do include a couple of more data points as 5 data points.

You only need to calculate these f functions the values of their functions and can calculate it can improve your accuracy in the same program. So, the same program can be a little more can be added in order to calculate or including this 5 points to do a quadratic interpolation. So, this is written in a summation as i equal to 1 to n and we have a f and you have a x 0 all the way up till x i and then we have a product of functions which is j equal to 0 to i - 1 and x – x j. So this is the P n x that is the; so this is called as a Newton's interpolation polynomial okay.

(Refer Slide Time: 18:41)



So, this is what we need and let us look at a simple problem again. So, example table of log x is given okay, so this is the table that you have and okay. So, you have these are labels for the serial number then you have these values of x and then log of x i, so when i equal to so this is the 0 it is the marker for you know the counter rather the counter which counts the number of data point. So, it is 0 1 2 3 so there are 4 data points if x i equal to 1 then of course log of x i equal to 0 then it is 2 then it is .3010.

And when it is 2 when it is 2 then it is equal to 3 and then it is .4771 and when it is 3 when it is 3 this is 4 and this is .60 to 1 ok, so a number of x values have been given such as x equal to 1, 2, 3, 4 and the values the log of those x's are also tabulated. Now what you have to do is that find log of 2.5 which is not given here which lies between 2 and 3 doing a quadratic interpolation okay.

So, quadratic interpolation as we have told several times that it requires three values a linear interpolation needs 2 values. so, this 3 values can be any 3 if you wish to actually incorporate 1, 2 and 3 because 2.5 is subtended in that case you can go ahead and include 0, 1 and 2 and do this quadratic interpolation such that the 2.5 actually lies beyond what we have taken and do a quadratic interpolation and get the value.

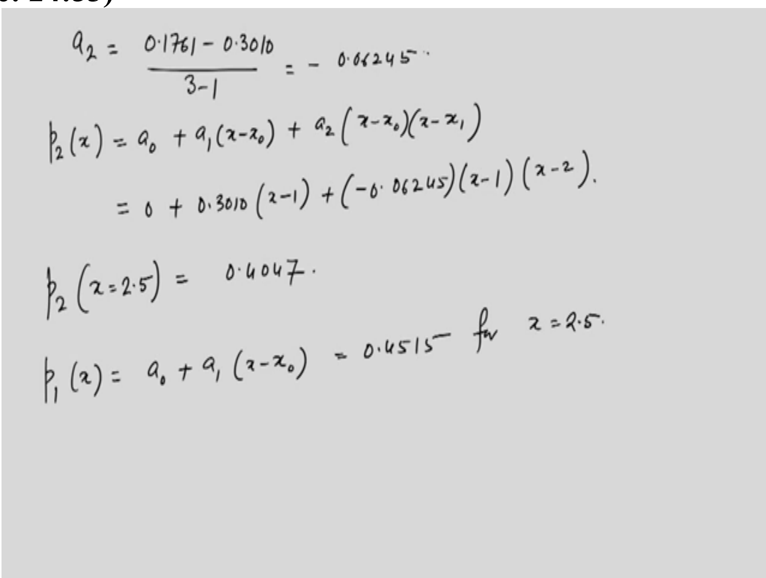
Let us see so a 0 which is nothing but f at x 0 which is 0 okay because your x 0 point is this your a 1 equal to f x 0 x 1 which is equal to f x 1 - f x 0 divided by x 1 – x 0 which is equal to 3010 divided by 2 - 1 this is equal to .3010, so this is the value of the slope which is 0.3010 which is which is calculated from this value of the function which is of course because the first term is 0 so we get just 3010 and we have a 2 - 1 all right.

So, a 2 which is nothing but f x 0 x 1 and x 2 we have f of x 1 x 2 - f of x 0 x 1 divided by I am just following the definitions that have been given earlier. Now f x 0 x 1 we have already calculated f 3 0 which is a value of 0.3010. Now we need to calculate f x 1 x 2 which means it is f x 2 - f x 1 divided by x 2 – x 1 all these values are given here. So, we can use them and calculate this value and then this value is already known these this is known and one can get a value which is 0.4771 I am giving you up till 4 digits of accuracy.

If you wish more please go ahead 3010 and this is like 3 - x so sorry this is x 2 – x 1 because; so this is this is x 2 – x 0, so now this is we will just calculate it the intermediate such that, so let us just calculate f x 1 x 2 which we have not calculated the other one we have calculated. So, it is f x 2 - f x 1 divided by x 2 – x 1 which is equal to 0.4771 - 3010 divided by 3 - 2 which is equal to .1761 so that is the value of this one here, there is a value of this one which we have calculated here.

And this one is anyway the value that we have obtained here now we have both the values. Now we can go ahead and calculate a 2.

(Refer Slide Time: 24:53)



So, a 2 becomes equal to .1761 - .3010 divided by 3 - 1 which is – 0.606245, so we can construct a quadratic polynomial using all our a 0 a 1 and a 2 now we know, I am so sorry a 1 has to be multiplied because that is a linear function that has to be multiplied with the x linear in x term and a 2 which is x - x 0 and x – x 1 and will keep up to this a 0 is of course 0 and this is 3010 and this is x - 1 + this is 0.0624 5 into we have gone one digit more here x – x 1, so x 1 is nothing but this value 2 and sorry x – x 0 is 1, and x – x 1 is 2.

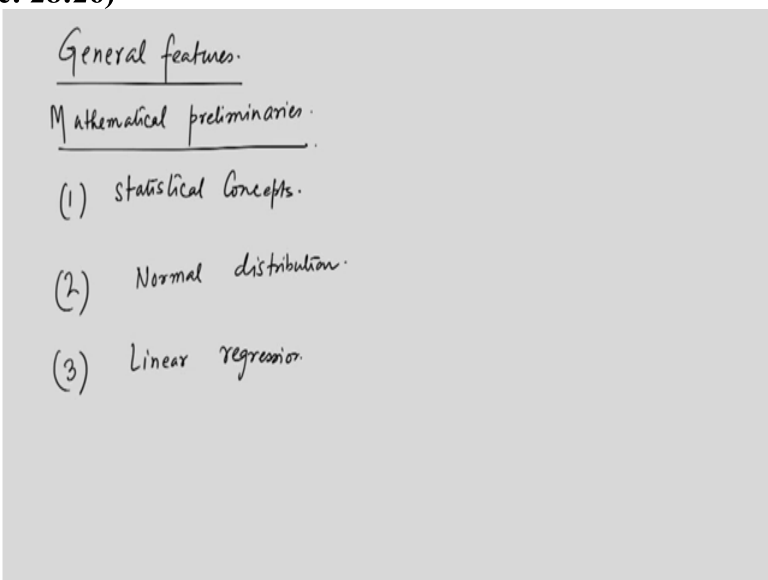
So, if you simplify this and then put x equal to 2.5 which is where you need the value of the algorithm and then you get a value which is 0.404 okay, you can check what the exact value is.

And if you use a linear polynomial that is had you stopped at this a 0 + a 1 x – x 0 then you would have gotten a value which is 0.4515 for x equal to 2.5 ok. So, you see that we are much better off by doing it has gone down by a by .05 which is a large number as far as in numerical accuracy goes.

And quadratic gives you a value which is much better of course that you would understand if you actually calculate the law to 0.5. Now the important point is the following that as I told here that if you have if you want to include one more term that is this term that is which we have left out which is 4 that is x 3 if you wanted to have a x 3 you simply had to write down or if you want to take into account 3 different terms that is 1, 2 and three then you have to calculate the a 1 a 2 and a 3.

And then you could simply do it by computing the divided difference these functions. So, this is a mostly about curve fitting but let me tell you on a general ground what we need about curve fitting.

(Refer Slide Time: 28:26)

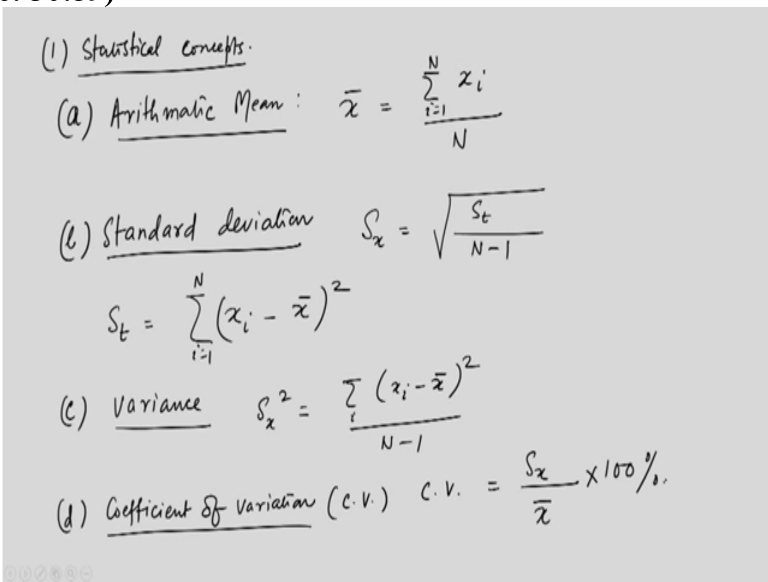


So, general features of course we could have told earlier but we will talk about a few general features and basically the mathematical preliminaries that we need for doing curve fitting of a set of data points that we have obtained from an experiment. But at the same time we need a few mathematical details that are needed and because we are dealing with a very large number of data points. It is neither possible to show in the class by considering a large number of data points and keep calculating it because that defeats the purpose of showing the utility of the method.

But it is best done using a computer programming and writing down either a code or using some software's for the fitting part. But if you use a software instead of writing a program you should still know that what are the methods that the software is using. The software is made by human beings and since the human beings have to follow certain algorithm one needs to know that what are the algorithms that have been you know followed.

So and let us just start with a little bit of mathematical preliminaries will not talk much mostly we will talk about the linear regression methods and how we can talk about even just for a linear fitting how can we sort of take into account the errors that are that could arise from a set of data points. So, in this as I told let us talk about a few key statistical concepts let us talk about normal distribution and let us talk about the linear regression.

(Refer Slide Time: 30:59)



So, if we agree on this then the simple statistical concepts which you already are aware of I am just reminding you your engineering economics course or your steam tables in thermodynamics course or if you are a teacher and dealing with a class of large number of students the marks etcetera they all require certain amount of statistical analysis. And the reason that it is required is that you need to show either your teacher or your evaluator or your you know the head of the department that how your performance in teaching was.

And in order to do that there has to be some quantitative aspects of teaching and the quantitative aspects are that what is the average or what is the mean same goes for any other you know data set that you are talking about. Maybe you are calculating a few of something relevant in particular physics experiment or you are doing some chemistry experiments and you have generated a table and you need to know a few things.

And those few things the first one starts with an arithmetic mean which is the easiest and the arithmetic mean is defined as x bar which is equal to i equal to 1 to n x i divided by n and we have so there is the mean of n data points we have taken just the arithmetic mean that is we have summed them up and divided by the number of data points this of course causes no or rather this involves no biasing of data points.

Because in some experiments or in some you know measurements there could be certain amount of data or more relevant than others. And then you need to wait that data point we it means attach a certain larger weight to those data points which are of you know importance more importance than the others. And this is called as a stochastic sampling or rather they have different names you know in a particular context.

But we are simply talking about the arithmetic mean and we can also talk about the standard deviation which is an important quantity by itself why standard deviation important is that if the deviation is large from the mean value then of course the data are too scattered say the marks of a particular class of students in a given course. If they are distributed over very large number of values then one has to understand that there are really some very good students and some poor students in the class and that that causes a large discrepancy of values.

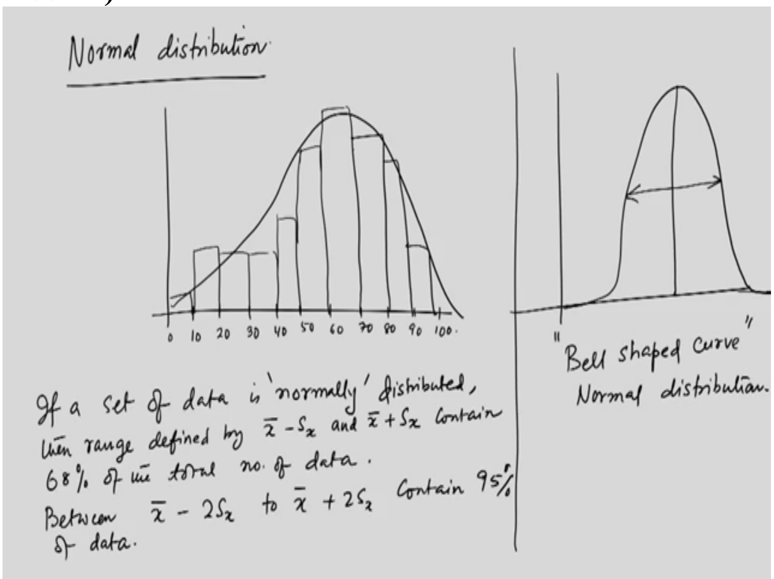
But if you have this if you have a distribution or if you have a standard deviation that is low which means that more or less the students have gotten numbers or marks which are close to the mean value which means that there's a on an average the teacher has done well. So, this is defined as say s x is equal to s t divided by n - 1 where s t is equal to the quadratic you know from the mean value the difference the square of the difference of each of those data points from the mean value.

And so this is the standard deviation and sometimes we actually talk about the square of this because there are square roots involved and when you take square roots there are farther you know errors that could be introduced. So, these are removed when you talk about a variance which is simply S x square which is equal to sum over i x i - x bar square divided by n + n - 1 so that is the variance that we often talk about okay.

There is also another thing that is said it is called coefficient of variation and one writes a CV for that so our CV is defined as s x which is defined here as a standard deviation divided by the mean and then you take the percentage of that that is called coefficient of variation. And so in principle you should take a set of data points which could be anything we will find a large number of data points which could be leading to I mean which could be you know you will find it in Internet that it is about you know results of an election or the coefficient of linear coefficient of expansion of you know solids or metals.

Or it could be the steam tables in thermodynamics as I told well this could be anything you know marks obtained in a 101 course in a in a university and various number of things so you one can calculate these all these quantities and understand what those set of data points mean.

(Refer Slide Time: 37:12)



Next let us talk about a normal distribution and this normal distribution is nothing but when you have a set of data the distribution of the data is important as I told that we need to understand what is the distribution of such a data point. Is this data point centered about certain things. So, there are you know bimodal distributions that is the centered about two such values and so on. so, let us just talk about just why they are centered about one value and one what one can do is that.

Let us say you have a class of maybe 500 students and they have the mark maths marks are out and it is put on a sheet of paper and one is given the task to analyze that. So, what you can do is that one can make bins of how many people got 0 to 10 how many people got 10 to 20 how many people got 20 to 30 and 30 to 40 and so on 50, 60, 70, 80, 90 and 100. It usually happens that you know there are very few very poor students and very good and very few say excellent students very which who are really very good.

So, one can actually get a distribution which is like this and so on and just drawing and it starts speaking about certain you know value and then of course these are and what I am drawing is just a freehand histogram okay and this okay say it is like this. So, if the data points if the number of students are very large in a given class then you can actually discretize it even smaller bins that is 0 to 5 marks and 5 to 10 marks and so on and then eventually you land up with the distribution.

So, we are talking about a distribution like the now if this distribution is like this that is it has a mean and it is it has a kind of variance then this has a name, it is like a bell-shaped curve pretty much what I wanted to draw here that is one would be termed as a very good teacher if it is really a sharp Bell because if the you know the average is quite high and most of the students have gotten values near the average.

And very few people who have got very low values or almost none and similarly maybe the values which are like the marks such as 100 etcetera or close to 100 maybe there are only a few people or there are none. But even then one is termed as a good teacher if the majority of the students have done in the say average or above-average range. So, this is called as the normal distribution.

Now it is in fact one can test it that so if a set of data is normally distributed normally means there is no abnormal there but if this follows a normal distribution then distributed then the range defined by x bar – S x we have defined S x here which is the standard deviation. So, it is on either side on the left side or the lower side of the average value and the right side of the average value.

So, these contain something like 68% of the data of the total number of data okay. If it is a normal thing and between x bar - 2 s x 2 s bar sorry not as far it is x bar + 2 S x we have contained 95% of the data okay. And of course there are deviations from that there could be deviations like the one that I have drawn is that there is a; this is called as a Lorentzian curve which has a bigger tail towards lower values.

And it is slightly you know skewed towards or rather this is not like exactly a normal curve but it is a curve that is cued towards one side that is towards larger values and so on. And so these are the data points that we actually talked about in experiments there is also a Gaussian distribution. Where the Gaussian distribution corresponds to which spreads from x equal to - infinity to x equal to + infinity.

If you are talking about centered about 0 but in real experiments we cannot have values extending up to + - infinity but of course we can talk about suppose we talk about velocities of molecules in an ideal gas. So, there is the velocities can actually go from minus infinity to plus infinity if we remember Maxwell Boltzmann distribution and calculation of the average velocities and the root mean square velocities we have actually integrated from minus infinity to plus infinity for the velocities.

Which means they can there is equal probability for them to go to minus take values minus infinity to plus infinity on either sides. So, those are called Gaussian distributions. So, we let us pause here for a moment and we will talk we will carry on with linear regression and maybe some more better approximations than that in order to feed our data. But the ones that we have talked about that is the Lagrange's interpolation formula.

And Newton's interpolation formula or the polynomials that we have talked about in that connection these work pretty well for fitting large number of data points.