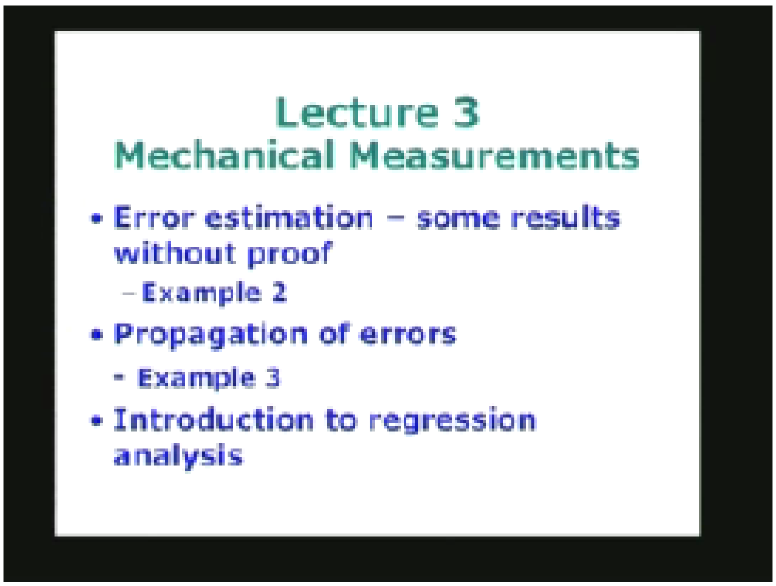
Mechanical Measurements and Metrology
Prof. S. P. Venkateshan
Department of Mechanical Engineering
Indian Institute of Technology, Madras
Module - 1
Lecture - 3
Errors in Measurements (Continued)

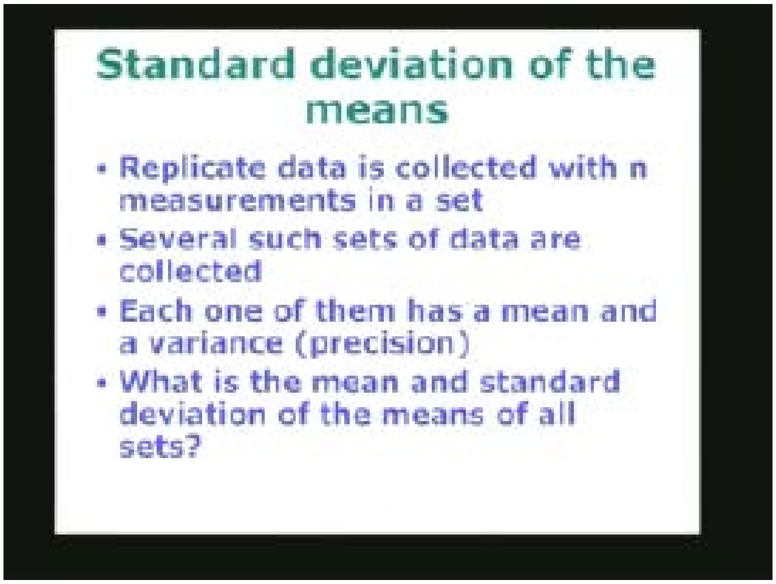
Welcome to lecture number 3 in our ongoing series on Mechanical Measurement. The topic I am going to cover today is error estimation and this will be basically a discussion on statistical principles.

(Refer Slide Time: 1:13)



Broadly speaking, they will be given without proof but I will give some physical explanation for the plausibility of the results, which I am going to give without proof. The subsequent topic, which I am going to look at in more detail is the question of propagation of errors and in both the cases, I am going to give some examples, 1 example each to bring out the meaning of what I am going to talk about. And if time permits I will introduce regression analysis as the last topic. I am not sure there will be enough time but if there is some time, we will do that one. So the question which I am going to look at is going to appear in the following fashion.

(Refer Slide Time: 2:10)



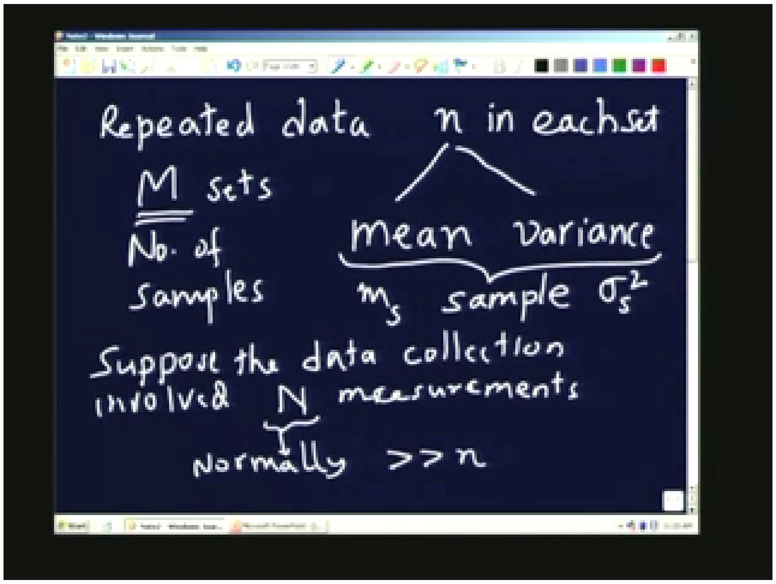
Suppose I am doing a certain experiment, I have set up the experiment. I am going to repeat the experiment. Physically it is not possible to do the experiment a large number of times and therefore I will be collecting a certain set of data. That means I will repeat the experiment may be a couple of times, a few times, and I am going to collect replicate data from these measurements. Suppose I repeat the measurements n times, that means I am going to set up the experiment at a particular configuration and in that particular configuration, I am not going to change the variables. I am going to make the measurements repeatedly.

The idea of repeatable measurement is to get at the statistics of what is going to happen to the errors, which are associated with the measurement. So what I can do or what one can do is to do this replication of data n times, n data collection at a time and we will call it a set. Then I will do the experiment again and repeat the experiment n times and get a set, which I will call set number 2. So in principle, I can collect a large number of sets, may be m sets. So the question now is the first set, second set and so on up to the last set that I have collected, each by itself contains a certain number of data, n. So each set of data got has its own mean. So let us just look at what we have (Refer Slide above).

Each one of them, that means, I am talking about the set here; each one of the sets has got a mean and a variance, variance represents the precision of that particular set. So the question is, is there a relationship between the mean and the standard deviation of the means? That means, I am measuring by collecting data in the form of sets, each having n data points in the set. I am measuring again and again, so if I measure m such sets, I will get m means. There will be m means and each one of them will have its own precision or variance, therefore m variances I am going to have. I have got several means and several variances so what I will do is I will indicate how it is going to be done on the tablet.

So what we have is repeated data n in each set and we have M sets so each set will have its mean and its variance. We will refer to each one of the sets also as a sample. This is a sample; so I have M number of samples of replicated data or replicate data or repeated data, each with its own mean and its own variance. That means the sample has got a mean denoted by m_s as the sample mean and variance as sigma s square for the sample.

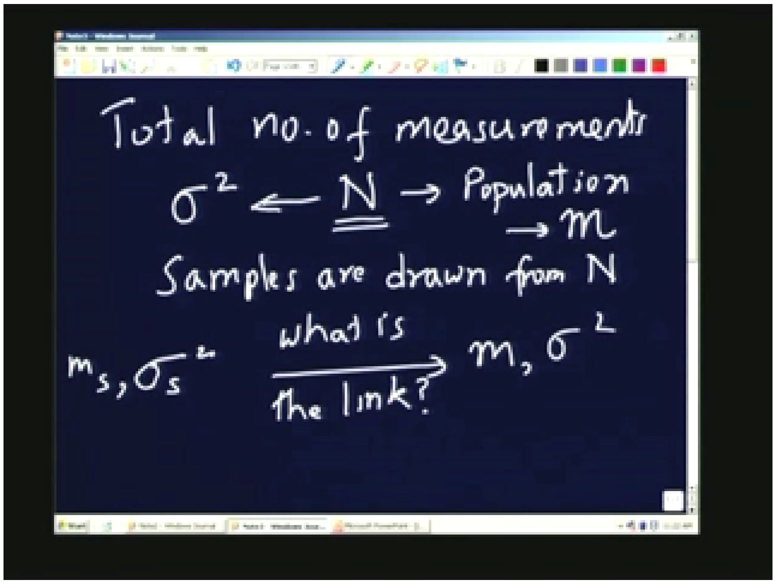
(Refer Slide Time: 6:10)



Now suppose the data collection involves let us say N measurements. This N, which is the total number, normally or usually is greater than n. Suppose the data collection involved N number of measurements, which is very large compared to n—of course, I have not done this—I am going to assume that I am able to do it or repeat the data N number of times, which is a very large number. Then what I will get is shown in the next slide.

Here I have total number of measurements N. I will refer to this as the population; the terminology is all borrowed from statistical analysis and when we talk about the sample which has got its own mean and variance and when we refer to the population here the samples were just a few of the data which are contained within N. So the samples are expected to be or are samples drawn from the population N. So I can assume that the population has itself got a mean, the mean of the population is denoted by the symbol m. Let us assume that the population has also got a variance which I will call sigma square; if I don't use any subscript it means that it represents the population. So the question we are going to ask is what is the relationship between them? So we have sigma s square and m_s, we have m and sigma square. What is the link between these two? That's the question I am going to ask.

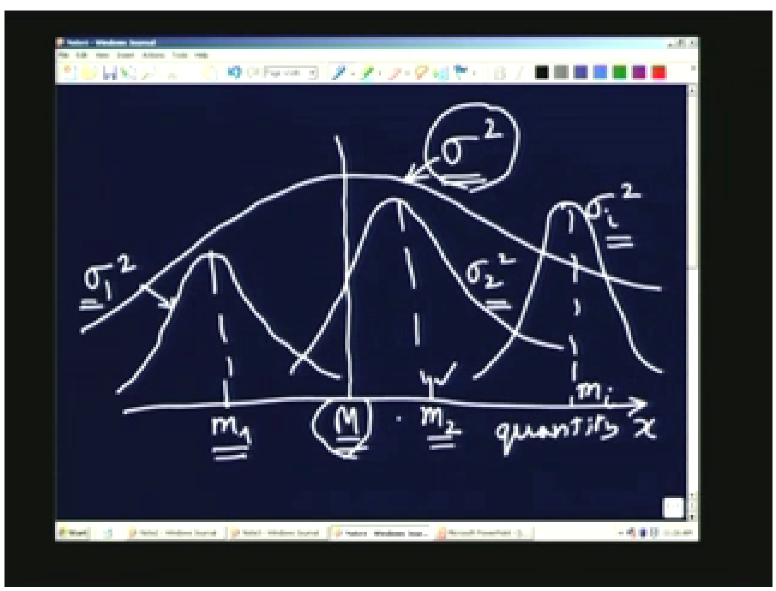
(Refer Slide Time: 10:02)



So the thing I am trying to describe is what is called sampling theory. I am not going to give any proof for any of the things I am going to describe today because that is going to take too much time and in this course we don't have enough time to describe in detail. Nevertheless, we are interested in the outcome of sampling theory because it is going to be a very important input when it comes to estimation of errors, which we are going to do constantly in our measurement process. Now I can also describe in the form of a graph whatever I have described in words. So let me just draw a graph here. This is the axis of the quantity we are measuring. Let me just call it x and I am going to measure it again and again.

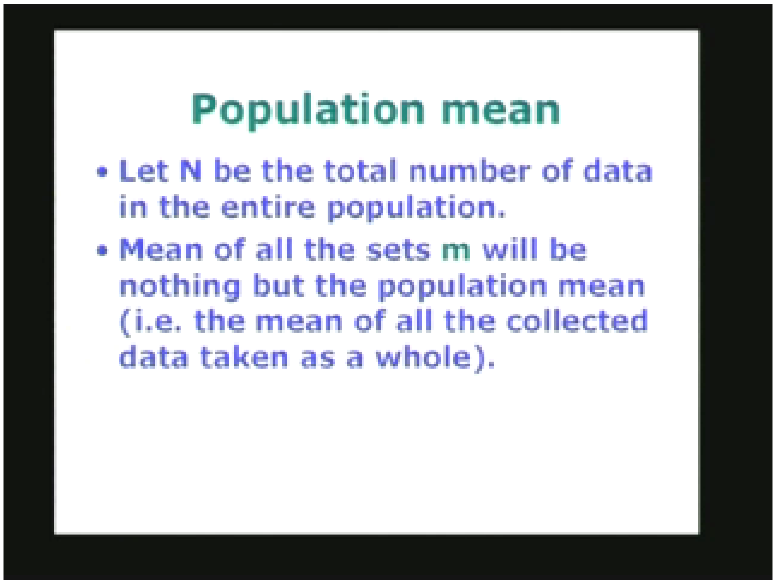
Suppose I made a sample as I said. Let us say n is equal to 5; I make the measurements five times and I get a sample and I can characterize the sample by its own mean and its variance, so let us say this is the mean of the sample, mean 1 (m_1) and if I were to indicate the variation of that, I will get probably something like this. This is the distribution, which has got a variance given by sigma 1 square. I will repeat it and next time I do again another experiment where I am going to measure it an equal number of times and I will get a separate, different mean. It will have its own sigma 2 square. So m_1 and m_2 are two samples and each one has got sigma 1 square variance, for sample number 1, sigma 2 square is the variance for the sample number 2. So I can do any number of times I can show 1 more or generally I can show this is m_i and it has got sigma i square. So the question I am asking is, suppose the mean of the population is M. This is for the population as I have already introduced the notation earlier, and it has got its own distribution. This is sigma square so I want to know the link between this and these quantities and I want to see what is the link between these quantities that is m_1, m_2, m_i etc.

(Refer Slide Time:: 14:08)



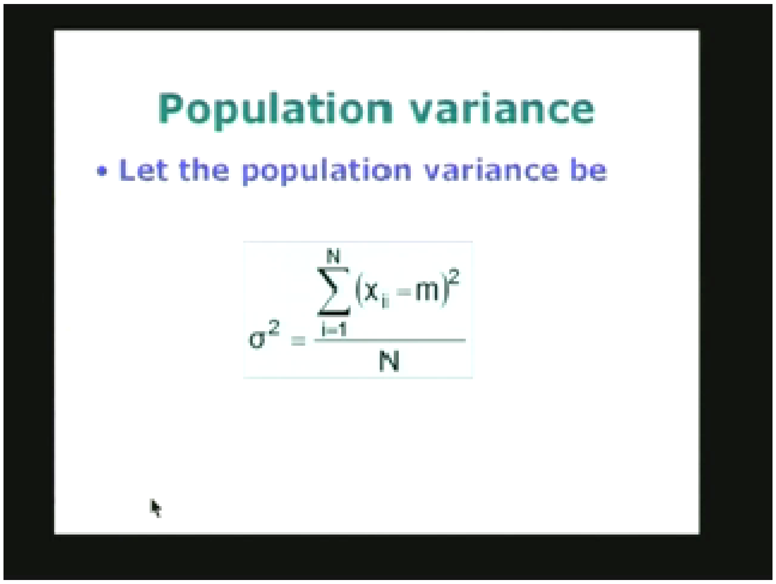
I am going to find out what is the relationship between them and the mean of the population and I am also going to look at the variance of the entire population and the variances of the each one of the samples. Why are we concerned with this? Imagine you are conducting the experiment and may be I am going to just do it once. I am going to collect 1 sample. Let us say this is 1 sample. Let us say this is the sample: m2 is the only 1 sample I am going to collect, only 1 sample and it will have its own mean and variance. Now what I want to get at is the value indicated here and I want to find out what is the precision if I were to repeat the measurement a large number of times. So it is a very important question from measurement theory and practice and therefore we are concerned with the question of linkage between these two. So what I am going to do is to look at this question and that is done in the subsequent slide.

(Refer Slide Time: 14:26)



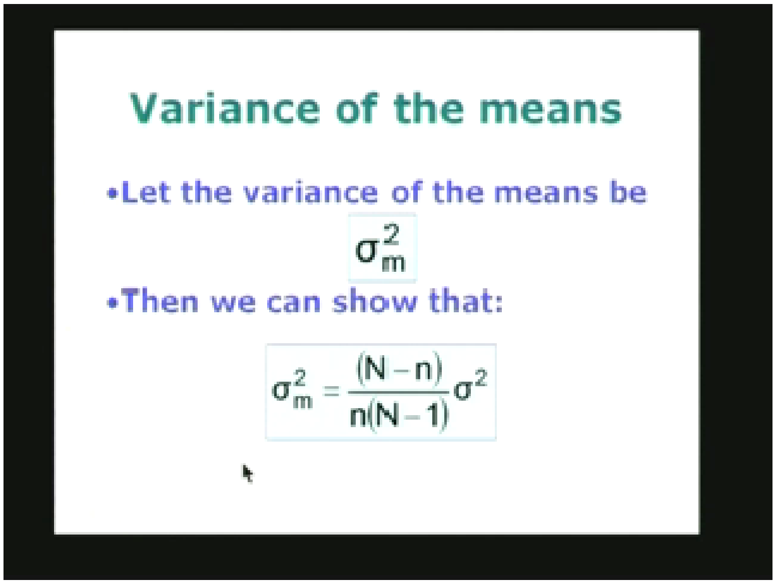
So I am just repeating whatever I have said earlier. So we have N, the total number of data in the entire population and the mean of all the sets we can show, which I call as m without any subscript, will be nothing but the population mean that I used as the M earlier. That is, the mean of the collected data taken as a whole will be actually equal to the population mean. So that is number 1 observation.

(Refer Slide Time: 15:05)



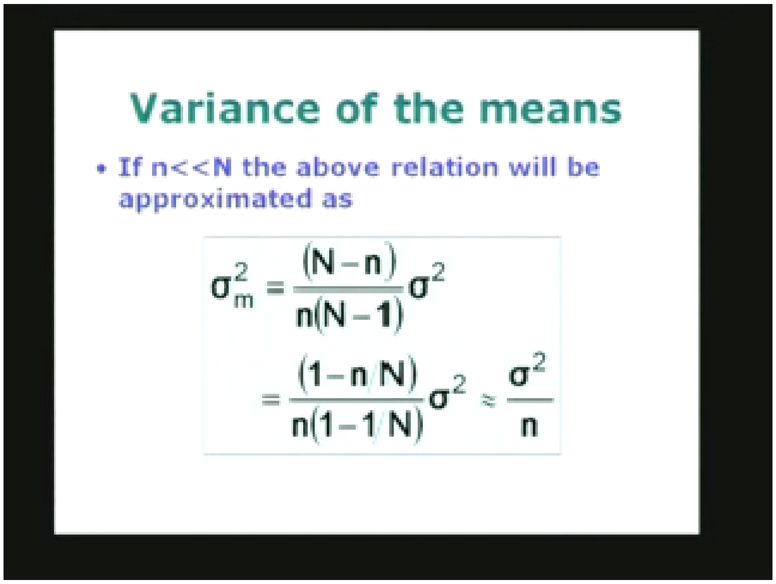
The second observation is if you look at the variance of the population I can define it in this particular fashion using the definition which we gave in our earlier lecture. Sigma square for the population, equal to sigma i equal to 1 to N, where N is the total number, x_i minus m whole square divided by N is the definition.

(Refer Slide Time: 15:28)



Now let us look at the population values; Sigma m square is the variance of the means; the means are distributed in their own way m_1, m_2 etc. They have their own mean and their own variance. Regarding the mean I have already explained in the earlier slide, mean of those means must be equal to the population mean. Now let us look at the variance of these means. So if I indicate by the symbol sigma m square, we can again show, this is without proof I am giving, that sigma m square the variance of the means is given by the formula N minus n divided by n into N minus 1 sigma square. That means the variance of the means is related to the population variance. So sigma m square is related to sigma square through a factor. If I look at this factor, you see that it contains N minus n in the numerator, in the denominator it contains 2 factors: the lower case n is the number of experiments performed in each sample and N minus 1 where N is the total number of experiments. So now what I am going to show is what is going to happen in actual practice; that is shown in the next slide.

(Refer Slide Time: 16:48)



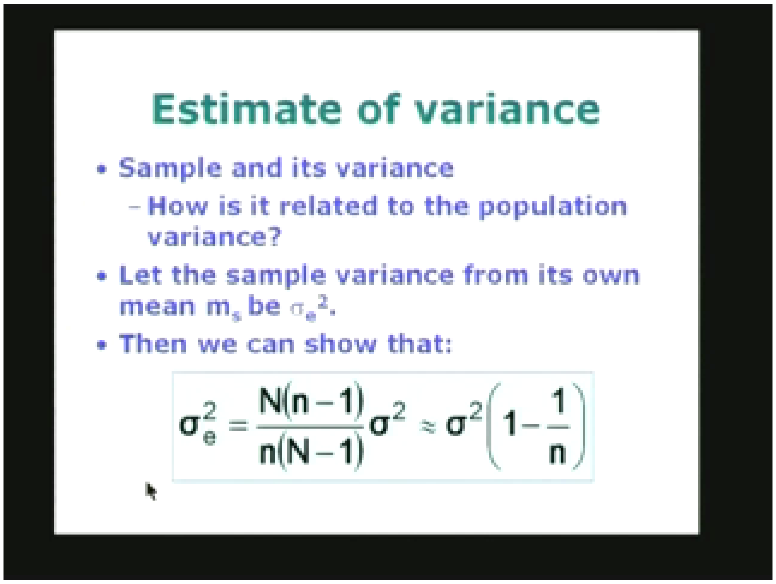
$$\sigma_m^2 = \frac{(N-n)}{n(N-1)}\sigma^2 = \frac{(1-n/N)}{n(1-1/N)}\sigma^2 \approx \frac{\sigma^2}{n}$$

So if n is very small when compared to N that means I am doing the experiment only a certain small number of times, which is a lower case n, that means I am going to take only 1 sample; i equal to 1, 2, 3... etc. I was talking about several samples; I am taking only 1 of those samples because that is the only one that is available in our measurement. We have made only 1 measurement containing the sample; let us say small n number times measured. So I have 1 set with n data available in that set, and this is much smaller than the total N, which should have been done if I wanted to understand totally what was happening to the errors. So the above relation, which is given in the previous slide, is given here: N minus n.

So I am going to take N outside so it becomes 1 minus n divided by N the ratio of number of data in a sample divided by the total number divided by actually the total number; the number of m such samples multiplied by N will be the number. So in the denominator, I have got n. I will take N out. It will give 1 minus 1 over N and those 2 Ns got cancelled; therefore I got this. Now if I assume n is small compared to N. That is the assumption we made, which is what is possible practically. You see that this factor is very small compared to 1 and this factor is also small compared to 1. Therefore these two can be neglected and therefore I can say approximately that sigma m squared, that is, the variance of the sample I have collected is equal to sigma squared the variance of the population divided by n. This is a very important formula because knowing the variance of the sample we have collected, we can say something about had we repeated the measurements very, very large number of times.

What would have been the difference?
You can see that 1 over n is coming here; sigma squared divided by n is equal to sigma squared m or if you want to find the variance of the population, sigma squared is equal to n times sigma m squared. So you have to multiply sigma m square by a factor of n to get the variance of the sample. So with this background, let us look at the next question which I am going to ask. The next question is about the sample and its own variance. If you go back to the last slide, you will see: sigma m squared is the variance for that particular sample. I am going to slightly change the symbol and I will call the sample variance from its own mean m_s as sigma e squared. This symbol e is used as an estimator; I am going to use the sample variance as the estimator or estimate for the mean of the population.

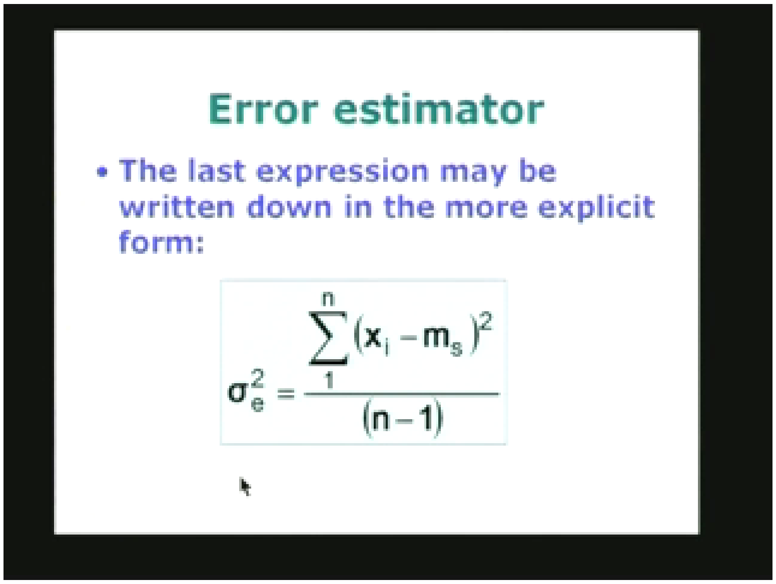
(Refer Slide Time: 19.22)



So the question is how it is related to population variance that means sample and its variance. We have 1 sample containing some number of measured values and its own variance I have calculated. I want to find out how this is related, that means, how sigma e squared is related to the population variance. Actually without proof I am going to show the expression sigma e squared, the value which I calculated from 1 sample and the variance of that sample is given by N (n minus 1) divided by n (N minus 1) into sigma squared and if I take N outside, it becomes 1 minus 1 over n and you see that n is small compared to N. Therefore I can neglect that term. I can write that as sigma squared into 1 minus 1 over n. So sigma squared 1 minus 1 over n is equal to sigma e squared, or you can also see sigma squared is equal to 1 minus 1 over n into sigma e squared is the factor, which is coming. If I have not taken into account the fact that a sample variance is different from the population variance, I would have been ignoring this 1 over n quantity. This 1 over n is the quantity which is going to be extra.

So going back to the expression we gave here: sigma e squared is equal to sigma squared into 1 minus 1 over n, I am recapitulating the definition.

(Refer Slide Time: 21.50)

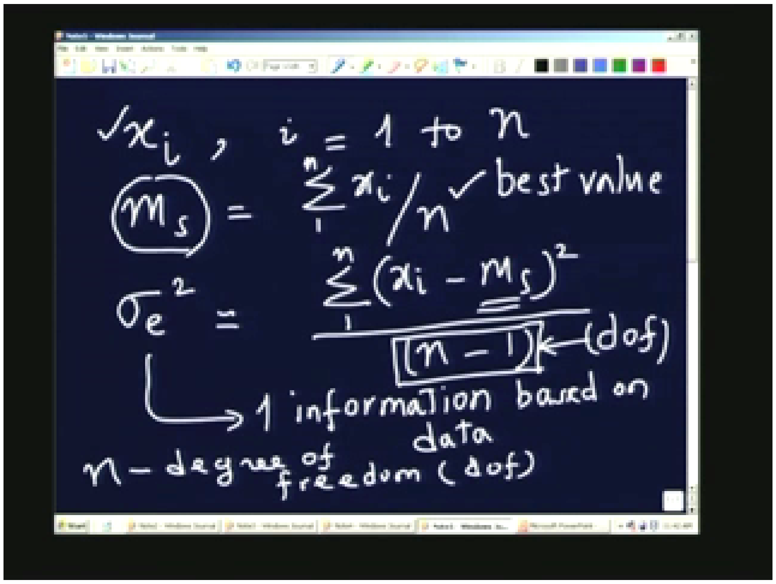


The last expression may be written in a different form, more explicit form. I just want to make sure we understand what we are doing. I have got a certain number of measurement x_i where that is equal to 1 to n. I am going to calculate the variance; I am translating this expression (Refer Slide Time: 19.22) sigma e squared equal to sigma squared into 1 minus 1 over n to this definition here. So sigma e squared is equal to sigma 1 to n because that 1 minus 1 over n is coming.

Here 1 minus 1 over n is nothing but n minus 1 divided by n and that n will cancel off with number of summations in the numerator and therefore n minus 1 is what I am going to get here. That n will cancel off; 1 factor n from the numerator and 1 from the denominator, which came from there. Therefore essentially what I have is the variance, which estimates the error in a single sample is given by sigma 1 to n into x_i minus m_s the sample mean whole square divided by n minus 1. If you recollect, previously we had n in the denominator but now I have n minus 1. Without proof we have derived this relationship based on sampling theory. So I would like to make it physically plausible. Therefore what I am going to do is I am going to make a note on the board so that we can understand or give a physical explanation on the board.

So let us look at the physical explanation. To recapitulate, we have x_i, i equal to 1 to n and the mean m_s equal to sigma x_i 1 to n divided by n. There is no confusion here; however, when I want to calculate the sigma e squared, I am going to write it as sigma x_i 1 to n minus m_s the whole square divided by n minus 1.

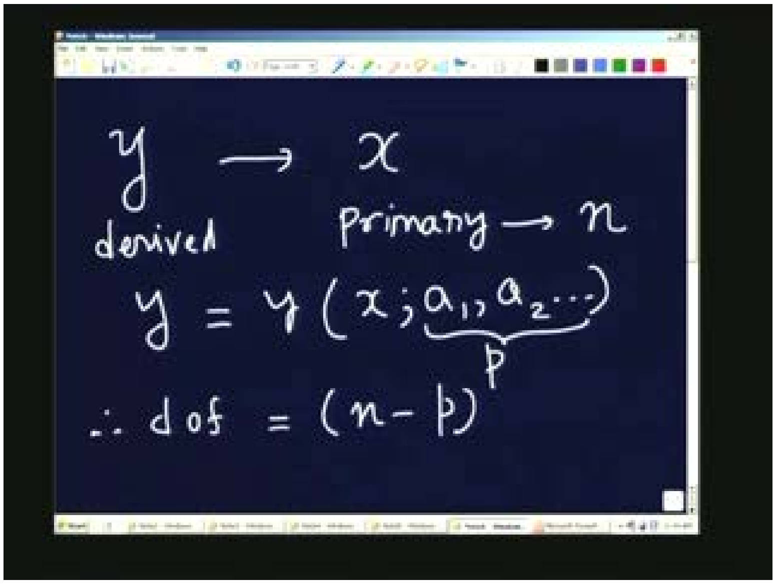
(Refer Slide Time: 26.34)



So the physical argument is like this. So in calculating the mean or estimating the best value if you remember, best value was proved earlier. So the best value is nothing but ms equal to sigma x_i 1 to n divided by n. What is this best value? The best value is based on the measurements available here, x_i i equal to 1 to n and when I am calculating this, I am using m_s, which is calculated using this formula here; that is sigma e squared. So in calculating this m_s, I am using 1 information or 1 unit of information based on data. So I have already used 1 information based on data in the form of m_s. Therefore when I am evaluating sigma e squared I will have only n minus 1 information with me. So we call this n is the degree of freedom to start with, or we'll call also dof, degree of freedom as we use this terminology again and again. Now we can say that this is the degree of freedom available to us.

In other words, the error estimator has to take into account that 1 piece of information has already been derived based on the sample and therefore, to that extent, the sample is already used and therefore the number of degrees of freedom was n. That means n data were available. It is as though n minus 1 data is available to me because 1 information has been obtained using this data. We can in fact generalize this to a case. Let us say I am measuring data which is in the form of a relationship between 2 different quantities.

(Refer Slide Time: 29.02)

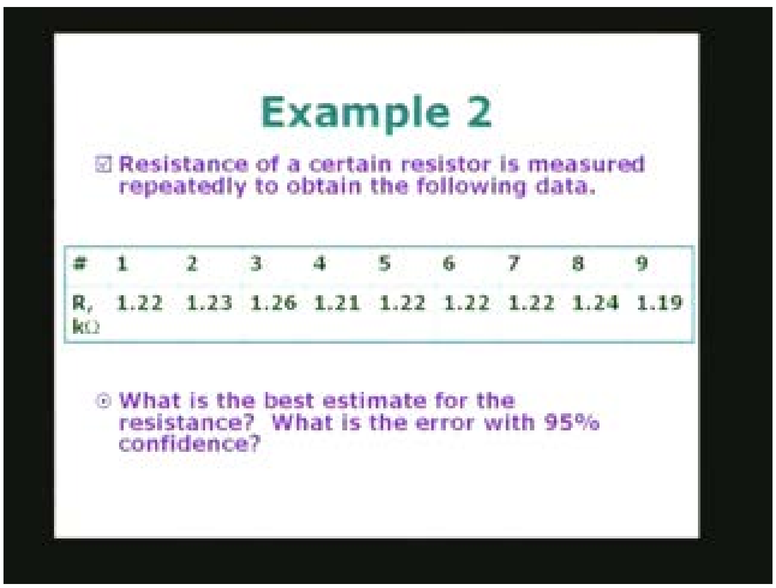


Suppose I have a derived quantity y. x is the primary quantity. The idea is now to generalize what we have said till now to measurement of 2 quantities related to each other. So when I measure x, I can find out y and this can be symbolically written as y equal to y of x, and I will write a set of parameters a_1, a_2. So we can have the number of parameters as p, number of primary measurements, and n. That is the sample I am talking about. In the previous case p was exactly equal to 1; I just calculated the mean of the values. p was 1 single parameter, which was derived by using n values. In this case I am going to use the information given to determine p number of parameters. Therefore degree of freedom is n minus p; in the earlier case it was 1, y was actually the mean of the values. In the previous case y could have also been something else, which is dependent on x.

For example, I may want to determine the square of x, I may want to find out the estimate of the square of x. So I will be getting another parameter, which will be square of the parameters. Therefore each time you are going to use the data to obtain some parameter which characterizes it, I will be losing 1 degree of freedom and therefore if there are p number of parameters as shown here, a_1 to a_p, then the number of degrees of freedom that is lost is p and therefore you get the degree of freedom lost as n minus p. Therefore, when you use the variance formula you have to divide it by n minus p. Therefore, the sampling theory shows or says that if you evaluate the number of parameters p, with n number of data available, after repeated measurements the variance is actually bigger than you think and that is what you will see.

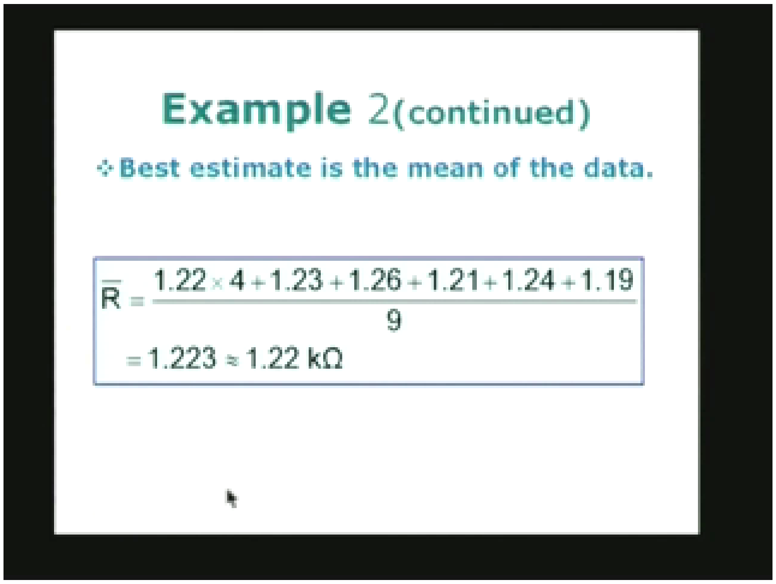
In fact I am going to take an example, which clearly indicates this. In fact I am taking an example, which was taken in the previous lecture. The same example where I measure the resistance of a certain resistor again and again: the data is exactly the same, no difference. But I am going to interpret now in terms of what we learnt from sampling theory. So with the resistance a certain resistor is measured to repeatedly obtain the following data. Number represents the experiment number in this case 1, 2, 3 and each one of these, I have measured the resistances as 1.22, etc. I have already discussed this example; therefore we need not go through all of them so 1.22, 1.23, 1.26, 1.21, 1.11, 1.22, 1.22, 1.22, 1.24, 1.19. These are the individual values of the resistance obtained in the experiment. So what is the best estimate?

(Refer Slide Time: 30.31)



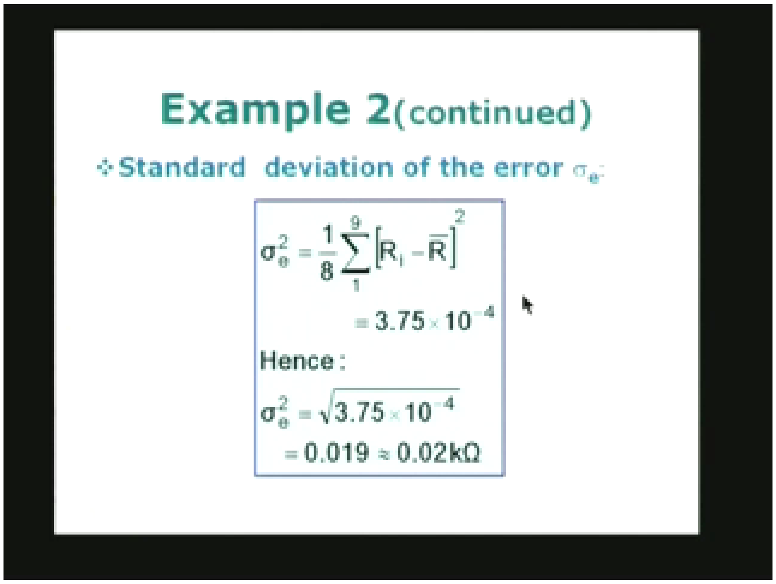
Of course we know that the best estimate is nothing but the mean of all these values and that doesn’t change from the previous lecture to this lecture, but what is the error with 95% confidence? That is the one which is going to change. So the best estimate is the mean; it was also obtained in the last example.

(Refer Slide Time: 31.39)



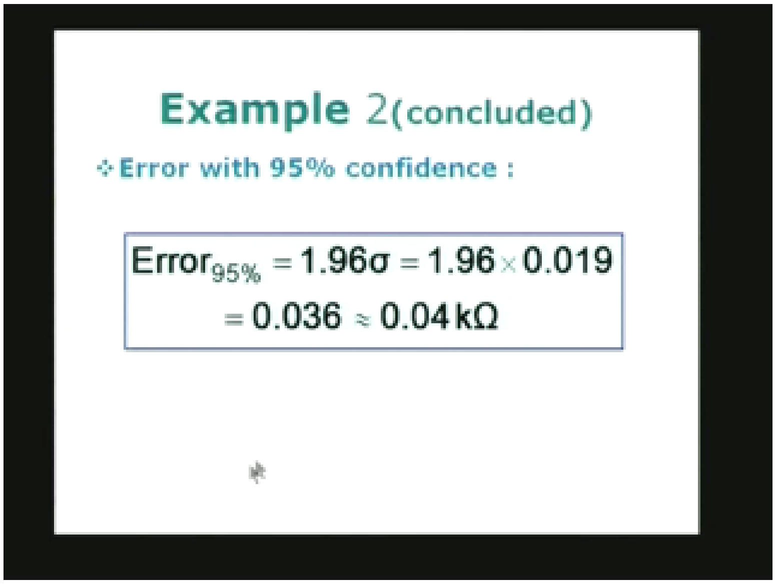
Example 1 contains the same numbers. In fact I have taken the slide from the second lecture so the value of R bar is 1.22 kilo ohms. Now I have changed the symbols here. I have used the estimator sigma e and then sigma e squared is now 1 over n minus 1 n is 9, 9 minus 1 is 8, 1 over 8 of sigma 1 to 9 instead of 1 over 9. I have taken 1 over 8 sigma 1 to 9 into R_i minus R bar the whole squared. R bar is nothing but the sample mean in the previous terminology R_i minus R bar whole squared. This gives you slightly more: 3.75. In the earlier case it was 3.33. It has increased to 3.75 into 10 to the power of minus 4. Therefore I can obtain the sigma e; it is given as sigma e squared, which is equal to 0.02 kilo ohms.

(Refer Slide Time: 32.01)



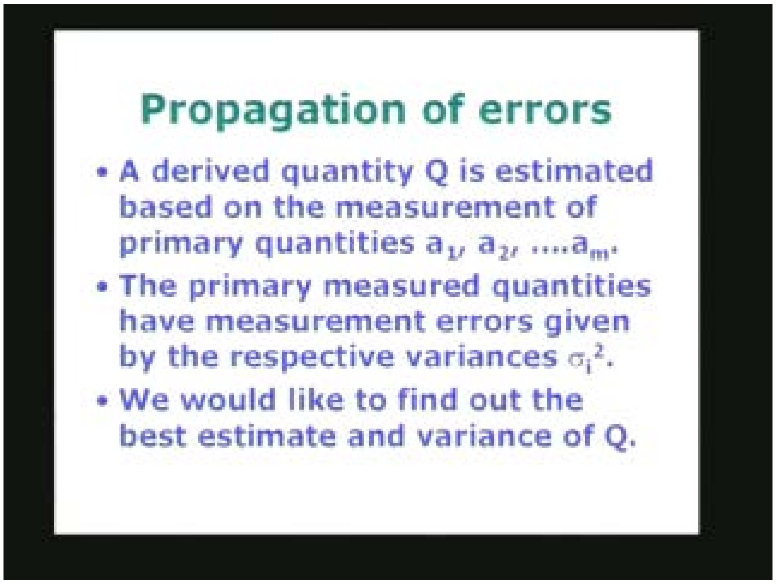
So sigma e is the square root of this number, which is 0.02, it is 0.19, which I am rounding it out to 0.2. In the previous case also, I have rounded it off to 0.2, which was smaller than this. Therefore the final answer does not change, 1.96 sigma, is roughly equal to 0.4. Actually it was 0.36 in the earlier case; it was slightly smaller than this.

(Refer Slide Time: 33.11)



Therefore, you see that the estimated value of the error is larger when you take into account the results from the sampling theory. Just to round off what we should do is, in the future whenever we give an example, we will use the results derived from the sampling theory. That means we are going to divide the variance by the degrees of freedom but not the number of samples, number of measurements. It will be n minus p, where n is the number of times we repeated the measurements. p is the number of parameters estimated. In fact when we go to regression analysis, it will be more clear as to what we mean by number of parameters, but right now let us remember division is by n minus p but not n itself. That means the error is actually larger than what you would imagine. So let us look at the second thing which we mentioned in the first slide, the question of propagation of errors. Let me explain briefly what is happening here.

(Refer Slide Time: 34.19)

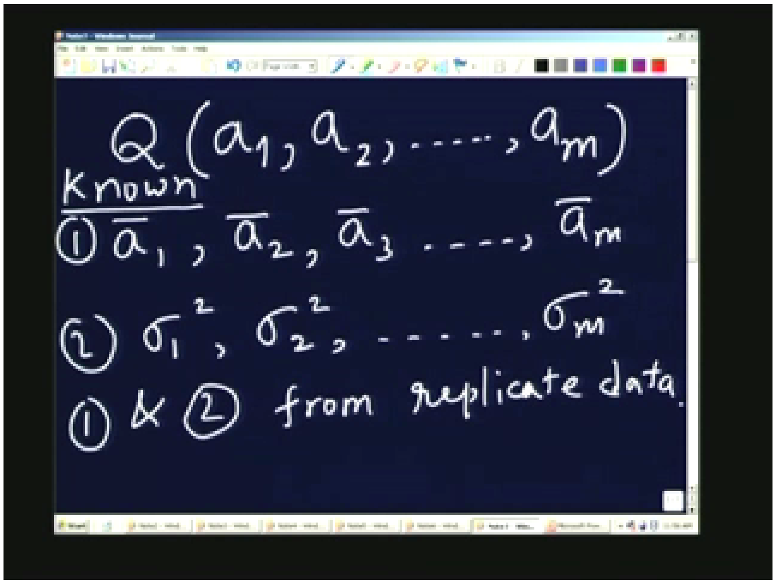


A derived quantity Q is estimated based on the measurements of primary quantities a_1, a_2, etc., a_m. The number of primary quantities measured, a_1, a_2, a_3 etc. a_m; each one of these is repeated again and again. That means a_1, a_2 etc. a_m are all measured again and again and we have some idea about the statistical variation of these quantities also. What I want to know is to find out how the statistical quantities, a_1 up to a_m are going to affect the statistical variation of the Q. So the primary measured quantities have measurement errors that have already been characterized given by the respective variances, sigma i squared a_1 has got sigma 1 squared, sigma 2 squared corresponds to the error in a_2 and sigma m squared corresponds to the variance in this particular parameter I am measuring.

Sigma 1, sigma 2 ... are the standard deviations of the measurements of each one of these quantities, a_1 to a_m. Now the question is, what is the best estimate and what is the variance of Q? The question is very simple. I have got a_1, a_2 etc. measured again and again; therefore, I should be having the mean values of a_1, mean value of a_2 and mean value of a_m. That means after doing them large number of times or using the measurements large number of times, I have got mean value for a_1, mean value for a_2 … and so on up to a_m and I also have, as given here, variance for each one of these quantities. So the propagation of errors means there is certain error in the value of a 1. How is it going to propagate or make itself felt in the case of Q? There is certain error in a_2. How is it going to affect the error in Q? There is certain error in a_m. How is it going to affect Q? So the error in each one of these is going to influence or affect the measured value of Q. Therefore this process of error migrating or moving or shifting or propagating from a_1 to Q or a_2 to Q or a_m to Q is what is called the propagation of errors.

So what I will do is I will use the board to work out the details of the relationship between the error in the measured quantities or primary quantities and the error in the derived quantity. So the end product of this exercise would be a formula, which will help us calculate the error in Q. Let us just recapitulate what we have done. So we have measured a quantity Q, which actually depends on several primary quantities: a_1, a_2, etc. a_m. I also have a_1 bar, a_2 bar, etc. a_m bar; these are also known. We also know sigma 1 square, sigma 2 square, and sigma m square. How do we know these things? We know 1 and 2 from repeated or replicate data.

(Refer Slide Time: 39.35)



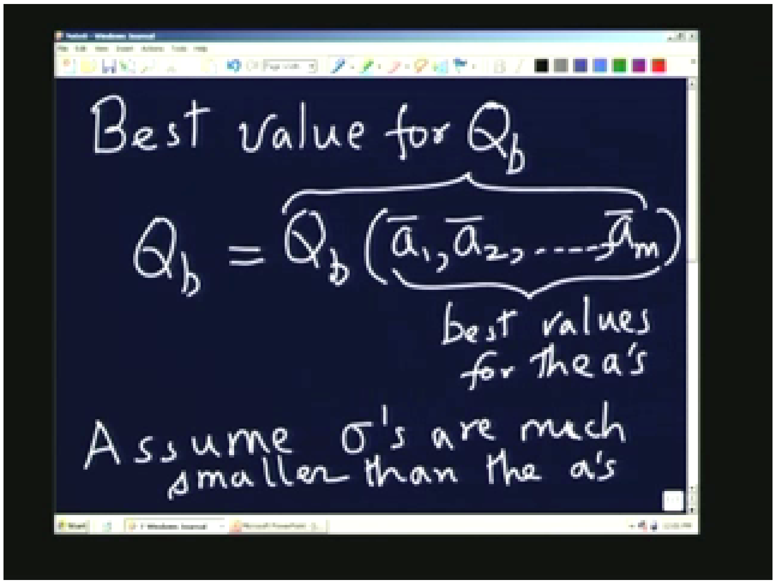
I want to digress a little bit here and talk about these errors, sigma 1 square, sigma 2 square, and so on; what is going to happen in natural practice? That's why it is sometimes confusing when I say something mathematically, then I go to the laboratory and do something else totally different. So this is a problem for which there is no solution; only little bit of thinking about it and understanding what is going on would probably help us. In practice what happens is, I am not going to repeat the measurement again and again.

Sometimes I may take only 1 reading or may be 2 or 3 readings at the most and then I would like to find out what is happening. But in making the measurement, I am making use of some instrument and we know that the instrument can resolve some, basically the smallest quantity I can measure. I can resolve between 2 values, which are only that close and no closer. So sometimes what we do is even though it may not be mathematically correct, what we do is instead of using the errors obtained from replicate data we may use the maximum error due to just the instrument itself. Instrument itself cannot resolve a certain quantity better than some value and therefore the resolution of the instrument may be used sometimes for the estimation of errors. So sigma 1 square, sigma 2 square, and sigma m square may be taken by the behavior of the instrument.

This is not strictly like doing experiment again and again and understanding the statistic, we are replacing the statistical behavior of the error by looking at the errors introduced in the measurement process due to the limitation of the instrument. This may be justified in a physical sense that the measurement process involves the process of taking a reading. You may take the reading yourself using your eye sight or you may ask somebody else to take the reading or it may even be gathered by data acquisition system. The measurement process can be any thing; in each one of the cases, there are certain minimum resolutions, the smallest quantity with which we can resolve and we cannot resolve better than that and therefore we can say physically that probably that is the kind of error we are going to introduce in measurements. Therefore I am going to replace the statistically determined errors sigma 1 square, sigma 2 square etc. By the errors due to the measurement process, which involves a certain instrument and certain process of getting the data and there is an error introduced in that, which is considered as accidental. So it is a certain approximation involved in this certain hand waving is involved but I think we all try to do that all the time.

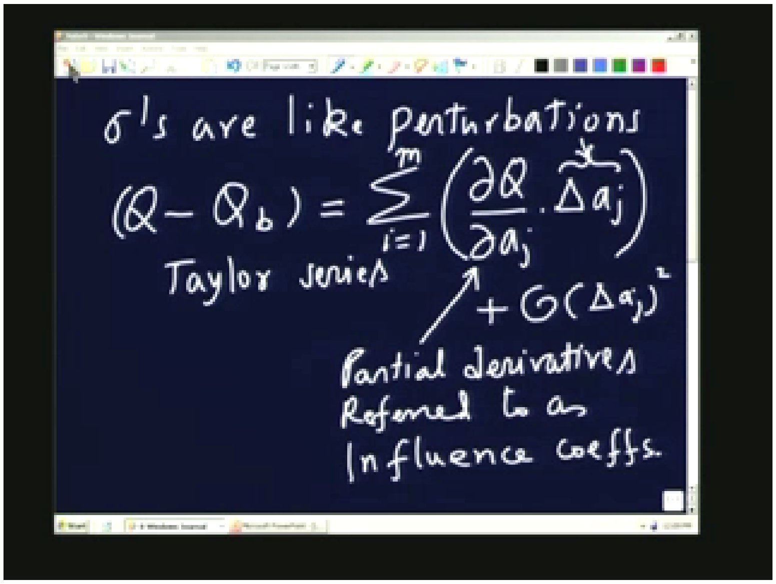
So with this background, the question I am asking now is how to find sigma Q square? So, to understand this problem, let us indicate how we estimate the best value for Q. For determining the best value Q_b, I will say Q_b is the best value based on a_1 bar a_2 bar and a_m bar. That means I am going to use the best values a_1 bar a_2 bar and a_m bar are the best values of the a's. There is no need to prove this because we know that a_1, a_2 etc. a_m are the mean values or the best values for these, so I am just assuming and guessing or claiming that the best value of Q is nothing but the value obtained by using the best values for each one of these measurements.

(Refer Slide Time: 44.00)



There is no proof for this, we are just asserting this and this may be any function relationship, Q_b related to a_1 etc. through some mathematical expression. Now let us assume, it is a very important assumption, sigmas are much smaller than the a's. Most of the times we can justify this. We assume that the errors in the measured values are not very large compared to the values of the measured quantities themselves. This is a very important assumption. If they are not, then whatever we are going to derive is not going to be applicable: as simple as that.

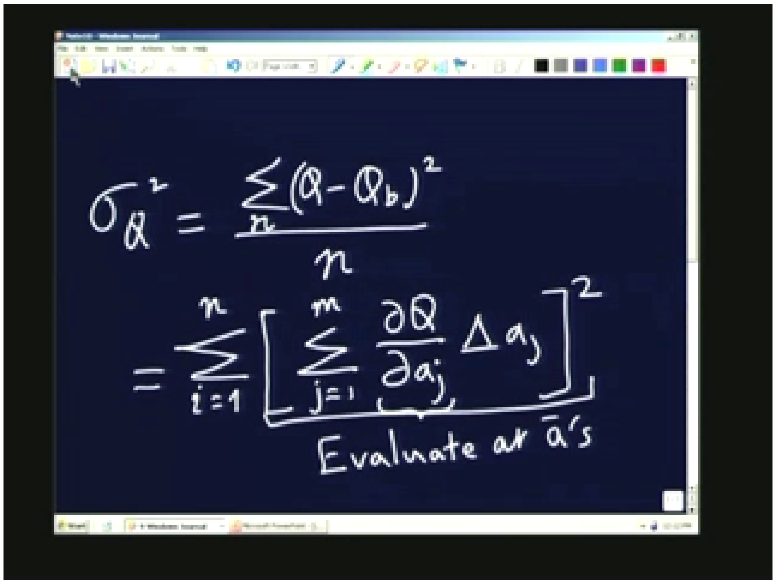
(Refer Slide Time: 48.30)



So with this assumption I am going to assume that the perturbations, sigmas are like perturbations and therefore I can use a Taylor expansion to decide what is going to be the perturbation of the value of Q. So if I take Q minus Q_b, Q_b is the best value which is obtained by the values of the as given by a_1 bar a_2 bar etc. and now I am going to perturb the value because whenever a_1 is perturbed Q will be perturbed from the value of Q_b. When a_2 is perturbed Q will be perturbed from the value of Q_b. Therefore this can be written as a Taylor expansion. This is a Taylor series; of course Taylor series is valid for infinitesimal changes and that can be written as sigma i equal to 1 to m partial of Q with respect to a.

In fact I am showing only the first order terms, of course, plus terms of orders delta a_j square. So these are the perturbations and these are the partial derivatives because you notice that the function Q is a function of several variables and therefore the Taylor expansion will contain derivatives with respect to each one of these variables in turn and therefore, you get partial derivatives. If you write this for a single variable you would not get a summation here and these perturbations are the ones given to us. We will see how it is incorporated later on. These are partial derivatives and we also refer to them as influence coefficients. The reason they are referred to as influence coefficients is delta a_j is some perturbation, Q minus Q_b will be perturbed by a product of delta a_j multiplied by this coefficient *doh* Q by *doh* a_j. So the perturbation is multiplied or enhanced by the magnitude of *doh* Q by *doh* a_j: the larger the influence coefficient, the larger the magnification; therefore the larger the influence of a_j on Q. That is why it is called as the influence coefficient.

(Refer Slide Time: 50.51)

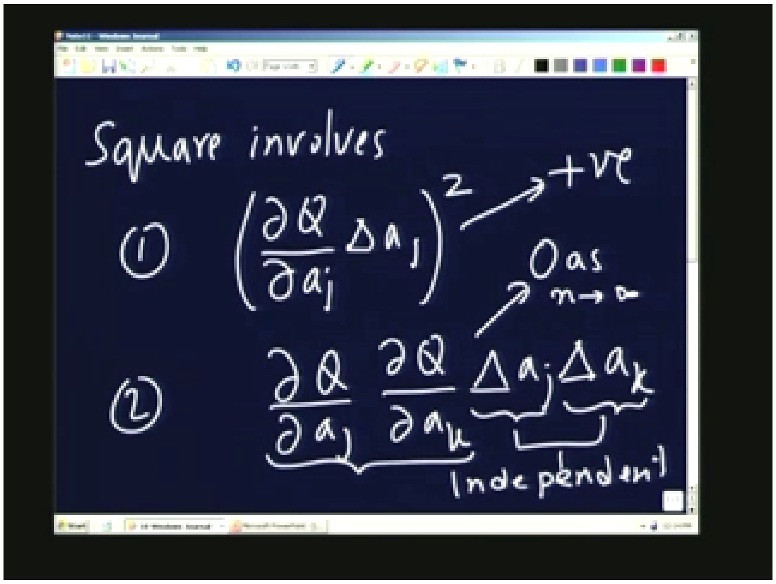


Now what we have is Q minus Q_b is equal to this and what I am interested in finding out is the variance of the Q and we will just use the definition of the variance which is given by sigma q squared must be equal to sigma Q minus Q_b the whole square divided by the number of times the measurements were done. So I will say sigma over n for concise. And this can be written as sigma; I will just keep the sigma outside. This is the experiment number; so I can say i equal to 1 to n Q minus Q_b the whole square.

I will write the expression derived earlier, so sigma j equal to 1 to m *doh* Q by *doh* a_j multiplied by delta a_j whole square. So we have 2 sums, one is the number of times the measurements have been done, another one is over the number of quantities a_j, which are involved in the process. And the question is how do we evaluate the influence coefficient evaluate at a bar's *doh* Q by *doh* a_j? This will be evaluated because that is the point which is known. The derivatives are all evaluated at the best value so I am going to replace the derivative also at the best value. That means *doh* Q by *doh* a_j are going to be evaluated at the set of values a_1

bar, a_2 bar etc. up to a_m bar. Now, if I square this, I am talking about this quantity (sigma Q square value) the square will involve 2 types of terms like x plus y plus z the whole square is x square plus y square plus z square plus 2xy plus y, etc. There are two types of terms: one type involves square of individual term and the other one involves the products of the term, so let us look at the two types of terms.

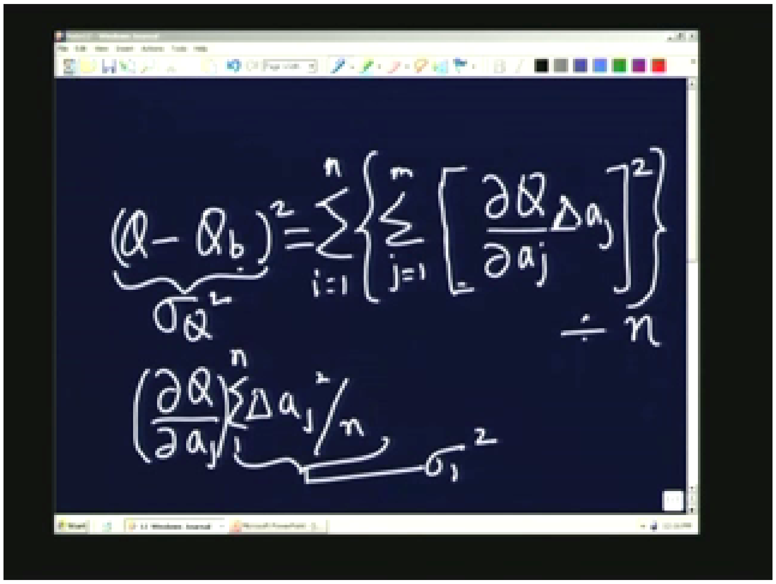
(Refer Slide Time: 53.39)



The square term involves for example *doh* Q by doh a_j whole square is one. The second one will involve delta a_j, delta a_k. This is simply a coefficient, a certain number. When we evaluate at the best point this is some number or quantity or some value. These are the individual errors. Now I am going to make the following assumptions that the error in delta a_j is not related to error in delta a_k, and that a_k and a_j are independent of each other.

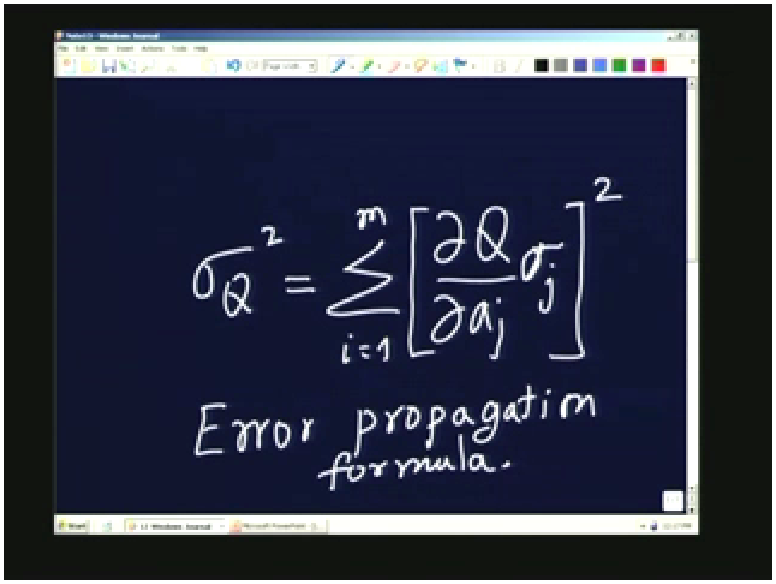
For example, in practice I may use a voltmeter to measure the voltage. a_k may be the voltage; I may be measuring the length using the Vernier caliper to measure delta a_j. So how are they going to be affected by each other? These errors are independent. So what will happen when I have the product of 2 quantities which are not related to each other? This will tend to 0 as n tends to a large number. If I measure again and again the product of 2 quantities, which are not related to each other, it will tend to 0 whereas the square quantity is always positive.

(Refer Slide Time: 53.39)



Therefore with this we can see that what we are going to do is to have Q minus Q_b the whole square is equal to sigma i equal to 1 to n the number of experiments. This is the square of the error and if I divide this by n, I should get the variance. So I have to divide by n on the right-hand side. So what is this quantity, this whole thing *doh* Q by *doh* a_j delta a_j whole square? I can take one of those terms *doh* Q by *doh* a_j delta a_j square and actually this is also a summation sigma j equal to 1 to n divided by n, because this n is coming from there. This is nothing but sigma 1 squared. So I can relate it to sigma Q squared and therefore I can say that sigma Q squared equal to sigma i equal to 1 to m *doh* Q by *doh* a_j multiplied by sigma j whole squared. This is called the error propagation formula.

(Refer Slide Time: 57.04)



I think we will have to stop here and we'll resume from here in the next lecture. We will take an example, which will be given at that time, and then move on to the question of regression analysis as a sequel to this detailed look we are having at the question of errors in measurements. How to characterize them? So we have come quite far from where we started. Now we understand the sampling theory. Its results are known to us. We are able to look into propagation of errors and legitimately the next part is to look at the relationship, which means regression analysis which is description of this. Thank you.