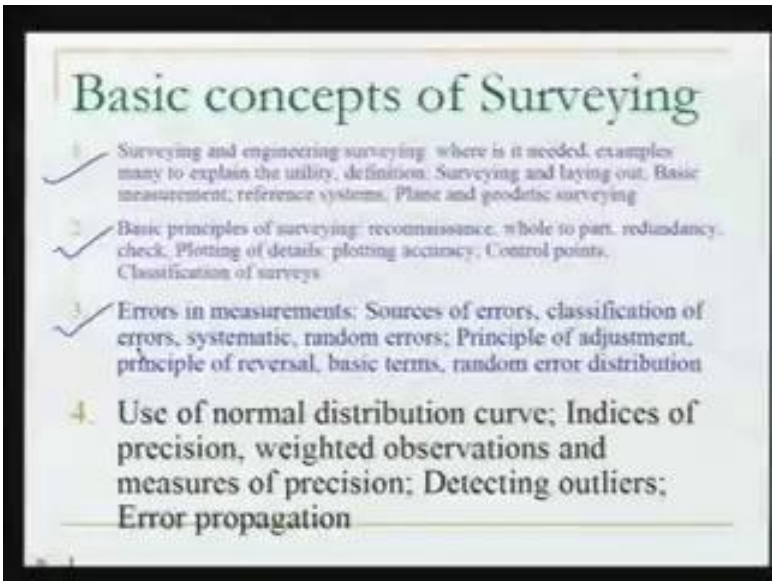
Surveying
Prof. Bharat Lohani
Department of Civil Engineering
Indian Institute of Technology, Kanpur

Module - 2
Lecture - 4

Basic Concepts of Surveying

Hello! Welcome again to our this video lecture on basic surveying. Today, we are in module 2 and in lecture number 4, and the name of the module is 'Basic Concepts of Surveying'. Well, in the entire structure of our this video lecture, we are here at the moment in the module number 2.

(Refer Slide Time: 00:54)



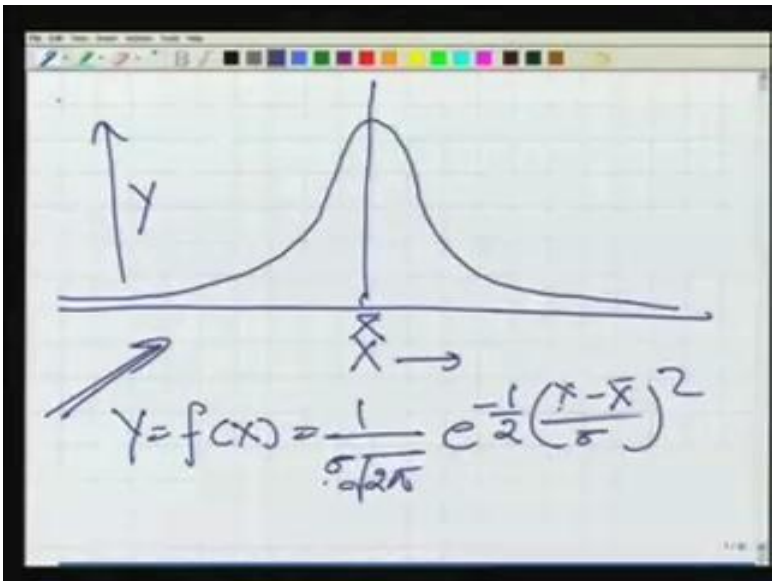
What we have done so far, we have discussed lesson number 1, 2 and 3, and there in the three, we have talked about various things, particularly about the errors in the measurement, because surveying, as we said, is about measurement, and whenever we are taking a measurement, we are introducing - there is some errors coming in to the observation. So, we need to understand about those errors. What are their sources - this is what we have discussed - we saw the sources could be natural, could be the human-induced - the personal one, also the instruments which we are using may result in some kind of errors. We saw some examples also of that. Then in the classification of errors, the errors could be, you know, the blunders which we commit - we saw it -

instead of writing, maybe, let us say, 31.5, what I am writing, I am writing 3.15. So, things like this may happen, and these are the blunders; mistakes. Then, the second one: there are various physical forces which are acting - the temperature may be more, the gravity is working, the wind is blowing, and many of these factors, they contribute some errors, again, in the observation. But about these errors, we can quantify them; you can write a mathematical model; you can write - because we know the physical law behind these errors. So, from our observations, we can eliminate those errors. These errors which can be eliminated by writing a mathematical law are called systematic errors. Then, next we discussed was about the random errors. From your observation side, you have eliminated mistakes or blunders - taken away them. We have also taken away the systematic error component, but still your observations are not true values. We saw that it is never possible to observe the true value. So what we saw? At this stage, the errors which are there in the observations are random errors. We saw the distribution, what kind of distribution is there for the random errors - they are equally in positive and negative sides; small errors, the small errors have got high probability, while the large errors on the side of the curve - we saw it - has got less probability.

Then we talked about principle of adjustment, when the errors are there in the observations many times, you know, our network there in the field. For example, we had an example of the triangle. So, the geometry from the field says the sum of those three angles in a triangle which you have measured should be 180, but we saw it is not so, because in each and every theta which we measured - the angle - we had some error. So, we need to adjust our observations so that they conform to the geometry which is there in the field. Well, we did not talk about this principle of reversal - we will talk about this principle of reversal later on, when we are talking about some instruments; adjustment of the instruments. Then, we saw some more basic terms: the true error; the percentage error - relative error; one very important concept we looked at was the most probable value - a most probable value is a value which has maximum chances of being nearest to the true value. Then, we saw how we can find it; we proved it that the most probable value for a set of observations is the arithmetic mean. Well, towards the end of the lecture, we were looking at the random error distribution, and we saw that the random errors are normally distributed; we saw the bell-type curve.

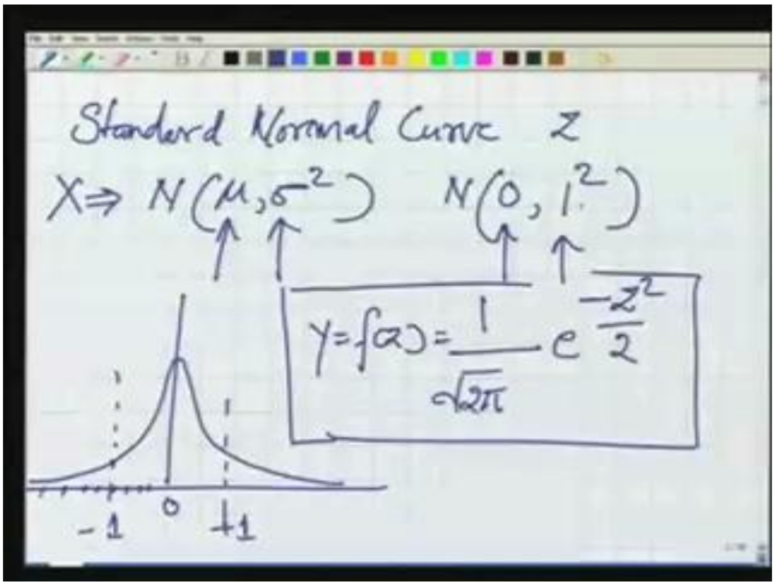
So, what we will do today in our today's lecture, we will again look into a little detail of random error distribution, while mostly, we will be talking about use of normal distribution curve because we know our errors follow this pattern; we can assume it very safely. So, we will try to make use of this normal distribution curve - we will have some indices for the precision; also, we will go for the weighted observations, what are the weighted observations and their measures of precision, then we will try to see - if the time permits, in this lecture - that how we can detect the outliers, and what is here? - what happens in case of the error propagation. So this is all we are going to talk today. Now, what we will do, we will look a little bit - of course, we have seen it before also - into the normal curve.

(Refer Slide Time: 05:44)



The normal curve, as we saw yesterday - or maybe in our last lecture - was like this (Refer Slide Time 05:48). Any variable 'x' may have an arithmetic mean which you can determine, and the probability of occurrence of this x is 'Y'. So, this Y is written as 'f(x)', which is given by 1 by sigma 2 pi - you can determine this sigma, standard deviation, we saw the formula for that - e raised to power 1/2 of x minus X bar upon sigma square (Refer Slide Time 06:07). Okay, our observations in surveying, our errors, random errors - and this is true only if, from our observations, the blunders, mistakes had been taken away; those observations which were blunders or mistakes, they have been taken away, and also from our observations, the systematic errors have also been taken away. Then only - if only the random errors are there, then we can safely assume that our observations will follow this pattern or this curve. Now, what we will do, we will try to see how we can make use of this.

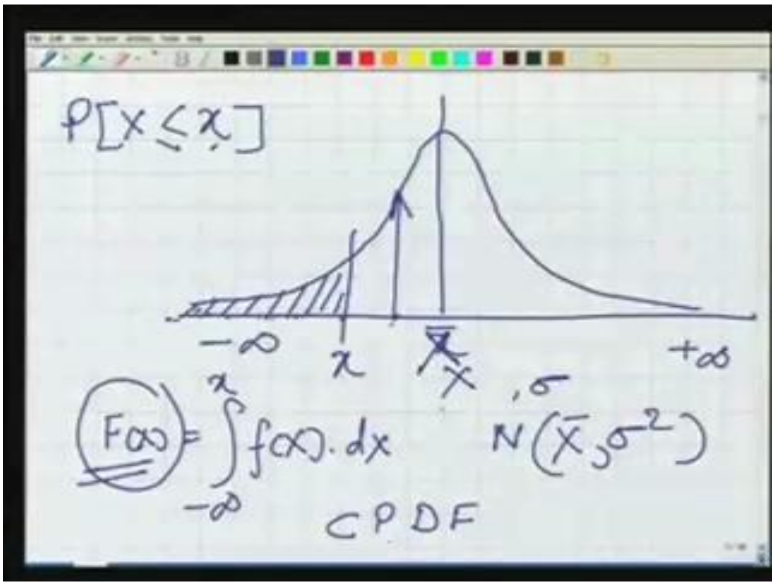
(Refer Slide Time: 07:17)



Before that, what we will discuss, we will discuss a 'standard normal curve' or 'standard normal distribution'. Now, what is this? Generally, a normal distribution is written as 'mu sigma square' (Refer Slide Time 07:29). The meaning is, the distribution for a variable that we say'x' is normal, and the normal distribution is given by two parameters: the mean, which is the most probable value, and the standard deviation. So, these two, they define the normal distribution curve. For a standard normal distribution, we write it as 0 and 1^2 (Refer Slide Time 07:58). So, you can understand, for a normal - a standard normal curve, if the variable is 'z', the mean of the variable is 0 and the standard deviation is 1. Well, I can plot it also - if you plot it like this (Refer Slide Time 08:21), the meaning is, this value is 0, while the standard deviation, we will see in a while, is -1 and +1 (Refer Slide Time 08:29) here. We will see these sigma terms - how they are coming here, what this point is - we will look into this in a while. So, for a standard normal curve, what we can do, we can do slightly more - we can write it as 'f(z)', as in the figure here. The equation of that: 1 by sigma is 1, and root 2 pi, e is to power $-z^2$ by 2 (Refer Slide Time 08:53), because in this case, the mean is 0. So, it reduces to this (Refer Slide Time 09:17). What we will do, we will make use of this standard normal curve, because for a standard normal curve we can compute the probability of occurrence of these values here.

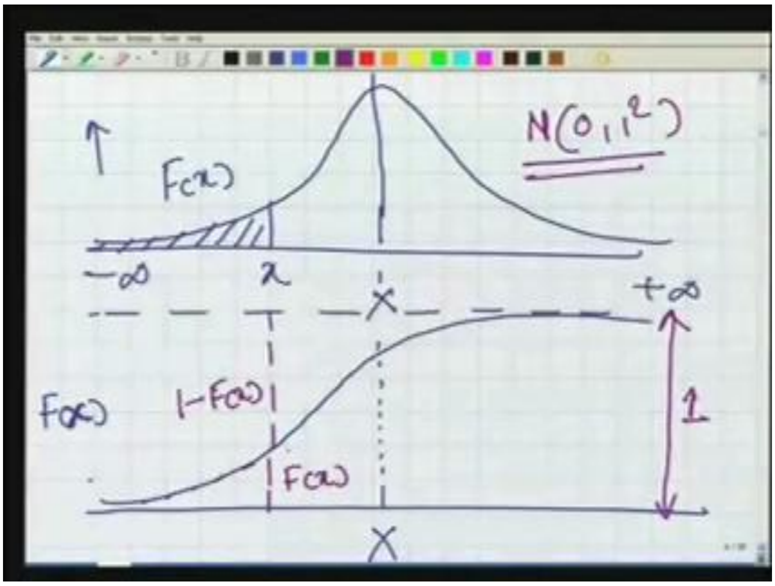
There are some standard tables - we will try to make use of those tables in our computations, so this is why we need to discuss this.

(Refer Slide Time: 09:52)



But before that, what I will do, I will write a term here. For example, let us say there is a variable X for which the mean is - sorry the mean is X bar, and the standard deviation is sigma. So, we can write this normal distribution as X bar and sigma square (Refer Slide Time 10:03), and this - we will show like this, where this X may vary from minus infinite to plus infinite (Refer Slide Time 10:12). Now, I write a term here: I write that probability, that when my variable X is less than or equal to a value x (Refer Slide Time 10:28). What is the meaning of this? I am looking if this small x, this value, is somewhere here (Refer Slide Time 10:44). What I am looking at? I am looking at: now, by writing this, what is the probability of occurrence of my variable X so that it is smaller than the value x? Now, as you know, in this curve, at any point here, this is the probability of occurrence of this particular value (Refer Slide Time 11:06). What we are looking here for? We are looking here for all the values which are less than and equal to x. So, what is that probability? That probability is basically the area of the curve here (Refer Slide Time 11:22). So, this area, we can also find as - I can write it as, let us say, F(x) is integration of this curve from minus infinity to x, f(X) is my this curve, into dx (Refer Slide Time 11:28). So, what it will do, it will give me the area of the curve in between these two. Now, let us look at this thing in a slightly different way - what is this really? This is telling me the 'cumulative probability' - you know, the probability of occurrence of all the variables, all the values, from minus infinite to x, is this value. So, this is actually the sum of all these probabilities starting from here to there (Refer Slide Time 12:21). It is starting from minus infinite to here at x, because we are taking this area. So, we say this as 'cumulative probability distribution or density function' (Refer Slide Time 12::31), because this is the probability density function, and this is now - if f(X) is being plotted with x, that will become cumulative probability density function. Many times, this is also called as simply 'distribution function'. Distribution function; this f(X).

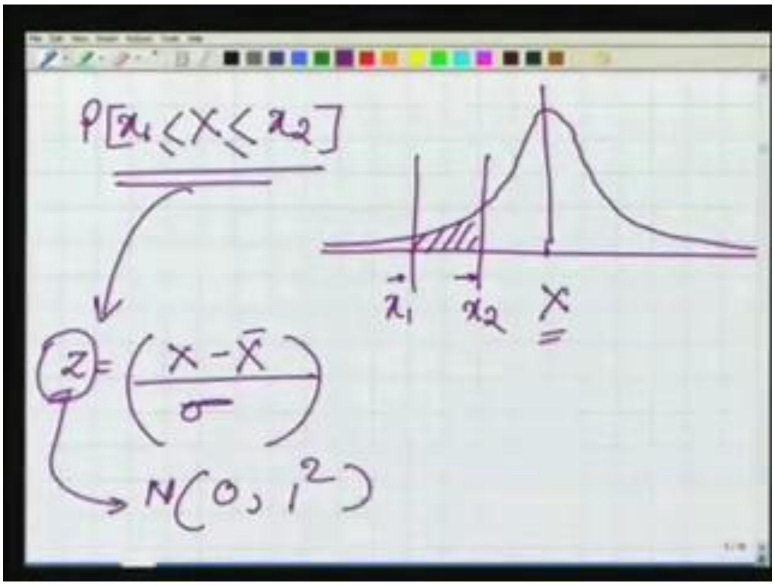
(Refer Slide Time: 13:17)



Now, we will try to plot these two together - I am plotting my normal curve first (Refer Slide Time 13:16) - here is the normal curve from minus infinite to plus infinite for my variable X, this is the probability of occurrence of any value here (Refer Slide Time 13:27). Now, what I am doing, I am plotting, along with this, in the same line here, my another curve (Refer Slide Time 13:35), and this curve is called F(X) for - again, the value is X . Well, what we had done in our last slide, we had computed the value of F(X) at x - this small x - and that was this area (Refer Slide Time 14:07), and the value was F(x), which we said the cumulative probability density function, or cumulative probability density, or cumulative probability as such, we can say. Well, if I plot this x against this X here, what we will get, we will get a curve like this (Refer Slide Time 14:33). Now, you can very well tell me - you can guess it, now - that what will be the value of this (Refer Slide Time 14:46) total one.

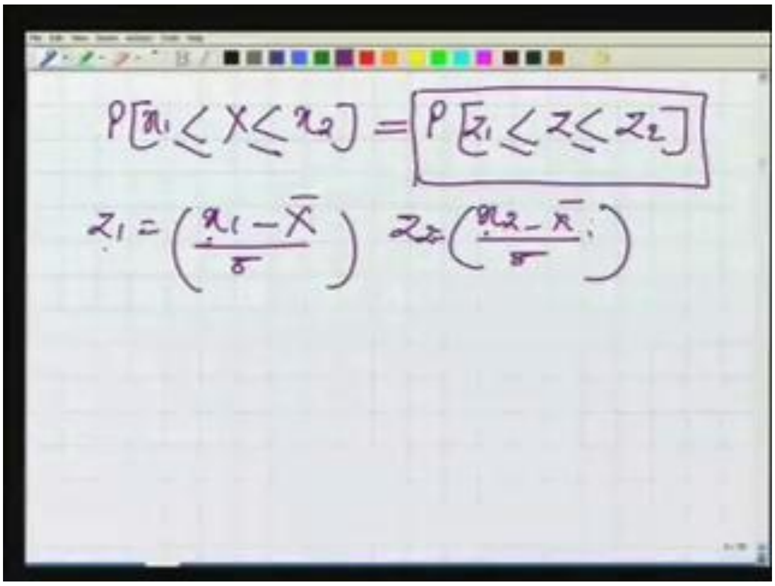
What this total one is? This total one should be 1 - why? Because the probability of occurrence of all these variables - all the values starting from minus infinite to plus infinite - all of them, they will occur; what is the probability of that? The probability is 1. So, the cumulative probability will be 1, while at this place (Refer Slide Time 15:13), this particular value is F(x) at small x, and this value will be 1 minus F(x) (Refer Slide Time 15:22). So, for our standard normal curve, what we are looking at? N - standard normal curve - 0, 1^2 (Refer Slide Time 15:31). These values are known; there are some standard tables where all these values are known against the variable. We will make use of those in a while.

(Refer Slide Time: 16:00)



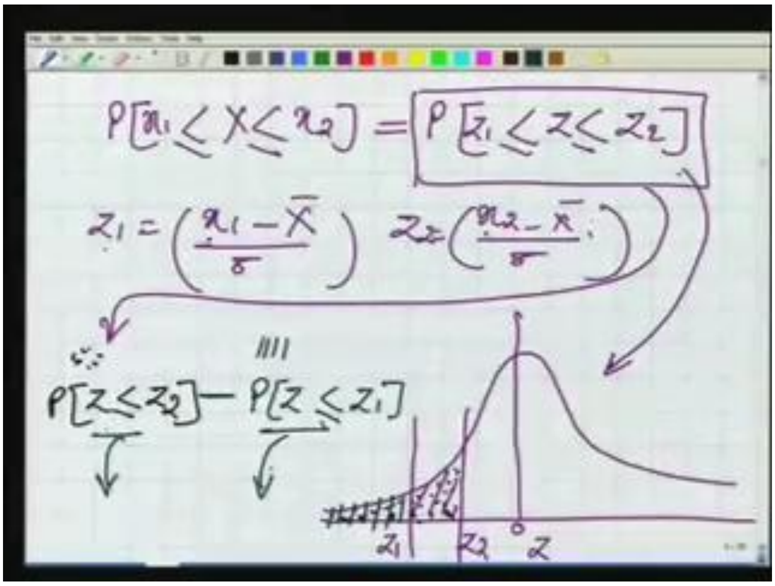
Well, in order to make use of that, I am writing here now, again, for a variable - let us say it is capital X - and its distribution is normal distribution (Refer Slide Time 16:04). The meaning is, from our observations, all the systematic errors and blunders, they have been taken away - only we have got random errors. This is why this distribution will be there. Well, what I try to find, I try to find now, between two values x_1 and x_2 - so, these are two realisations x_1 and x_2 of my random variable X. I am trying to find the probability that my variable is between x_1 and x_2 (Refer Slide Time 16:36). Now, what is the meaning of this, first? Let us understand the meaning. The meaning of this is, what is the probability of occurrence of all the values which are between x_1 and x_2 , and this, as you can see here, is actually this area. So, we need to integrate our curve starting from this point to this point (Refer Slide Time 17:13), and to do that, we should know the characteristics of the curve. Now, what we will do, we will make use of standard normal curve in order to find this value. Now, to do that, what we do, we find as - we have written earlier also - Z. Z is the normalized variable. What we are doing, we are transforming this variable x here (Refer Slide Time 17:40) and this transformation is X minus the variable X bar - X bar is the mean of this variable - divided by sigma - the standard deviation of this variable (Refer Slide Time 17:44). So, using this, we can transform our variable capital X here into the standard variable or into the standard normal curve variable. Now, this particular variable Z has the normal distribution, which is 0 and 1^2 (Refer Slide Time 18:17). Why we did it; what is the use of this?

(Refer Slide Time: 18:34)



To see the use of this: the probability that X -capital X - between x_1 and x_2 - the probability of occurrence of capital X between these two variables is actually same as the probability that the normal standard variable occurs between z_1 and z_2 (Refer Slide Time 18:31). Now, what these z_1 and z_2 are? Obviously, you can find z_1 - again, the transformed variable - and the z_1 will be the transformed variable, x_1 minus X bar by sigma (Refer Slide Time 19:07) and similarly z_2 : x_2 minus X bar by sigma. So, you can transform your values now - the values x_1 and x_2 into z_1 and z_2 (Refer Slide Time 19:14). So, if you have done so, you would like to determine the value of this (Refer Slide Time 19:31). Because we are saying that we - it is possible to find this value from the standard normal curve table. How to find this value?

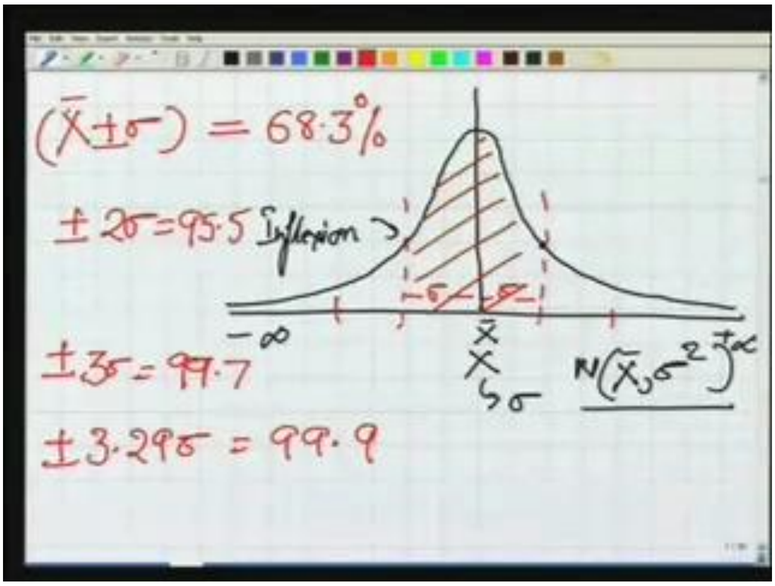
(Refer Slide Time 19:35)



This value, we can write further, and to explain that, I am drawing, again, my standard normal curve here (Refer Slide Time 19:53)- this is for my Z, where it is 0 (Refer Slide Time 19:58). Let us say z_1 and z_2 are here (Refer Slide Time 20:01) - this is my z_1 and this is my z_2 corresponding to my x_1 and x_2 . Basically, what I am interested in, I am interested in this area (Refer Slide Time 20:14). The moment I write this (Refer Slide Time 20:18) is actually the area here, the shaded area; the shaded part. How can I determine the shaded part? To determine the shaded part, well, let us say, if I can determine this entire part which has the dots (Refer Slide Time 20:32), what this will be? The one which has the dots will be the probability that my Z is less than or equal to z_2 (Refer Slide Time 20:44)- isn't it? Then, if I can also determine the area here - if I am drawing by these lines - can I determine this one (Refer Slide Time 21:06)? Let us say the area of this one is the probability that my Z is less than or equal to z_1 (Refer Slide Time 21:09). Now, having known this and this, if I can differentiate, you can find the difference of these two - so this total area minus this area here will give me the area under the curve here (Refer Slide Time 21:27). So, by determining this and this (Refer Slide Time 21:34), I should be able to find this probability. Well, these (Refer Slide Time 21:42) are the values which you can get from the standard normal curve, or from the standard normal curve table, for any value of standard variable. Now, this is standard variable; these are the standard tables of z_1 and z_2 . You can find these two values, you can find the differences, and by finding that, in fact, what you are finding? You are determining the probability of occurrence of X in between x_1 and

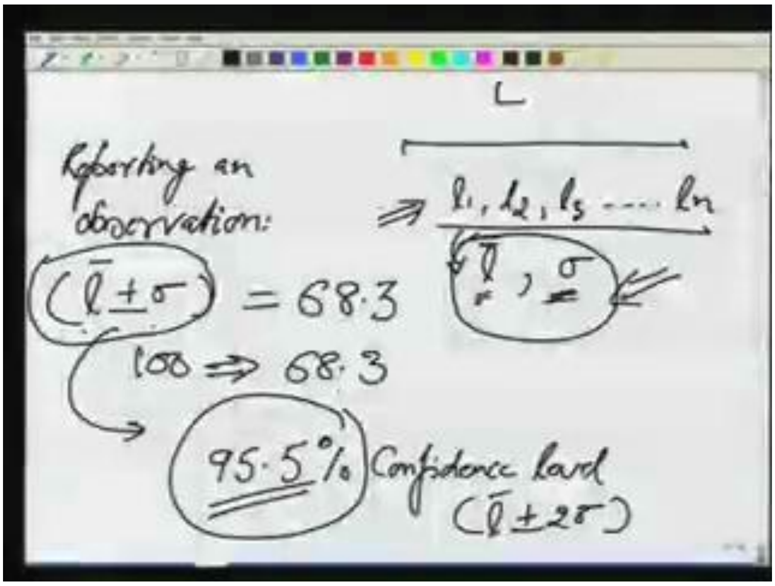
x_2 . What we will do, we will make use of this now. How it is useful to us; why we are talking about it? So, we will make use of this.

(Refer Slide Time: 22:28)



Well, to make use of this, what I will do, I will draw again the normal curve for a variable X, and of course, this is the mean (Refer Slide Time 22:22), and we also know for this X, the corresponding sigma. Then we can say this is normally distributed with the mean of x bar and sigma square (Refer Slide Time 22:40). Actually, this is the conventional way, the standard way, in which the normal distribution is represented by these two parameters; that is why I am writing it every time. Then, our normal curve, the bell shaped curve, which we say from minus infinite to plus infinite (Refer Slide Time 22:57). Now, we want to make use of what we saw just in the previous slide; we want to make use of the - can we - the idea is, can we really find what is the probability of occurrence using the normal curve between two values? Now here, in this curve, this point and this point (Refer Slide Time 23:29) are the points of inflexion - points of inflexion - and these points, they are corresponding to the values of sigma - sigma on this side and sigma on this side (Refer Slide Time 23:43). Well, can I really find the area of the curve which is within mu (Refer Slide Time 23:58) - sorry, this X bar - plus minus sigma (Refer Slide Time 24:06). What is the meaning of this? The meaning is - X bar, minus sigma on this side, plus sigma on this side - so this area which is shaded by red colour here. So, what we can do, we can make use of the standard normal curve and again, we can find this area, and this area comes out to be 68.3 percent. If we are considering the entire area within the normal curve to be 100, then it is 68.3. Similarly, we can also get the area corresponding to plus-minus 2 sigma, because our two sigma will be here (Refer Slide Time 24:49) somewhere, and that comes out to be 95.5 (Refer Slide Time 24:45). We can also find this area corresponding to 3 sigma, which is 97.7 - sorry, 99.7 (Refer Slide Time 24:59). For plus-minus 3.29 sigma, the area is 99.9 (Refer Slide Time 25:14). What is the meaning of these things what I am writing, why we need it, how we can make use of this? We should see this thing now.

(Refer Slide Time: 25:37)

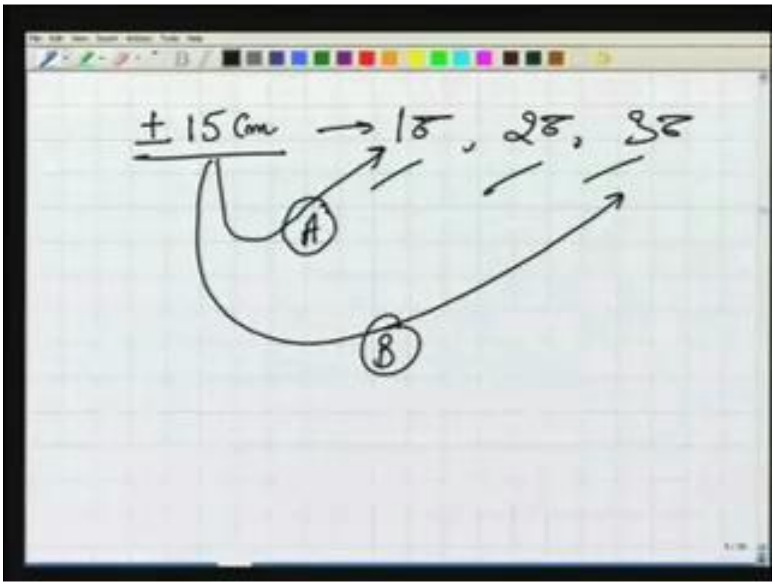


So, to see this thing, what I will do, I will give you one little example here. Let us say we have a length; the true value of the length is L (Refer Slide Time 25:39) - we do not know this; we cannot measure it. It is not possible to measure this. What we do in turn, we take several observations of this length L. For example, let us say l_1, l_2, l_3 and so on, up to l_n (Refer Slide Time 25:54). We take many observations. From these observations I can find 'l bar' - the mean of this - and, as well as, I can find the sigma, the standard deviation (Refer Slide Time 26:07) - you know how to compute the standard deviation. So, once we know these two, also - the important point: also, we are assuming here, the observations - the meaning is, from these observations, the outliers or the mistakes have been taken away. Also, it means, from these observations, if at all there was some error component - systematic error component - it had been also taken away. So, these observations have got only and only random errors within them. Now, in that case, because our assumption is, if only random error is there, then our observations will follow the normal curve; the normal distribution.

Well, after finding this sigma and the mean, how to report this observation? Reporting the observation (Refer Slide Time 27:10) - how do we write it; what do we say about it? Something which we have found, how do we present it? Well, I say 'l bar plus-minus sigma' is my value (Refer Slide Time 27:21) - what does it convey? If you tell - if you measure this length here (Refer Slide Time 27:30), after measuring this length, you find the mean of that, you find the sigma of the observations or standard deviation, and you say 'Well, this is my mean, this is my standard deviation'. What is the meaning of this? What - you know, someone, a third person - what should he understand out of this? What you as an observer, you as a surveyor should understand out of this? The meaning is - because we knew already that 68.3 percent or 68.3 was the area within the curve - our standard normal curve - which was falling within plus-minus sigma. What is the meaning now? The meaning is, we are sure now, we are sure that out of 100 observations if we take, 68.3 observations will be within this limit (Refer Slide Time 28:29). This is what the normal curve says - the normal curve gives you the probability of occurrence of any observation. And we have seen that within plus-minus sigma, from the mean, the probability of occurrence of all the observations in that is 68.3 percent - that, we have seen. So, if I am reporting my measurements like this (Refer Slide Time 28:53), so the meaning is, if I am taking 1000 observations, then - if the observations have got only the random error - 683 observations out of 1000 will be within plus-minus 1 sigma of that.

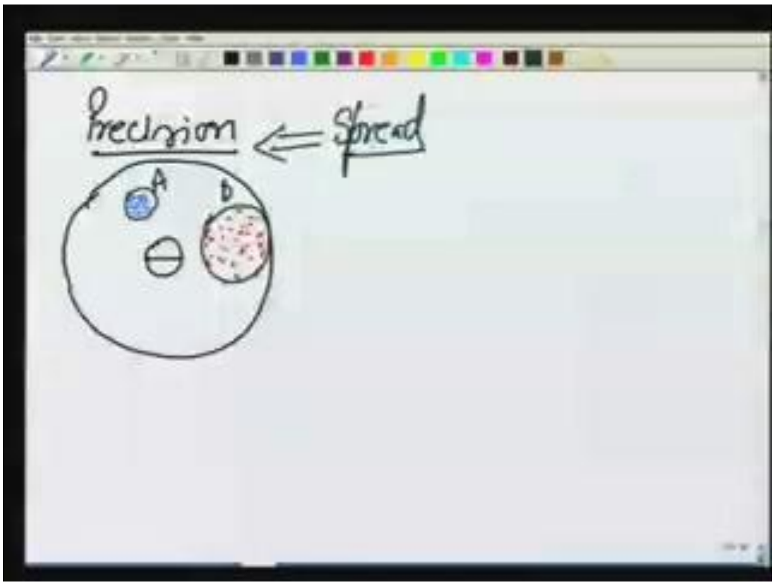
So, this gives you a kind of confidence about your values. For example, let us say, I can say at 95.5 percent confidence (Refer Slide Time 29:18) - this is the confidence level now - at this level, I can say that my observation is 'l bar plus-minus 2 sigma' (Refer Slide Time 29:31). What is the meaning? So many times (Refer Slide Time 29:38), we are sure. For, example, if you are taking a little set - for a population, l bar and sigma (Refer Slide Time 29:50) are the values which you know; the population is defined, or the normal distribution curve of the population is defined, by these two. Then, if you take a little set from that population, the mean of that set, of that sample, will have - in 95.5 percent times, it will be within this (Refer Slide Time 30:15), while this is for the population. So, for any set you take, this tells you about confidence about your observation and this is how, you know, we report our observations. So mostly, you will find - any instrument - if there is an instrument, you are going to buy an instrument, for the instrument, let us say, a simple scale - well, what is the accuracy by which the scale can measure?

(Refer Slide Time: 30:50)



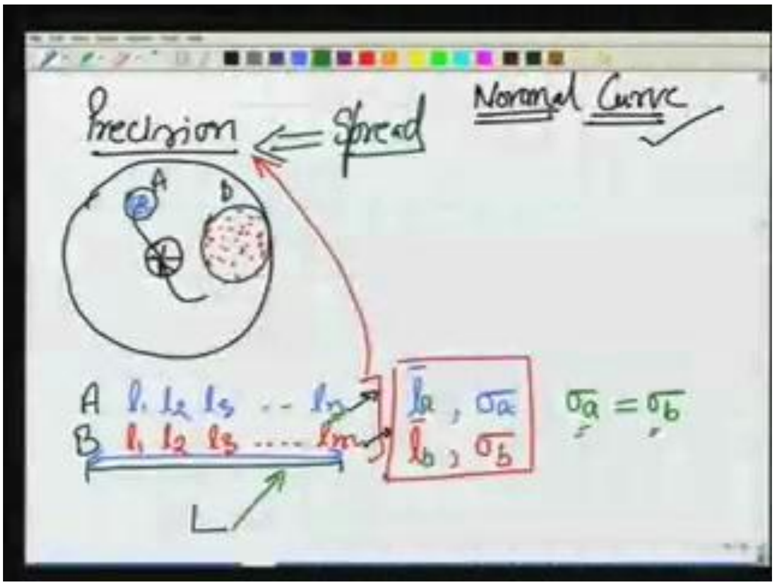
So, the manufacturers, they will say, the accuracy of my scale is, for example - anything - this is plus-minus 15 centimetre (Refer Slide Time 30:48). You have got some instrument which can measure the length, and the instrument says the accuracy of the measurement is plus-minus 15 centimetre. You should ask at that time: what is this 15 centimetre - is this 1 sigma, 2 sigma, 3 sigma or at what sigma (Refer Slide Time 31:07)? This is important because depending on this, you will come to know what kind of instrument is that; how good it is for your purpose. You know, at what - because if this accuracy is given at 3 sigma; also, in the second instrument, it is given as 1 sigma - there are two instruments instrument A or manufacturer A, and manufacturer B (Refer Slide Time 31:34) - they are building the same kind of instruments; the manufacturer A, he says that his instrument can give at 1 sigma this accuracy, while he (Refer Slide Time 31:48) is saying that at 3 sigma, you can get the same accuracy. What is the meaning; which one is more accurate? You will go for one instrument in which case this accuracy is available in more number of observations, which is here. So, we should understand how we are making use of this standard or normal curve: our concept, our observations, are - or our observations are normally distributed; our errors are random errors. So, we should know all these things and then only we can start making the judgement about the observations, about the observations which are reported, because this particular observation or this particular error is reported. So, we can start making use of this - so, this is important.

(Refer Slide Time: 32:54)



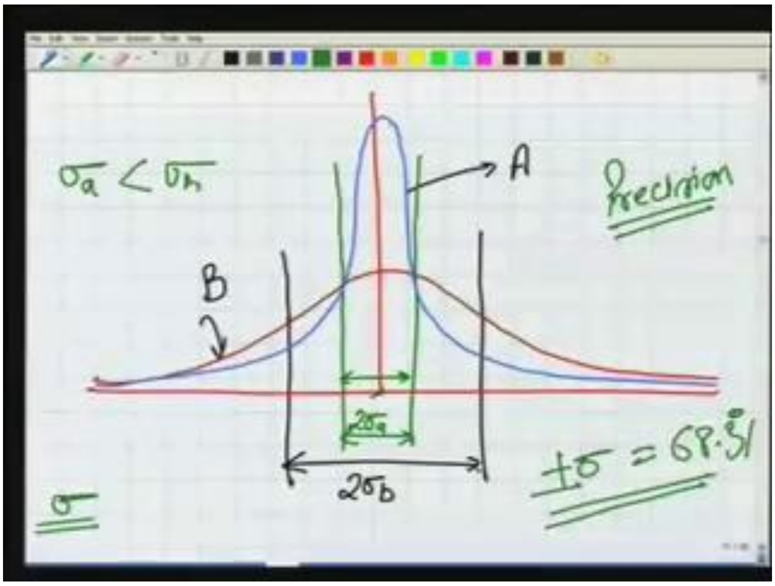
Now what we will do, we will continue our discussion in the same line. We saw in our last video lecture a term called ‘precision’. Now, we have seen so far that what is the meaning of the precision, and I gave you one example where you are firing bullets in the bulls-eye (Refer Slide Time 33:02). What you are doing in that case? Let us say, there is a fireman and he is very good; he fires all his bullets very closely, here (Refer Slide Time ddd 33:15). His observations are, or his firing is, not accurate - obviously, because it is away from the bulls eye. But why is it happening? It is happening because of some forces which are working there; some systematic errors, and because of those systematic errors all the observations are clubbed here. Well, there is another person who is not a very good fireman, and he also fires. Let us say, in his case, all his observations, all his bullets, are this way (Refer Slide Time 33:47). This is also not accurate because it is also away from the bulls-eye. Again, in his case, some systematic error is taking place, but if you look at these two, you will appreciate this person A for his very, very precise firing than the person B. Just by looking at this diagram, we are able to say, well, the person A is a better fireman than the person B. Now, how; what made you think that? Because somewhere, you are talking of precision. That precision, over here, in these bullets, is shown by the spread; the spread of bullets - are they spread widely or in very narrow area. So this spread is an indicator of the precision.

(Refer Slide Time 35:05)



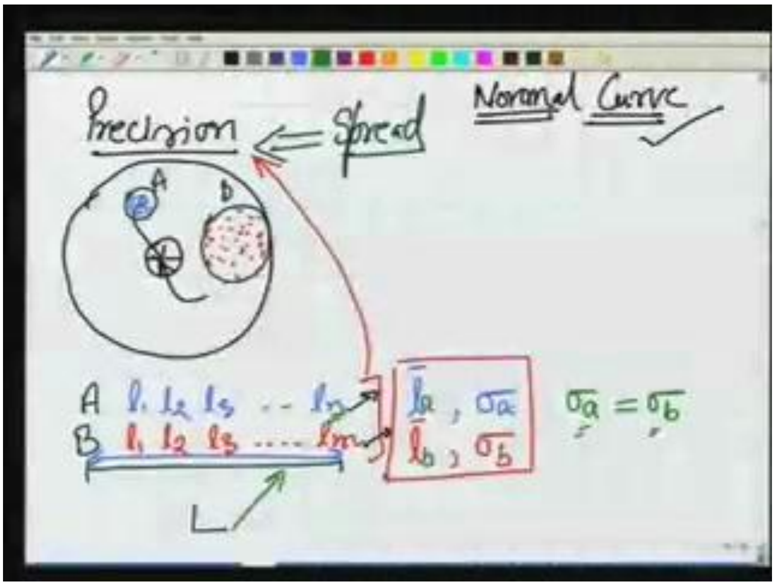
Now, what we do, we will try to take the same concept into the surveying observations. Let us say there is a length. ‘L’ is the true length, which we do not know (Refer Slide Time 35:06). There are two observers - A and B - both of them, they measure this length. A also comes here, and he takes several observations, for example, he takes l_1, l_2, l_3, up to l_n, (Refer Slide Time 35:25) and the B, B also takes several observations, for example, let us say l_1, l_2 - I am writing them by red in order to differentiate - l_3, l_n - or let us say l_m (Refer Slide Time 35:36)- different number of observations - whatever. Now, just by looking at these observations, can you speak about the precision - who is more precise out of these two? Well, what we do in this case, we will make use of, again, our normal curve. As in this case, even if you eliminate the systematic errors from your observations, all your observations shall shift here (Refer Slide Time 36:18), and they will still form a close group. Here also, if you eliminate the systematic errors all the observations will shift somewhere here (Refer Slide Time 36:29), and again they will form a widely spread group. So, what we are doing? Maybe from these observations also, we have taken away the systematic error component - it has been taken away. So, in your observations, only the random error is there now, and this is why I am saying that now, henceforth, because it is only random errors, we can make use of the normal curve. So, we will make use of the normal curve. Well, for this first set of observations, I find the ‘l bar’ and the sigma - let us say, I should write it as ‘sigma a’ for this set of observations. B - from the B, I find again ‘l bar’ and I find ‘sigma b’ (Refer Slide Time 37:03) - you can compute these. Of course, we are assuming in our observations that there is no systematic error component, there is no outlier, or there is no any mistake or blunder. Well, how to make use of these two in order to see which observation is more precise? What I will do, I will try to plot it now.

(Refer Slide Time: 37:49)



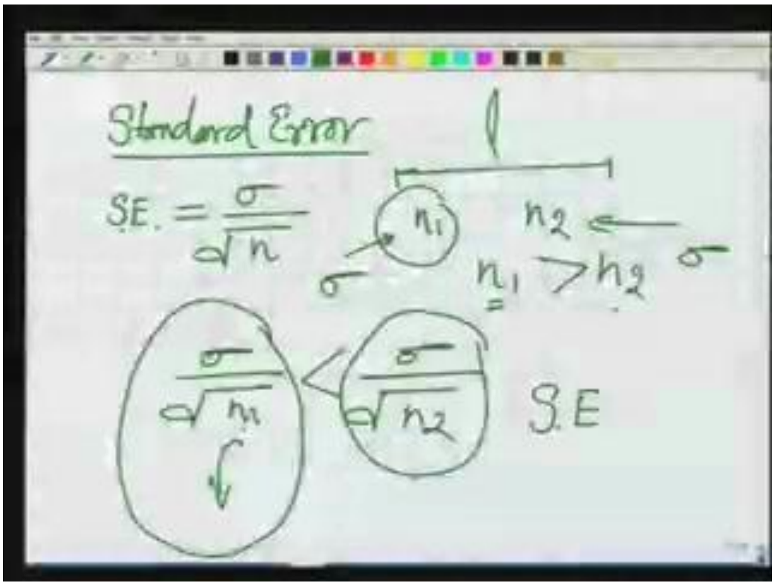
The plot of these observations may look like - for B, let us say for B, the plot looks like this (Refer Slide Time 37:53), and for A the plot looks like (Refer Slide Time 38:01) - it is quite possible. Now, what is the meaning of this? The meaning of this is, for B, the standard deviation will be the point of inflexion somewhere here (Refer Slide Time 38:19), and that is the value of the standard deviation, which we have already computed. This total (Refer Slide Time 38:24) is ‘2 sigma b’ - we have already computed sigma B from the observations. For observer A, the point of intersection should be somewhere here (Refer Slide Time 38:44), and we have already computed ‘sigma a’, so this is ‘2 sigma a’. Now, look at these two curves - do they say anything about precision; are there any kind of indicators of the precision? Just look at that what we saw earlier: we said, within plus-minus sigma, there are 68.3 percent observations (Refer Slide Time 39:14). So now, we are trying to relate these two curves with the precision. This is the stick (Refer Slide Time 39:32). We are talking that, within plus-minus sigma, 68.3 percent points or observations will be within plus-minus sigma. What is the meaning of that? The meaning is, here, in this case, the deviation of the observations from the mean - how many observations? - 68.3 percent observations is only within this limit. So, the deviation is less - the meaning is, observations are closer; they are nearer to each other. While in the case of B, in the case of B, the deviation of observations from the mean - because that is our sigma (Refer Slide Time 40:16) - 68.3 percent observations are spread in this area, in a wider area. So, the deviation of the observations from the mean is more here. This is something like, you know, a big spread, while this particular one is like a smaller spread. So, what we are doing? We are relating our standard deviation with the precision - standard deviation is an indicator of the precision. If you can find the standard deviation - so here, in this case, the standard deviation of A is smaller than standard deviation of B (Refer Slide Time 40:48). So, we can say these observations, the A observations, are more precise than the observations B, which is very obvious from here.

(Refer Slide Time: 41:11)



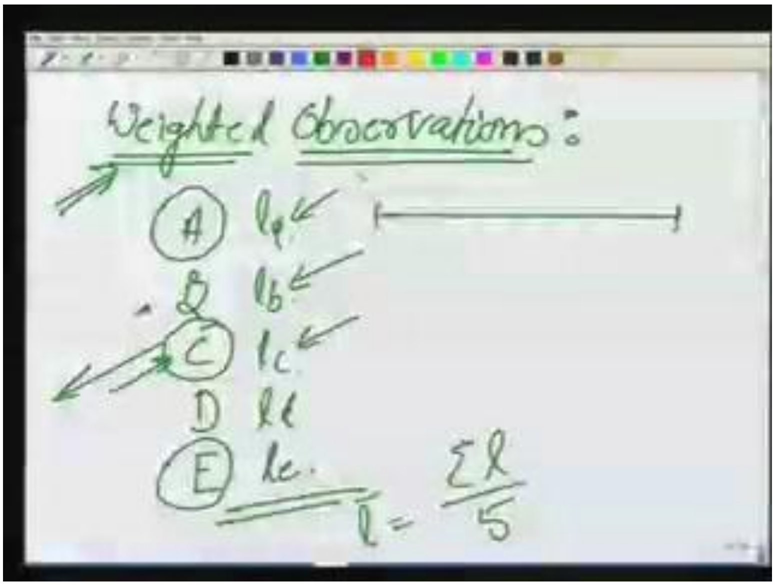
Now, there is one question. The question is, for example, here only, in this slide, both of them, they find for this length some mean values. Let us say, I am writing the this mean value here as l_a and l_b (Refer Slide Time 41:20). Both of them, they have found some standard deviation also, and by the way, incidentally, the standard deviations are equal (Refer Slide Time 41:30). Now what is the end of this? Two observers, they went to the field, they measured one length - one has done the observations n number of times, the other has done the observations m number of times - and both of them, they come to the laboratory, they present their results, they have one mean, one standard deviation; another mean, another standard deviation. Incidentally, both the standard deviations are same. Now, out of these two means, which one is a better estimate of this length? The standard deviations are coming to be same; can you say that the mean which is found from these observations is equally precise? Are the means same - are the means, you know, with the same confidence, can you say so? The one thing which is different here - that is about the number of observations. If we are taking only 10 observations in this n, and we are coming with these two (Refer Slide Time 42:43); in another case - let us say in case of B - we are taking 100 observations, and we are coming again with this (Refer Slide Time 42:50). Are these two means equally confident; can you be equally confident about these two means? So, in order to solve this problem - because this really leads to the kind of confusion - there is a term which is called ‘standard error’.

(Refer Slide Time: 43:07)



What we do, we find the standard error as the standard deviation divided by the root of number of observation (Refer Slide Time 43:13). What we have done now, in this case? If, let us say, for the length which we measured, we measured it n_1 times and n_2 times, while n_1 is more than n_2, and in both the cases - for case A, case B - we had the same sigma (Refer Slide Time 43:29). So, what we get, we get sigma by root of n_1 and sigma by root of n_2 (Refer Slide Time 43:47). Because n_1 is more than n_2 - so, this is more, this value will be less (Refer Slide Time 44:00) - the standard error in case of A, in case of this (Refer Slide Time 44:06), is less, because of more number of observations, though the standard deviation is same. What is the meaning of this? The meaning of this is, the mean which is associated with this observation set (Refer Slide Time 44:19) has more precision - you are more confident about that mean than the mean which is coming from here (Refer Slide Time 44:29). So, the standard error is a way in which - because in case of the standard deviation, it does not depend, you know, it does not take into account the number of observations. Two sets of observations may have same standard deviation, but that is wrong; the standard deviation - yes, it may be same, but when we start talking about our mean - the value which you are finding from it; our result - then we should not rely upon standard deviation alone. We should talk about standard error, because here, we are considering also the number of observations which you have taken there in the field, and the number of observations which are resulting in this standard deviation. So, this is important, and we should keep this in mind whenever we are reporting; whenever we are judging the observations.

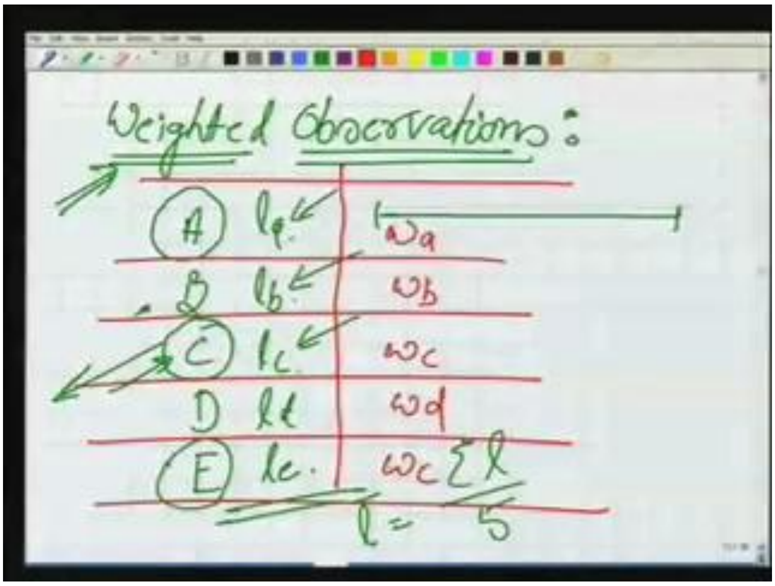
(Refer Slide Time: 45:44)



Now, after this, we will start talking about, now - weighted observation. What is the meaning of this? I will pose one question first. The question is, let us say, for a line there are 5 persons: A, B, C and D and E (Refer Slide Time 45:59). All of them have done the measurements of this line. What they did, they went to the field with a certain tape and then measured this length, you know, maybe on 5 different days, and they come out to the lab with 5 sets of observations. Are all these observations of same reliability? - that is the question. Let us say, right now, these 5 persons, they go to the field and measure 5 lengths - l_a, l_b, l_c, l_d and l_e. Now, question number one - question number one, will all these - l_a, l_b, l_c, l_d and l_e - will be the same, or will be the different? - question number one. Now, you should be able to answer this question where the five individuals who are going to the field, with maybe the same instrument - but will all these be same? No, they cannot be same, because there are many varying factors. The person A, he looks at the instrument in a slightly, you know, careless way; he does not look exactly the way the instrument should be read. The person B - maybe he is committing some blunder there. While person C is working, the temperature is more; there is some error in this instrument - it may be the systematic error. Even if the way you put the instrument there in the ground in order to measure the length, slightly some changes are there, what we saw earlier, we can never measure the true values. Only if you can measure the true values - then only all these observations would be same. So, all these observations will have to be different, because the amount of the random error or systematic error which is there in this observation is different. Even if you eliminate your systematic error, still, these observations will be different, because the amount of the random error which is there in the observations will be different.

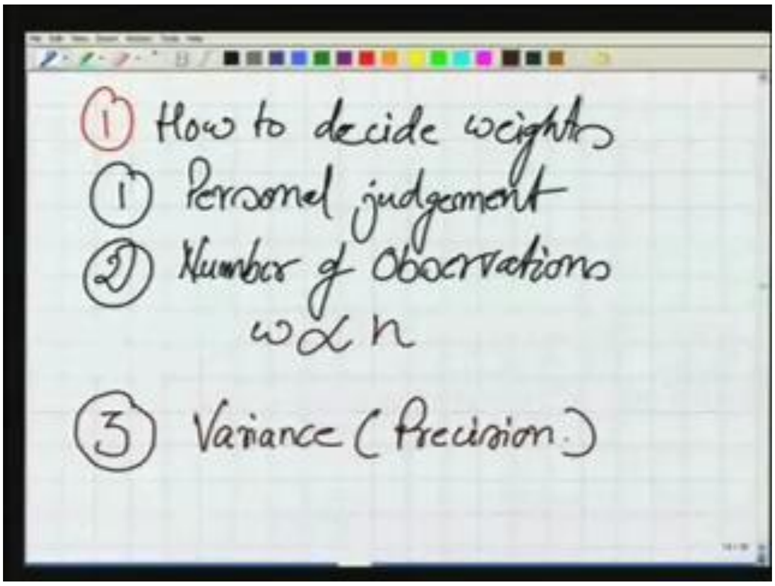
Well, so you have got five sets of observations; five measurements of this length. Now, we want to take a mean; you want to find the mean of this using these five. Should we, straight away from this 'sigma l by 5' (Refer Slide Time 48:39)? No, we should not do that. Why we should not do that? Because if you are doing it, we are considering each and every observation to be having the same weight; same reliability. You knew the person A is very sincere, while you also knew that person E is not that sincere. You know, when person C was working in the field, the weather was very bad; the working conditions were very, very poor. If that is so, each of these observations will be not equally reliable; they will have different reliabilities. So, this reliability of observations, we say as 'weight' - the weight of the observations. An observation which is more reliable is having higher weight; an observation which is having less reliability will have less weight.

(Refer Slide Time 49:40)



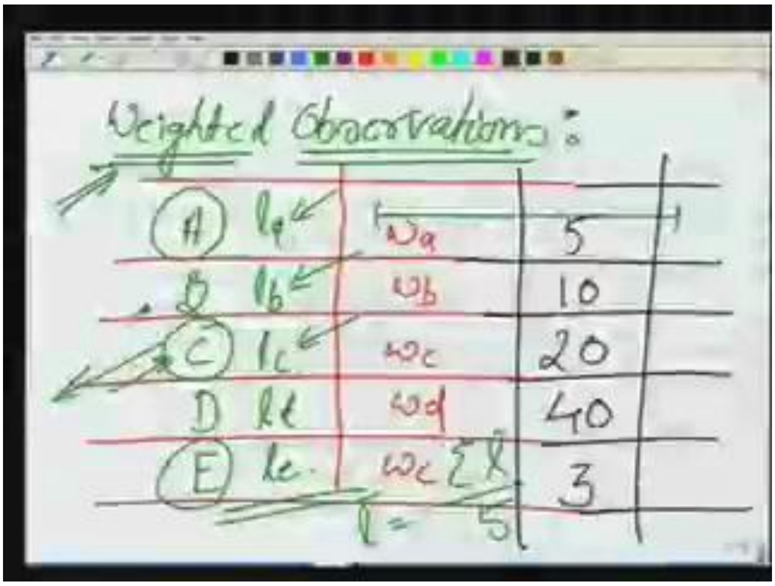
So, along with these observations, what we do, we also write - we also make - a table. Now, along with these lengths, we also write w_a, w_b, w_c, w_d, w_e, and these are the weights of these observations. Well, if our observations are weighted, how can we proceed further? Now, each and every observation is having a different weight.

(Refer Slide Time: 50:20)



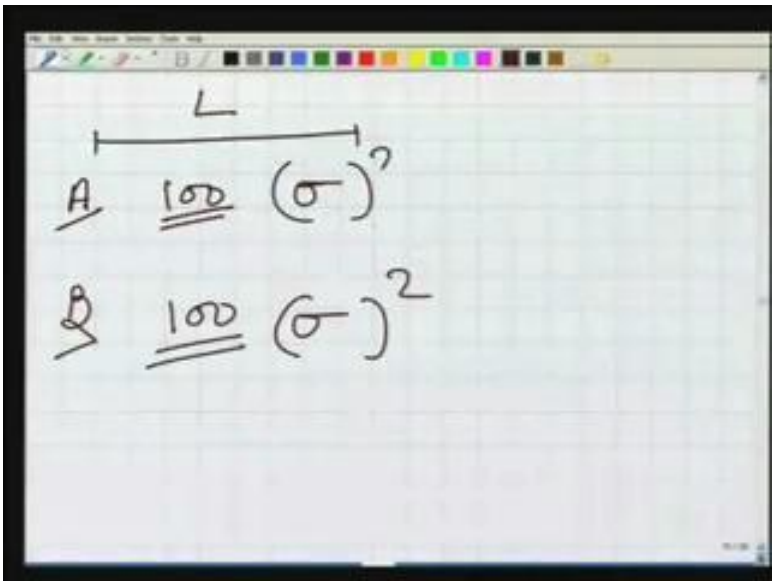
Number one thing - that is, how to decide or give weights. How do you want to give the weight? There should be some methods. Well, in the methods, number one is: by your ‘personal judgement’. What is the meaning of that? Well, here, you need the personal judgement. I said A - the person A - is a very, very reliable person; he is very sincere, he cannot commit mistake - I give him higher weight. Person D - he is not that good in working; I give him less weight. Person C, when he was working in the field, the weather conditions were very, very poor - I give him also less weight. Person B is - he is all right, so I give him, you know, weight. It is just like that, the personal judgement - you make the personal judgement and you assign the weight. Can there be some other methods? Well, the other method is ‘number of observations’.

(Refer Slide Time 51:55)



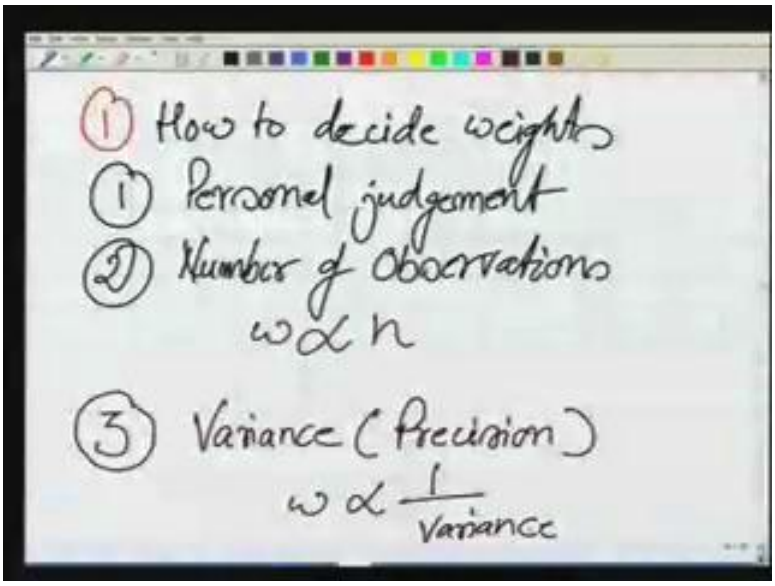
What is the meaning of this number of observations? Well, here only, if I extend my this table further (Refer Slide Time 51:41), now, this length was measured by these observers 5 times, 10 times, 20 times, 40 times, 3 times (Refer Slide Time 51:50) . We assume - let us say it is given to you, all these observers A, B, C and D are equally sincere, and they have done the observations of the same line different number of times. Obviously, which observation has got more weight? This observation (Refer Slide Time 52:24), having 40 number of observations, because a single thing we have measured 40 number of times. So, we can give definitely give more weight to that. So, the second method of giving weight is proportional to the number of observations.

(Refer Slide Time: 53:06)



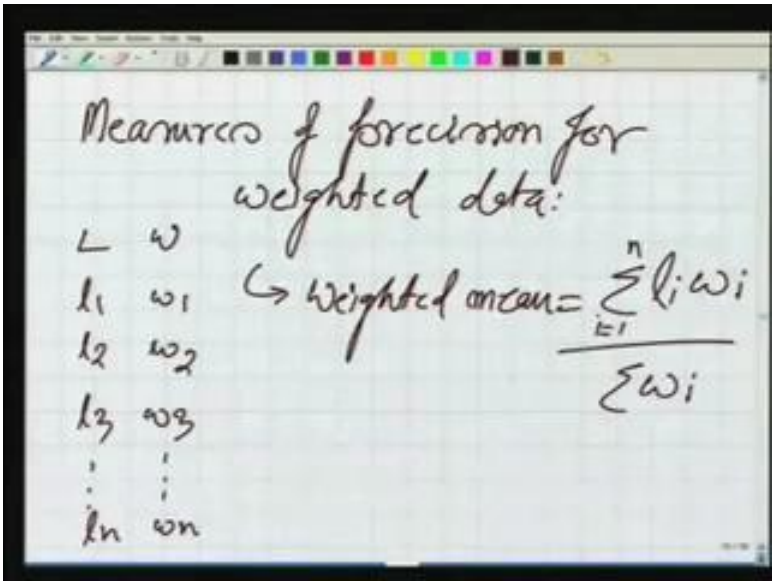
Can there be any better method? The third one - we give this weight as by the ‘variance’, or we can say, the precision. We know - for example, now, I will draw a diagram here: the same length is being measured now by two observers A and B (Refer Slide Time 53:06). A has done, for example, let us say 100 measurements. B has also done 100 measurements for this length L. Both of them - A and B - are equally sincere; the weather conditions when they were working in the field are almost same - if that is the case, whom should we give more weight? So what we do, for this 100 measurements and for this 100 measurements, we can find the sigma, and the square of sigma is the variance. Now, why we are giving the weight as per the sigma; what is the reason? Because it gives us the closest number of observations. The observer who is more sincere, who is working very well, the errors - the sources of the errors in the field are less - the chances are, all the observations will be very near to each other; very, very close to each other. So, if the observations are closer, you want to give them more weight.

(Refer Slide Time 54:23)



So, the weight is given now by variance. Inversely proportional to the variance (Refer Slide Time 54:24) - less is the variance, less is spread, more is the weight - obviously. So, these are the methods by which we can give weight to our observations.

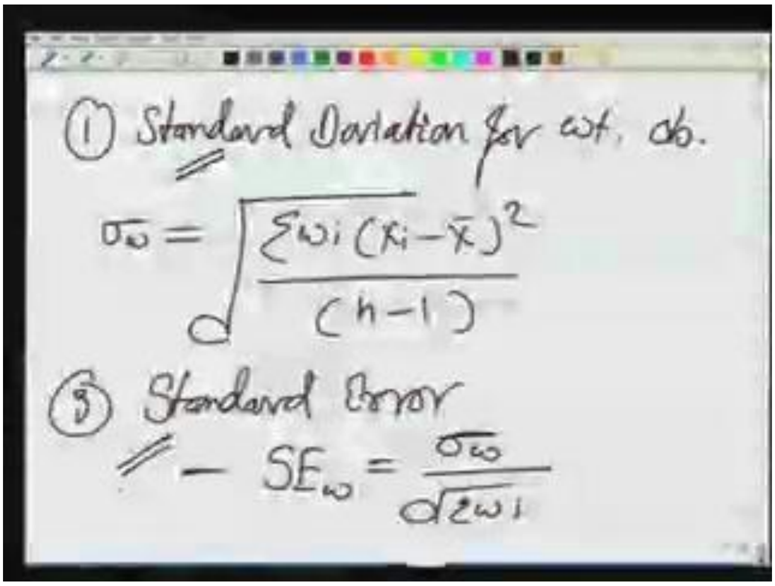
(Refer Slide Time: 54:52)



Now, what we will do, we will look for measures of precision for weighted data or observations. Now, in case of the weighted data, what this weighted data is? You have a length L measured as l_1, l_2, l_3, l_n (Refer Slide Time 55:10). Then, you have some weights: w_1, w_2, w_3 and w_n to each of these observations. Well, how can you find -

number one thing - the weighted mean? Weighted mean - because, we will need the weighted mean very often whenever we are working with the data. The weighted mean of this observation is ‘sigma l_i’ - where ‘i’ is 1 to n - divided by - you can guess it very quickly now - that is, ‘sigma w_i’ (Refer Slide Time 55:40). That gives you the weighted mean; the weighted arithmetic mean.

(Refer Slide Time: 56:06)



Now, how about the measure of precision? The measure of precision - number one - we say, ‘standard deviation for weighted observation’ or weighted data - so the same thing. We just put a little ‘w’ here in order to indicate that this is the weighted (Refer Slide Time 56:22). This is, as we seen before, also ‘sigma w_i’, ‘x_i’ minus ‘x bar square’ - of course, this x_i and x bar are the mean and the observation, and w_i is the weight of the observation - divided by n minus 1 (Refer Slide Time 56:27) . So, this gives you the standard deviation of the weighted data. Third: ‘corresponding to the standard errors’. The standard error - as we have written ‘SE’ earlier - weighted, is (Refer Slide Time 57:08)… divided by…‘i’.

So, all these are equivalent, in terms of their physical meaning, to what we have already discussed. We have discussed the standard deviation - what it means; standard error - what it means, where we use it; so all of these measures are also similar to that. So, what we have done, we have discussed about the errors, now - the normal distribution; how we make use of that, standard normal curve; how we make use of that, if we are taking the observations how to report it, what is the meaning of plus-minus sigma. Then, we saw the precision; how we can give the precision. This is all what we have discussed in this video lecture. Now, in our next video lecture, again, we will continue with the errors in the measurement. So, the sixth lecture of this module will be again in the same. So we end this lecture today and now - thank you very much.