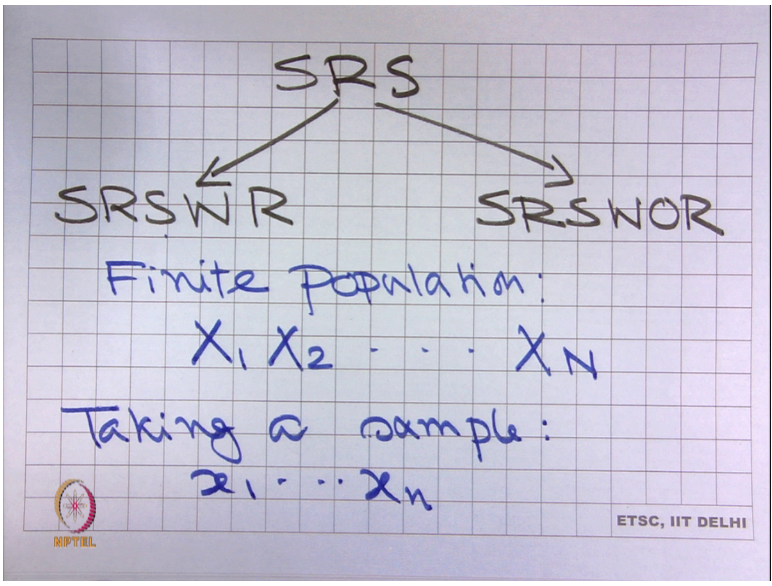
Statistical Inference
Prof. Niladri Chatterjee
Department of Mathematics
Indian Institute of Technology, Delhi

Lecture – 03
Statistical Inference

Welcome students to the third lecture of the MOOC's series on Statistical Inference. In today's lecture I will start from what I was discussing towards the end of the lecture 2.

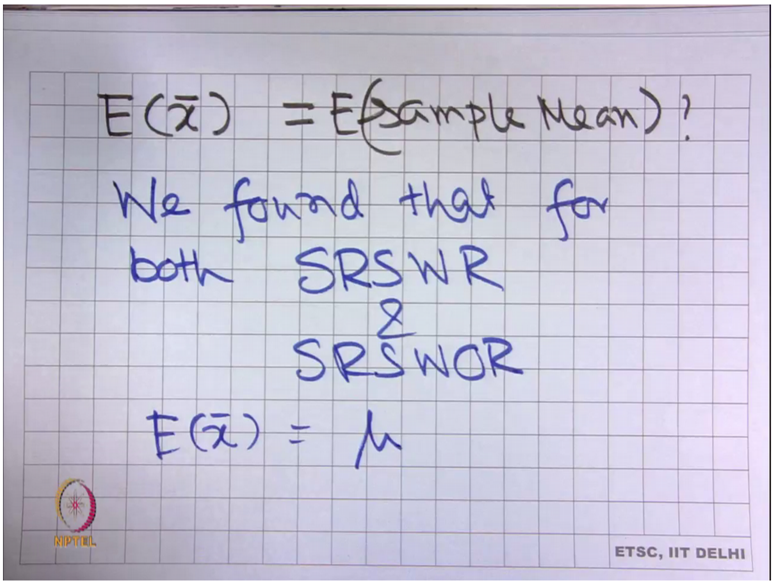
(Refer Slide Time: 00:47)



If you remember what I was doing is simple random sampling. In fact, we have talked about 2 ways SRSWR, which is simple random sampling with replacement and SRSWOR, which is simple random sampling without replacement. What we have found that if we take it from a finite population, let me call it X 1, X 2 up to XN, and we are talking about a small sample.

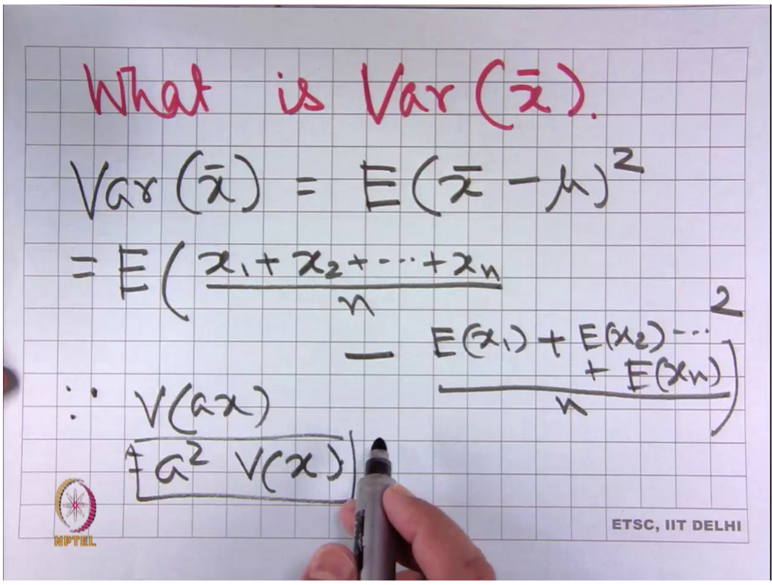
So, each small x i is the random variable and what we are looking at expected value of x bar is equal to sample mean, and we have found that for both SRSWR and SRSWOR expected value of x bar is equal to mu.

(Refer Slide Time: 02:07)



In the last class we started discussing about what is the variance of x bar.

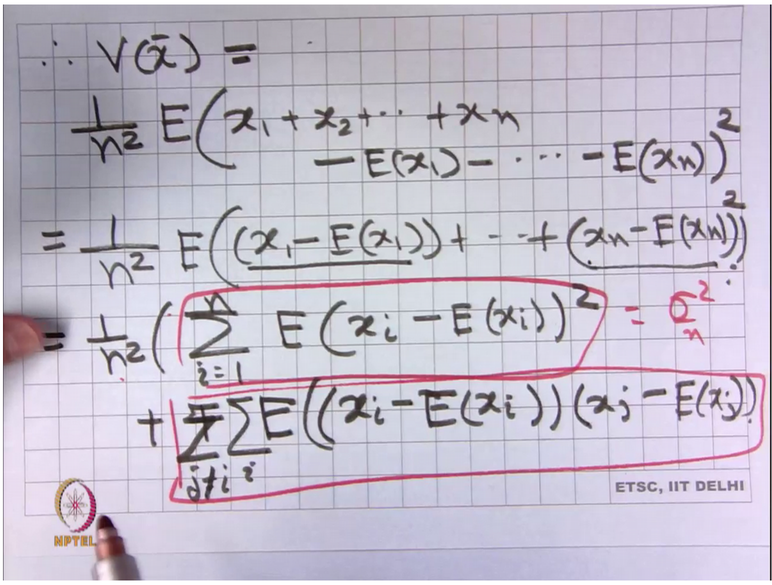
(Refer Slide Time: 03:00)



Today, I will start with that again, we know from definition variance of x bar is equal to expected value of x bar minus mu whole square; is equal to expected value of x 1 plus x 2 plus xn by n minus expected value of x 1 plus expected value of x 2 plus expected value of xn divided by n, this whole square. We know that variance of ax is equal to a square times variance of x.

Therefore, we can take 1 by n to be the constant multiplier, and we can take it out and therefore, variance of x bar becomes 1 upon n square into expected value of x 1 plus x 2 plus xn minus expected value of x 1 minus whole square; is equal to 1 upon n square into expected value of.

(Refer Slide Time: 04:34)

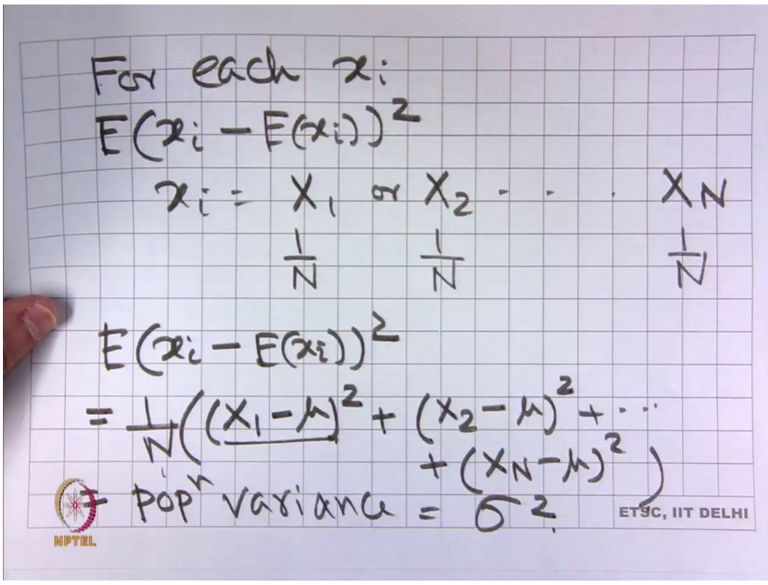


So, what we have done we have combined xi and expected value of xi together,.

So, this is equal to the square of n terms. This we can write it as 1 upon n square into sigma i is equal to 1 to N expected value of x i minus expected value of x I whole square. Because each individual term will be squared and contributed to the overall square, plus for each pair we are getting expected value of x i minus Exi into xj minus Exj. And this we are summing over all i and all j not equal to i.

Now, let us look at these 2 terms separately. This term and these term. So, what is the term one?

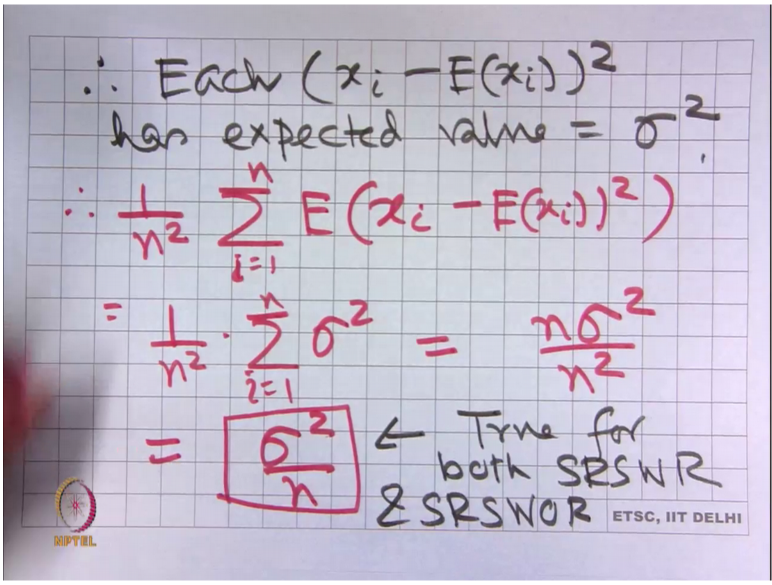
(Refer Slide Time: 07:10)



Expected value of xi minus expected value of xi whole square; xi is equal to X 1 or X 2 or XN each with probability 1 by N 1 by N 1 by N. Therefore, the expected value of xi minus expected value of xi whole square is equal to 1 by N into x 1 minus mu whole square plus X 2 minus mu whole square plus XN minus mu whole square.

Which is expected value of actually you are looking at the values and subtracting the mean and you are taking the square. So, this is the whole thing is the sum of square of deviation from the mean, and that divided by N is equal to population variance is equal to sigma square.

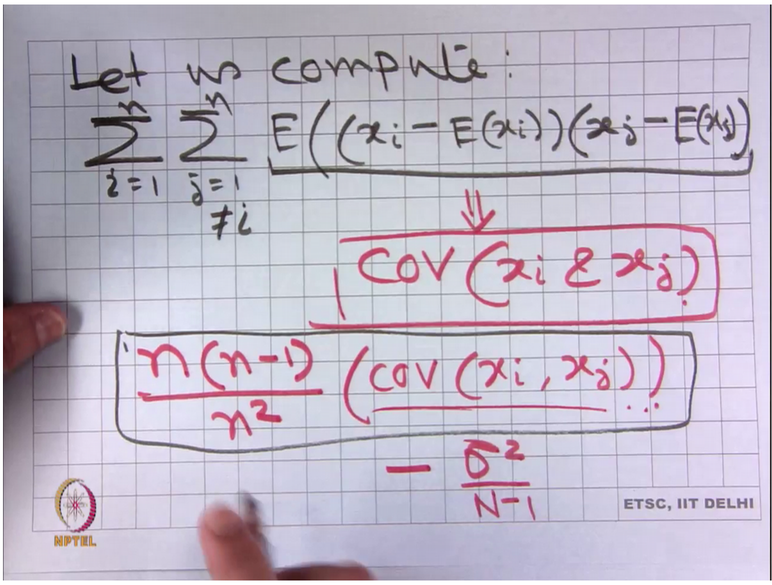
(Refer Slide Time: 08:53)



Therefore, each xi minus expected value of xi whole square has expected value is equal to sigma square. Therefore, 1 by n square sigma 1 to N expected value of xi minus expected value of xi whole square is equal to 1 upon n square into sigma i is equal to 1 to N sigma square, which is equal to n sigma square upon n square which is equal to sigma square by n.

So, if we go back to my old slide, you see this first term when multiplied by 1 upon n square is actually giving you sigma square by n. We need to now find out what is the contribution of second term. Note that this is true for both SRSWR and SRSWOR.

(Refer Slide Time: 10:46)

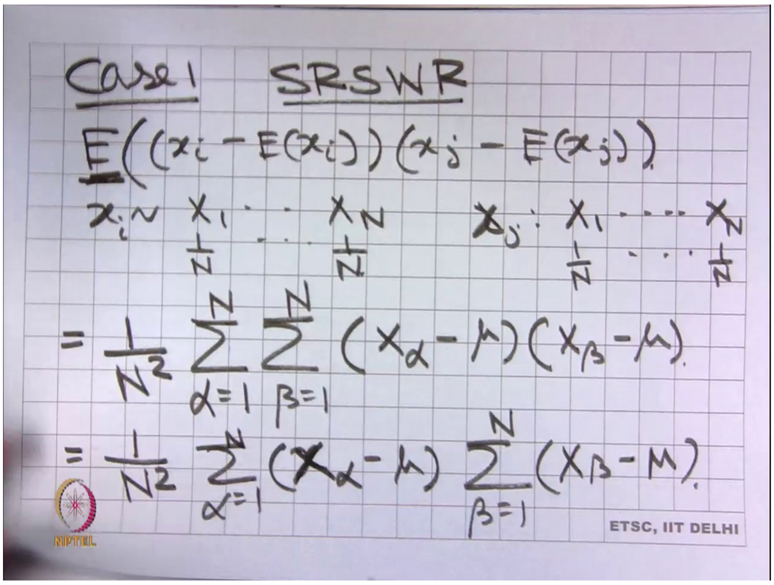


Therefore, now our target is to compute sigma over i is equal to 1 to N sigma over j is equal to 1 to N j not equal to i, expected value of xi minus expected value of xi into xj minus expected value of xj.

Now, what is this term? This is nothing but the covariance between xi and xj. So, in effect, we are summing the covariance between xi and xj, and how many such pairs are possible? The number of pairs is n into n minus 1, right because I can take all the values from 1 to N corresponding to each i j can take other remaining n minus 1 values. Therefore, it is n into n minus 1 already we have a factor 1 upon n square.

Now, what we are trying to write here is covariance of xi xj, all right?

(Refer Slide Time: 12:49)

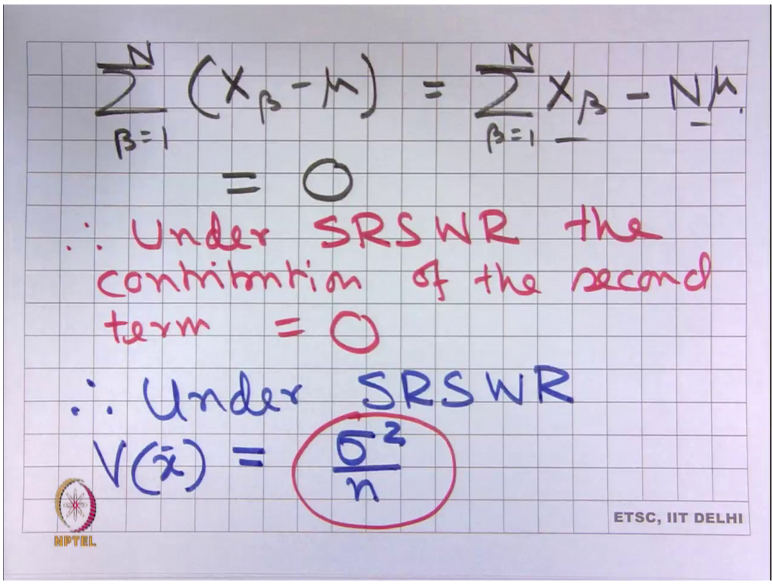


So, this is the term that we need to figure out. Here there will be a difference between SRSWR and SRSWOR. So, what is the expected value of xi minus expected value of xi into xj minus expected value of xj? If it is SRSWR, then xi can take X 1, X 2, Xn each with probability 1 by N up to 1 by N.

Similarly, Xj you also take and take X 1 X 2 Xn and each one with probability 1 by N 1 by N 1 by N. Therefore, this expected value that we need for computing the covariance, actually giving probability 1 upon N square into sigma alpha is equal to 1 to N sigma beta is equal to 1 to N into X alpha minus mu X beta minus mu. Because each possible pair X alpha beta, there are n square many such pairs, can occur with probability 1 upon N square.

And now I am taking the expectation so, for each X alpha and X beta I am looking at the product terms, and then I am summing it up over beta is equal to 1 to N alpha is equal to 1 to N. And what we are getting 1 upon n square sigma alpha is equal to 1 to N X alpha minus mu sigma beta is equal to 1 to N x beta minus mu.

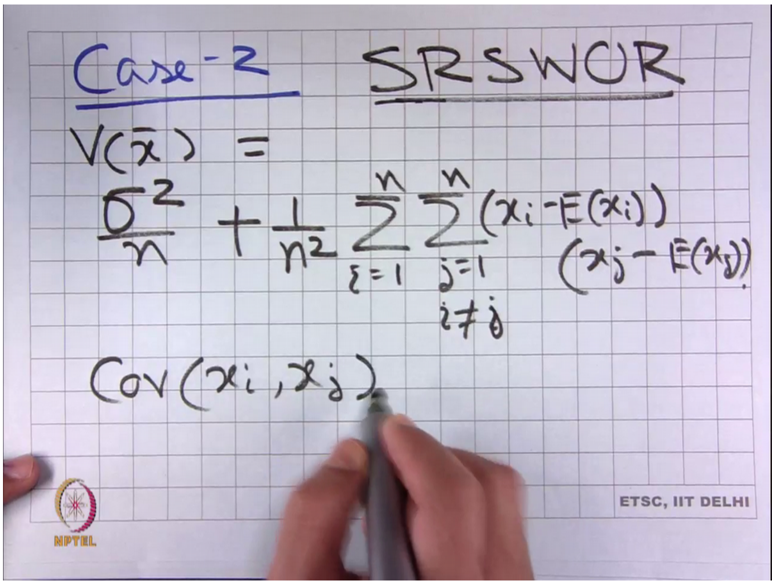
(Refer Slide Time: 15:21)



Now, consider just sigma beta is equal to 1 to N, X beta minus mu is equal to sigma over beta is equal to 1 to N; X beta minus N mu. And we know that mu is the population mean; so, therefore, N mu is the total of all the values of X beta, therefore, these 2 terms are equal is equal to 0. Therefore, under SRSWR the contribution of the second term is equal to 0.

So, if we look at the slide, these part gave us sigma square by n, this part gave us 0. Therefore, under SRSWR variance of x bar is equal to sigma square by n. So, if we are sampling from the population which replacement, then the sampling variance of x bar is equal to sigma square by n, where n is the sample size. What does it mean? It means that as n increases the variance of x bar also decreases with increasing n.

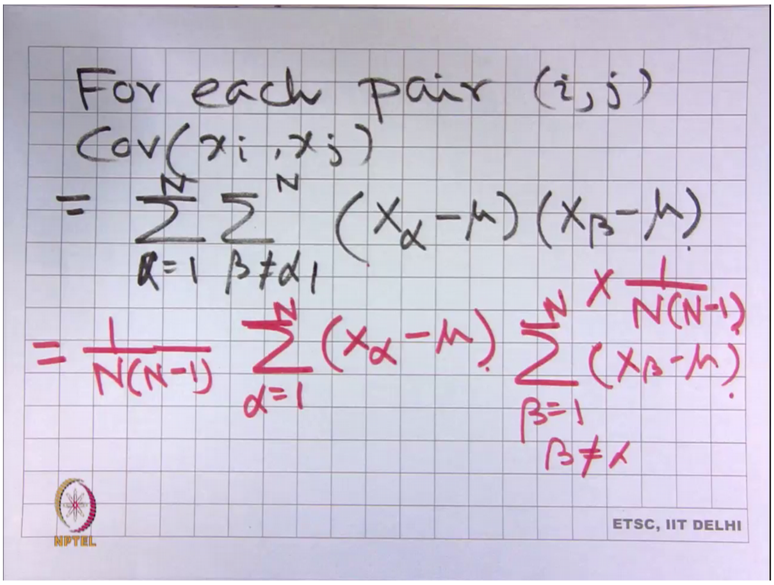
(Refer Slide Time: 17:43)



Now, let us look at case 2 what is that? Case 2 is SRSWOR so now; we are looking at simple random sampling without replacement. What is therefore going to happen? When we have chosen the first sample, we cannot take it again from the population. So, each element can occur at most once in the sample. So, going back to the equation of variance, this term still remains sigma square by n. Therefore, this variance is going to be from the earlier term, it is sigma square by n plus 1 by n square sigma over i is equal to 1 to N sigma over j is equal to 1 to N; j not equal to i xi minus expected value of xi multiplied by xj minus expected value of xj.

Here there is a difference; because these 2 are not independent therefore, depending upon the value of xi there will be some change in this term. Therefore, we need to look at it in a more careful way. Again as before, we can see that essentially you are looking at covariance between xi and xj.

(Refer Slide Time: 19:59)

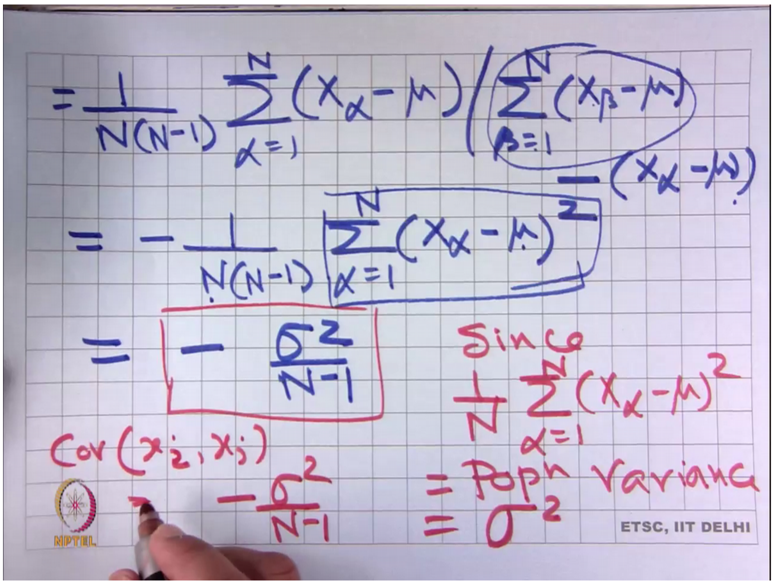


As before for each pair i, j covariance between xi and xj is equal to sigma over alpha is equal to 1 to N sigma over beta not equal to alpha from 1 to N X alpha minus mu into X beta minus mu.

And what is the probability? We can have m into n minus 1 many possible pair. Because alpha can be between 1 to N and beta is going from 1 to N, but not equal to alpha. So, alpha can be chosen in capital N ways, beta can be chosen in n minus 1 ways given alpha; therefore, number of pairs is 1 upon N into N minus 1. Let me take X alpha out of the summation, and let me write it as beta is equal to 1 to N; beta not equal to alpha X beta minus mu.

Now, I will do some jugglery on that one, sigma alpha is equal to 1 to N X alpha minus mu.

(Refer Slide Time: 22:08)

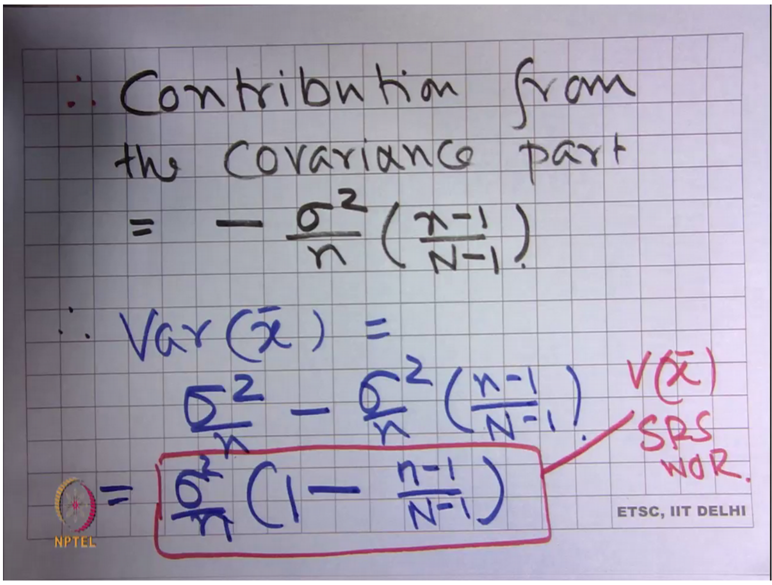


Now, this summation I write it as sigma beta is equal to 1 to N X beta minus mu which also includes alpha. Therefore, I subtract the term X alpha minus mu; so, basically what I have done? Since beta is not equal to alpha; I have added and subtracted X alpha minus mu to this term; the advantage thereby we get is that this term becomes 0; because I am summing over 1 to N, all the elements of the population and subtracting n mu from there therefore, this is going to be 0.

So, we are left with only this term, and we know that sum of square of the deviation from the mean divided by n is equal to sigma square. So, this term is sigma square upon n minus 1, since 1 upon n sigma alpha is equal to 1 to N X alpha minus mu whole square is equal to population variance is equal to sigma square. So, we have got that covariance between xi and xj is equal to minus sigma square by N minus 1. As we started with therefore, with respect to SRSWR when we are looking at the second part we knew that this is minus sigma square upon N minus 1.

Therefore, what we are getting is this n cancels with this square.

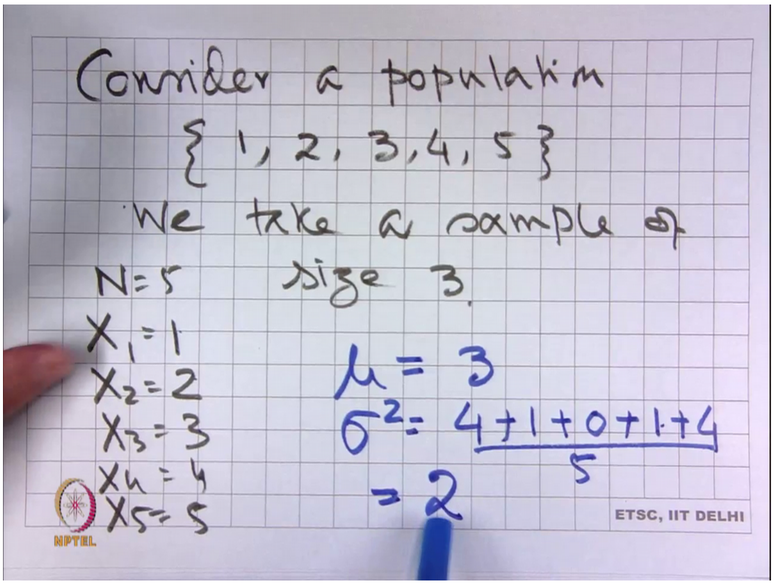
(Refer Slide Time: 25:12)



Therefore, the second part that is the covariance part is equal to minus sigma square by n into n minus 1 upon N minus 1. Therefore, variance of x bar is equal to sigma square by n minus sigma square by n into n minus 1 N minus 1 is equal to sigma square by n into 1 minus N minus 1 1. So, this is the variance of x bar when SRSWOR, ok.

So, when you are looking at a SRSWOR, we find that there is a small factor that affects the variance of under SRSWR which is sigma square by n; it is modified by a small factor there. If small n is very small in comparison with capital N, then slowly this becomes 0, and we get something that is very close to sigma square by n, ok. So, let me now do a small problem for you to understand what is happening.

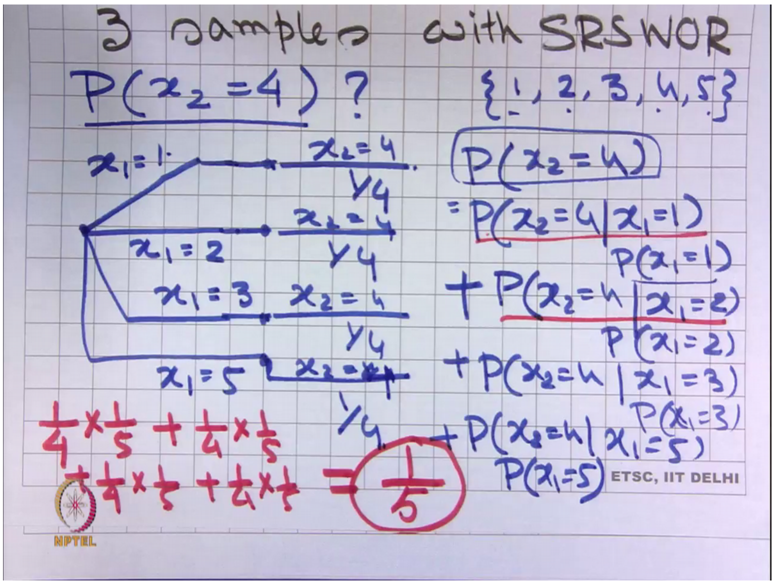
(Refer Slide Time: 27:16)



Consider a population with 5 values 1, 2, 3, 4, 5. And suppose we take a sample 3. We take a sample of size 3, ok. So, basically my X 1 is equal to 1, X 2 is equal to 2, X 3 is equal to 3, X 4 is equal to 4, X 5 is equal to 5, capital N is equal to 5. And we can easily calculate that mu is equal to 3, 1 plus 2 plus 3 plus 4 plus 5 is equal to 15 divided by 5 is equal to 3. And sigma square is equal to 1 minus 3 whole square which is 4, plus 2 minus 3 whole square which is 1, plus 3 minus 3 whole square which is 0, plus 4 minus 3 whole square which is 1, plus 5 minus 3 whole square which is 4 divided by 5 is equal to 4 plus 1 is equal to 5 plus 4 plus 1 5 is equal to 10 by 5 is equal to 2.

Therefore, in this case my population mean is 3 and population variance is 2.

(Refer Slide Time: 29:22)



Suppose I take 3 samples with SRSWOR; I am dealing with SRSWOR, because SRSWOR is comparatively simple, because it is doing from the same population so, it is easily understandable. Now my question is what is the probability that second sample is equal to 4? Suppose I want to calculate what is the probability that the second sample is equal to 4. I am doing it for the second, because first in any case we understand that if I am taking one from this population, then each one of them under simple random sampling has probability 1 by 5. Now I want to show that these probabilities also 1 by 5.

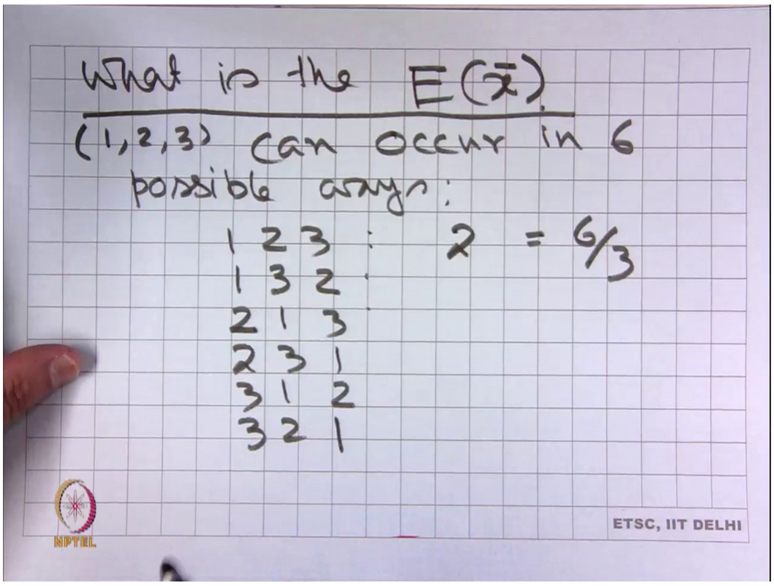
Now, second sample is equal to 4 means, the first sample is not 4, it is one of the remaining things. Therefore, it can happen in the following ways. The second sample can before if the first sample is either 1 or 2 or 3 or 5. Now after taking the first sample to be one the probability the second sample is 4 what is this probability? After one is taken there are 4 elements therefore, the second element is 4 that probability is 1 by 4.

Similarly, here x 2 is equal to 4 that has probability 1 by 4. Here also x 2 is equal to 4 that probability is 1 by 4, and in this case in a similar way this probability is probability x 2 is equal to 4 is 1 by 4. Therefore, second sample can be 4 if the second sample is 4 given that the first sample is one multiplied by the probability the first sample is 1 is in plus probability x 2 is equal to 4 given x 1 is equal to 2 into probability x 1 is equal to 2 plus probability X 2 is equal to 4 given X 1 is equal to 3 into probability X 1 is equal to 3 plus probability X 2 is equal to 4 given X 1 is equal to 5 multiplied by probability X 1 is equal to 5.

So, these are the cases that leads to the event X 2 is equal to 4. So, what is going to be the probability. These probability is as we have calculated is 1 by 4 and what is the probability? That x 1 is equal to 1, that is 1 by 5 plus this is 1 by 4 multiplied by 1 by 5, plus in a similar way multiplied by 1 by 5 plus 1 by 4 multiplied by 1 by 5 and therefore, when I sum this up it is 4 times 1 by 4 into 1 by 5, therefore, it is 1 by 5.

Therefore, what I show that even probability second sample is equal to 4 is 1 by 5 or which is same as the probability first sample is equal to 4. In a similar way we can show it for other values of the sample.

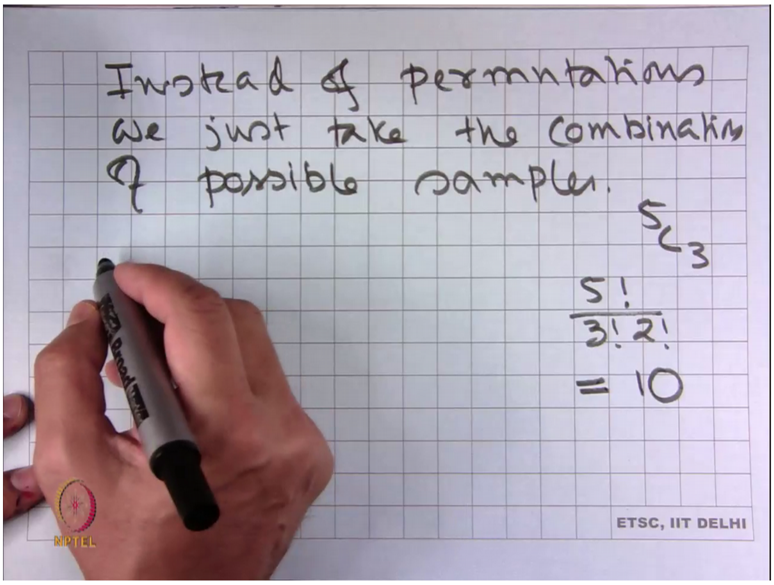
(Refer Slide Time: 34:44)



Now, let me go to slightly more difficult thing, what is the expected value of x bar. We know that when we are sampling without replacement, the same triplet 1, 2, 3 suppose can occur in 6 possible ways, right? Because of the permutation that we know, but for your clarification, I may take them in this 6 different orders. For each one of them x bar is coming out to be 2 is equal to 6 by 3.

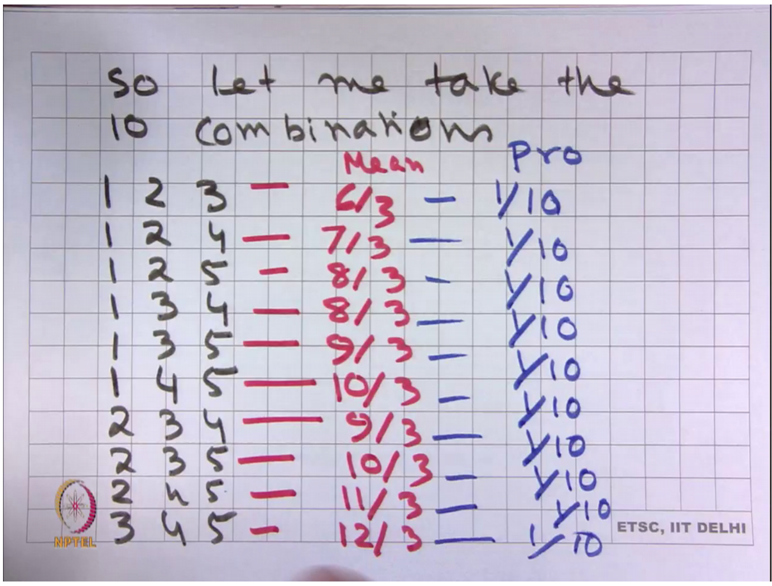
Therefore, in order to make the calculation simple, in my calculation instead of permutations we just take the combinations.

(Refer Slide Time: 36:07)



And we know that we can have the following, what is 5 c 3 number of ways of choosing 3 out of 5, it is factorial 5 upon factorial 3 2 factorial 2 which is equal to 10. So, I know I will get 10 different triplets.

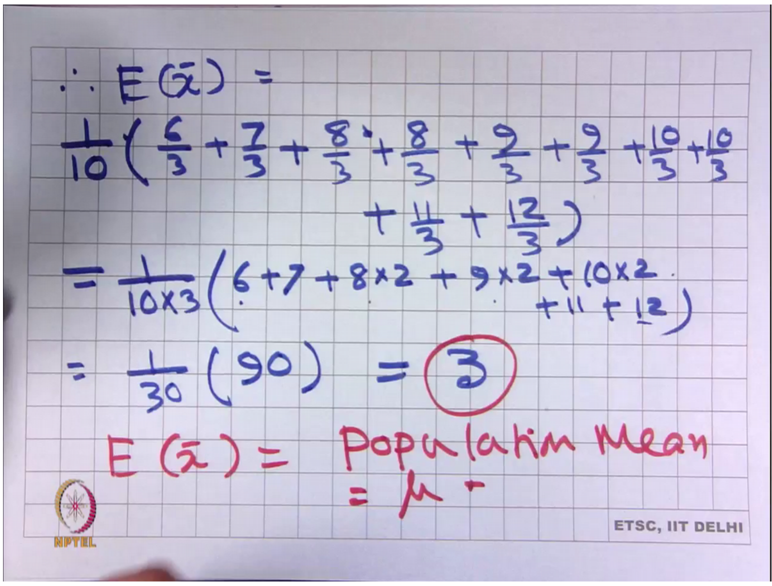
(Refer Slide Time: 37:07)



So, let us these are 1 2 3, 1 2 4, 1 2 5, 1 3 4, 1 3 5, 1 4 5, 2 3 4, 2 3 5, 2 4 5 and 3 4 5. So, 1 2 3 4 5 6 7 8 9 10; so, we get the 10 different combinations. What are going to be their mean? The means are 6 by 3, 7 by 3, 8 by 3, 8 by 3, 9 by 3, 10 by 3 9 by 3, 10 by 3 11 by 3 12 by 3, right?

And each one of them have probability equal probability. So, this is have probability 1 by 10, this has 1 by 10, this has 1 by 10, this has 1 by 10.

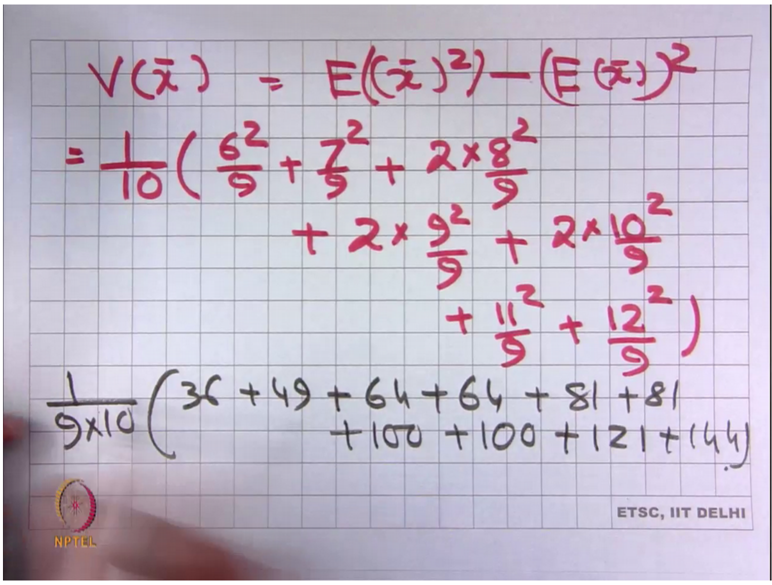
(Refer Slide Time: 39:39)



Therefore expected value of x bar is equal to 1 by 10 into 6 by 3 plus 7 by 3 plus 8 by 3 plus 8 by 3 plus 9 by 3 plus 9 by 3 plus 10 by 3 plus 10 by 3 plus 11 by 3 plus 12 by 3. Which is equal to 1 by 10 into 3 into 6 plus 7 plus 8 into 2 plus 9 into 2 plus 10 into 2 plus 11 plus 12 is equal to 1 by 30 into 6 plus 7 is equal to 13, 13 plus 16 29, 29 plus 18 47, 47 plus 20 67, 67 plus 11 is 78, 78 plus 12 is equal to 90.

Therefore, expected value of x bar is equal to population mean; which we have already established. That this is going to be mu, and in this case the value is 3.

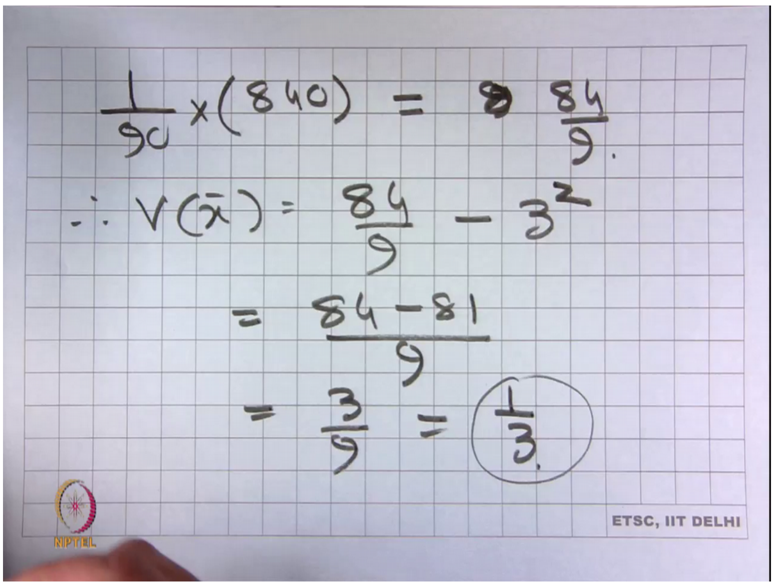
(Refer Slide Time: 41:39)



Now, let us calculate variance of x bar is equal to expected value of x bar square minus expected value of x bar whole square. And therefore, this is equal to we have got the x bars to be 6 7 8 8 6 bars. So, this square we have to take.

So, this is going to be, right? This is equal to we can now calculate easily it is 1 upon 9 into 10 6 square is equal to 36, plus 49, plus 64, plus 64 because it is twice plus 81, plus 81 plus 100 plus 100 plus 121, plus 144. This is equal to.

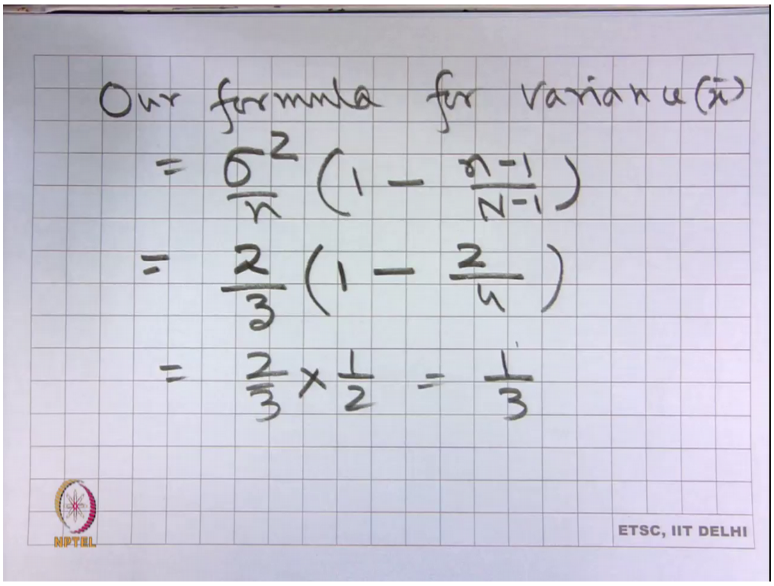
(Refer Slide Time: 43:42)



If you calculate you are going to get 1 upon 90 into 840 is equal to 84, 849.

Therefore, variance of x bar is equal to 84 by 9 minus 3 square is equal to 84 minus 81 upon 9 is equal to 3 upon 9 is equal to 1 upon 3. Therefore, under SRSWOR the variance is coming out to be 1 upon 3.

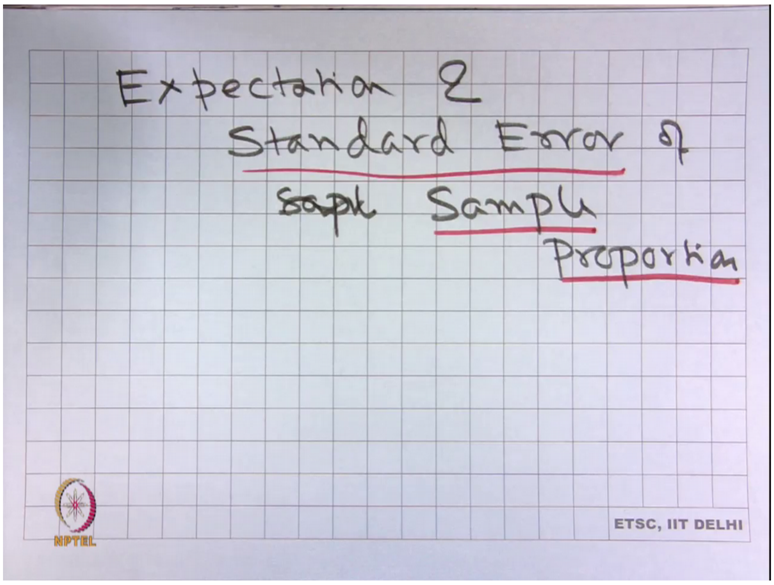
(Refer Slide Time: 44:33)



Our formula for variance is equal to sigma square by n into 1 minus N minus 1 upon n minus 1. Here, sigma square was 2 that we have already calculate population variance into n is equal to 3 1 small n is equal to 3, therefore, small n minus 1 is equal to 2, capital N is 5. So, this is 4 is equal to 2 upon 3 into half is equal to 1 upon 3.

Therefore, the variance of x bar that we have calculated to be 1 upon 3 we also get from the same formula, it is coming out to be 1 upon 3. It is not a proof of anything.

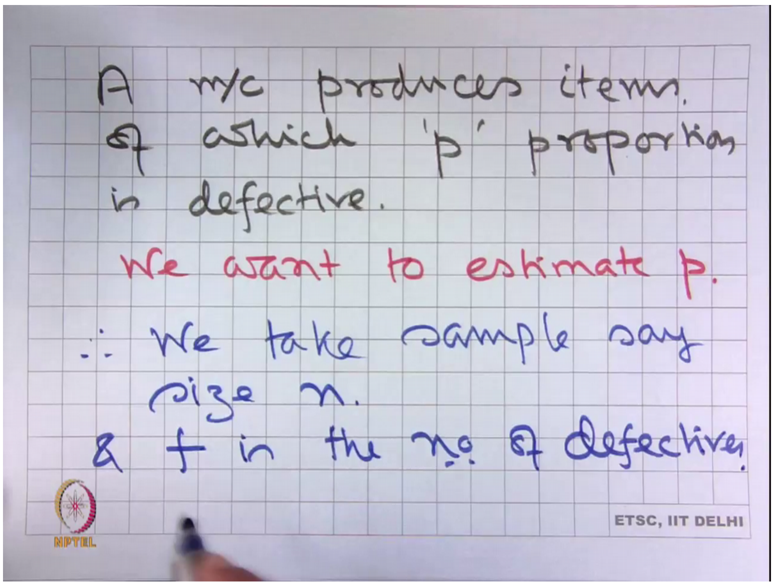
(Refer Slide Time: 45:53)



I have just validated the formula that I have derived, and by working out on such an example many of your doubt should be clear. Now let us look at the following, now I introduced a new term, standard error; standard error is nothing but the standard deviation of a statistic, ok.

So, in this problem I am focusing on sample proportion.

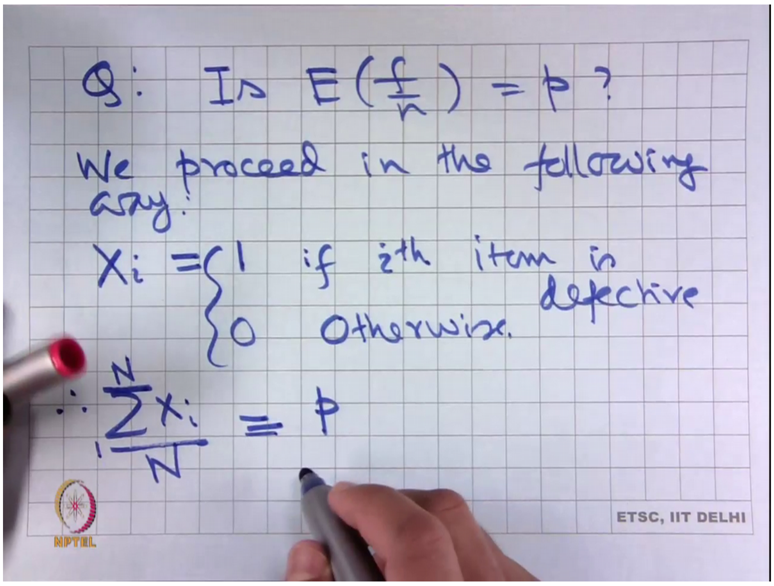
(Refer Slide Time: 46:53)



Consider the problem that a machine produces of which p proportion is defective. In reality we often like to estimate p. Of course, we can test an item produced to see it is defective, but complete enumeration is often a problem. And secondly, sometimes the tests are destructive.

So, if we destruct all the elements produced then there is nothing that is remaining for further use. So, what is done we take a sample say size n, and f is the number of defectives our aim is to estimate p. So, question is expected value of f by n equal to p? So, f by n is the sample proportion of the defective.

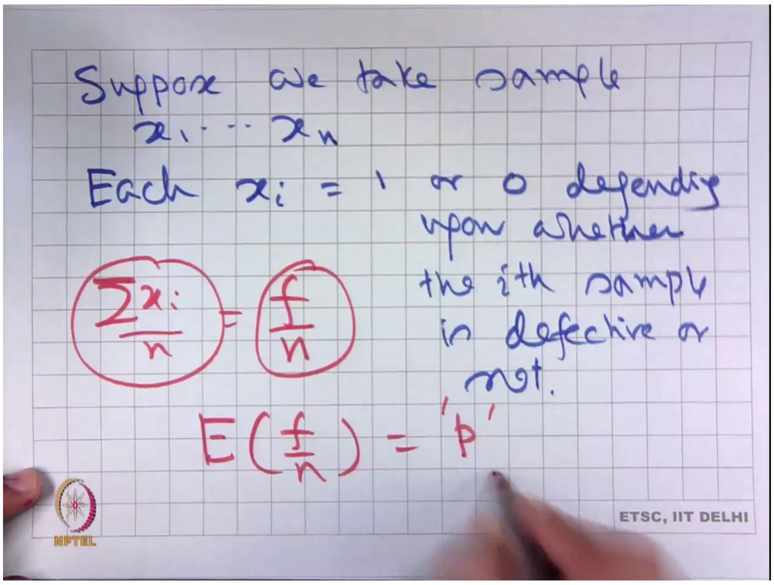
(Refer Slide Time: 48:42)



So, whether is the expected value is same as the population proportion and also our what is this variance or standard deviation who is just now we have said is called standard error. We doing we proceed in the following way, we define a random variable xi which is 1 if I th item is defective and 0 otherwise.

Therefore, sigma xi divided by N is equal to p, because sigma xi 1 to N will count the number of defective items in the population. Therefore, that divided by N will give you the proportion of that one.

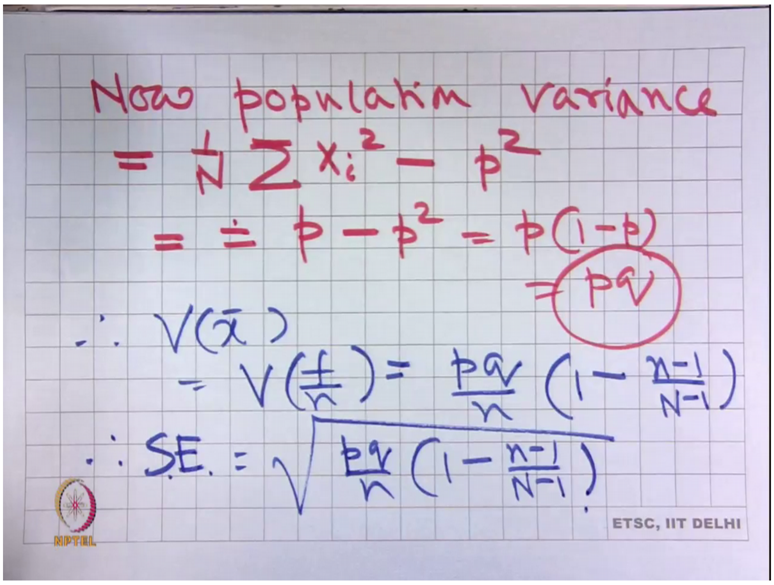
(Refer Slide Time: 50:12)



And suppose we take a sample each xi is equal to 1 or 0 depending upon whether the I th sample is defective or not. Therefore, now what is the sample mean? Sample mean is actually f by n and sigma xi upon n is equal to actually f by n and what we are looking at proportion.

So, instead of what we have done earlier with respect to population mean and sample mean, now we have converted that by changing the variable into the problem of proportion, and we know that the sample mean is actually estimate of population mean therefore, sigma xi by n which is f by n is expected value is truly the population mean. Therefore, the sample mean can be used as an estimate for the population mean.

(Refer Slide Time: 51:53)



Now, population variance is equal to 1 by N sigma xi square minus p square is equal to sigma xi square. If xi is equal to 1 it is still 1, and if xi is 0, it is 0 therefore, sigma xi square is same as sigma xi. And which is actually therefore, this is p minus p square is equal to p into 1 minus p is equal to pq.

Therefore, population variance is equal to pq. Therefore, variance of x bar is equal to by our earlier result is equal to pq upon n into 1 minus n minus 1 upon N minus. Therefore, standard error is equal to root over pq upon n into 1 minus N minus 1, friends with that I stop here today.

Thank you.