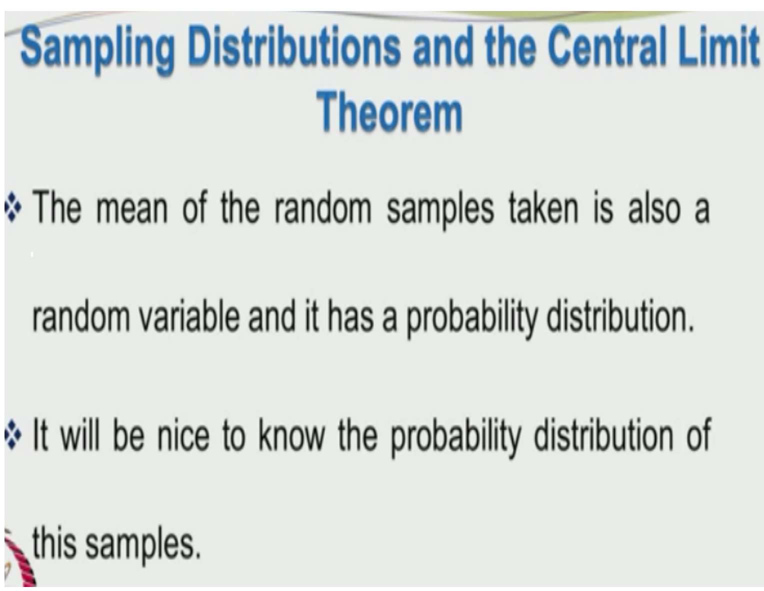
Statistics for Experimentalists
Prof. Kannan. A
Department of Chemical Engineering
Indian Institute of Technology – Madras

Lecture - 13
Sampling distributions and the Central Limit Theorem

We are now looking at the sampling distributions of the mean. We started originally with the random variable X, then we started looking at the sample mean X bar. We collected the random variables into one group or set and we created the sample mean X bar, X1+X2 so on to Xn/n and this is also a random variable. It is also having the distribution, but what is the type of the distribution of the random variables forming the sample mean.

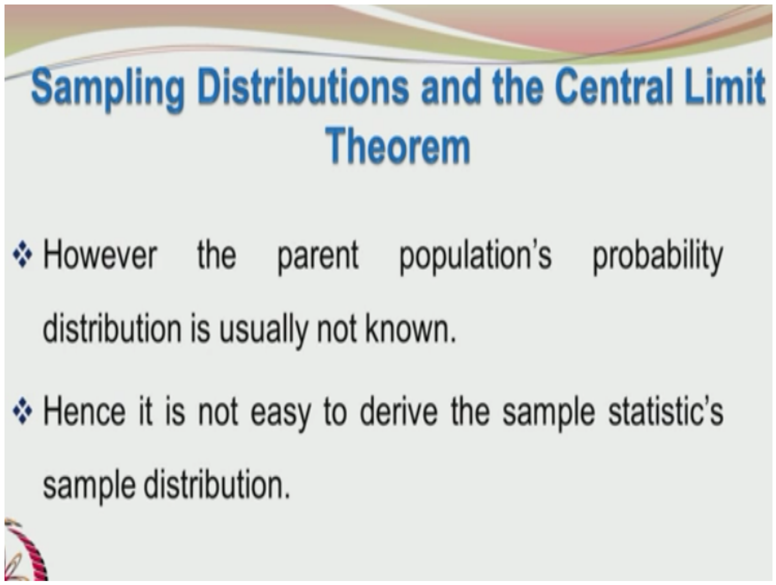
We are unaware of it, but rather than working with random variables X, as I said earlier we are now going to work more with the random samples. We are going to work with a collection of random variables and we are going to look at the sample means and use them to draw certain inferences. So we should also know the distribution of the sample means, what population they follow.

(Refer Slide Time: 01:38)



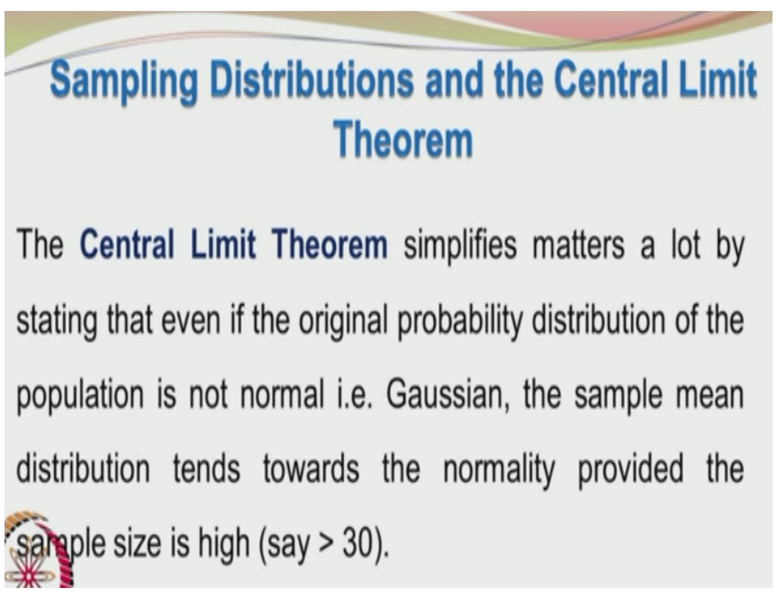
The question is we do not know about the population itself. We do not know about the original population. We do not know whether this is normal or gamma or variable or what are the type of distribution. We do not know its parameters, but all we have is only the random samples and they themselves are forming another distribution. Fortunately for us, the central limit theorem comes to our rescue. What is the central limit theorem? That is going to be the focus for the next half-an-hour or so.

(Refer Slide Time: 02:19)



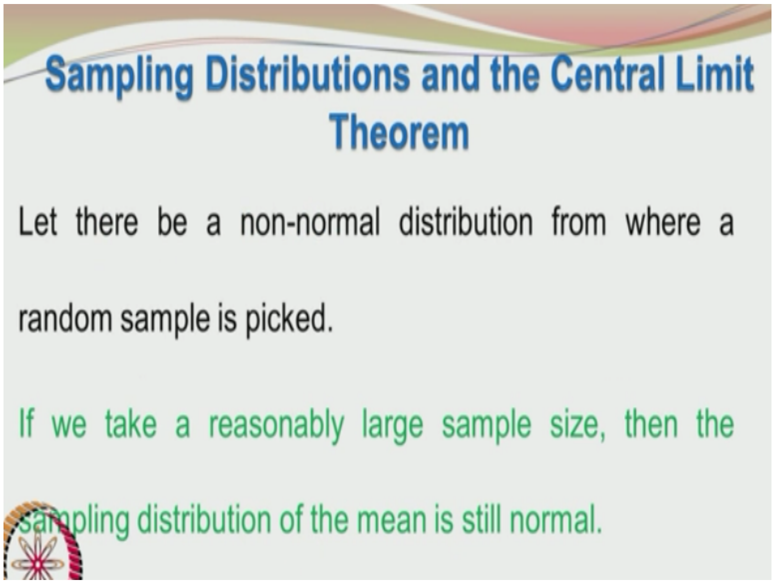
Since the parent populations probability distribution is usually not known. We cannot also say directly what is the sample statistics sampling distribution.

(Refer Slide Time: 02:41)



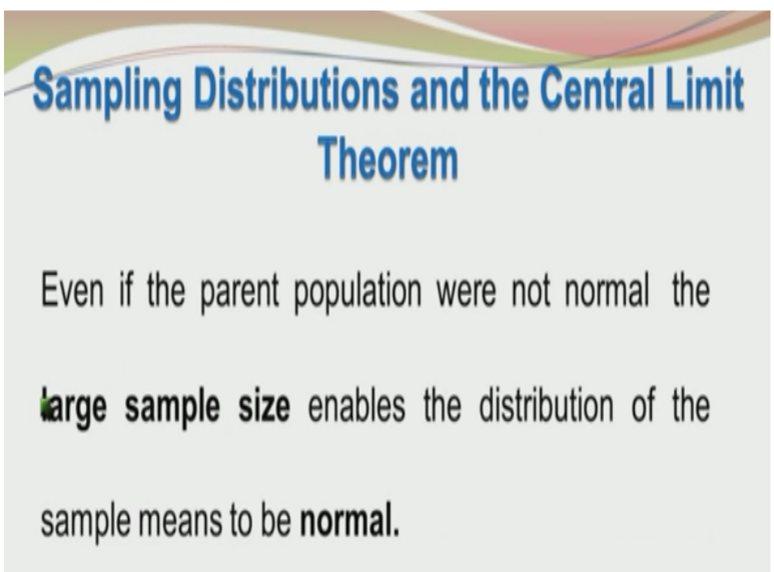
The central limit theorem simplifies matters a lot by stating that even if the original probability distribution of the population is not normal, i.e., it is not Gaussian, the sample mean distribution tends towards the normality provided the sample size is high (say > 30).

(Refer Slide Time: 03:02)



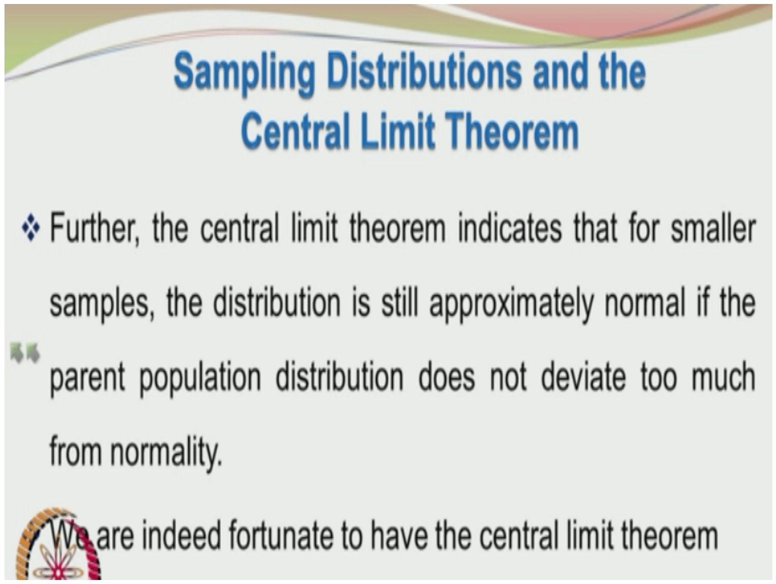
If there is a non-normal population from where a random sampled is picked. If we take a reasonably large sample size, the sampling distribution of the mean is normal. So the important thing is the sampling distribution of the mean is tending towards normality, provided the sample size is reasonably large.

(Refer Slide Time: 03:34)



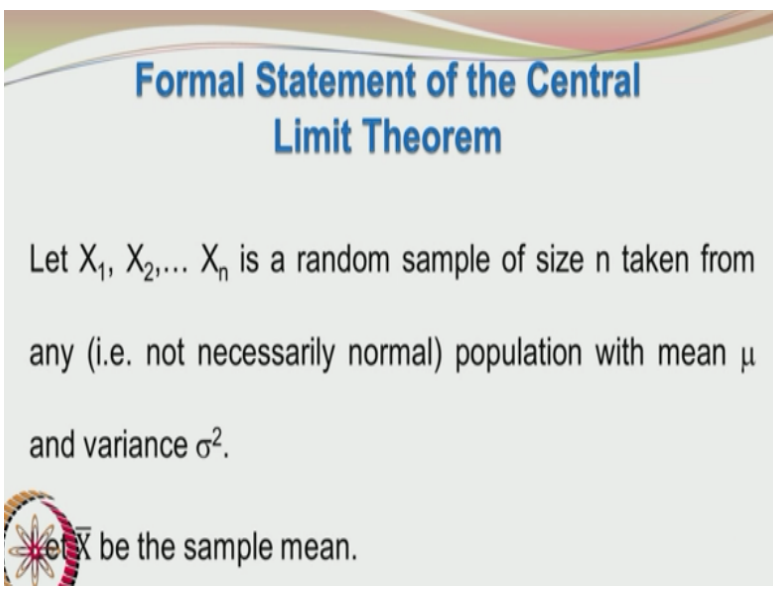
Further even for smaller samples, the distribution is still approximately normal if the parent population distribution does not deviate too much from normality. Even if you have a small sample, the sample size is small, the distribution is still approximately normal if the parent population distribution does not deviate too much from normality. So it is indeed fortunate that we have the central limit theorem.

(Refer Slide Time: 04:11)



So now let us make the formal statement of the central limit theorem. Let X1, X2 so on to Xn be a random sample of size n taken from any not necessarily normal population with mean mu and variance sigma square. Let X bar be the sample mean.

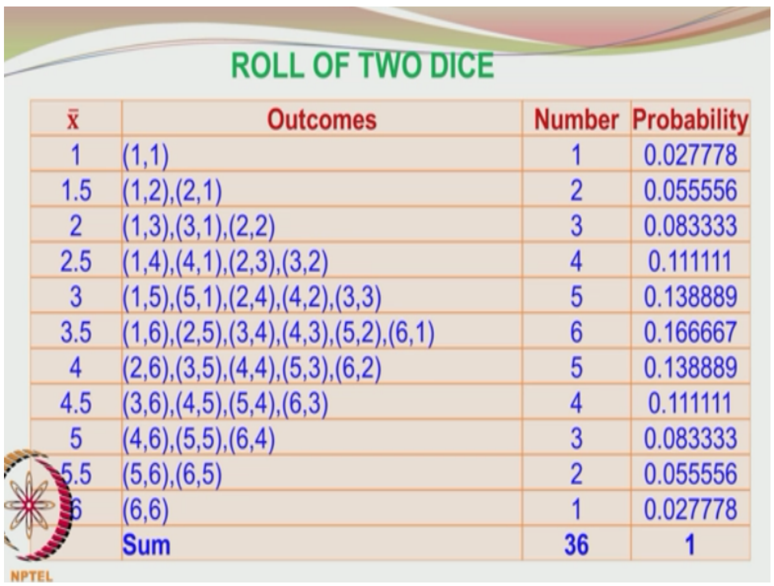
(Refer Slide Time: 04:29)



The limiting form of the distribution of the standard normal variable z=X bar-mu/sigma/root n as n tends to infinity is the standard normal distribution. Since it is a standard normal distribution, we are using the symbol z. So z=X bar-mu/sigma/root n. the limiting form of the distribution of z=X bar-mu/sigma/root n, as n tends to infinity is the standard normal distribution. What we are doing is we are creating a new random variable z by expressing it or defining it in terms of X bar –mu whole divided by sigma/root n.

Here X bar is the sample mean, mu is the population mean, sigma is the population standard deviation, n is the sample size. When n tends to a large number, then this random variable tends towards a standard normal distribution. Please recall that the standard normal distribution is something which is having mean 0 and variance of unity.

(Refer Slide Time: 05:54)



ROLL OF TWO DICE

$\bar{x}$	Outcomes	Number	Probability
1	(1,1)	1	0.027778
1.5	(1,2),(2,1)	2	0.055556
2	(1,3),(3,1),(2,2)	3	0.083333
2.5	(1,4),(4,1),(2,3),(3,2)	4	0.111111
3	(1,5),(5,1),(2,4),(4,2),(3,3)	5	0.138889
3.5	(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)	6	0.166667
4	(2,6),(3,5),(4,4),(5,3),(6,2)	5	0.138889
4.5	(3,6),(4,5),(5,4),(6,3)	4	0.111111
5	(4,6),(5,5),(6,4)	3	0.083333
5.5	(5,6),(6,5)	2	0.055556
6	(6,6)	1	0.027778
	Sum	36	1

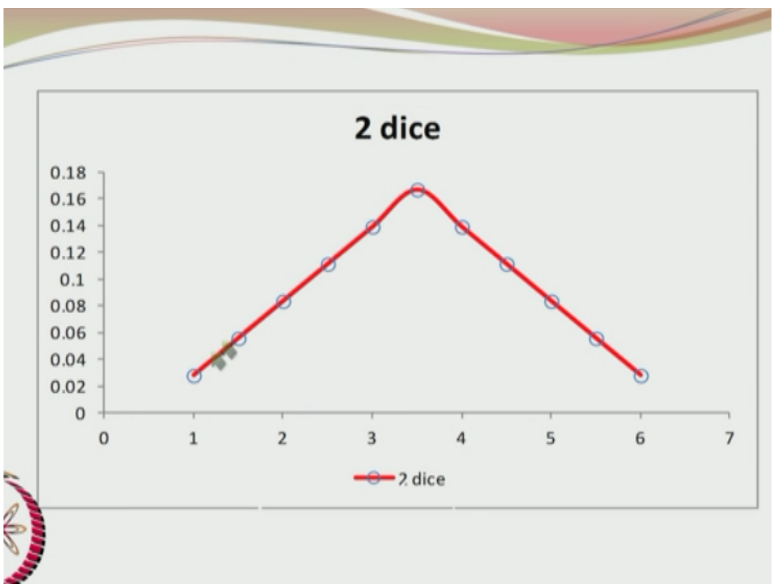
Let us take some interesting examples. The first one involving the role of 2 dice. Lot of theory has been built from games. We will be demonstrating the central limit theorem with a couple of simple examples. We have the outcomes tabulated here, all the possible outcomes are tabulated here. You have 1 and 1 that means the first dice is showing 1, the second dice is also showing 1, the average of 1 and 1 would be 1 and the number of such outcomes is only 1.

When you have the dice showing numbers 1 and 2, the first dice may show 1 and the second dice may show 2 or the first dice may show 2 and the second dice may show 1. So there are 2 possible outcomes. So that is why the number 2 has been put and the average of 1+2 and 2+1 would be both 1.5. Similarly, you have for other cases. For example, if you have 3.5 as the average that may be formed by the combinations, (1, 6) (2, 5) (3, 4) (4, 3) (5, 2) and (6, 1).

In such a case, you can have 6 base of getting the average 3.5. So these are the 6 possibilities through which we can get an average of 3.5. Similarly, we can do for the other averages also. You cannot get an average of 1.33 with 2 dice or you cannot get an average of 5.25 with 2 dice, we can get only these as the possible outcomes and the numbers are here and so the numbers will add up to 36. There are 6 ways in which the first dice can throw up the results.

There are 6 ways in which the second independent dice will throw up the results, so you have 6 x 6, which is 36 possible outcomes and the probabilities are calculated based on the number divided by the total number 1/36, 2/36 and so on and these are the probability values and they sum up to 1. This can be represented on a graph. We can plot the probability and X bar.

(Refer Slide Time: 08:53)



So when you plot the probability versus X bar, you can see a kind of a hat. This is definitely not a bell shaped curve, but it is more of a hat shaped curve and these are the probabilities. We are talking about discrete probability outcomes and so we can directly mark the probability against the outcome. So 1 was about 0.03, which is 0.027 and that is what you have here. So the probabilities are marked against each of the averages that are possible.

(Refer Slide Time: 09:43)

ROLL OF THREE DICE

outcome	Mean	Possible Outcomes					
3	1	111					
4	1.333	121					
5	1.666	221	113				
6	2	114	123	222			
7	2.333	115	124	133	223		
8	2.667	116	125	134	224	233	
9	3	126	135	144	225	234	333
10	3.333	136	145	226	235	244	334
11	3.667	326	335	344	425	461	515
12	4	156	246	354	444	525	633
13	4.333	166	256	346	355	445	
14	4.667	266	356	446	455		
15	5	366	465	555			
16	5.333	466	556				
17	5.667	566					
18	6	666					

Let us see what is going to happen when you have three dice. It becomes slightly more cumbersome. You have more occurrences of the mean. The discrete probability distribution tends towards the continuous one or appears to be continuous as you increase the number of dice. You can now get more possibilities of the mean, you can get 1, you can also get an outcome of 4, the sum of numbers appearing on the dice.

The sum of numbers appearing on the 3 dice can be 4 and the mean value would be 4/3, which is 1.333. How can you get the number 4. The dice will have numbers 1, 2, and 1 and the 3 dice can roll in a such a way to get 1, 2, 1 in three different ways. It can be (1, 2, 1) (1, 1, 2) and (2, 1, 1). There can be three ways in which the number 1, 2, 1 may arise. Similarly, when you look at the outcome as 5, it can be 2, 2, 1 or 1, 1, 3. You can get 2, 2, 1 in three different ways.

You can get 1, 1, 3 in three different ways. The average is 5/3, which is 1.667. So when you do like that for all the possible cases, you can see that the averages can range from 1 to 6 and there is a finer division of the interval between 1 to 6 because you are having 3 dice.

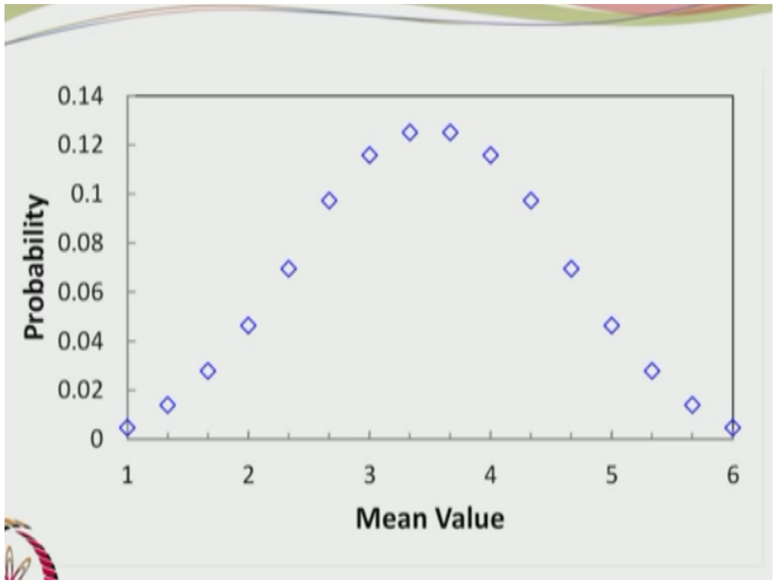
(Refer Slide Time: 11:34)

outcome	Mean	Frequency of Occurrence						Total	Probability
3	1	1						1	0.00463
4	1.333	3						3	0.013889
5	1.667	3	3					6	0.02778
6	2	3	6	1				10	0.04630
7	2.333	3	6	3	3			15	0.06944
8	2.667	3	6	6	3	3		21	0.09722
9	3	6	6	3	3	6	1	25	0.11574
10	3.333	6	6	3	6	3	3	27	0.125
11	3.667	6	3	3	6	6	3	27	0.125
12	4	6	6	6	1	3	3	25	0.11574
13	4.333	3	6	6	3	3		21	0.09722
14	4.667	3	6	3	3			15	0.06944
15	5	3	6	1				10	0.04630
16	5.333	3	3					6	0.027778
17	5.667	3						3	0.01389
18	6	1						1	0.00469
								216	1

So you look at the frequency of the occurrence and 3 will occur in only one way, the outcome of 3 or a mean of 1 can occur in only one way. An outcome of 4 or a mean of 1.333 can occur in 3 ways. Like that you can see the number of occurrences for all these outcomes or all these means, both are equivalent. You can see that the numbers can be recorded in this table and they can be counted and that would be total ways in which a number 3 can arise.

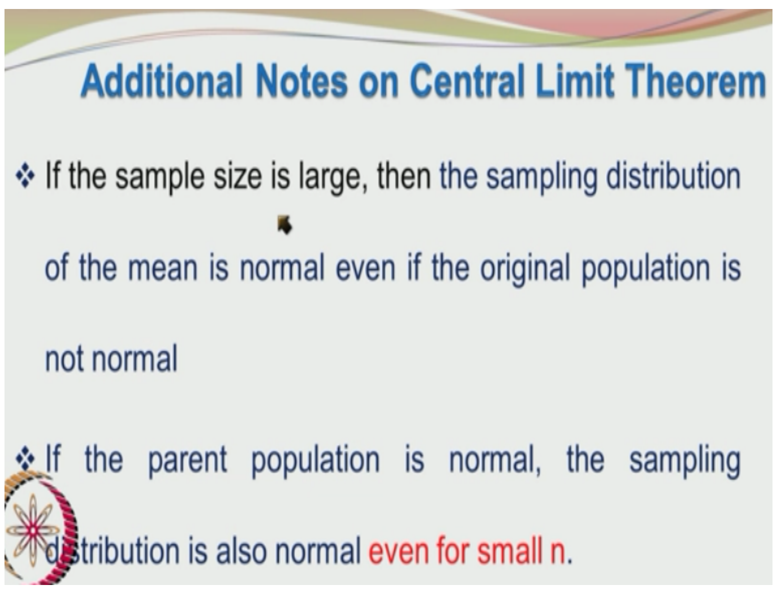
The sum of numbers on the 3 dice are average of 1 can realize. There can be 3 ways in which a number of 4 can totally appear on the three dice or a mean of 1.333 can arise and so you can have all these possibilities. Since you are talking about 3 dice, the number of possible outcomes are 6*6*6, which is 216. So you have 216 here. The probability can be obtained by dividing 1/216, 3/216, 6/216. So you can have all these probabilities and they can add up to 1.

(Refer Slide Time: 12:55)



No even with 3 dice, you are taking an average based on n=3. Sample size=3. You can see that the distribution is tending towards normality. It is appearing more bell shaped. For n=2, you had a hat shape, now you are getting slightly a broader peak.

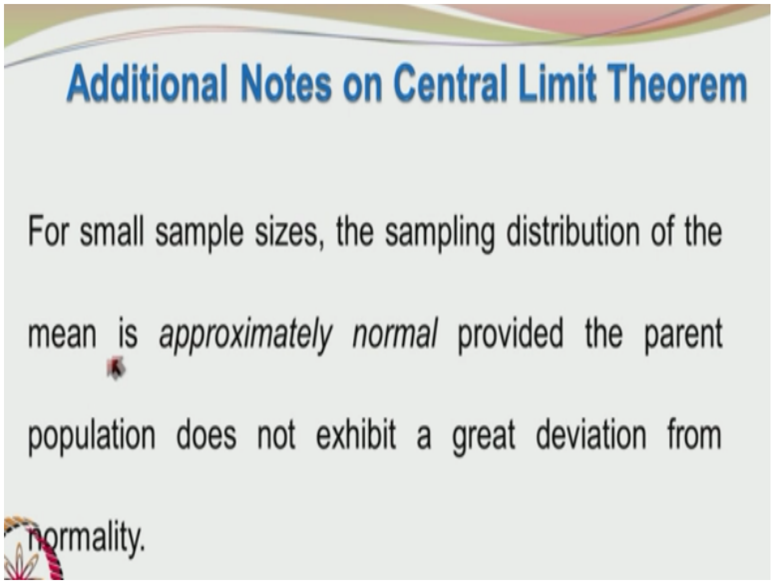
(Refer Slide Time: 13:32)



So if the sample size is large, the sampling distribution of the means is normal even if the original population is not normal. If the parent population is normal, the sampling distribution is also normal even for small n. There are 2 distinct cases. The parent population is not normal, but the sample size is large. The resulting distribution of the sample means is normal. In the second the parent population itself is normal.

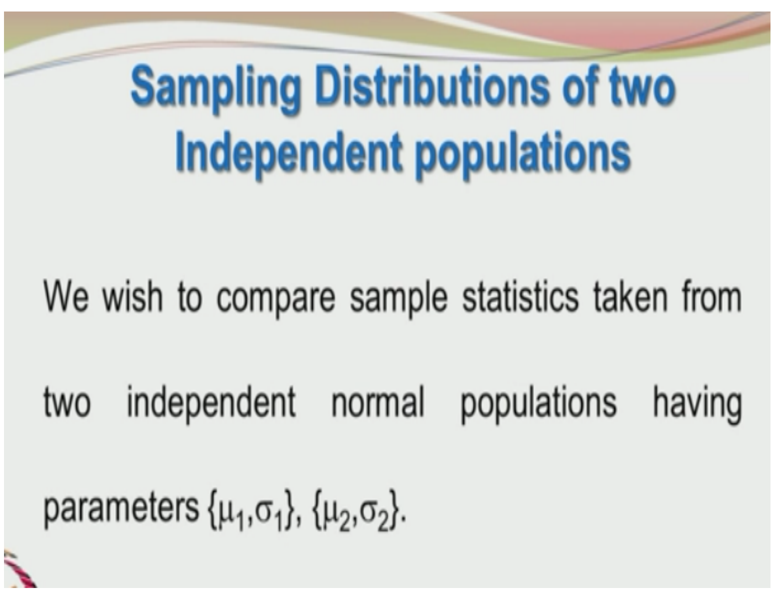
So even if you take a small sample from such as population and you look at the distribution of the sample means, you will find the sampling distribution of the means is also normal, even for small n.

(Refer Slide Time: 14:31)



For small sample sizes, the sampling distribution of the means is approximately normal provided the parent population does not exhibit a great deviation from normality. Even if the parent population was not normal, it was only slightly deviating from normal and you have a small sample size. A sampling distribution in such as case involving the small sample size would also tend to be approximately normal.

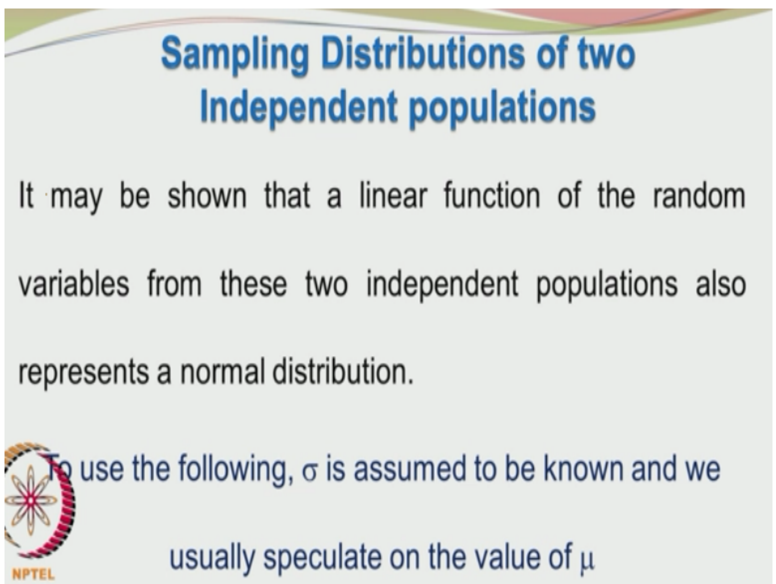
(Refer Slide Time: 15:04)



We were looking at variance of X1+X2. We were also looking at variance of X1-X2 and expected value of X1+X2, expected value of X1-X2. The reason for doing that is in our statistical applications, we may wish to compare sample statistics taken from 2 independent normal populations. Let us say that we are taking the sample statistics from 2 independent normal populations.

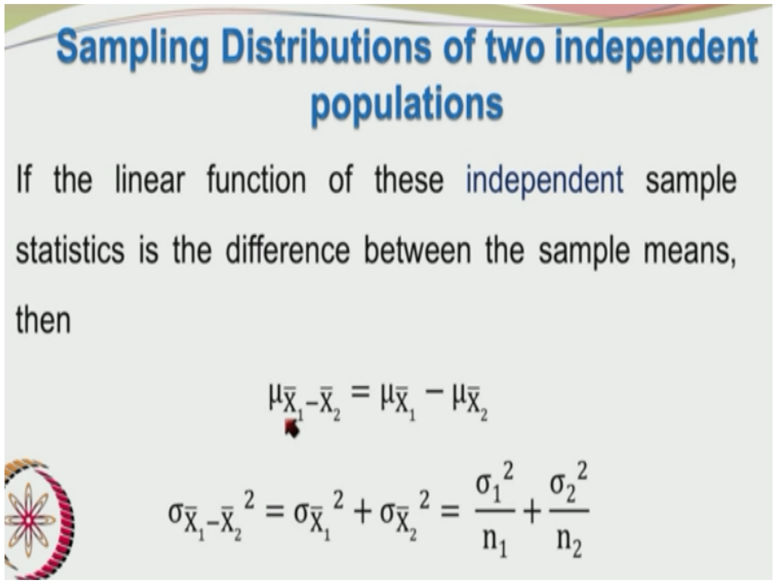
Both the populations are normal, from where the samples are taken and sample statistics are calculated. Let us say that the 2 normal populations have different parameters and they are mu1, sigma 1 for the first population, mu2, sigma2 for the second population, mu1 is different or may be different from mu2, sigma1 may be different from sigma2 and that is what I meant by 2 populations which are belonging to the same type, but they are having different parameters.

(Refer Slide Time: 16:18)



We know by now that the linear function of the random variables from these 2 independent populations is also a normal distribution, because the original random variables were from normal distribution themselves. Let us assume before we go to the most general case, let us assume that sigma is known and so we do not know only the value of mu.

(Refer Slide Time: 16:49)



Now let us consider a linear function of the independent sample statistics. Let us define the linear function as X1 bar-X2 bar, X1 and X2 are random variables. X1 bar and X2 bar are also random variables. X1 bar-X2 bar would also be a random variable and that would be having a probability distribution. What is the mean mu of such a distribution X1 bar-X2 bar. This can be written as expected value of X1 bar-X2 bar, which is expected value of X1 bar-expected value of X2 bar.

By now, you should be familiar with this. That is why, I am not giving you the steps and that can be written as mu of X1 bar-mu of X2 bar. The mean of the first sampling distribution of the mean-the mean of the second sampling distribution of the mean. This is interesting and this is important. So X1 bar-X2 bar is a random variable. It is having a probability distribution and it will have its variance.

What is the variance of the distribution formed by the difference of the 2 sampling means X1 bar-X2 bar. What is the variance? That would be sigma X1 bar square + sigma X2 bar square. We saw that variance of X1+X2=variance of X1+variance of X2. X1 and X2 can be any random variable. In the present case, X1 is X1 bar, X2 is X2 bar. Do not look at it as X1 and X1 bar as very different quantities. X1 is a random variable, X1 bar is also a random variable.

X2 is a random variable, X2 bar is also a random variable. So when you are trying to find the variance of difference of any 2 random variables, it would still be the sum of the variances of the

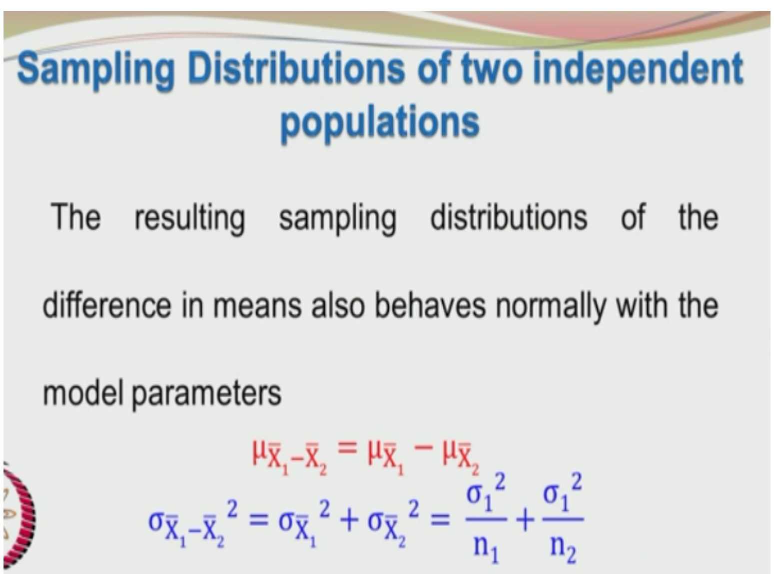
2 random variables in question, provided the 2 random variables were independent and that is why we are talking about 2 independent populations. So we are having sigma X1 bar square + sigma X2 bar square. Now we have to ask ourselves, what is sigma X1 bar square?

What is the variance of the probability distribution formed by X1 bar? What is the variance of the sampling distribution of X1 bar? The variance of the sampling distribution of X1 bar would be sigma1 square/n1. The variance of the sampling distributions of the mean X2 bar is given by sigma2 square/n2. Sigma1 square is the variance of the first population. Sigma2 square is the variance of the second population.

Sigma1 square/n1 is the variance of the sampling distributions of the mean corresponding to X1 bar. Sigma2 square/n2 is the variance of the sampling distributions of the means corresponding to X2 bar, n1 and n2 are the samples sizes for X1 bar and samples size for the X2 bar. This is a very important concept. I request you to think it over, understand it and try to write down the combinations properties on a paper after thinking about these concepts and see whether you are able to understand.

Otherwise you again go through the lectures and see where you did not understand.

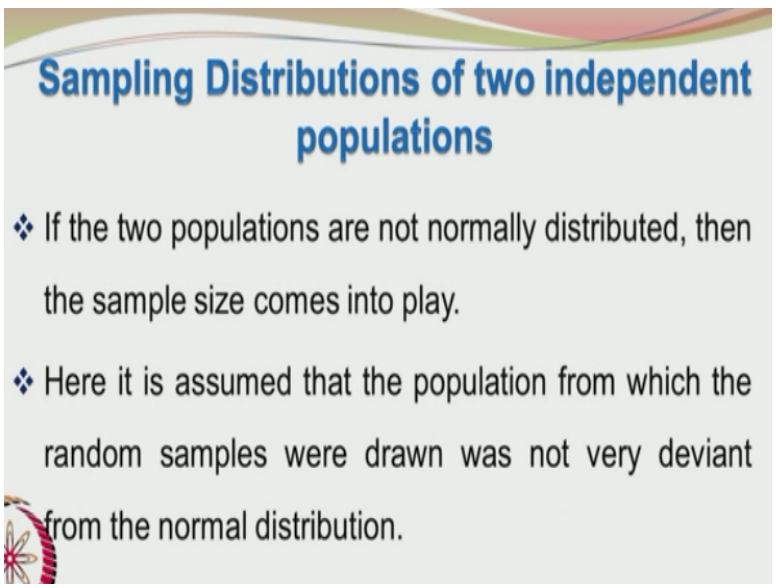
(Refer Slide Time: 21:18)



So if the 2 parent populations were normal in addition to being independent, then the resulting distribution formed by the difference of X1 bar and X2 bar would also be normal and the parameters would be mu of X1 bar-mu of X2 bar. What is mu of X1 bar? What is the mean of the sampling distributions of X1 bar? In other words, what is the expected value of X1 bar. We know by now; it should be mu1.

Similarly expected value of X2 bar or mean of the distribution formed by X2 bar would be mu2. So you will have mu1-mu2. Similarly, the variance of the distribution formed by the difference of the 2 samples means X1 bar and X2 bar would have sigma1 square/n1+, there is a typo, I will correct it, sigma2 square/n2. So the variance of the distribution formed by the difference between the 2 sample means X1 bar and X2 bar would be sigma X1 bar square + sigma X2 bar square, which is nothing but sigma1 square/n1+sigma2 square/n2.

(Refer Slide Time: 23:00)



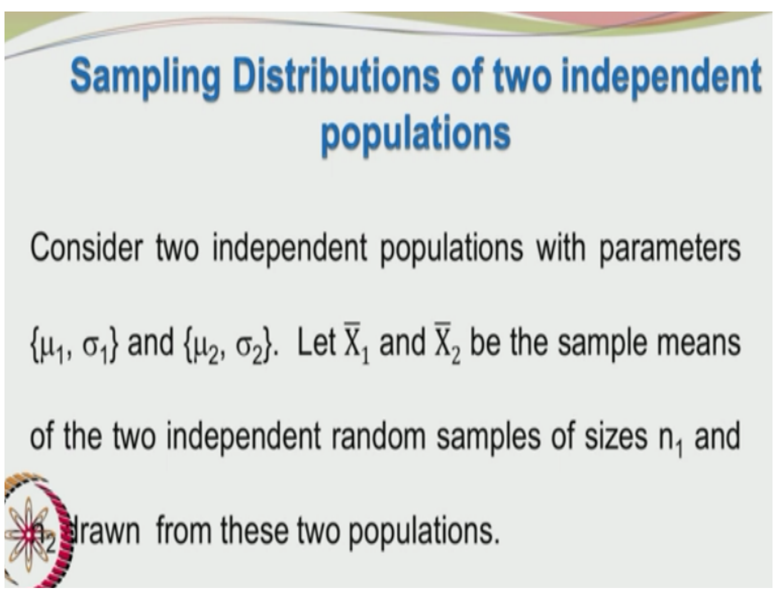
If the populations are not normally distributed, then what can you say about the resulting sampling distribution of the mean. It would depend upon the sample size. If you assume that the population from which the random samples were drawn was not very deviant from the normal distribution, then for samples sizes >30, the 2 independent sampling distributions are approximately normal and a linear combination of them would also behave approximately normal.

Here what you are doing is quite important. We are now talking of difference between 2 sample means X1 bar and X2 bar. X1 bar and X2 bar have been taken from 2 different populations 1 and 2. Please do not confuse with X1 bar and X2 bar being taken from the same population. Now we are talking about 2 different populations and we are taking samples from these 2 populations and we represent them by X1 bar and X2 bar.

Now we are looking at the resulting distribution we will get based on the difference between the 2 sampling means and what are observing is, if the sample sizes are >30 in both the cases, the sample taken from the first population is having the size >30, the sample taken from the second population is also having the size >30, and according to the central limit theorem, the 2 independent sampling distributions would be behaving normally and hence a linear combination of them would also behave approximately normally.

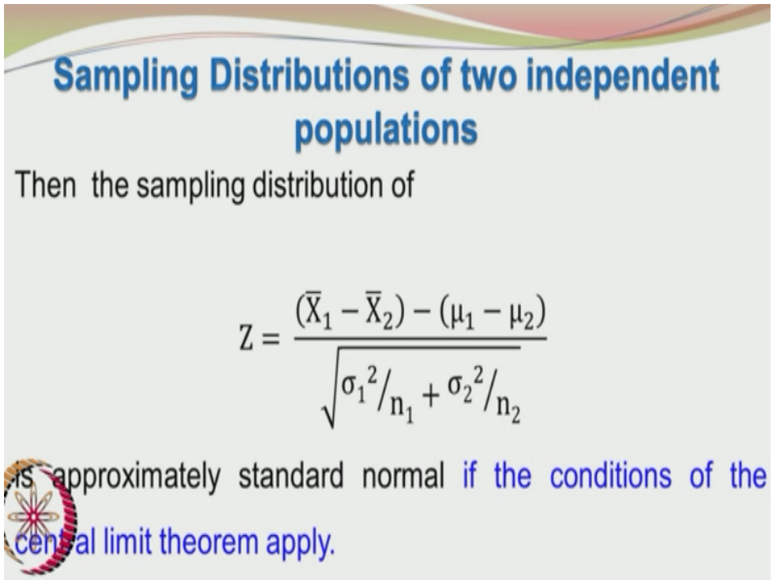
So according to the central limit theorem, since the sample size was >30, X1 bar would behave in a normal manner. The sampling distribution of X2 bar would also behave in a normal fashion. Under linear combination of them, here X1 bar and X2 bar would also be behaving approximately normally.

(Refer Slide Time: 25:54)



Consider 2 independent parameters mu1 and sigma1 and mu2, sigma2. Let X1 bar and X2 bar be the sample means of the 2 independent random samples of sizes n1 and n2 drawn from these 2 populations.

(Refer Slide Time: 26:13)



So now we are going to define a new random variable based on the difference between the 2 random sample means. These are 2 independent random samples drawn from 2 different populations of mean mu1 and mean mu2 and variance sigma1 square and variance sigma2 square. I am talking about the 2 populations of means mu1, sigma1 square and mu2, sigma2 square. Now we are having the sample means, X1 bar and X2 bar.

We are taking the difference of them. Then, we subtract this quantity X1 bar-X2 bar with mu1-mu2, also note that the expected value of X1 bar would be mu1, expected value of X2 bar would be mu2. As far as the original population as well as the sampling distributions go, their means are identical. The mean of the sampling distribution of the means is = the population mean, but the same thing is not true with the variance.

The sampling distribution of the means will have a variance sigma square/n, where n is the sample size. So as far as the variance is concerned, the sample size comes into play. So we know that the variance of X1 bar=sigma1 square/n1, variance of X2 bar, the variance of the sampling distribution of the means for X2 bar would be sigma2 square by n2. The variance of X1 bar-X2 bar=sigma1 square/n1+sigma2 square/n2.

When you are making this combination, we are not arbitrarily choosing our mu1 and mu2, we are not arbitrarily choosing sigma1 square/n1 and sigma2 square/n2, you may recollect the standard normal variable was defined as Z=X-mu/sigma, where mu was the mean and sigma was the standard deviation of the population from where X was chosen. So we are taking X1 bar-X2 bar and we are looking at that corresponding distributions mean, which is mu1-mu2 and the variance sigma1 square/n1+sigma2 square/n2.

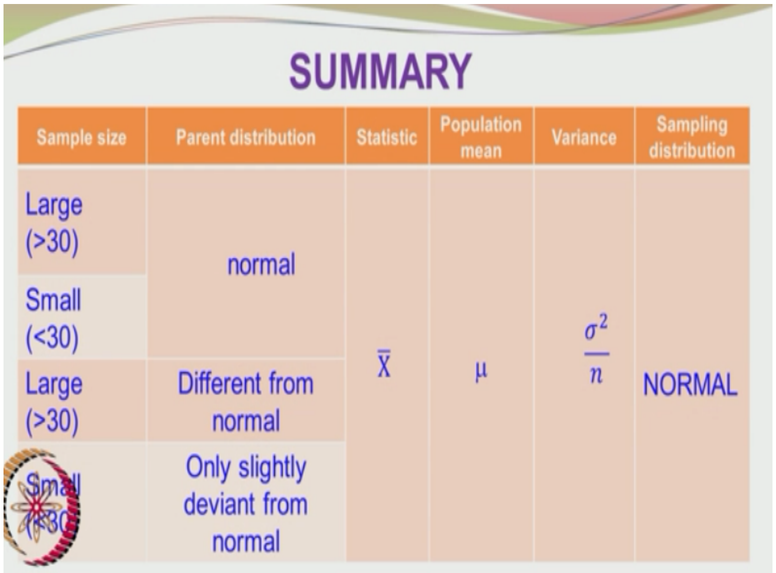
So that sigma is square root of that would become the standard deviation. So we are defining a standard normal variable, because of the central limit theorem, the X1 bar-X2 bar was behaving approximately normally owing to the large sample size. Because of the large sample size for X1 bar, because of the large sample size for X2 bar, both of the according to the central limit theorem would tend to exhibit normal behavior.

A linear combination of the 2 random variables X1 bar and X2 bar would also tend towards normal behavior and so we are creating a standard normal variable Z for this particular situation and that standard normal variable is given by (X1 bar-X2 bar) – (mu1-mu2)/square root sigma1 square/n1+sigma2 square/n2. If the two populations are normal, right now we are looking at two original normal populations.

Then irrespective of the sample size, you are not constrained by a small sample size, in such a situation, (X1 bar-X2 bar) – (mu1-mu2)/square root of sigma1 square/n1+sigma2 square/n2 will be a standard normal. In the previous case, the original populations 1 and 2 were not normally distributed, but large samples were chosen. So the sampling distribution of the difference in means also behaved normally and for large sample sizes n1>30 and N2>30, we had the standard normal variable.

In the easier case, where both the populations are coming from normal distributions even if the sample sizes for both the sample means X1 bar and X2 bar are small, even then the resulting distribution of the sampling distribution of the means X1 bar-X2 bar would be normal. Because the parent populations were themselves normal.

(Refer Slide Time: 32:04)



SUMMARY

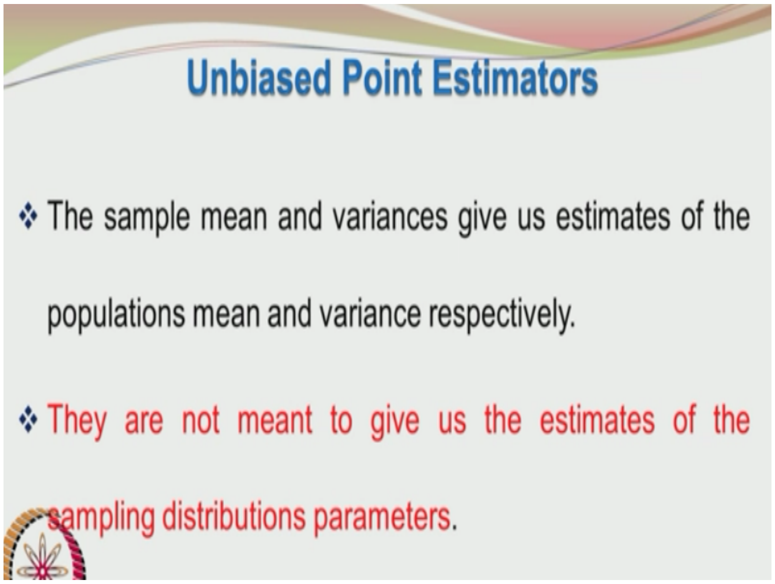
Sample size	Parent distribution	Statistic	Population mean	Variance	Sampling distribution
Large (>30)	normal				
Small (<30)					
Large (>30)	Different from normal	$\bar{X}$	μ	$\frac{\sigma^2}{n}$	NORMAL
Small (<30)	Only slightly deviant from normal				

So this is what, I am summarizing here. You are having a large sample size >30, parent distribution is also normal, the statistic involved is X bar, the population mean is mu, variance is sigma square/n, the sampling distribution is normal. If the sample size is small and the parent distribution is normal, does not matter, he resulting sampling distribution of the mean would be normal. You have a large sample >30. The parent distribution is different from normal.

Nothing to worry, central limit theorem will help us and the sampling distribution of the mean would be normal with mean mu and variance sigma square/n. The population mean would also be equal to the sampling distribution mean. The population variance is sigma square, but the sampling distribution variance would be sigma square/n. So the sampling distribution variance is sigma square/n and you have a large sample size.

The parent distribution is different from normal. The resulting distribution of the sample would have mean mu and sigma square/n owing to the central limit theorem, it would be normal. If you have a small size, < 30 the parent distribution is only slightly deviating from normal, then also you can assume that the sampling distribution of the mean would be approximately normal with mean mu and variance sigma square/n.

(Refer Slide Time: 33:53)

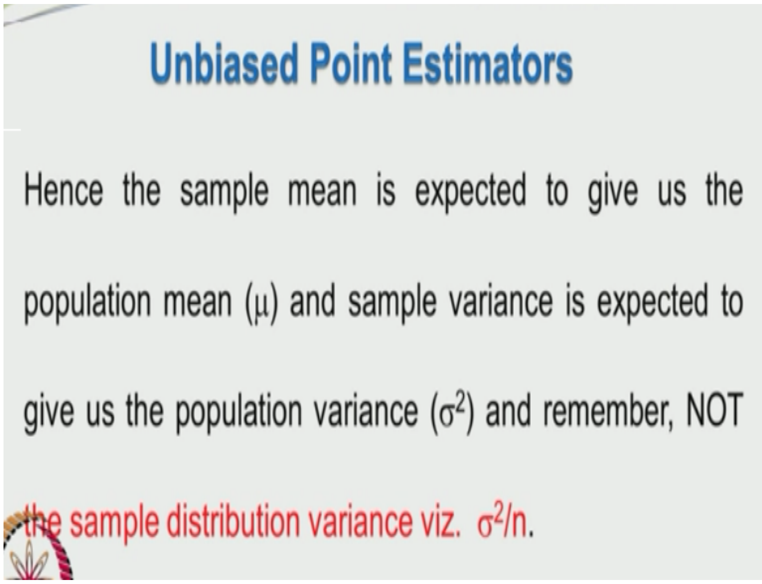


Now let us look at the desirable properties of the point estimators. We have seen that we are estimating the population parameters mu and sigma square by using sample statistics. We are using the sample mean X bar and the sample standard deviation S to get good point estimates of the population mean mu and population standard deviation sigma. We are talking about good point estimators. We will qualify it even further by saying them as unbiased point estimators.

The sample mean and sample variance give us estimates of the population mean and variance respectively. They are not meant to give us estimates of the sampling distributions parameters. We are talking about samples taken from a population. The samples have been taken from the population to get idea about the population parameters. We are not using the sample estimators to help us to find the sampling distribution parameters.

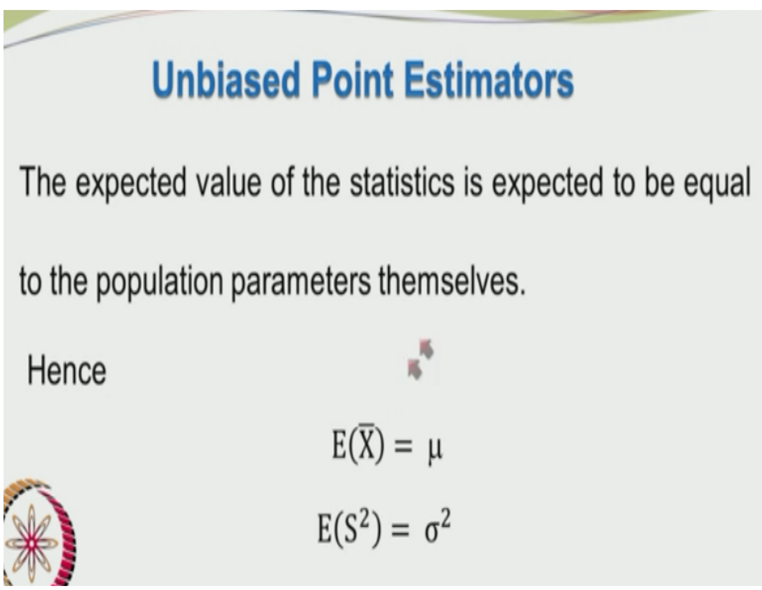
This is an important difference, which we should be aware of. We are using sample estimators X bar and S square to know about mu and sigma square of the original population. We are not using X bar and S square to get us estimates of the sampling distribution properties. Once the information of mu and sigma square is estimated, then it would be helpful for us. How, we will be seeing some examples in the future.

(Refer Slide Time: 36:02)



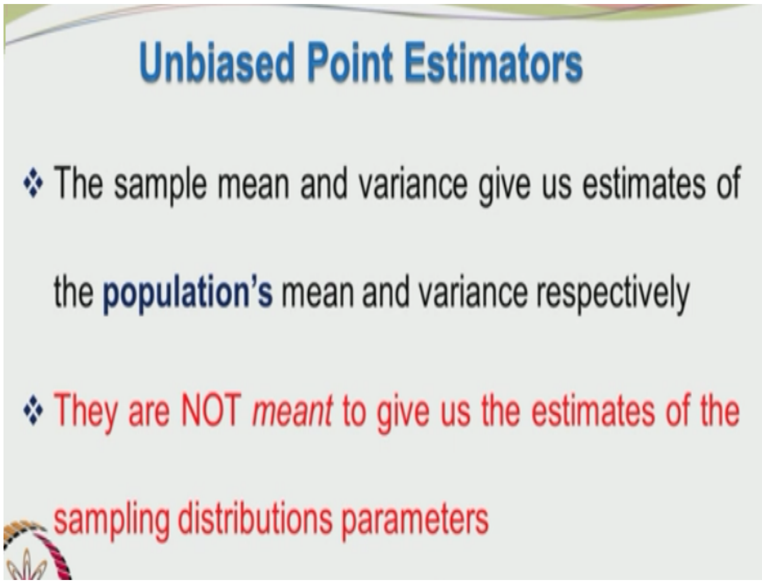
A sample mean is expected to give us the population mean mu and sample variance is expected to give us the population variance sigma square and remember not the sampling distribution variance sigma square/n. I know the sample size n, I know the sigma square, estimated from the sample variance; however, conceptually we are querying the population through the random sample and most cases have only one random sample taken.

(Refer Slide Time: 36:54)



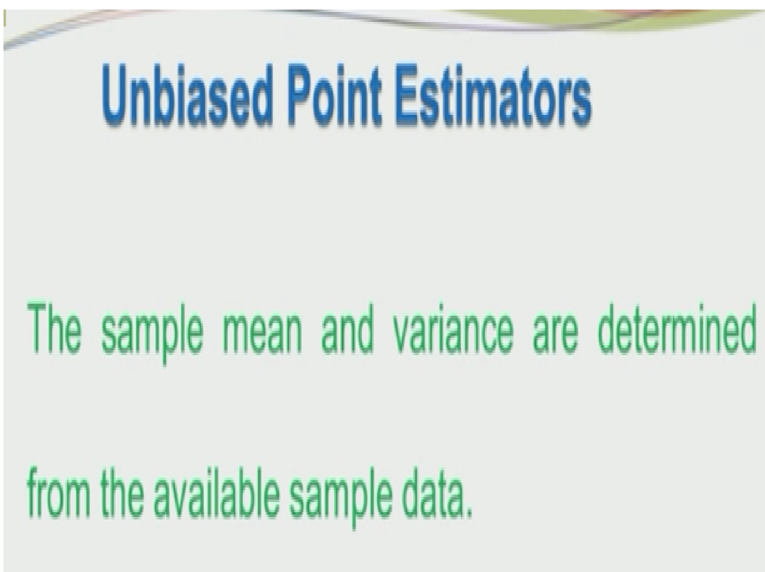
For unbiased point estimators, the expected value of X bar will be equal to mu and the expected value of S square will be equal to sigma square. What this means is the expected value of X bar that means the mean of the sampling distribution of the means would be equal to mu and the expected value of S square = sigma square. The sigma square is the population variance and the expected value that S square will take is also equal to sigma square, we can prove them.

(Refer Slide Time: 37:48)



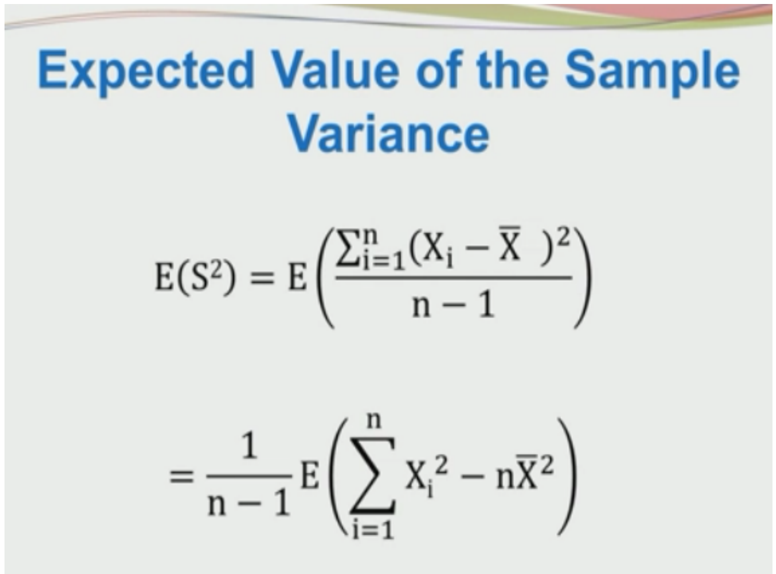
This I have already told you.

(Refer Slide Time: 37:54)



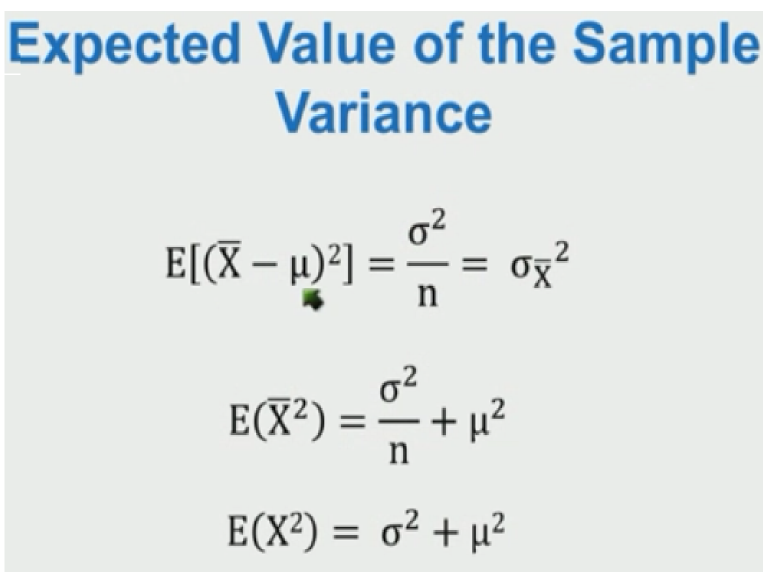
The sample mean and sample variance are only determined from the available sample data, sometimes the available sample may be only 1 and it may be also small in number. So whatever we have, we have to make do and draw the appropriate estimates.

(Refer Slide Time: 38:20)



This is very interesting. We have to prove that expected value of S square = sigma square, so I am just substituting the definition for S square here. So since n-1 is constant, we can take it out and you are essentially having expected value of sigma=1 to n, Xi-X bar whole square. We have already seen that this will reduce to sigma=1 to n X1 square – nX bar square. I request you to carry out the calculations on a paper on your own. If you are stuck, you please look at some of the earlier examples we have covered.

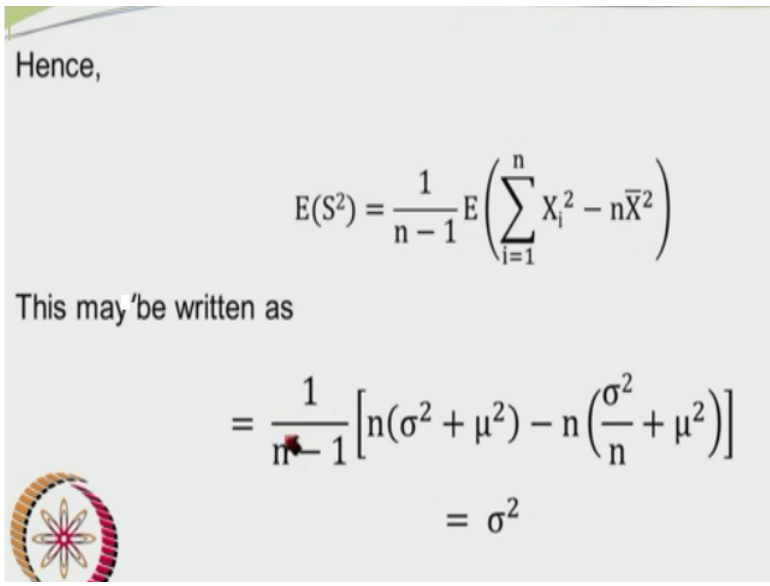
(Refer Slide Time: 39:12)



So expected value of X bar-mu whole square is the variance of X bar and that we get as sigma square/n. we also know that the expected value of X bar square is equal to sigma square/n+mu square. Previously, one of the first example set problems, we saw that expected value of X square was sigma square + mu square. The same concept, I am applying for expected value of X bar square.

Instead of sigma square, which was the variance for X, I am using sigma square/n, which is the variance for X bar and the mean of X was mu and the mean of X bar is also mu. So expected value of X bar square=sigma square/n + mu square.

(Refer Slide Time: 40:29)



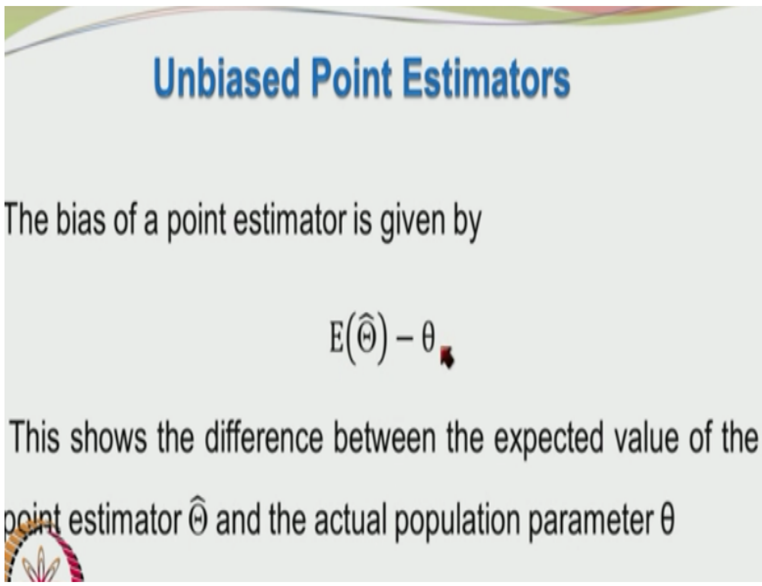
Hence we can write 1/n-1*expected value of sigma Xi square can be written as n*sigma square + mu square. Since all the random variables were identically distributed for each of these Xi squares we will write sigma square + mu square, then add it up, i=1 to n, sigma square + mu square n times, so that will become n*sigma square + mu square, then we write for the expected value of this n*X bar square. So we will have n and we use the previous result.

Expected value of X bar square=sigma square/n + mu square, we plug it in here and we have 1/n-1, n to sigma square + mu square-n*sigma square/n + mu square and so we get n-1 sigma square in the bracket, this n and n will cancel, you will have -1 sigma square, so this n mu square will cancel this n mu square, you will have n-1 sigma square resulting. That n-1 will cancel out with this n-1 and you get sigma square.

So by defining our variance, the sample variance in terms of n-1 makes it possible for us to have the sample variance S square as the unbiased estimator of the population variance sigma square. If we had n in our definition for the sample variance, this expected value of S square would have been different. That is not the same as the population variance sigma square, just by making the definition properly in terms of the degrees of freedom given as n-1 for the sample variance.

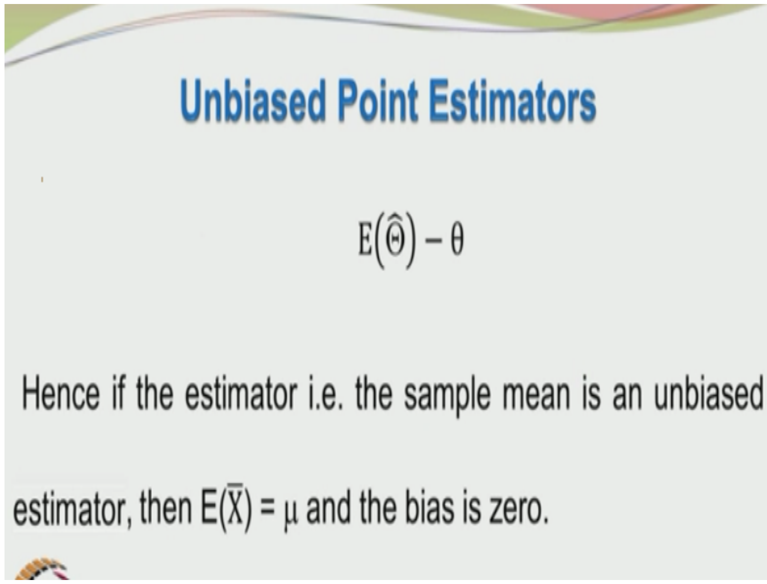
We can see that the expected value of S square is sigma square itself, hence S square is an unbiased estimator for the population variance sigma square.

(Refer Slide Time: 42:56)



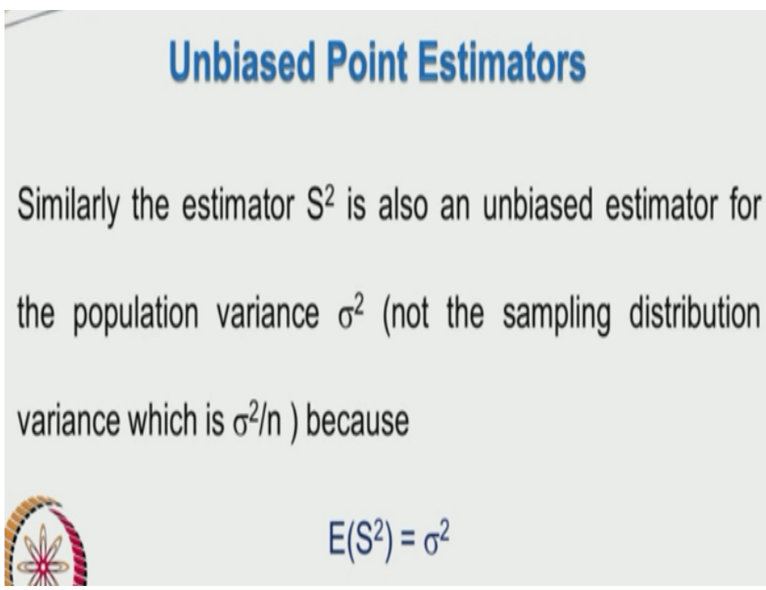
So the bias of a point estimator is given by the expected value of the estimator – the actual population parameter theta. We want the bias to be 0. We want the expected value of the point estimator to be theta itself so that we can get theta-theta=0, so that the bias disappears.

(Refer Slide Time: 43:22)



When you have X bar, which is the point estimator for the population mean, we are using the random sample mean as the point estimator for the population mean mu, expected value of X bar was mu and theta was also mu, mu-mu=0. So the bias has become 0. We can confidently proclaim that the sample mean is an unbiased estimator of the population mean mu.

(Refer Slide Time: 44:00)



Similarly, we saw that expected value of S square=sigma square, so we can proclaim that the S square, the sample variance is an unbiased estimator of the population variance sigma square. So concluding, we have seen the point estimation process. We were looking at random samples, the sample means, the sample means also behaved as random variables, it exhibited the full fledged probability distribution and the complication was we do not know about the population parameters mu and sigma.

We do not know the nature of the population whether it was normal, log normal or viable, but even with so many uncertainties by carefully choosing a sample and by using the sample statistics like the mean and sample variance, we were able to generate estimates of the population parameters mu and sigma square respectively and we are also able to show that these X bar and S square sample statistics where unbiased estimators of the 2 population parameters.

We also talked about the central limit theorem and the central limit theorem is a boon to us, because if we choose an adequately large sample size, say $n>30$, the sampling distribution of the mean behaved in a normal fashion even if the original distribution did not belong to the normal classification. So we have covered quite a lot of important ground here and these definitely form the bases for design of experiments and analysis of statistical data.

I would request you to revise the portions up to this point and be clear with the concepts. You do not have to remember the formulae or the rules. It is important for you to understand the concepts, assimilate the concepts and then the remaining part of the course would not only be easy, but also enjoyable. You will be able to directly relate to what we have covered up to this point, with what you are learning from now on. Thank you.