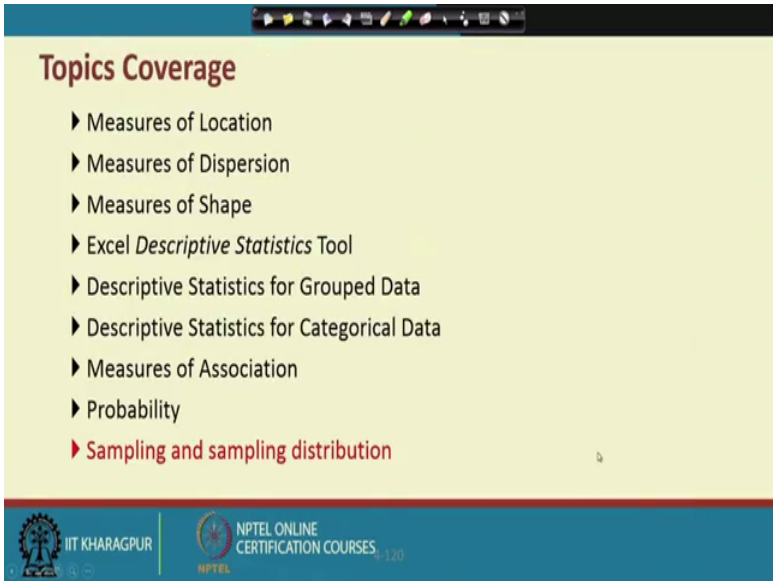
**Business Analytics for Management Decision**
**Prof. Rudra P Pradhan**
**Vinod Gupta School of Management**
**Indian Institute of Technology, Kharagpur**

**Lecture - 15**
**Descriptive Analytics (Contd.)**

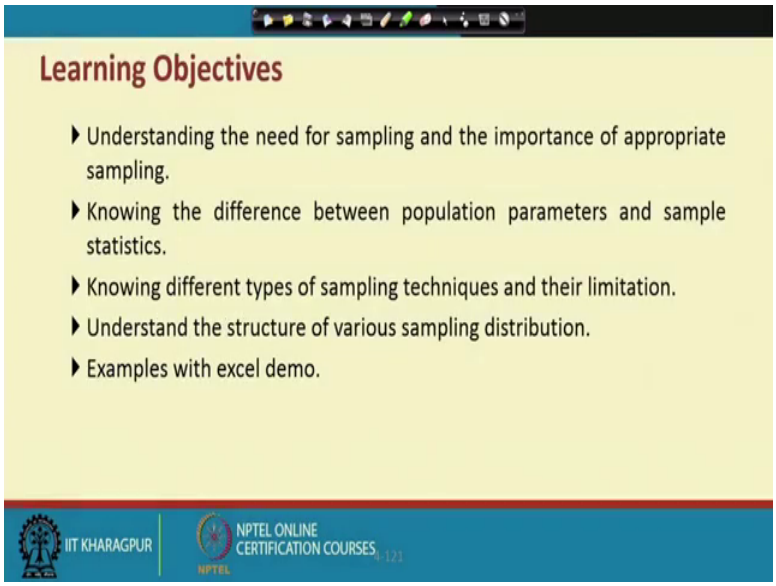
Hello everybody and this is Rudra Pradhan here. Welcome you all to BMDA course, and today we are here to discuss about the descriptive analytics. So, in the last couple of lectures, we have discussed various you know issues or you know and requirements, in fact, we have discussed the descriptive statistic, association statistics, the concept of probability and probability distributions.

(Refer Slide Time: 00:51)



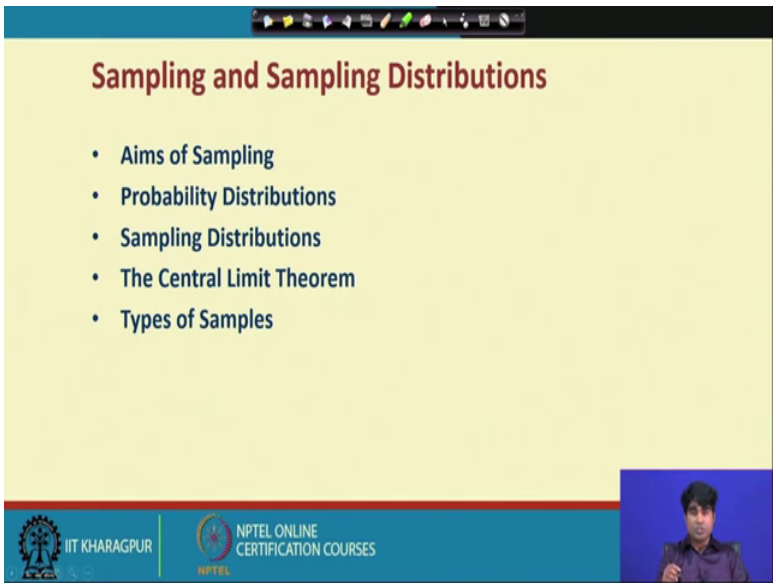

So, now connecting to all these lectures, so we will discuss now the concept called as you know sampling distributions. So, sampling and sampling distributions.

(Refer Slide Time: 01:00)



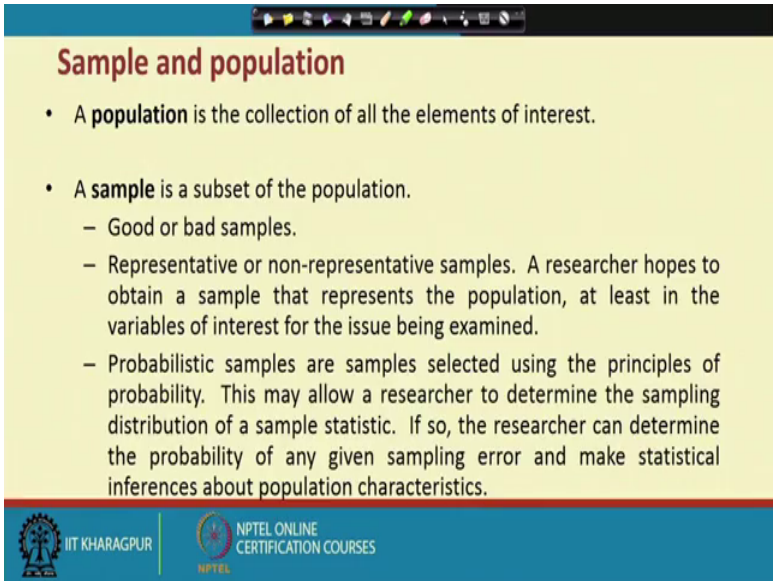

So, this is what the objectives behind this particular you know lectures that to sampling and sampling distribution. And the structure is to understand the need of sampling and importance of appropriate sampling, knowing the difference between population parameter and sample statistics, knowing different types of sampling techniques and their limitations, understand the structure of various sampling distributions. And finally, I will show you some kind of you know examples with the excel spreadsheets. So, this is what the discussion is all about.

(Refer Slide Time: 01:37)



And the kind of requirement is like this. What is actually it means we like to address the typical structures m of sampling, connection with the probability distributions, connection with the sampling distribution then we like to connect with the central limit theorems, and then types of you know samples. This is the typical structures or highlights of this particular you know lecture.

(Refer Slide Time: 02:04)




So, what we have discussed earlier. So, in any kind of you know empirical process or you know in problem investigation, so there is always you know issue about the sample and populations and what I have already mentioned the every time the experiment or you know empirical processing will be sample specific. So, you want to know what is the kind of you know population structure and corresponding to the population structures you have to draw a samples and then the investigation will be it done with the help of you know this is sampling only.

So, a sample is a subset of the population which you have already highlighted. So, it may have a good samples, it may have a bad samples. So, since it is a kind of you know basket, where you can pick up various samples; then if the sample is not actually good then you can actually reject then you pick up another sample and then test as per your requirement. So, in fact, the particular investigation process is very you know kind of you know continuous in nature, so if your theory is very accurate and the problem is well

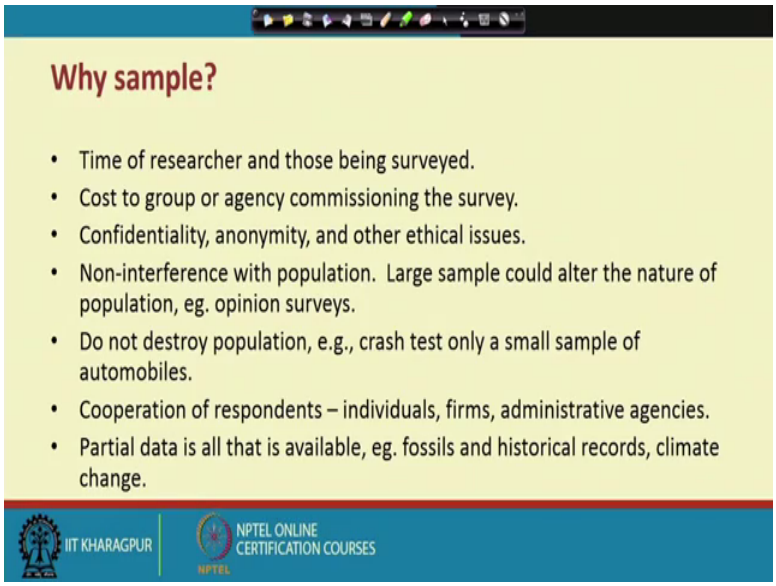
connected then sometimes with the help of you know samples, so you have to actually validate the particular fact or you know observations.

And if the sample will not be supportive that means typically if the sample is a bad one then obviously so it may not actually support to your theory and model. So, that is why so we typically follow sensitivity analysis and robustness check to validate whether you know the particular sample is a very accurate to predict certain kind of you know situation or to test particular situation. So, the sample will be good representative or sometimes non-representative samples. A researcher how to obtain a sample that represents the population so that means, the idea is actually so you have to try to you know pick up one after another samples, and then you know try to give conclusion or you know kind of findings which can general in nature and that will be population specific.

So, that means actually it is a kind of you know method of induction. So, you increase one after another process so that means, you can increase the sample also with a given kind of you know structure. So, starting with you know a 100 observation, 150 observation, 200 observation then you check the radiation. Actually if you know it is more or less you know coming similar kind of you know results then obviously so this will be very dirty or you know theory and the kind of you know models, so that is the beauty of actually the game between sampling to you know population.

So, that is how you have to understand what is the specific difference between in sample and population, but most of the times our structure is actually sample specific. So, I mean we have to empirically test the particular process or you know the kind of model through in the process of you know sampling only.

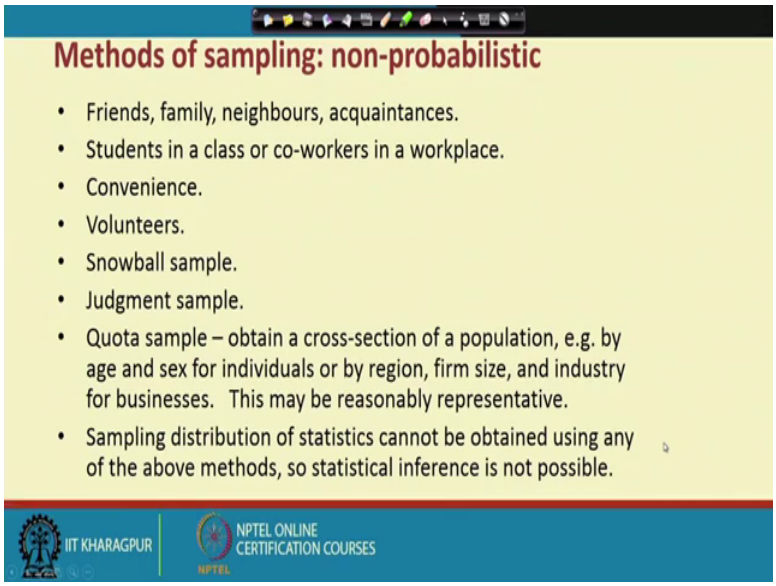
(Refer Slide Time: 05:14)




So, now the question is why samples because the means it is a very tricky kind of you know situation, because it is not possible for a researcher or you know any kind of you know organization or any kind of you know individual to gather all the sample which can be represented by a kind of population. Since, you are not in a position to touch all the samples so as a results, so you cannot say that this is a kind of you know population. Since you are less than all samples by default this particular structure is a sample specific; it cannot be population.

If you say population means all these sample points will be covered altogether. Since, nobody can you know in a position to collect all the samples then your every time the particular experiment or the particular testing will be sample specific that is how. So, these are the typically you know understanding, so why you need actually sample not the kind of you know population, because it is not possible for researcher together or kind of you know information.

(Refer Slide Time: 06:24)



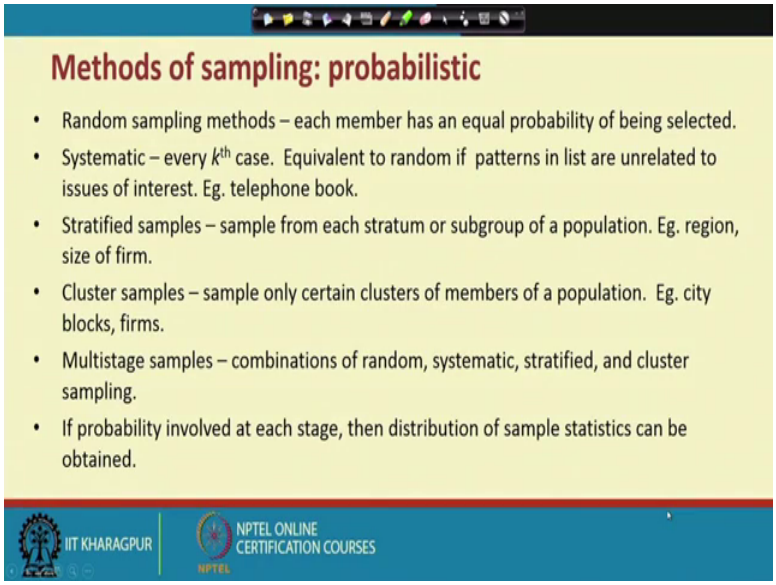

So, methods of sampling, so there is a probabilistic kind of you know structure and there is a kind of non-probabilistic structures. So, you can have the sample from friends, family, then you know students in a class or something kind you know coworkers in your workplace, then a convenient sampling volunteers. So, these are you know judgmental sampling and then you know these are all ways you actually you have to collect the samples. It means what I have mentioned actually when you collect the information, you must have a kind of you know description. So, the kind of you know reporting will be cross sectional specific and you know time specific.

But you know what is the particular cross sectional structure, and what is the particular you know time series structure that itself will give you some kind of you know clarity and understanding between the sample versus you know populations. But before you start the particular process, you must be very clear that you know what is exactly the sample and what should be the structure of the population. Of course, you are not in a position to say what is the exact you know population, but you should know what is the population basket altogether then you can pick up a samples.

And there should be some kind of you know convergence between sample and population. Every times our approach is actually sample to populations, so that you know you can generalize the particular you know findings or generalize the particular structures and accordingly the management decision can be applied. So, that is how it is

a kind of you know continuous process till you get the kind of you outcome as for your you know requirement.

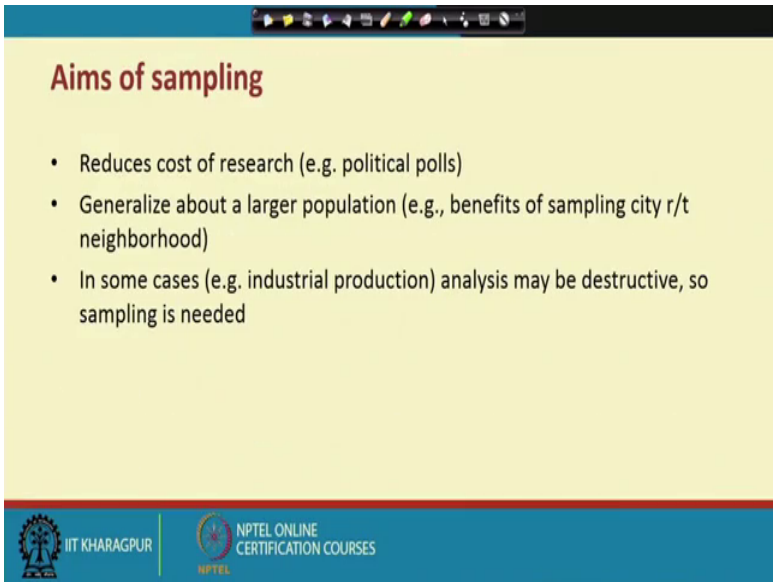
(Refer Slide Time: 08:01)




So, methods of sampling is again you know the compared to previous ones, this is the probabilistic kind of you know situations. And most of the cases, we use probabilistic samples for you know kind of you know empirical testing. So, like the previous lecture we have discussed the concept called as a random variables, so here also there is a structure called as a random samplings. So, each number has an equal probability of being selected. Then there is a systematic structure, stratified sampling, cluster sampling, multistage sampling and means these are all you know various ways to you know gather the information and mean say the idea is that you know it is a kind of art actually. If you follow the particular structure then obviously the accuracy of the data and the accuracy of you know sample would be very perfect.

If you are you know structure will not be specify as per the requirement, then by default the inputs are not correct then obviously, you do not expect the kind of you know output and the kind of findings. So, what is actually the kind of requirement that you know you should actually focus on a kind of samples which is a as per your need that means, you know you follow the best kind of sampling structure, so that your sample or your data will be very effective and very efficient as far your requirement is concerned.

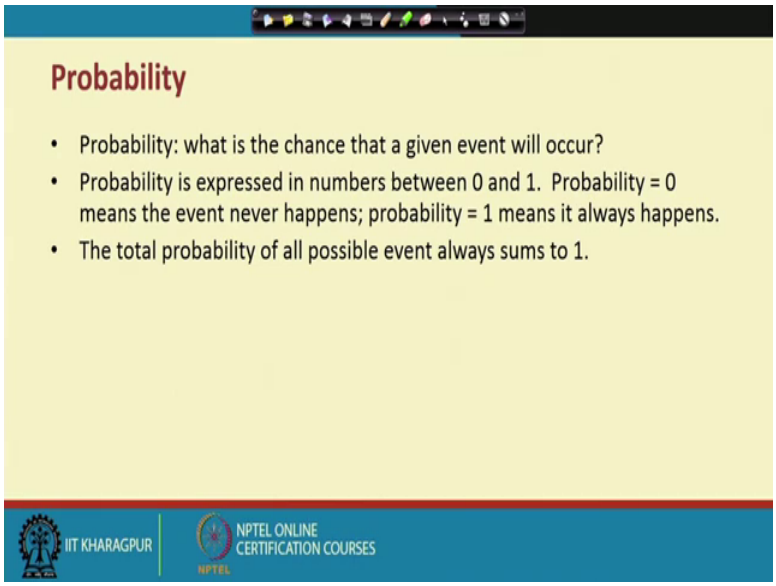
(Refer Slide Time: 09:38)




So, now so m sub sampling actually we sometimes you know with the kind of you know limitation subject to cost time etcetera, so you have to be a you know do some kind of you know compromise. But still the suggestion is that you must have always you know accurate sampling as per the requirement, and try to have more number of sample points, so that the kind of you know findings will be efficient and accurate.

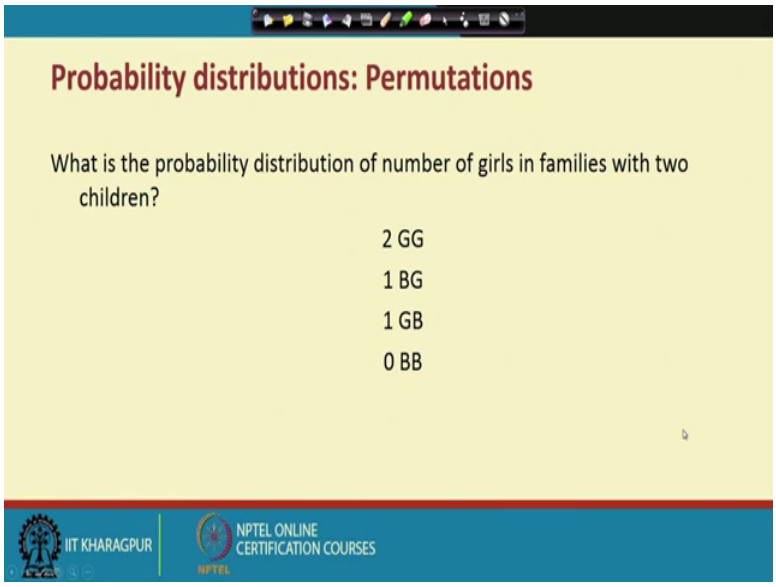
If you are you know sample is not good one and the sample size is not good one then obviously this will affect the kind of you know analysis by using any kind of you know complex analytics tools, and then the findings will be not efficient. So, that as a result you are supposed to follow a particular structure where your sampling or you know sample point should be very efficient one. This will give you better predictions. So, these are all you know probability concept which you have already discussed. So, I am not going in details.

(Refer Slide Time: 10:34)




So, the value of the probability will be 0 to 1, and then these are you know the kind of you know sampling structure particular probability sampling will follow this particular you know principle.

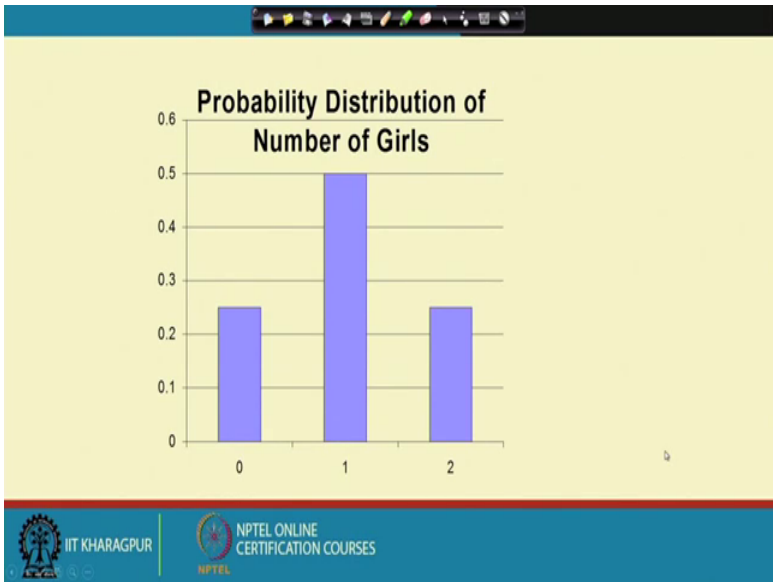
(Refer Slide Time: 10:50)




And now I will give you a little bit a hint here actually what is the probability distribution and number of girls in family with it two children. So, these are all actually possible samples means you should know the trick how to you know gather the kind of you know sample points right. So, we have already discussed in details in the case of you

know probability and probability distribution with so many probability distribution we have highlighted. And this is the series of you know samples you know possible cases you have to detect and then you have to predict the particular requirement as per the need of the business.

(Refer Slide Time: 11:27)



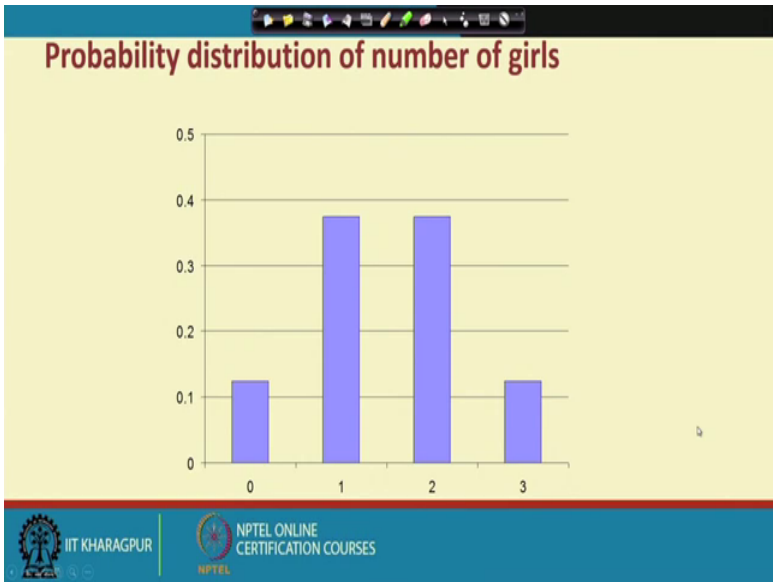

So, this is the kind of with the available data you have to predict the kind of you know structures and try to find out the particular pattern you know probably distribution pattern so that you know your prediction can be effective you know again as per the business requirement.

(Refer Slide Time: 11:44)


So, again you know how about family of three, so when you are you know number of girls. So, these are the kind of you know possible cases. And this can be generated samples and then on the basis of the these are all actually how to apply the trick and gather the information; once you gather the information then you are in a position to address the problems right. So, it may be for instance, number of girls in three family, it can be start with a 0, it can be start with a 1, it can be start with a two and so on.

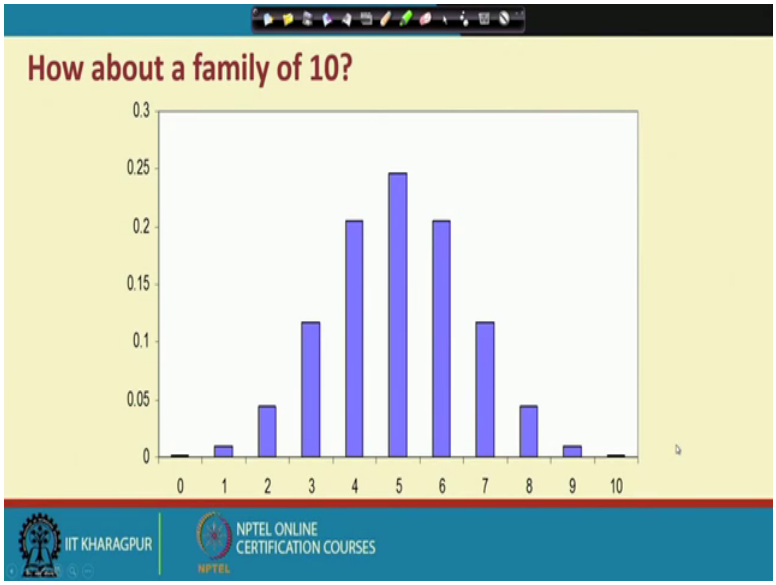
So, if it is 0 then that means, all the family having actually boy, boy, boy. So, if it is then you know it may be a kind of you know in the case of first one, it may be case of second one or it may be case of you know third one only. So, likewise you know it is the kind of you know structure through which you have to generate the sample. So, that means, you have if your problem is very accurate, understanding is very accurate, so there is no problem to generate the samples.

(Refer Slide Time: 12:46)




So, like this, so this again if you plot then it will give you the kind of you know distributions of this particular you know sampling. So, once you know the distribution accordingly you have to apply the particular distribution and predict the kind of you know environment.

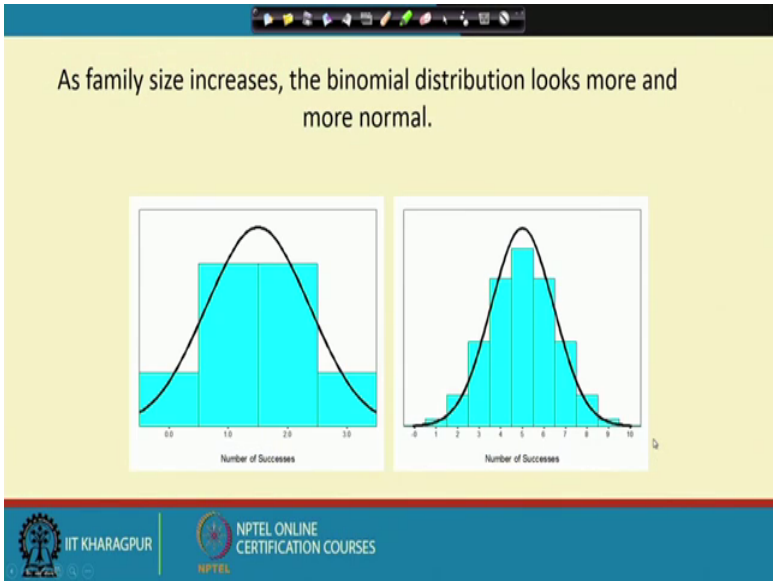
(Refer Slide Time: 12:59)




So, how about a family of 10, so that means I am just you know giving you the cool of you know indication, once you add up you know something more and more, more than

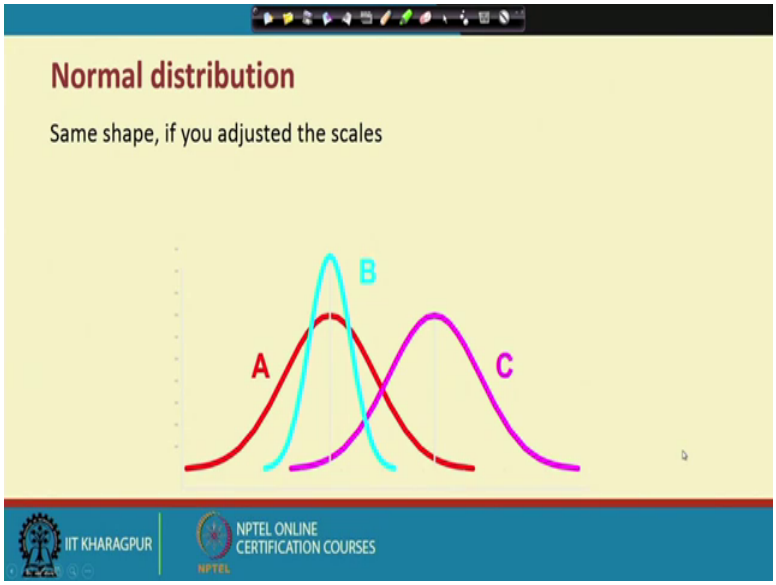
the kind of you know complexity will start increasing. And that is how you have to follow their particular you know pattern which will be as per your requirement.

(Refer Slide Time: 13:17)




Then a family size increases. Then you know I mean simple starting with the probability distribution then you follows the kind of you know typical complex distribution like binomial distributions, then a normal distribution and something like that.

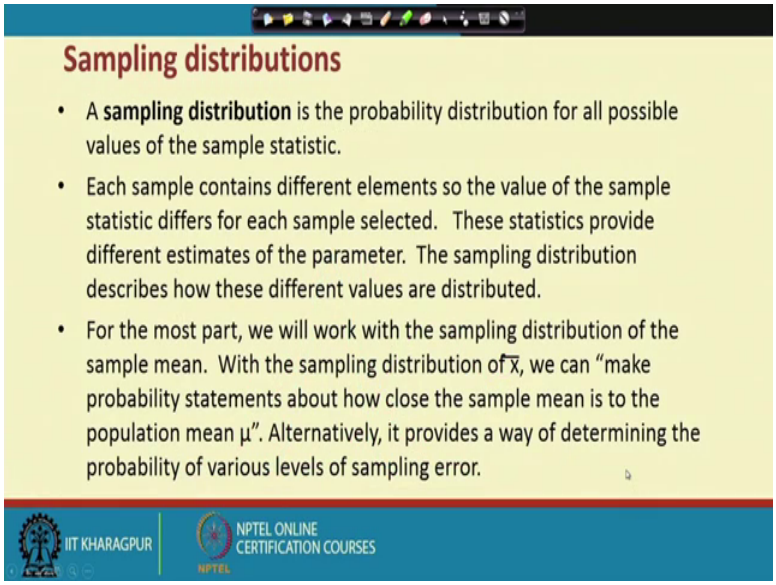
(Refer Slide Time: 13:43)




So, this is the kind of you know case you know which we have already highlighted, but the pattern of distinction only follow like this. And this is a kind of you know normal

distribution pattern and that means, once you have the samples you have to just you know plot and see how is the kind of you know shape. If your shape is not actually A accurate and then here you know analysis will not be accurate. So, you must be very careful how is the particular shape and as per the shape you have to structure the data and then you connect with your particular analytics tools to solve the business problem.

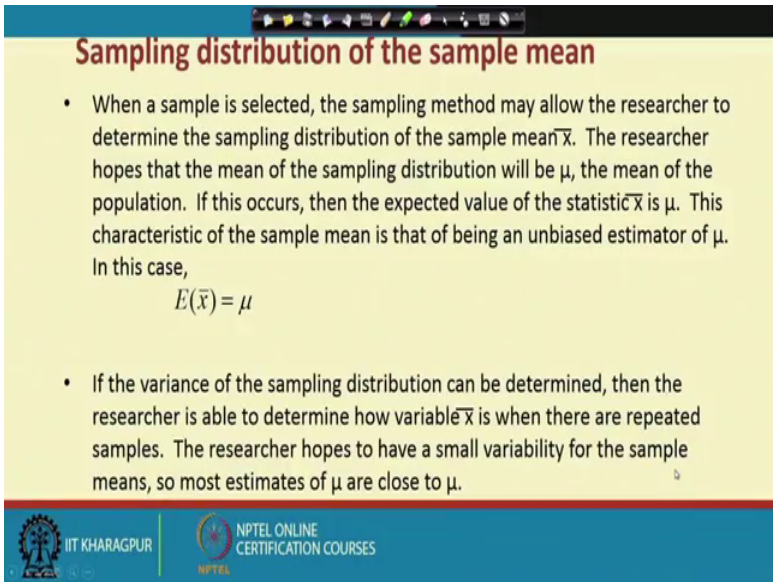
(Refer Slide Time: 14:02)




Then there is a concept called as you know sampling distribution. Actually we have a three different structure altogether sample sampling distribution and population so that means suppose you have a kind of you know basket, and in the basket you can call as a population. Then in the basket you pick up you know small, small samples let us say sample size of you know 50, 30, 150 like this then you try to find out sample statistics. Now, all the sample statistics will be having another kind of you know basket and that will create a kind of you know distribution, but in simple language you can called as you know sampling distribution. So, you have to follow.

So, when will be go for this particular you know process. So, some kind of you know normalization you are you know bringing. So, and every times the normalization or normalize data is always good fit for any kind of you know modeling. So, you try to bring the kind of you know accuracy and the kind of you know normality, so that you know your business it can be analyzed very effectively and that will be as per your you know requirement.

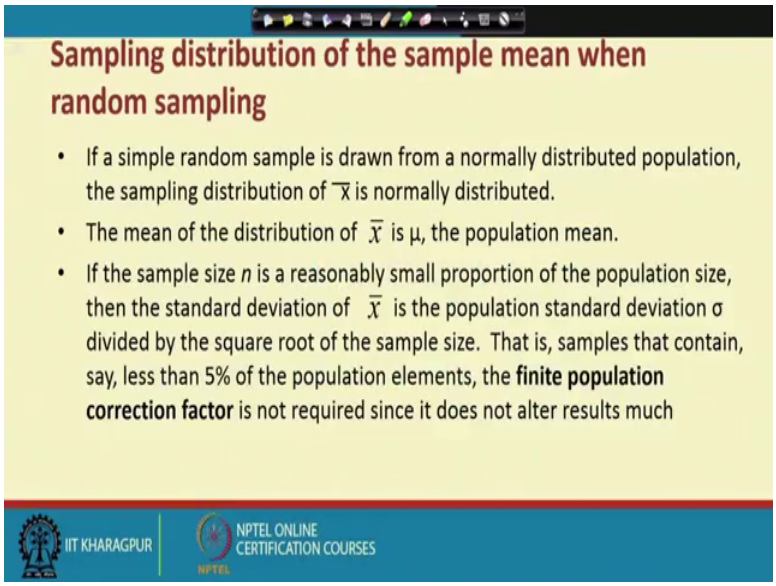
(Refer Slide Time: 15:18)




So, this is sampling distribution of you know so the kind of you know rules you have to follow and when you are you know picking up a particular you know sample. So, you have a sample statistic and you have a population statistic. And when you are calculating and I means when you are a picking up a particular sample, your sample statistics will be converged to population statistic. If the difference will be very high that means, the sample this particular you know a choice of samples may not be actually very accurate. So, your sample statistics you try to pick up the sample which is very close to population parameter. If it is deviating from the population parameter then the biasness will be very high and that biasness will be affect your you know business forecasting or you know business for predictions. As a result your management decision will not be very effective.

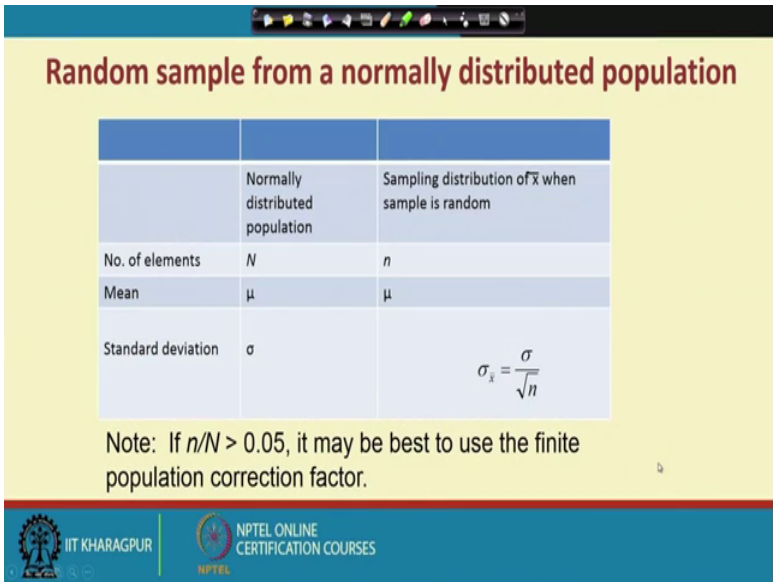
So, you must be very careful how you have to pick up this sample and the sample which you need pick up that will be you know very close to population parameters. Otherwise you know the you know if that is the that is not the case then there is a high chance that you know the populations parameters are actually be specified. So, you try to find out actually the structure through which you can the entire team a and entire structure will be again you know a rearranged as per your requirement.

(Refer Slide Time: 16:37)




So, this is another kind of you know sampling rule. So, I am just you know skipping this one. So, the thing is that you know some cases you know so for as you know population is concerned, it is called as a infinity population and finite population.

(Refer Slide Time: 16:54)

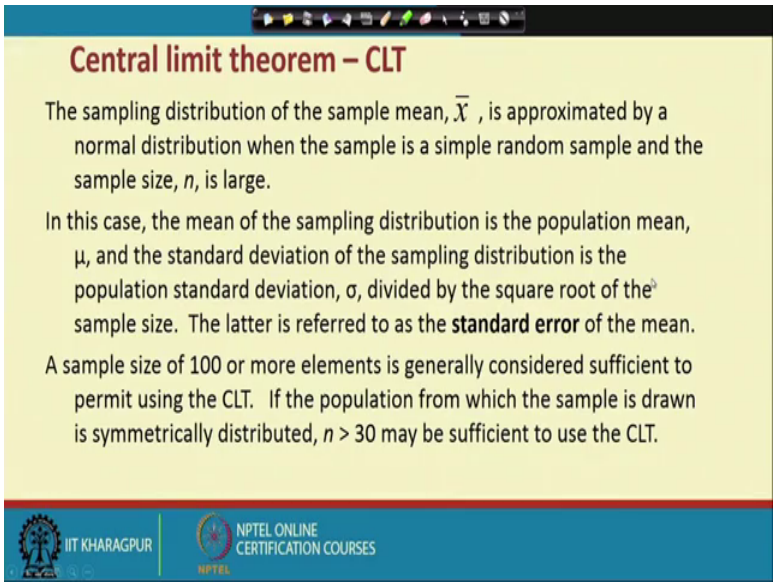



But whatever may be the case so we actually try to follow the kind of you know patterns that you know they pick up the sample which can be very close to the population parameters. So, here this when we are following a kind of you know structure random

sampling from a normally distributed population, so these are the typical structure you have to follow. And capital N stands for actually population parameter.

And this is what actually population parameter then this is a sample kind of you know structure. And you know a distribution will be called as you know efficient distribution here you know sample statistic will be converse to populations statistics. And then as you know sample mean will be equal to population mean, but sample standard deviations will be you know closely connected with you know population standard deviations. So, this is the kind of you know conversions you have to connect with you know when you are connecting sample with you know populations.

(Refer Slide Time: 18:03)



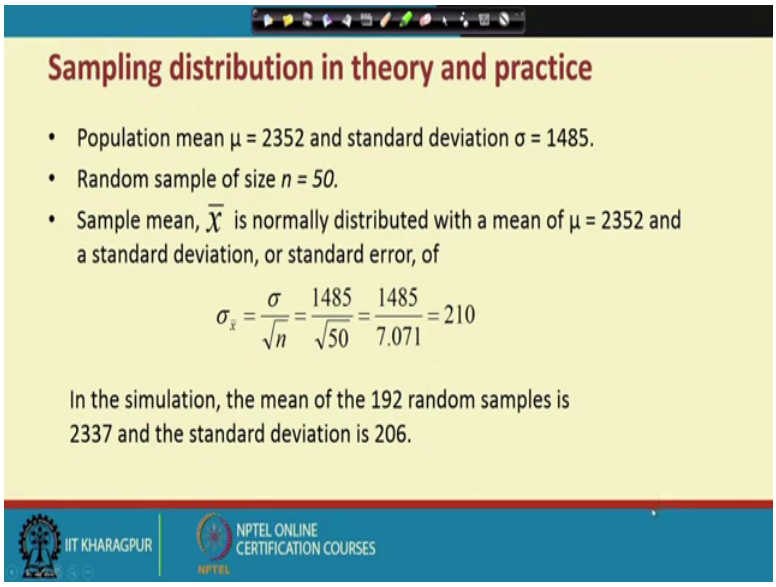

So, now, so I will give you the little bit more details about this particular you know structure. So, this is a kind of you know structure called as you know central limit theorem. And this theorem you know gives the idea that you know when you increase the sample size, you know one after another. Then the particular distribution will be you know close to normally a normal distribution, you know we have actually discussed couple of you know distribution, but best distribution in so for as you know prediction and forecasting is a concern is called as you know normal distributions.

Now, with the help of you know sample you know with the help of you know central limit theorem structures, so when we increase the sample size you know indefinitely then most of the distribution will be a you know follows the normal distribution. And normal

distribution once the data will be follow a particular normal distribution, then you know the outcome and the findings are you know the kind of you know decision making process will be very efficient and very pretty. Sometimes we will be normalize the data we will structure restructure the data and try to bring the kind of you know distribution which will be having actually kind of you know equal spreads. So, it should be like you know symmetrical in nature.

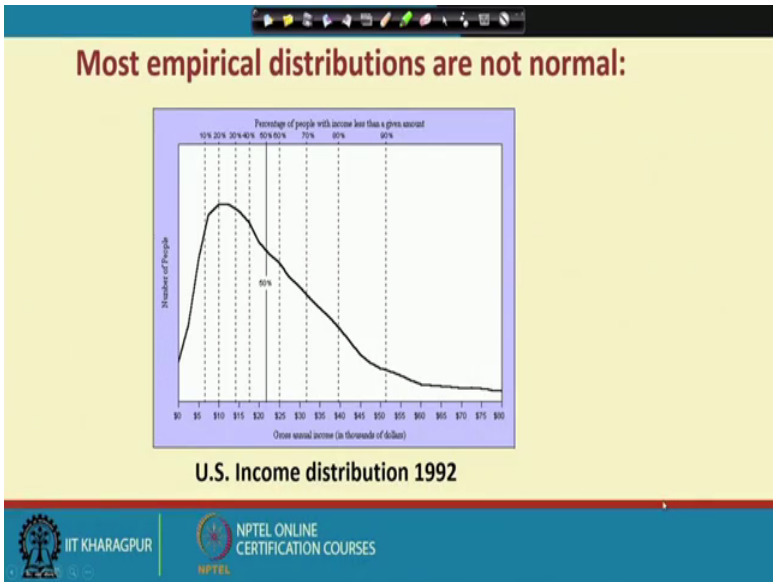
So, that is that is one of this strategy or skill you have to develop how you have to bring you know skew data to actually symmetrical data. So, that you know your business find you know kind of you know check or you know business requirement may be very effective so far as you know any kind of you know analytics tools are you know concerned. Some like this and this is this is the kind of you know structures.

(Refer Slide Time: 19:35)



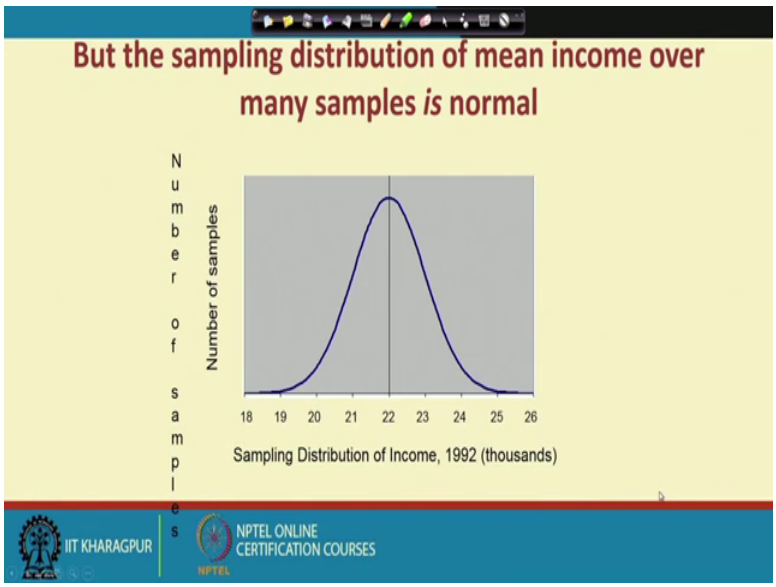

So, these are the kind of you know classic examples how we have to connect the central limit theorem and you know indicating that you know the particular distribution will be following actually normal distributions.

(Refer Slide Time: 19:48)




So, this is the kind of you know structure which actually followed a kind of a skewed distribution that is actually not normal.

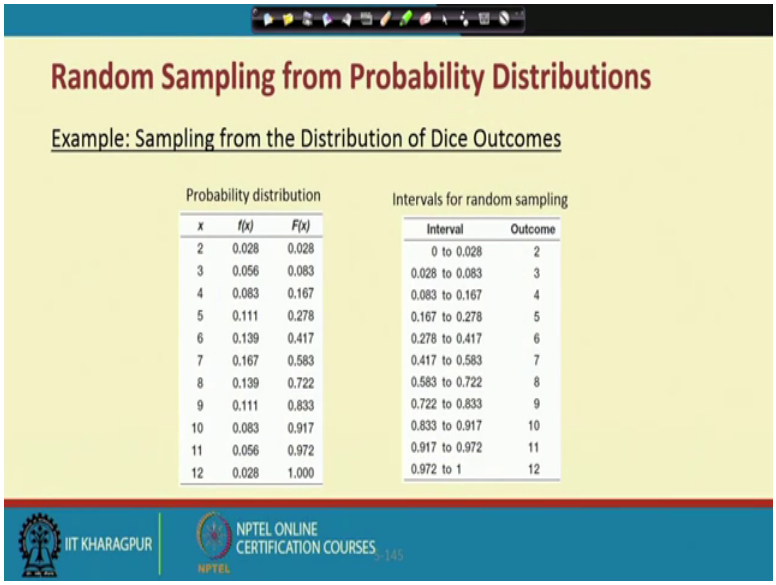
(Refer Slide Time: 19:56)




Then again with you know increasing samples, so now, the particular distribution will be following a normal distribution. So, that is why so what will you do sometimes you know increase the sample size that is as per the central limit theorem, so the distribution will be followed by a normal distribution or sometimes you can you know restructure the data or transport the data, so that the distribution can be little bit close to the normal

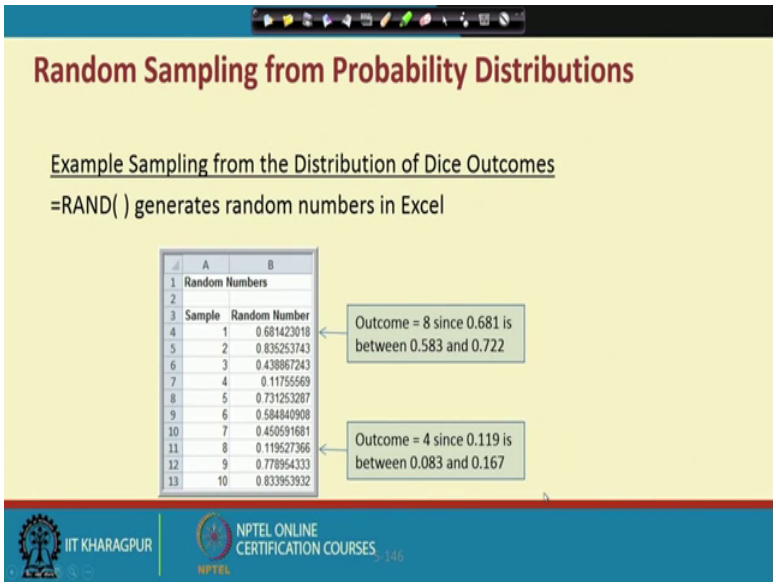
distribution. But the best structure to follow the normal distribution is the apply the you know increasing sample policy, so that you know the distribution will be follow normal distribution. And then once you get the kind of you know symmetrical spread or you know equal distribution, this will help you lot to take management decision as per your requirement.

(Refer Slide Time: 20:44)



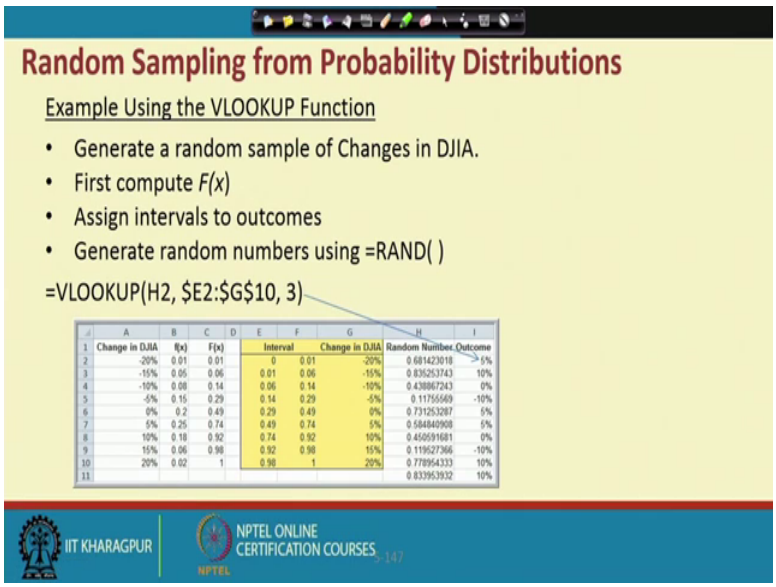

This is the classic example of you know probability distributions and followed by a you know simple stock probability structure. Then this is called as a interval probability structure which we have already discussed in unit two and in the case of you know probability distribution, so that is why I am not discussing in details. But you know you just follow up the kind of you know difference between simple probability distributions with a kind of you know absolute structure, and in the kind of you know interval structures.

(Refer Slide Time: 21:13)




And then in the particular you know structure you have to follow and you know generate the data. So, I will take you to the excel sheet and you may show you how these the particular you know sampling will be generated as per your you know requirement, as per your requirement.

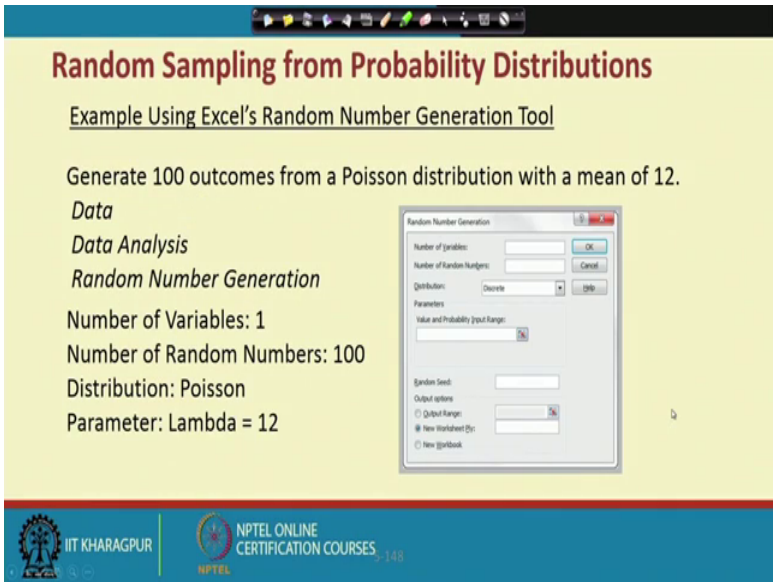
(Refer Slide Time: 21:28)




In the in the second unit we have discussed you know h loop v loop you know to indicate a particular you know structure as per your requirement. So, here in these structures with the help of you know probability structure and for you know probability and with relook

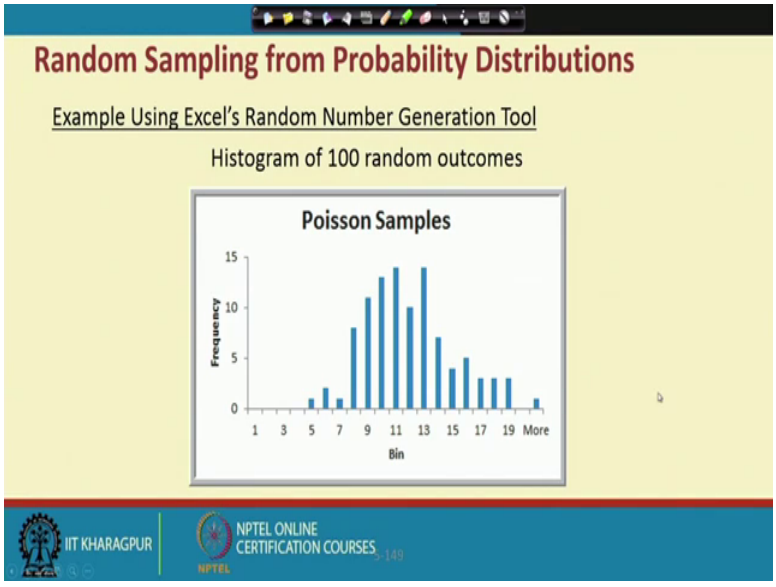
then you can you know specify the particular you know requirement. Again so the idea is that you know excel spreadsheet has a kind of you know advantage to predict the particular business as for the you know feasibility or you know requirement.

(Refer Slide Time: 22:03)

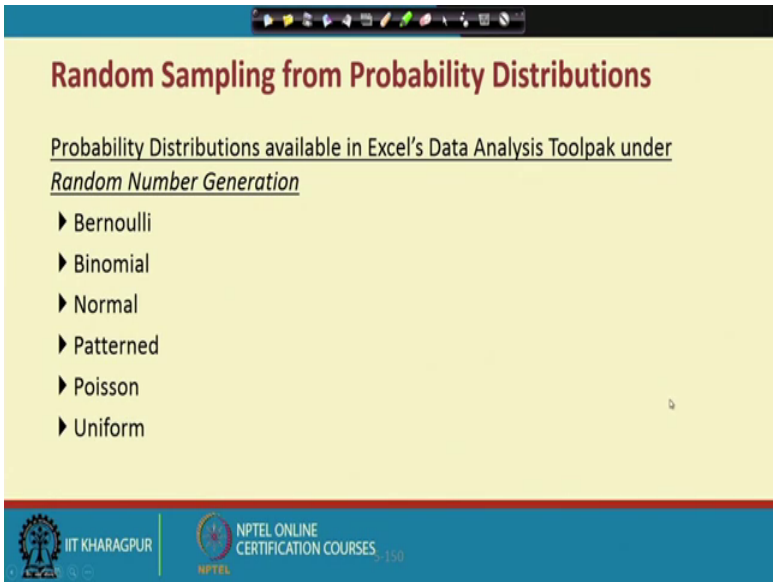



And this is a another kind of you know excel environment through which you know particular distribution can be generated.

(Refer Slide Time: 22:13)

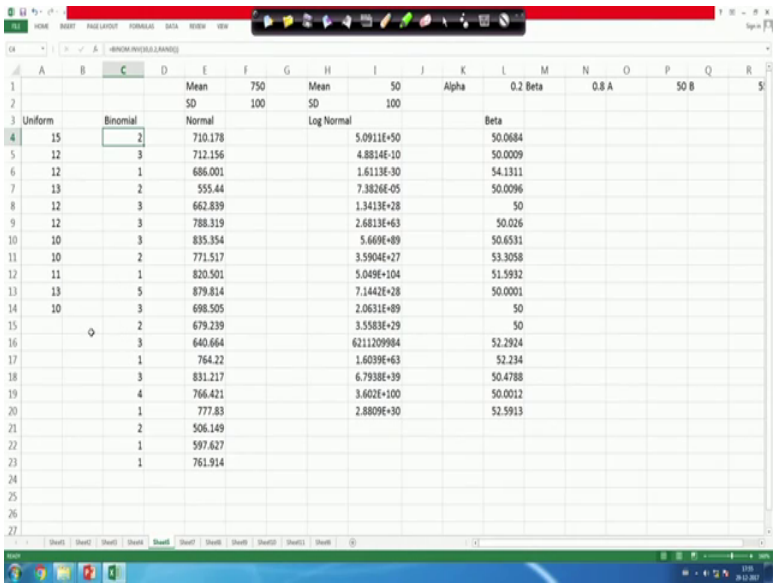


(Refer Slide Time: 22:16)




And let me take you to the kind of you know case where we can discuss you know many more you know a typical issues discussed various probability distributions.

(Refer Slide Time: 22:21)



| | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | | | | | Mean | 750 | | Mean | 50 | | Alpha | 0.2 | Beta | 0.8 | A | 50 | B | 5 |
| 2 | | | | | SD | 100 | | SD | 100 | | | | | | | | | |
| 3 | Uniform | | Binomial | | Normal | | | Log Normal | | | | Beta | | | | | | |
| 4 | 15 | | 2 | | 710.178 | | | | 5.0911E+50 | | | 50.0684 | | | | | | |
| 5 | 12 | | 3 | | 712.156 | | | | 4.8814E-10 | | | 50.0009 | | | | | | |
| 6 | 12 | | 1 | | 686.001 | | | | 1.6113E-30 | | | 54.1311 | | | | | | |
| 7 | 13 | | 2 | | 555.44 | | | | 7.3826E-05 | | | 50.0096 | | | | | | |
| 8 | 12 | | 3 | | 662.839 | | | | 1.3413E+28 | | | 50 | | | | | | |
| 9 | 12 | | 3 | | 788.319 | | | | 2.6813E+63 | | | 50.026 | | | | | | |
| 10 | 10 | | 3 | | 835.354 | | | | 5.669E+89 | | | 50.6531 | | | | | | |
| 11 | 10 | | 2 | | 771.517 | | | | 3.5904E+27 | | | 53.3058 | | | | | | |
| 12 | 11 | | 1 | | 820.501 | | | | 5.049E+104 | | | 51.5932 | | | | | | |
| 13 | 13 | | 5 | | 879.814 | | | | 7.1442E+28 | | | 50.0001 | | | | | | |
| 14 | 10 | | 3 | | 698.505 | | | | 2.0631E+89 | | | 50 | | | | | | |
| 15 | | | 2 | | 679.239 | | | | 3.5583E+29 | | | 50 | | | | | | |
| 16 | | | 3 | | 640.664 | | | | 6211209984 | | | 52.2924 | | | | | | |
| 17 | | | 1 | | 764.22 | | | | 1.6039E+63 | | | 52.234 | | | | | | |
| 18 | | | 3 | | 831.217 | | | | 6.7938E+39 | | | 50.4788 | | | | | | |
| 19 | | | 4 | | 766.421 | | | | 3.602E+100 | | | 50.0012 | | | | | | |
| 20 | | | 1 | | 777.83 | | | | 2.8809E+30 | | | 52.5913 | | | | | | |
| 21 | | | 2 | | 506.149 | | | | | | | | | | | | | |
| 22 | | | 1 | | 597.627 | | | | | | | | | | | | | |
| 23 | | | 1 | | 761.914 | | | | | | | | | | | | | |

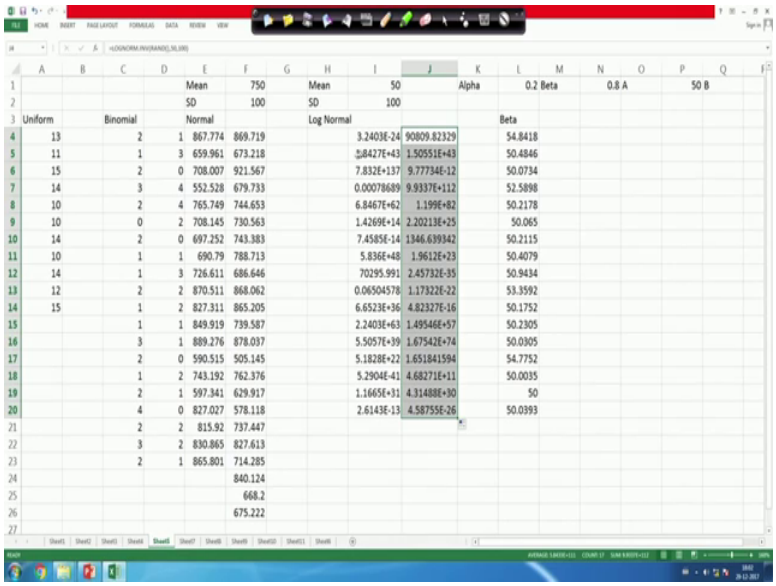
In fact, I like to highlight that you know how these you know sampling can be generated through a particular you know distributions last in you know lectures we followed you know kind of structures with a particular you know values of the random variables. And if the particular distribution only follow then what is the probability of that particular random variables. Here we are just following in a different kind of you know angles with

you know with a particular you know population specifications, then we have to generate this you know kind of you know values of the random variables.

So, now, you see here. So, this is the kind of you know uniform distributions, and I have actually already designed a kind of you know structures to justify that you know corresponding to the previous discussion in the case of probability and probability distribution, here the case of you know some sampling and sampling distribution. Here you have to generate these samples. In the earlier case, we have a sample specific then following a particular you know probability distribution, you are trying to find out what is the probability of a particular happening.

But here with the help of you know kind of you know population parameters, you are generating the values of the random variables that is the difference between you know probability distribution and the kind of you know sampling distribution. Like we have discussed various distribution here and some of the distribution I am highlighting here in the case of uniform distribution, so it is the kind of you know classic example here. So, I will take you to the particular you know examples let us put here, and this is the uniform distribution you put actually uniform distribution here.

(Refer Slide Time: 24:11)



| | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | | | | | Mean | 750 | | Mean | 50 | | Alpha | 0.2 | Beta | 0.8 | A | 50 | B |
| 2 | | | | | SD | 100 | | SD | 100 | | | | | | | | |
| 3 | Uniform | | Binomial | | Normal | | | Log Normal | | | | Beta | | | | | |
| 4 | 13 | | 2 | 1 | 867.774 | 869.719 | | | 3.2403E-24 | 90809.82329 | | 54.8418 | | | | | |
| 5 | 11 | | 1 | 3 | 659.961 | 673.218 | | | 8427E+43 | 1.50551E+43 | | 50.4846 | | | | | |
| 6 | 15 | | 2 | 0 | 708.007 | 921.567 | | | 7.832E+137 | 9.77734E-12 | | 50.0734 | | | | | |
| 7 | 14 | | 3 | 4 | 552.528 | 679.733 | | | 0.00078689 | 9.9337E+112 | | 52.5898 | | | | | |
| 8 | 10 | | 2 | 4 | 765.749 | 744.653 | | | 6.8467E+62 | 1.199E+82 | | 50.2178 | | | | | |
| 9 | 10 | | 0 | 2 | 708.145 | 730.563 | | | 1.4269E+14 | 2.20213E+25 | | 50.065 | | | | | |
| 10 | 14 | | 2 | 0 | 697.252 | 743.383 | | | 7.4585E-14 | 1346.639342 | | 50.2115 | | | | | |
| 11 | 10 | | 1 | 1 | 690.79 | 788.713 | | | 5.836E+48 | 1.9612E+23 | | 50.4079 | | | | | |
| 12 | 14 | | 1 | 3 | 726.611 | 686.646 | | | 70295.991 | 2.45732E-35 | | 50.9434 | | | | | |
| 13 | 12 | | 2 | 2 | 870.511 | 868.062 | | | 0.06504578 | 1.17322E-22 | | 53.3592 | | | | | |
| 14 | 15 | | 1 | 2 | 827.311 | 865.205 | | | 6.6523E+36 | 4.82327E-16 | | 50.1752 | | | | | |
| 15 | | | 1 | 1 | 849.919 | 739.587 | | | 2.2403E+63 | 1.49546E+57 | | 50.2305 | | | | | |
| 16 | | | 3 | 1 | 889.276 | 878.037 | | | 5.5057E+39 | 1.67542E+74 | | 50.0305 | | | | | |
| 17 | | | 2 | 0 | 590.515 | 505.145 | | | 5.1828E+22 | 1.651841594 | | 54.7752 | | | | | |
| 18 | | | 1 | 2 | 743.192 | 762.376 | | | 5.2904E-41 | 4.68271E+11 | | 50.0035 | | | | | |
| 19 | | | 2 | 1 | 597.341 | 629.917 | | | 1.1665E+31 | 4.31488E+30 | | 50 | | | | | |
| 20 | | | 4 | 0 | 827.027 | 578.118 | | | 2.6143E-13 | 4.58755E-26 | | 50.0393 | | | | | |
| 21 | | | 2 | 2 | 815.92 | 737.447 | | | | | | | | | | | |
| 22 | | | 3 | 2 | 830.865 | 827.613 | | | | | | | | | | | |
| 23 | | | 2 | 1 | 865.801 | 714.285 | | | | | | | | | | | |
| 24 | | | | | | 840.124 | | | | | | | | | | | |
| 25 | | | | | | 668.2 | | | | | | | | | | | |
| 26 | | | | | | 675.222 | | | | | | | | | | | |
| 27 | | | | | | | | | | | | | | | | | |

So, I will directly show you here in the case of you know you need a uniform distribution, but just you know put you know random between you know two different intervals then by default it will be giving you a kind of you know distribution pattern and

that will give you the uniform distributions. In fact, this is not actually very you know this is not a typical actually distribution which can be used for you know for a weak kind of you know problems or you know complex problem. But the other distribution will be very useful for you know kind of you know prediction and you know forecasting of you know business problem let us say binomial distribution.

So, we have already discussed the binomial distribution, now I am connecting here the kind of you know binomial distribution. So, let us say start with you know is same put you know equal to sign here then you put you know binomial command in the case of you know earlier we have a put you know binomial distribution. In this case, in the sampling distribution you put you know binomial inverse. Then if you put you know binomial inverse then it will ask you just click there and it will ask you the trials the trials. So, it will be let us say ten then it will be connect with the probability that is 0.2 and then it is a kind of you know alpha.

So, what did will be do here. So, let me give you another look. So, you know you have to put actually random value, so that you know otherwise it cannot be calculated. So, put equal to sign here. Then put binomial distribution this link. So, then it will be ask you to trials that will be put actually said here 10 like previous examples then probability value, so 0.2 and then and the kind of you know random click here and then close the loop and then you will be find here. So, this will this will give you actually kind of you know structure about the binomial distribution. So, you just you know drag it, then it will be generated like this.

So, this is actually the kind of you know sampling a kind of you know distribution which follows a kind of you know binomial distribution so that means, we have already discussed the kind of you know structure in the case of binomial distribution here. So, we like to generate the values of the random variable. Similarly, you come to the normal distribution case. So, in the normal distribution case, so these are actually values of the random variables.

So, now how you have to randomly generated actually, so now, in the same structure you have to follow here. So, you go to this particular you know equal to signs then put actually in normal distributions and again so this is the normal distribution follow a structure, but we need actually sampling distribution for normal distribution pattern.

Then this is actually probability distributions so probability distribution you put here a random then so close the loop then you put the kind of you know mean value and then you put the standard deviations as per the earlier you know kind of you know fixation.

So, then you just put enters, then this will be generated, and then you just drag it as per the kind of you know if you want you know you can you know increase any extent. So, that you know the sample will be by default will be generated. This is the sampling distribution so for as a normal distribution is concerned and similarly we have option for you know log normal distributions in the case of log normal distributions again the structure will be follow like this. So, I have already written here. So, similarly you just put you know randomized value, then give this specification about the mean and samples automatically it will give you some kind of you know value.

For instance, let us say you put here then you put the kind of you know log normal distributions then you just click here then put you know random structures. So, then close the loop then indicate the kind of you know mean let us say put 50 then you put actually 100 then again close the loop and enters it will be randomly generated. So, again scroll it and you know just like you know dragging, so you will get you know sampling distribution, so that means, you know depending upon the kind of you know structures. So, you have to follows you know kind of you know pattern that is you know sampling distribution pattern.

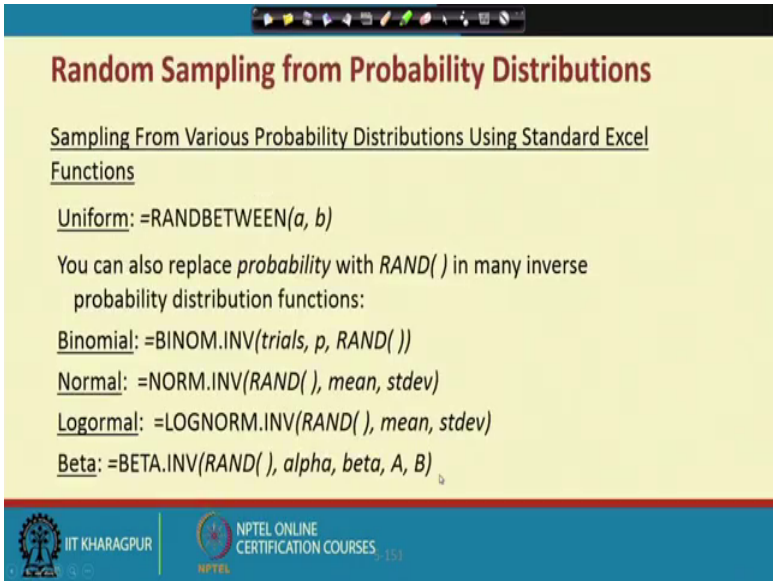
So, the whole idea actually is like this that you know what will do. So, corresponding to a particular distribution you just you know the parameter structure or the kind of you know mass function structure on the basis of that actually it excel spreadsheet will help you lot to you know create a sampling distributions. So, we have a structure you know called as you know probability distributions and corresponding to the probability distribution, you have a structure called as you know sampling distribution. So, this is actually well connected.

So, in the probability distribution case we are actually calculating values of the random variable for a particular case. But here we are actually generating the kind of you know values of the random variable subject to probability mass function is concerned, whether it is a uniform distribution or binomial distributions or normal distribution or log normal distribution whatever may be the kind of you know structure.

So, now take you to this particular structure. So, what I what I like to say that you know, so probability distributions and you know binomial distribution are very useful components so far as you know business analytics is concerned. So, until unless you clear about the probability concept probability distribution, sampling concept and sampling distributions, so you may not be in a position to do very good you know what we can say that you know effective experiment or effective you know empirical testing to get some kind of you know insight as per your you know management requirement. So, management decision altogether cannot be very effective or you cannot come to a perfect management decision until unless you know clear about the probability, probability distribution, sampling and sampling distribution.

whereas, probability - probability distribution, and sampling - sampling distributions are you know kind of you know requirements and you can say that you know it is a free conditions or you know without understanding or without to you know having sufficient knowledge you may not be in a position to analyze effectively any kind of you know business problem.

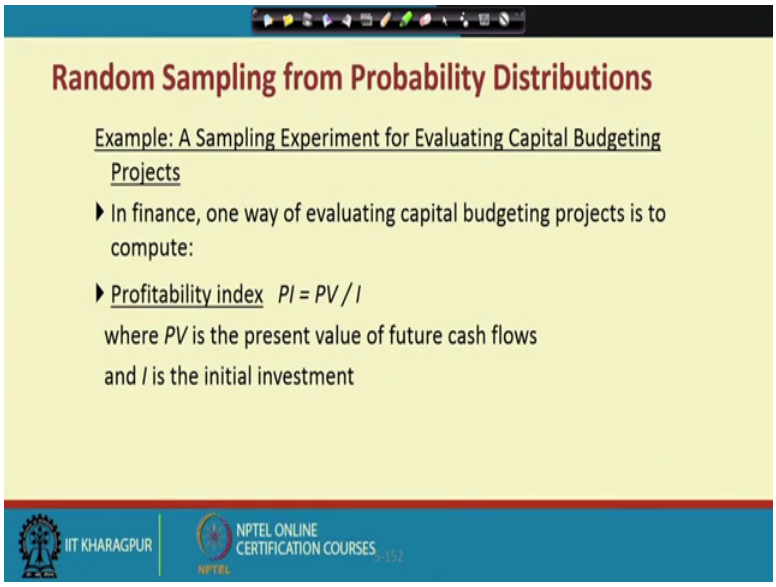
(Refer Slide Time: 31:39)




So, likewise we have already discussed several kind of you know structures. So, this is actually snapshot about the how you have to operate in the case of you know excel spreadsheet to generate the kind of you know sampling distribution corresponding to particular you know probability distributions. So, these are the commands which you can

apply, go to the excel spreadsheet and following a particular if you understand that this particular problem follows actually binomial distribution then you know randomly generate kind of you know series and that can be used for you know testing something like that. Similarly, if you know that this is your normally kind of you know structure then you can create a sampling distribution which can follow normal distribution patterns like this you know you have to follow a kind of you know structure.

(Refer Slide Time: 32:22)



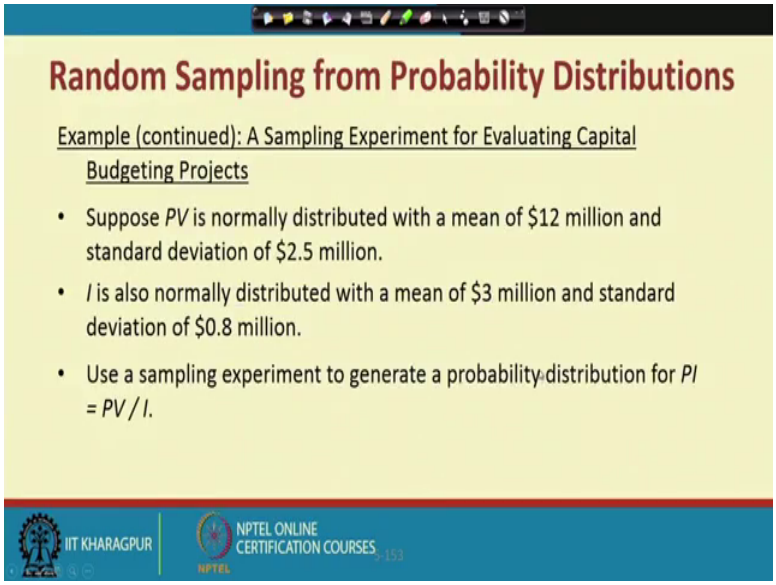

And the idea is that you know you have to understand the kind of you know concepts and means it is a kind of you know random sampling process all together and these are all very useful for any kind of you know forecasting. I will cite here couple of you know practical examples, where you know this probability-probability distribution, sampling- and sampling distributions are very useful for predicting certain events.

For instance we have discussed you know earlier the kind of you know capital budgeting that is you know project evaluations that is usually in the management. And that to in finance area one way of evaluating capital you know budgeting is a called as you know profitability index. Profitability index is nothing but you know present precipitant present value of you know future cash flows you know divided by initial cash flows that is you know initial investment.

And while doing all these things, you will be connected with you know probability distributions until unless you know probability then you are not in a position to calculate

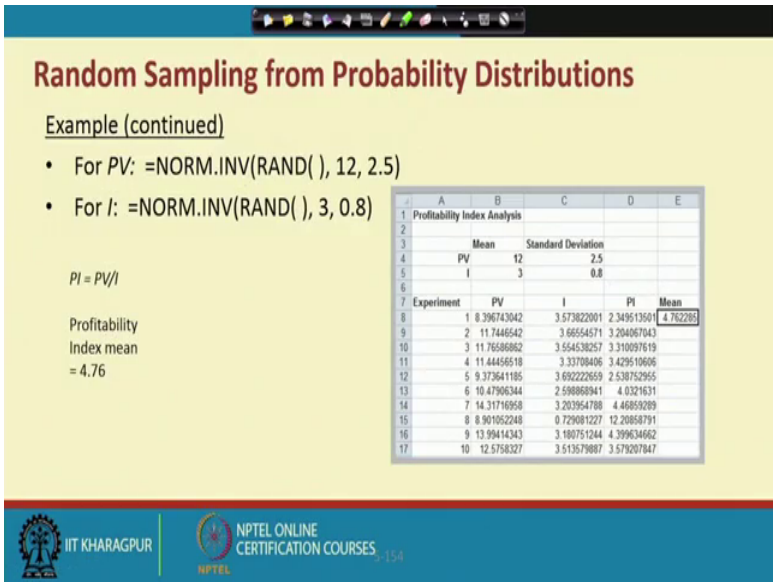
the kind of you know ah feasibility. For instance you know we have discussed a concept called as an net present value criteria which is nothing but actually the MNC, we are interested to know what is the present value of you know future of cash flows over the eight times. With respect to time so this is actually a connected with the cash flows with you know discounting rate; and the discounting rate; that means if you keep cash flows constant followed by the discounting rates that is nothing, but actually cash plus multiplied by 1 by 1 plus discounting rate then it will be continue over the time. So, 1 by 1 plus r follows a kind of you know structure which is nothing but called as a probability structures. So, this is actually called as a probability distribution.

(Refer Slide Time: 34:29)



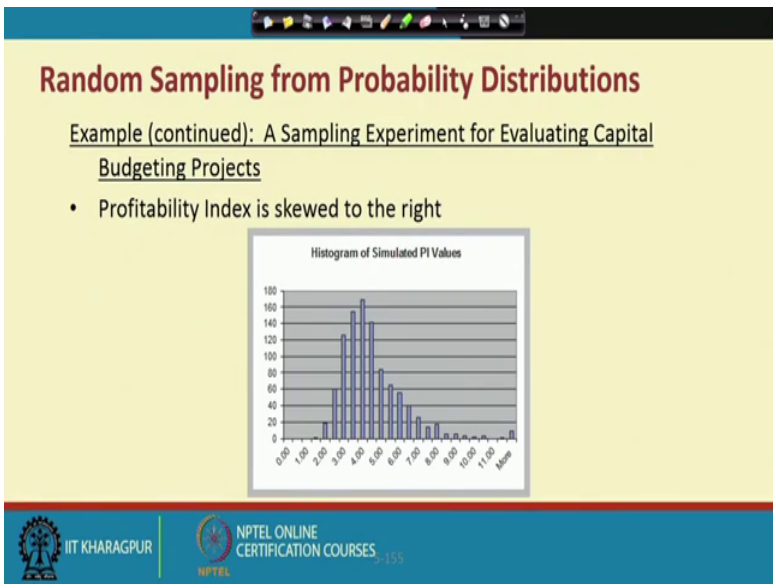

So, that means so what I like to say that you know probability distribution has a well connection with the all these kind of you know forecasting or you know prediction. So, that is how you know it is highly required for any kind of you know business analytics or you know business requirement.

(Refer Slide Time: 34:34)



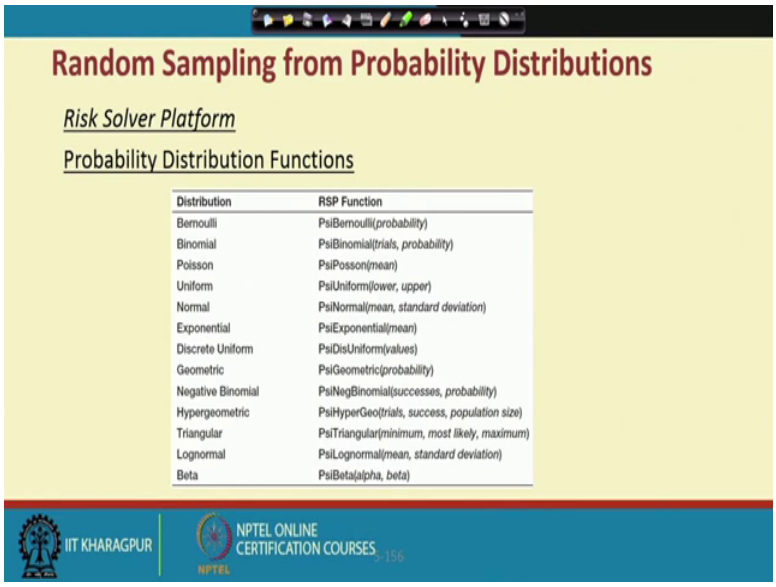

Similarly, so this is how the kind of examples which I have already cited. So, once you mean say same things you know it is you know case of you know probability-probability distribution, and sampling-sampling distribution if you know the kind of you know structure you can generate the samples. And if your random variables are there and you specify a particular distribution you can calculate a probability with your particular you know or with your particular specific case subject to you know specification of a random values of the random variable. So, this is how the kind of you know examples which you can you know follow and you know generate the series for you know capital budgeting.

(Refer Slide Time: 35:11)



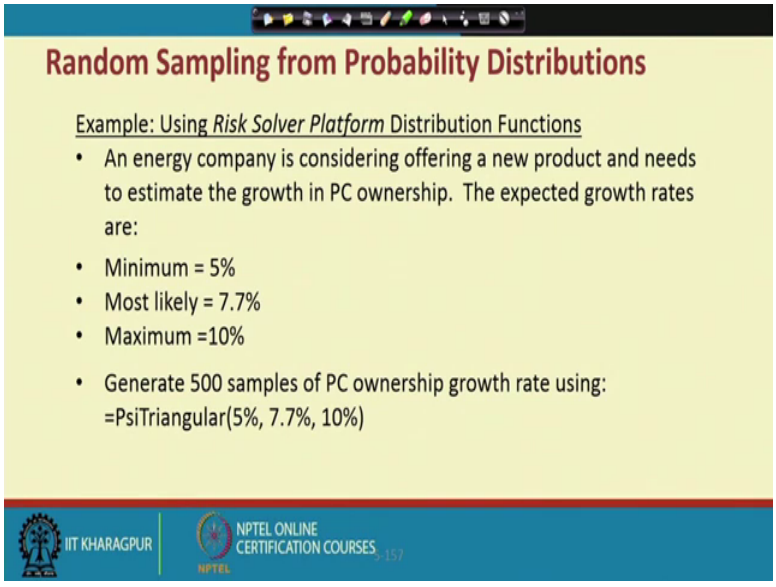
And then you can you know plot and you will find this will be follow a particular you know distributions right. And with this you know there are there is a actually this solver platforms where you know all kinds of you know distributions are there.

(Refer Slide Time: 35:17)



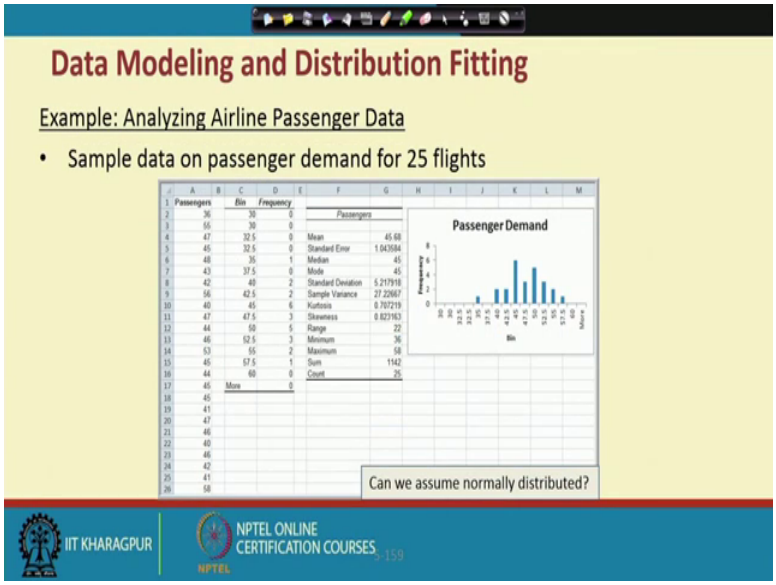

So, you will find here plenty of distributions are there and that will give you some kind of you know structure through which you know the business can be connected or you know business can be evaluated as per the kind of you know need and you know requirement.

(Refer Slide Time: 35:48)



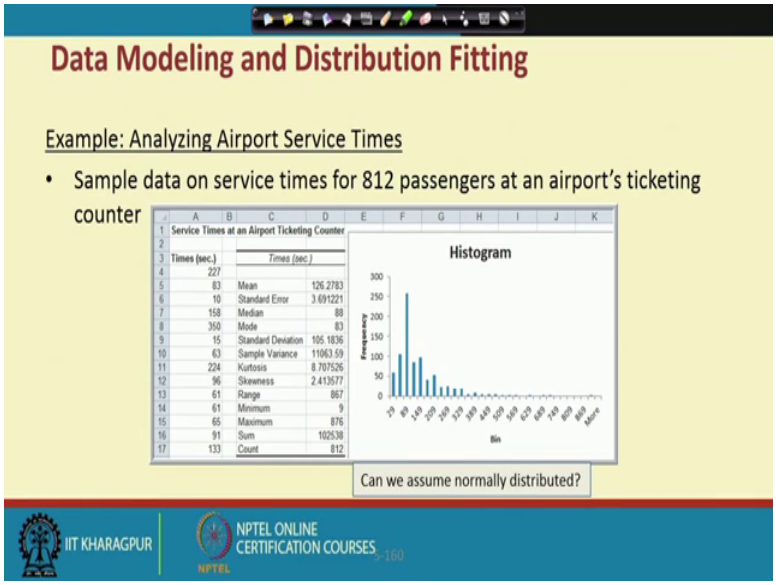
So, again the kind of you know summation is so you must be acquainted with all this you know concept, so that you know you can connect properly you know as per the need and then you can take some kind of you know management decision. So, this is another kind of you know examples and which you can you know solve through this particular you know modality.

(Refer Slide Time: 36:07)



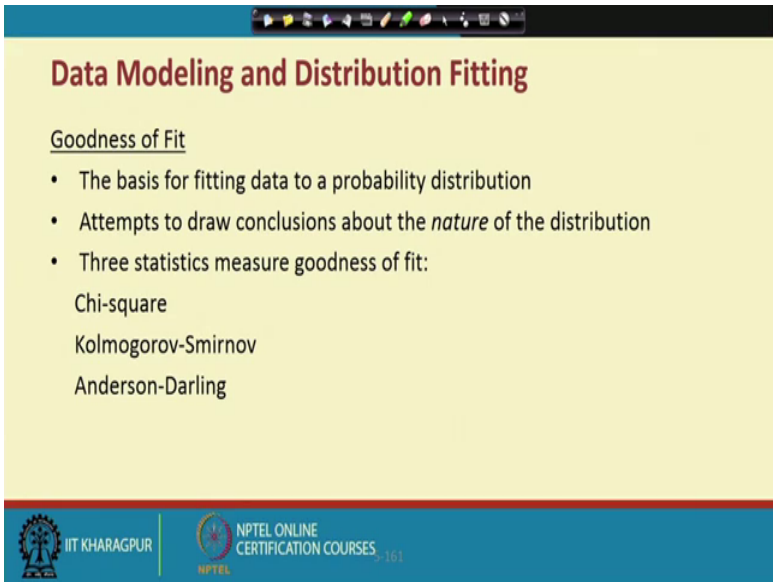

So, again so these are the kind of you know structure this is actually kind of you know airline passenger data.

(Refer Slide Time: 36:22)



And you can create a structures then follow you know kind of you know histogram then you have to check whether it is a normally distributed or not something similar kind of you know examples.

(Refer Slide Time: 36:25)

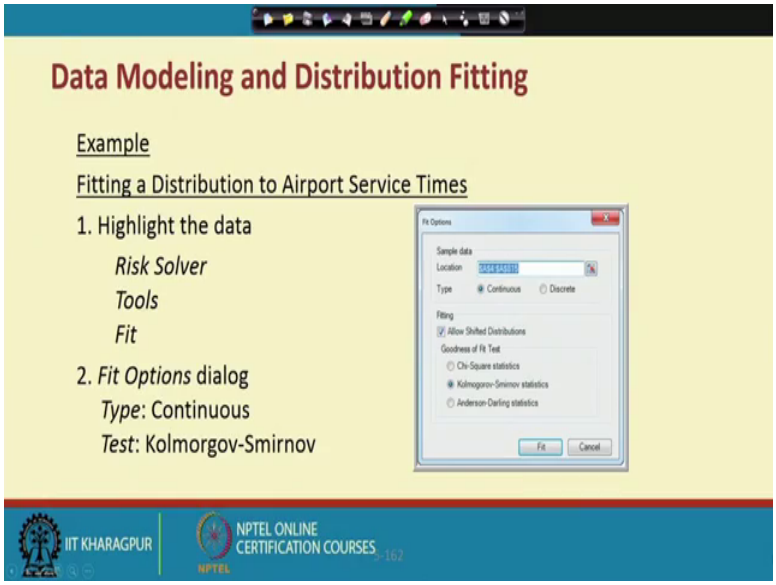



So, the data modeling and you know distribution fitting. So, this is actually why you need all these things you know we need to find out your particular you know efficient structures what we called as you know goodness of heat. So, all these things will help you to bring a kind of you know environment and that what we can called as you know efficient environment, where you know your prediction or you know kind of you know observation will be very effective as per your business needs. And until unless you will be acquainted with all these concept and you know the kind of you know structure and spreadsheets, so you will not in a position to analyze properly.

So, you will be actually give you effective kind of you know solution or you know best kind of you know solutions, once you know or you will be acquainted with you all these kind of you know requirements and you know spreadsheet use. So, because this will give you since it is a kind of you know continuous affairs, so you must have a kind of you know knowledge how you have to operate all these you know items depending upon the understanding of a particular you know structure and the kind of you know problem requirement. So, once you will be acquainted by default, it will give you a you know

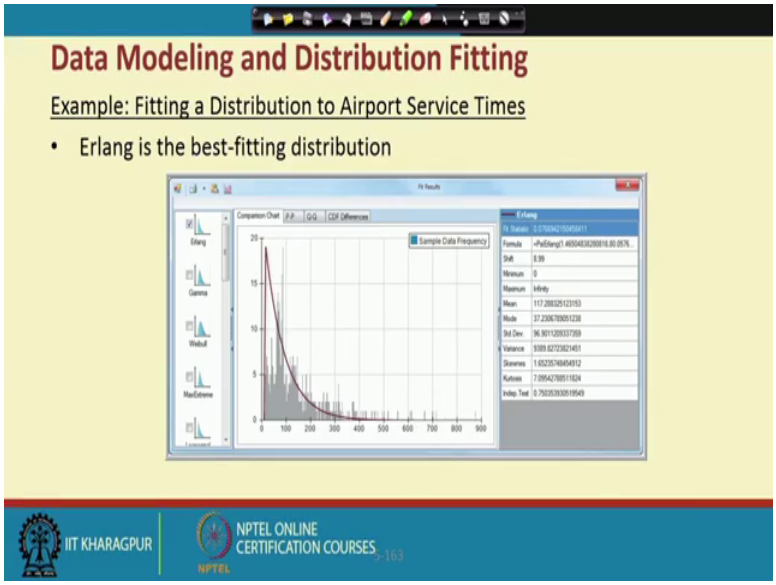
very useful findings or you know kind of you know insides that will help you lot to solve some of the business problems as for your need.

(Refer Slide Time: 37:48)

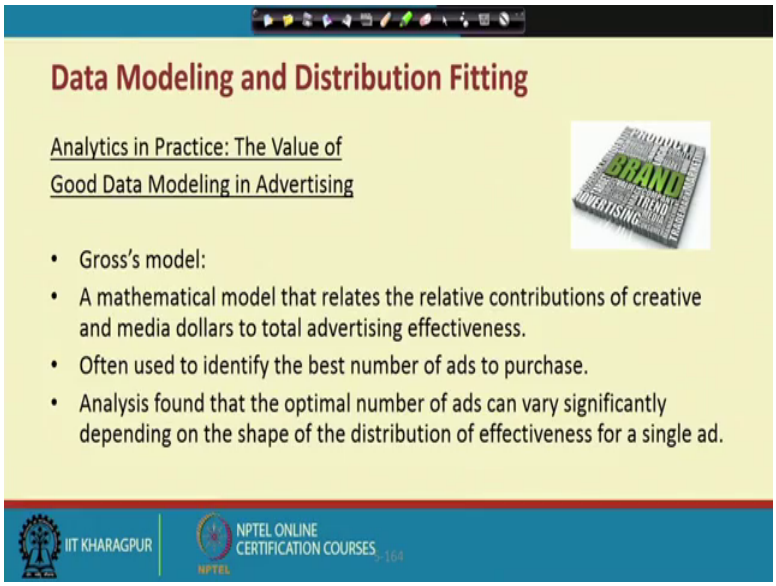



So, like this you know we have a plenty of such examples and you can go through this kind of know structure.

(Refer Slide Time: 37:50)

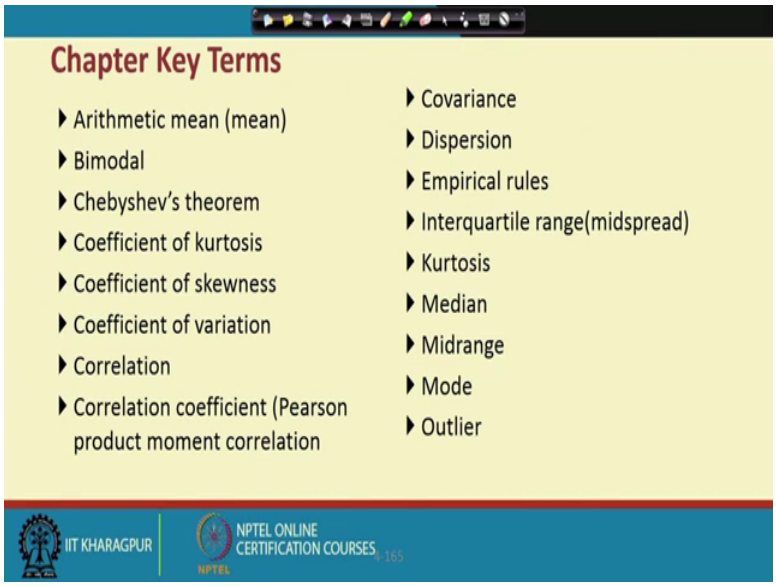


(Refer Slide Time: 37:53)




And these are all you know this particular you know unit key keywords.

(Refer Slide Time: 37:56)




And we have discussed all these kind of you know concept and by the way. So, you know till today we have discussed the unit one, unit two and unit three, and these are all actually basics actually and all these basics are very useful for future course of you know discussions particularly you know inferential analytics, predictive analytics and prescriptive analytics. So, my suggestion is that you know you must be very you must be very acquainted with all these concept and the kind of you know tools and spreadsheet,

so that you know the future discussion will be very handy and you can pick up very quickly. So, with this we will be stop here.

Thank you very much, have a nice time.