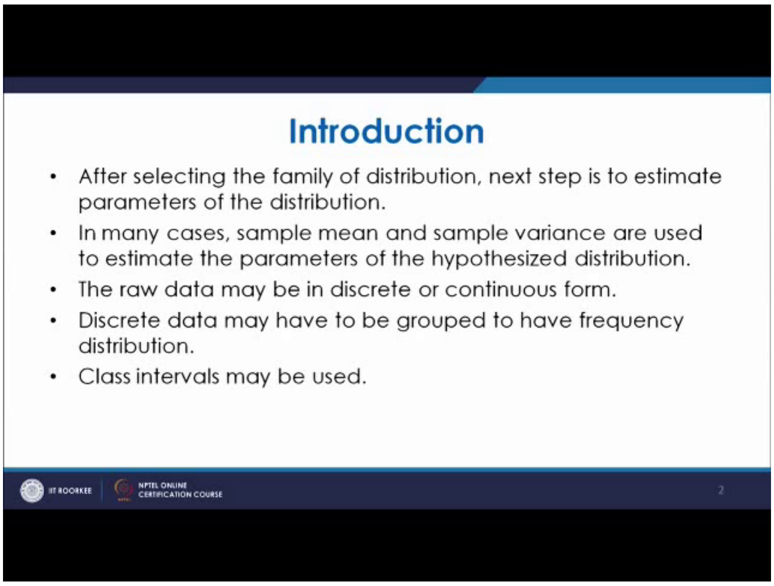
Modeling & Simulation of Discrete Event Systems
Dr. Pradeep K Jha
Department of Mechanical and Industrial Engineering
Indian Institute of Technology, Roorkee

Lecture - 22
Input Modelling: Estimation of Parameters

Welcome to the lecture on Estimation of Parameters. So, in this lecture, in the input modelling section we will talk about the Estimation of Parameters. So, we have seen that after selecting the family of distribution.

(Refer Slide Time: 00:38)



Now, we have to estimate the parameters, every distribution has certain parameter and that is to be estimated, and once you estimate that, based on that the nature of the distribution has to be found. So, you have the sample mean and sample variances, which are calculated, and based on that the estimation of the parameter is carried out. So, you have normally the data given to you; like reading or the entry arrival times reading of the time to failure. All these are the different data’s which you will be given.

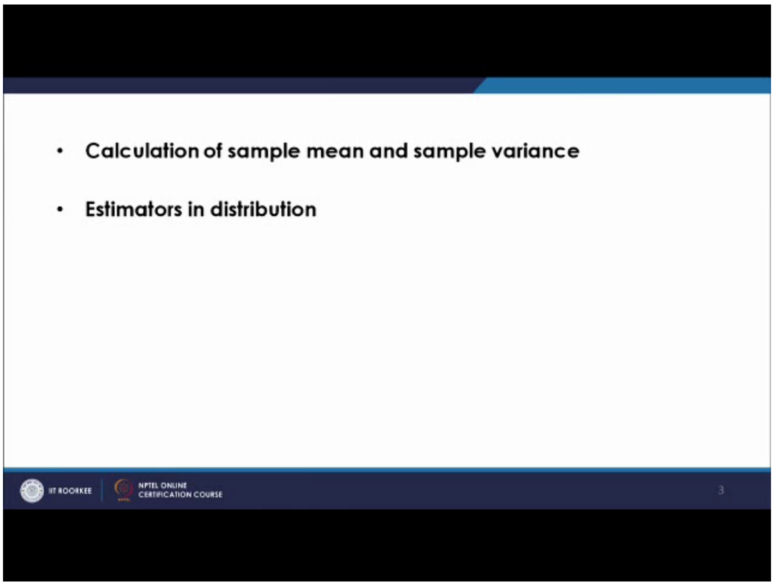
Now, based on that you have to find the mean and the variance; so normally the parameters which are used, are basically the sample mean and sample variance, and once you know that then every typical distribution has the different parameter estimation, also different parameters, like you have in gamma function, gamma distribution you have beta, and theta you have the alpha uniarrival rate in Poisson and all that. So, this way we will see in this lecture that how those need to be estimated.

Now, the raw data which we have in our hand, it may be either in discrete form or in continuous form. So, that is to be used, now many a times when you have the discrete data, you are directly taking these data and getting the sample mean and sample variance. Many a times you have to group them to have a frequency distribution. So, you have the discrete data, but you group it in.

I mean few of the data in 1 group, and then you find the frequency distribution. And then you find the sample mean and sample variance many a times you also use the class intervals. So, you have the continuous distribution in that many a times, you also use the class interval, you have the class and in that class that interval length is there, then it will have certain frequencies and the midpoint of the interval is taken at the point which is used for calculating, you know for the variance or so or even for the mean.

So, different type of data will be available to you, and in that case you have to use those data to find the parameters.

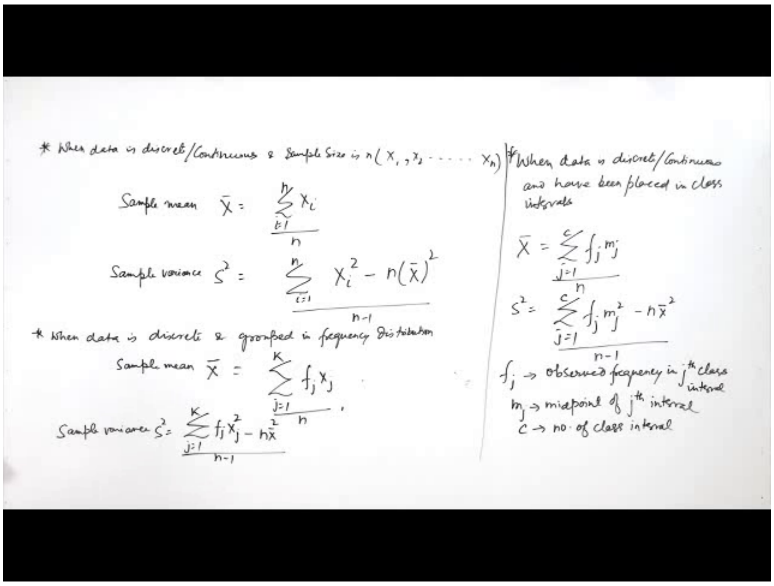
(Refer Slide Time: 03:15)



So, the thing is that, what is the normal process for getting the sample mean and sample variance. Sample mean or sample variance means you have the data, sample data, and form there you have to calculate the sample mean as well as sample variance, you have 3 types of data. You have sometimes the discrete or continuous raw data. So, for that there is 1 formula. Similarly you have you know data in the time of intervals and that also can be used, you have the grouping of data and it has the different frequencies. So, that time you will use different formula for getting this sample mean and sample variances.

So, if the data is discrete. So, if when the data is discrete or continuous and sample size is n; that is you have x 1 x 2 x n.

(Refer Slide Time: 04:03)



So, in those cases the sample mean, sample mean will be calculated as the summation of these values divided by the total number of observations. So, it will be x bar is summation of x i, i will be varying from 1 to n; that is sample size and then divided by n. So, this is how when you have the discrete data, raw data is there, you have discrete or continuous raw data is available. In those cases you find the sample mean like this. Similarly you have sample variance. So, that is S square or S square n we can also write and that will be again summation i equal to 1 to n, and then it will be x i square minus n x bar square divided by n minus 1. So, this is how you find the sample mean, and sample variance when the data is discrete or continuous in raw form.

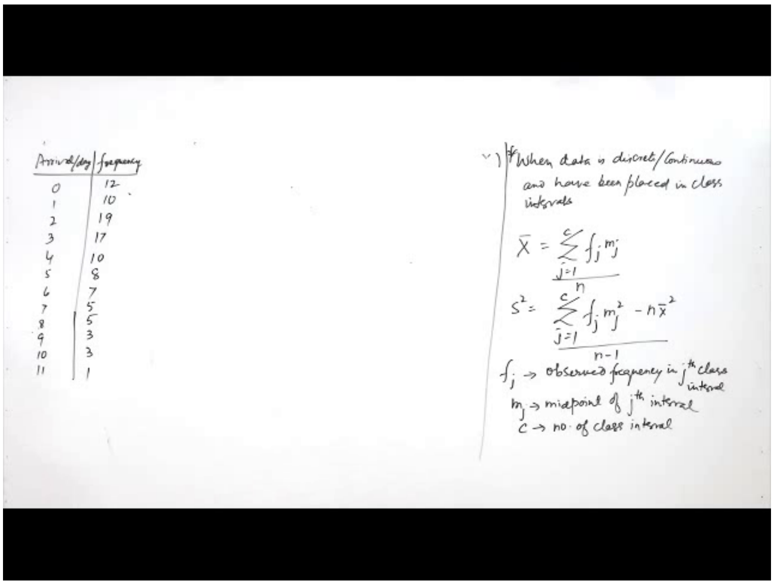
Now, the next case is, when the data is discrete and grouped in frequency distribution. So, you, the data is grouped and it has certain frequency distribution, like you have 3 or 4 data together, and it has certain frequencies. So, these are the normal formula in that sample mean x bar. It will be again j equal to 1 to k and f j x j upon n, and the sample variance. So, you have k number of classes, and you have sample variance coming as. So, we can write sample variance here. So, sample variance S square will be summation j equal to 1 to k f j x j square minus n x bar square again divided by n minus 1.

So, this is how the sample mean and sample variance will be calculated when the data is discrete, and grouped in frequency distribution. Now many a times you need to have the intervals. So, what to do when the data is in that form? So, when data is discrete or continuous and have been placed in class intervals: so when you have the raw data, and when the data is placed in class intervals, in those cases the sample mean will be summation j equal to 1 to c and then f j m j upon n, and the sample variance will be summation of j equal to 1 to c f j m j square minus n x bar square divided by n minus 1.

So, the f j is observed frequency in j-th class interval and m j is the midpoint of j-th interval, and c is number of class interval. So, in that case as we see many a times the data which is there, and particularly when the data is quite large, and its very difficult to group them, even in those cases you make the intervals class intervals, and in the class interval the interval which you take the midpoint of that interval is m j m subscript j c will be number of class interval. So, that is why you are summing over this c, number of class interval and then you use these formulas to find the sample mean and sample variances. So, this is how the sample mean and sample variance is computed.

Now, we can see how we compute the sample mean and sample variance, when you have some data available with us. So, suppose we are dealing with a problem, where we are in a particular time on a day for a fixed time, the arrival rate or numbers of arrivals are recorded and it is done for 100 suppose days. So, how many arrivals were reported in that particular time. Suppose that value is given to us, and that value is given as like you have arrival per day and you have the frequency.

(Refer Slide Time: 11:13)

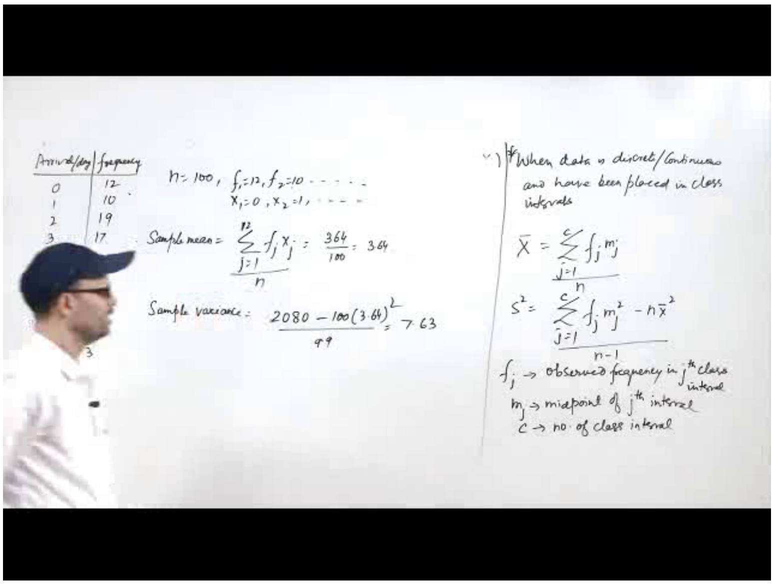


So, you have the arrivals starting. So, in some day you have 0 arrival in that particular time; say 5 minutes or 10 minutes or so.

So, in a particular time the number of arrivals may start from 0, and it may go to as high as 11 or even suppose, so more. So, you take from 0 11. So, from 0 to 11; suppose 0 1 2 3 4 5 6 7 8 9 10 and 11. So, suppose for 100 days in particular time, the number of arrivals were recorded and it was seen that on 12 days there were no arrivals. So, that is how the data will be there, that the arrival per day frequency was 12 when the data was 0.

Similarly, for 1 it was 10, 2 it was 19, 3 it was 17, 4 it was 10, 5 it was 8, 6 it was 7, 7 it was 5, then another 5, further 3, 3 and 1. So, if we add. So, 41 58 68 76 83 88 93 96 99 and this is 100. So, altogether you have 100 days of data. Now in this case, this is, the second equation will work, where we have or we have the data which has the frequencies.

(Refer Slide Time: 13:20)



So, we have to use the, to find the sample mean and sample variances you have. So, in this case you have n as 100 and you have f 1 as 12, f 2 as 10 and so on. So, this way and x 1 is 0 x 2 is 1 and all that. So, you have x 1 as 0 x 2 as 1 and. So, in that case if you try to find the sample mean. So, sample mean will be what? sample mean will be f 1 x 1 plus f 2 x 2 plus f 3 x 3, it will up to f 11 x 11. So, it will be 0 plus. So, it you will add. So, it will be summation, and it will be j equal to 1 to 11 and then you are adding. So, you will get f j x j.

So, you have. Now if we take f 1 and f 2. So, if we take x 1 as 0, and f 1 as 12. So, j will be basically varying from. So, we have to take j equal to 1, so fine. So, we are going up to. So, we have go up to. So, if it is 1 1 2 3 4 5 6 7 8 9 10 11 12. So, we have to 12 values. So, we are adding 12 times the values, and this of this comes out to be 364 and n is. So, this will be divided by n. So, it will be divided by 100 and it will be 3.64.

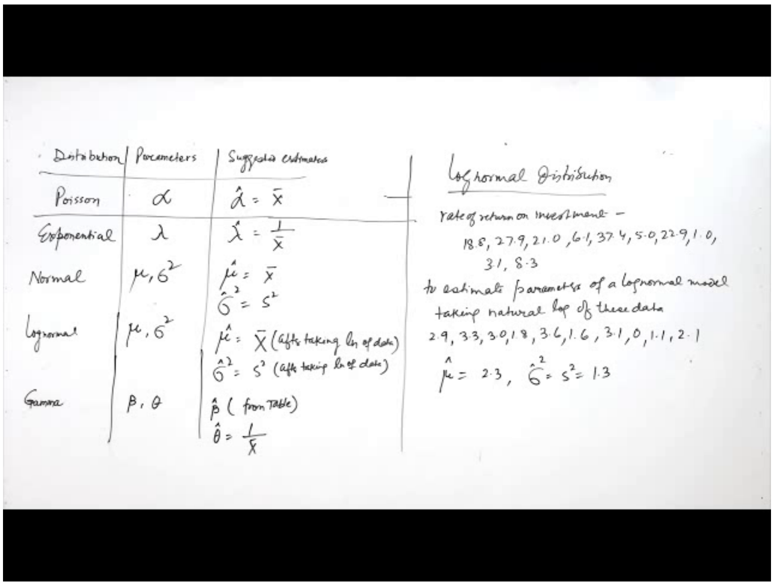
So, it is like 10 plus 38 48 plus 51, so 99, then 41 39 then it is 40; so179 plus 42. So, that is 221 plus 35. So, it will be 256 and then it is 40; so 296 plus 27. So, it will be 323, and then 33 53 plus 11, it is 364. So, 364 is coming and 364 by 11. So, that you get as 3.64.

Similarly, you can again use the formula for calculating the sample variance, and sample variance comes out to be. So, for sample variance we will have to calculate the f j x j square and x j square will be calculated, and then that will be multiplied by that then multi minus; so that f j x j square it will be basically 2080, summation of that will be 2080 and minus you have the sample mean, so 3.64 its square.

So, all this is n into x bar square and divided by n minus 1. So, it will be divided by 99 and this comes out to be 7.63. So, this is how when you have the data, you can calculate the sample mean and sample variance. Now what we have to see is, we have. So, this is how you may have the data for the class intervals. Sometimes you have suppose 50 data or 100 data, and you are putting in the class. So, in that you will have different classes, the class midpoint will be there, and then its frequency will be there, and then further using the midpoint value you can further use this formula to get the sample mean and sample variances.

So, this is how this sample mean and sample variance is calculated. Now we will discuss about the estimation of parameters. What are the different parameters which you required to know when you are dealing with particular frequency distribution? So, we will deal with the suggested estimators for frequency distribution, and the suggested estimators are like this.

(Refer Slide Time: 17:57)



So, we have the distribution. And similarly you have the parameters and you have the suggested estimators. So, as we know, if suppose you have the Poisson distribution.

So, for Poisson distribution the parameter is alpha and the suggested estimator is again. So, once we use the estimator you put this estimator sign; that is the unbiased estimator, we call it as, and that will be. So, whatever values are given whatever, suppose you have the number of arrivals per time per unit time interval is given. So, basically the mean value will be that alpha mean arrival rate. So, that will be alpha estimator will be x bar again for exponential. So, in exponential your parameter is lambda, and here the lambda estimator value will be 1 upon x bar. So, whatever value is given to you, you will have the reciprocal of that mean, and that will tell you the lambda estimator value.

Next comes is normal. So, as we know in the normal, you have the parameters as mu that is mean and the variance that is sigma square. So, mu and variance is given, you can have mu as the sample mean and variance you can compute from there. So, you have mu as sample mean and variance. So, this estimator value of the variance, it will be again the known as s square. So, there are the basically unbiased estimators. So, we assume that if the number of observations will be quite large, it will be closed to that value. So, we have to assume that, for getting any such value then you have log normal now. So, similarly you have log normal, as we know here we have to take the natural log of the values. So, in the log normal again, if the parameters are mu and sigma square in that case mu estimator will be again x bar after taking log of data.

Similarly, you have the sigma square estimator, and that will be again S square. Again you have to take the log then you have. Similarly you have all the other kind of distribution functions you have. Suppose gamma function gamma distribution and in that case the parameters are beta and theta. So, the value of beta estimator will be taken from the value I think it is given on the table at the end of the book. So, you have to see in that book, there is one parameter calculated; that is m, this m is calculated as the data from summation of x i and summation of summation of l n x i. So, based on that you calculate the m and then from that m. The 1 by m is calculated, and for that you have a table. So, from there you get the value beta.

So, beta estimator you will be seen from table, and you have the theta. So, theta will be 1 by x bar. So, you have the data. Suppose from that data you will have the theta, and for beta you will have from the table. So, this way we can calculate different types of, you know for different types of distributions, we will calculate the suggested estimators. You have again further suppose you have Weibull distribution in that also, you have the parameters as alpha and beta. So, in that case you will have to find many a times for the calculation of beta you have the alternative procedures.

So, that will be again calculated. So, let us see that how we calculate this parameter values for different kind of distributions. Now suppose you have the, we see the example which we have done earlier, the mean arrival rate where we calculated, and if it is said to be that Poisson distributed. So, suppose if we say for Poisson distribution, so for this Poisson distribution. Now here in this case the sample mean was calculated as 3.64. So, that is why alpha estimator will be 3.64. So, in that problem which we discussed, just earlier where you had the 0 number of arrivals for certain number of times 1 arrival in certain number of times and so.

Now, let us take the example of suppose log normal. So, in the log normal, suppose you have some data in the normal, we have the data you have, you can get the sample mean and sample variance by taking the mean of sample data, and then further once you know the mean, then you can use them for getting the sample variance. Now let us see for other type of distribution functions. So, suppose we get the log normal. So, suppose for log normal distribution, and if we are given some data then how to compute the suggested estimators for log normal. So, suppose you are given the rate of return on investment. So, rate of return on investment is given as.

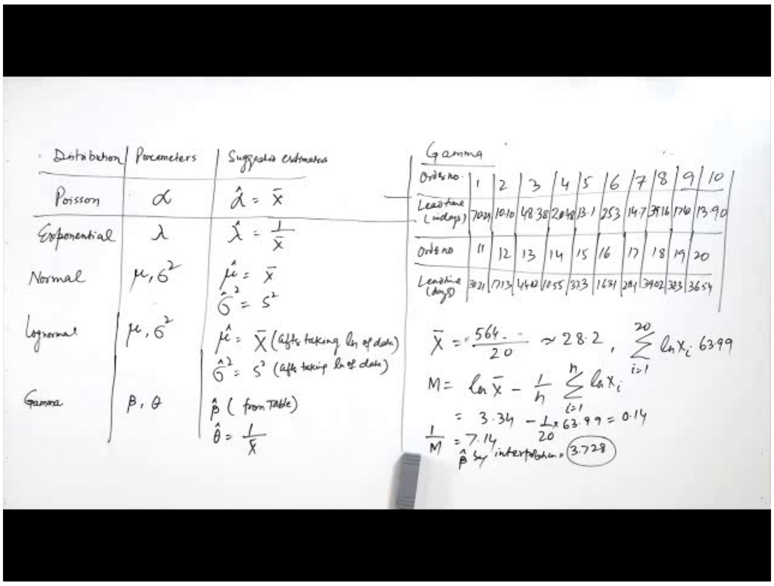
So, you have data like 18.827 7.921 6.130 7.4 5 22.9 1 3.9 and 8.3. So, these are the rate of return of the 10 investment. Suppose it is given and the. So, to estimate parameters of this log normal data of a log normal model, what we do is. We have basically we will get the mu and S square, but for that we have to take the l n of all that. So, you will have the data. Once we take the l n of these data; so taking natural log of these data. So, we will have the values. If we take the natural log of this data, it will be something like 2.9 3.33 1.8 3.6 1.6 3.10 1.1 and 2.1. So, when it is 1 it is 0 here. So, in that case, this is all the natural log l n of these numbers.

So, once we get that from there, we can get the mu estimator as. So, mu estimator estimated mean will be basically the sample mean of this. And sample mean if we calculate this, it will be equal to 2.3. So, some addition of all that divided by 10. So, it will be 2.3, and similarly we can have the estimated variance. So, estimated variance will be again sample variance. So, sample variances, if we compute it will 1.3. So, this is how you calculate the sample mean, and sample variance, if we have the data of a log normal model.

So, for the log normal model, we calculate the values of the estimated parameters; like that. Now let us see how to compute the estimated parameters for a gamma function gamma distribution. So, as we say see that in gamma distribution, you have 2 parameters beta and theta, and the estimated value of these for theta will be 1 by x bar. So, once you have the data, you get its sample mean from there you get the theta estimator, and this you get from the table. Now how to get that? So, let us see in case of a gamma distribution, you have been given certain data and how to compute it

So, in the gamma distribution suppose for an inventory model, in the every order, what is the lead time, and if it said to be gamma distributed the lead time. So, that data is given to you. So, suppose this data is given to you as this.

(Refer Slide Time: 29:05)



So, you have the order number and then you have the lead time in days. So, this is for gamma distribution. So, if suppose this data is given to you like 1 2 3 4 5 6 7 8 9 10 and 10. So, for the first it is 70.29, and then it is 10.10, and then it is 40 8 3 8, then it is 20.48 further 13.120 2.3.

So, 25.3 it is and further you have 14.7, then 39.16. Further you have 17.42 13.90. So, again order number and lead time in days. So, you have 20 data. Suppose then for 11, 11

12 13 14 15 16 17 18 19 and 20. So, and for this data again it is given 30.21, this is 17.13 this is 44.02, this is 10.55, this is 37.3, again 16.31 28.130 9.02 and 32.3 and 36.54.

Now, suppose you have this data, this 20 data which is said to be gamma distributed, and you have to find the parameters you know estimators of beta and theta; so for that as we know you get the theta as 1 by x bar. So, from these data what you get is, x bar that will be basically summation of all that divided by 20. So, summation is coming as 564.32 or. So, it is approximately that, and then you divide it by 20. So, it will be coming close to 28.2. So, it will, you can say that it is something close to 564 5 and then 564. So, it will be something close to 28.2.

So, that is your x bar. So, you can get theta estimator as 1 by x bar. Now what we have to see, we get m as what the formula for m, m is calculated as l n x bar minus 1 by n summation i equal to 1 to n l n x i. So, this is how m is calculated, while dealing with the gamma distribution. Now in this case w have computed x bar, we can also have the data for n summation l n x i i is varying from 1 to 20, we have 20 data. So, we get all the l n of these values, and then further you can add them. So, once you add them, it normally comes out to be 63.99, so closed to 64.

So, you can get m by putting in this. So, m will be l n x bar will be again 3.34, you have l n of 28.2. So, that is 3.34 minus 1 by 20 into summation of l n x i summation of l n x i is found to be 63.99. So, we get this value as 0.14. So, 1 by m we get as 7.14 now 1 by m is 7.14 we can refer to that table which is there.

(Refer Slide Time: 34:11)

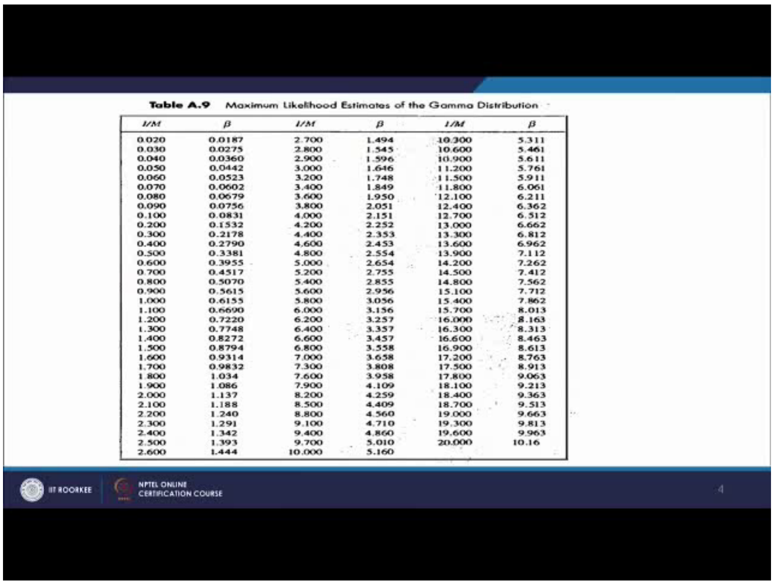


Table A.9 Maximum Likelihood Estimates of the Gamma Distribution

1/M	β	1/M	β	1/M	β
0.020	0.0187	2.700	1.494	10.300	5.311
0.030	0.0275	2.800	1.545	10.600	5.461
0.040	0.0360	2.900	1.596	10.900	5.611
0.050	0.0442	3.000	1.646	11.200	5.761
0.060	0.0523	3.200	1.748	11.500	5.911
0.070	0.0602	3.400	1.849	11.800	6.061
0.080	0.0679	3.600	1.950	12.100	6.211
0.090	0.0756	3.800	2.051	12.400	6.362
0.100	0.0831	4.000	2.151	12.700	6.512
0.200	0.1532	4.200	2.252	13.000	6.662
0.300	0.2178	4.400	2.353	13.300	6.812
0.400	0.2790	4.600	2.453	13.600	6.962
0.500	0.3381	4.800	2.554	13.900	7.112
0.600	0.3955	5.000	2.654	14.200	7.262
0.700	0.4517	5.200	2.755	14.500	7.412
0.800	0.5070	5.400	2.855	14.800	7.562
0.900	0.5615	5.600	2.956	15.100	7.712
1.000	0.6155	5.800	3.056	15.400	7.862
1.100	0.6690	6.000	3.156	15.700	8.013
1.200	0.7220	6.200	3.257	16.000	8.163
1.300	0.7748	6.400	3.357	16.300	8.313
1.400	0.8272	6.600	3.457	16.600	8.463
1.500	0.8794	6.800	3.558	16.900	8.613
1.600	0.9314	7.000	3.658	17.200	8.763
1.700	0.9832	7.300	3.808	17.500	8.913
1.800	1.034	7.600	3.958	17.800	9.063
1.900	1.086	7.900	4.109	18.100	9.213
2.000	1.137	8.200	4.259	18.400	9.363
2.100	1.188	8.500	4.409	18.700	9.513
2.200	1.240	8.800	4.560	19.000	9.663
2.300	1.291	9.100	4.710	19.300	9.813
2.400	1.342	9.400	4.860	19.600	9.963
2.500	1.393	9.700	5.010	20.000	10.16
2.600	1.444	10.000	5.160		

So, this is the table, this table can be referred 1 by m value, and 1 by m value which is coming out to be 7.14 if you look at.

So, in this case in between this 3.65 and 3.80; so from here you can get the beta by interpolation, and beta will be coming as 3.728. So, you can get the value of beta by interpolation, you can get as 3.728. So, you have got these two estimators beta and theta estimator, and this is how you compute these estimator values for different kinds of distributions. You can practice for different distribution functions, and get used to finding these suggested estimators.

Thank you very much.