INDIAN INSTITUTE
OF
TECHNOLOGY
KHARAGPUR

NPTEL
National Programme
on
Technology Enhanced Learning

Probability Methods in Civil Engineering

Prof. Rajib Maity

Department of Civil Engineering
IIT Kharagpur

Lecture – 19

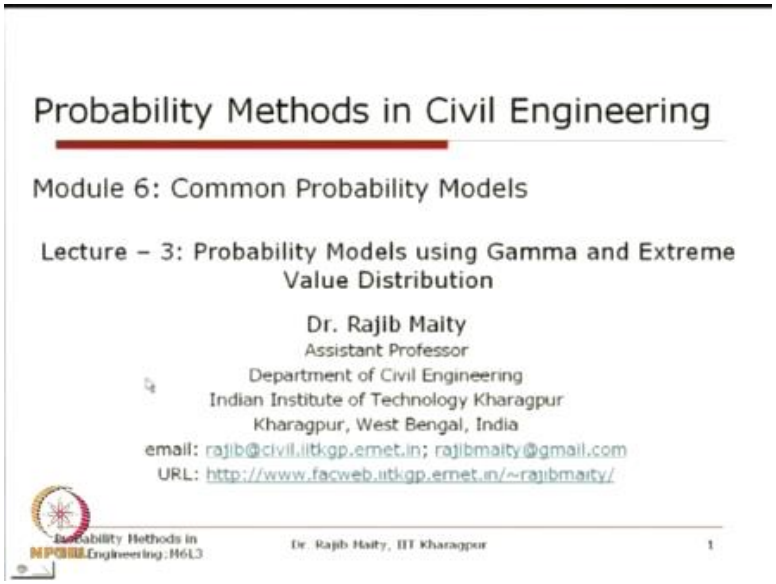
Topic

Probability Models using Gamma and Extreme Value Distribution

Hello and welcome to this lecture we are in now module six and where we are discussing about some of the probability models and using those standard probability distributions we are trying to solve some of the civil engineering real life problems and this is our third lecture of this module and in the earlier two lectures we have seen some of the models for example that normal distribution then log normal distribution and so this lecture will be taking some more continuous random variables for example we will take the gamma distribution then extreme value distribution.

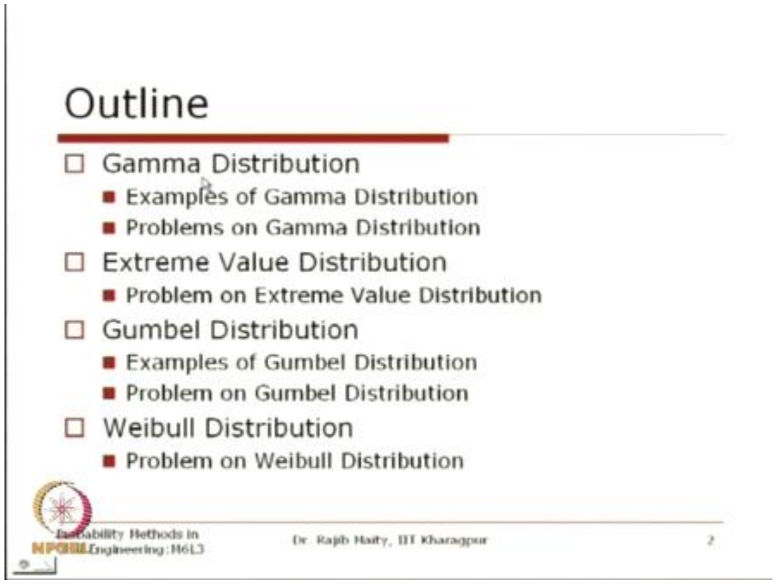
And all this distribution similar to the other engineering or other field of application in civil engineering area also there are tremendous application of this type of distribution so we will take up this in this lecture also we will take you through some of the applications of the civil engineering problems through these probability models.

(Refer Slide Time: 01:18)



So our today's lecture is on this probability models using gamma and extreme value distributions

(Refer Slide Time: 01:25)



And as we will start with the gamma distribution where we will first discuss about what are the basics of this distribution function and all you know that these distributions are discuss earlier in details now we will not go into that detail but we will just briefly overall we will mention those distributions and their few properties and we will basically this module is focused to its application through this type of probability models this type of models so we will go to some of the examples of this gamma distribution and there are some problems that we can address through this gamma distribution obviously this problems are related to the civil engineering problem that we will we will discuss.

After that we will take another distribution which is extreme value distribution and we will discuss some of this basic problems first on this extreme value distribution and we will discuss what are the different types of extreme value distribution there are basically three types are there So we will discuss which type is generally is use and we will also see what are the generalized extreme value distribution form and how we can take those distribution of this different types of information.

And how those different types of extreme value distribution can be applied to different types of problems then we will take that Gumbel distribution, which is also that extreme value type one distribution you also say so this Gumbel distribution having a very wide application particularly when we are talking about some extreme events in particularly in the area of the hydrology and water resource which is one of this discipline of the civil engineering that we will discuss through what are the different application problems that we can address with this distribution also we will see some of the examples using this distribution.

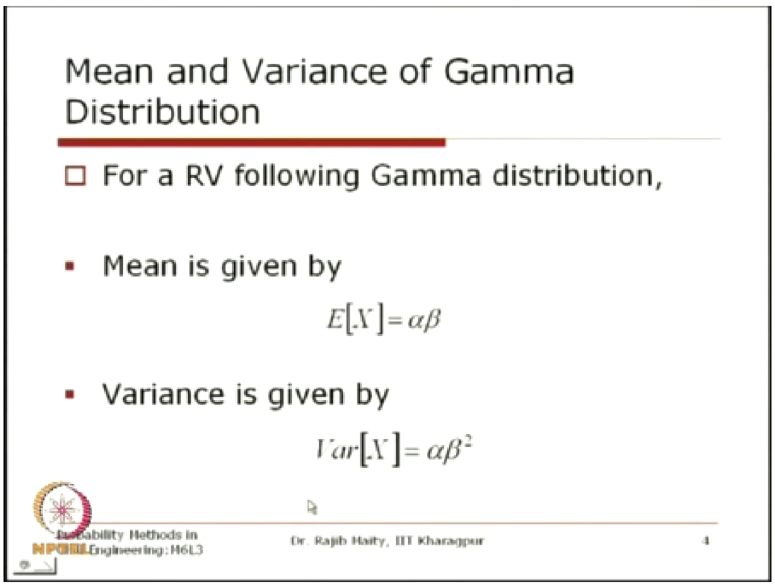
After that another important distribution is a Weibull distribution this Weibull distribution generally is used to address some of the structural engineering problem for it is failure and all and we will see we will discuss that distribution along with some of this problems also one thing, that this there is one distribution called the reverse Weibull distribution which is also that one of this type of this extreme value distribution this is known as the type three distribution we will see different types as I told that while when we are discussing when we will be discussing this extreme value distribution so, to start with we will start with this gamma distribution.

And we know from the earlier module and our earlier discussion that a random variable x is said to follow the gamma distribution if it is probability density function is given by this in this form which is that $f_x(x)$ = 1 by β^{α} $\gamma\alpha^{x\alpha\ -1}$ $e^{x/\beta}$ and you know that this support of this distribution is non-negative that is greater than equal to 0 that support of this random variable x and this α and β are non negative are greater than 0 so far as their this these are the two parameters of this distribution. This gamma function, this is this is a gamma function which is its notation is γ, α and which can be expressed through this one, there is from 0 to∞ $x^{\alpha\ -1}$ $e^{\ -x}$ dx.

So now this gamma distribution this is a pdf that is probability density function and the cumulative gamma distribution function is given by you know that we have to integrate it from this left extreme that is 0 to that particular value x. It will come like this 1/γα which is obviously, a constant and we can integrate this one to get its cumulative distribution now this integration that is this γα this generally available from this standard table. Now we also discuss earlier, that if α is an integer, then that γα can be expressed at this α - 1 x γα -1.

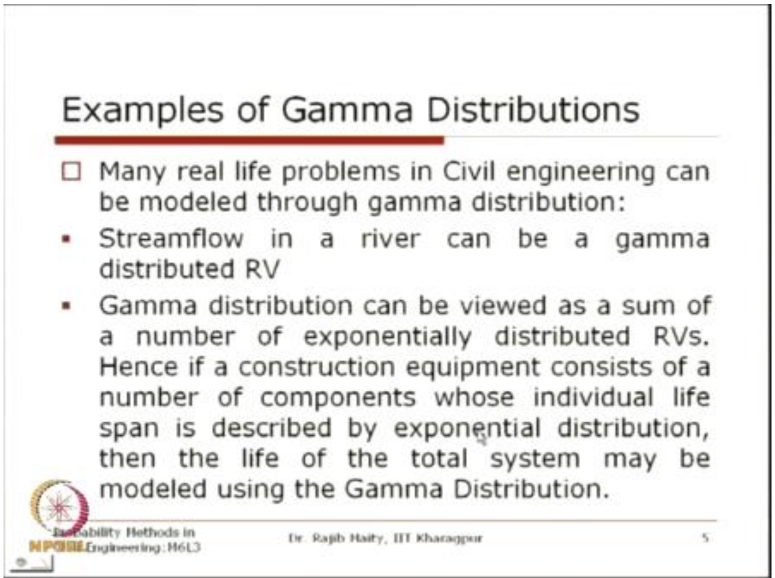
So in that way if we know that this values for the 0 to 1 only for this α then we can get any value of this gamma function for any number so that we can get from that table and this integration we can get from this either from this gamma distribution table from this standard text or from this numerically integration now there are different software's are available from which we can use this use we can use the numerical integration to obtain these this values.

(Refer Slide Time: 06:39)



Now from we have also seen that this parameters that is the mean of this random variable of which follows this gamma distribution can be expressed through its parameter that is mean is that the product of two parameters α and β and the variance is $\alpha\ \beta^2$

(Refer Slide Time: 06:59)



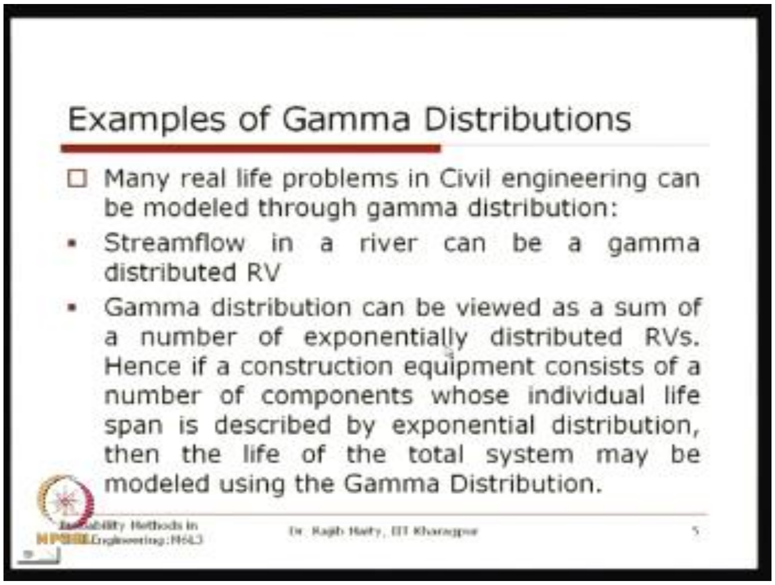
And the other parameters are also discussed earlier now we will see some of this as our main goal of this lecture is to just see how we can use this probability model for different problems of this civil engineering so first we will first we will discuss some of the real life problem which can be modeled through this gamma distribution so there are as I told that this is one of the very important potential distribution because of these two parameters and most of the cases when we can say that this random variable having a range from 0 to the ∞. So, the negative side is excluded.

Then the two parameters there is α and β are there, by changing their values it can take a wide range of safe of the distribution. So many random variable related to the civil engineering can modeled through this distribution just by adjusting their parameter and this we have discussed earlier also we have we have shown through the graphical presentation that how this the change of this parameters can change the shape of this shape of the distribution function so that a wide range of random variable can be modeled through this distribution.

For details you can refer to the module three when we discuss this distribution. So, many real life problems in civil engineering can be modeled through this gamma distribution for example that streamflow in a river can be a gamma distributed random variable so most of this large river basins where this you know this streamflow can is always a positive number 0 to ∞ so, this streamflow can be modeled through this gamma distribution again one another most important feature of this gamma distribution that we have also shown in the earlier that if you just adjust the parameter this is that exponential distribution is a special case of this gamma distribution.

Basically if we just take few of their few that exponential distribution and we can define another random variable which is a summation of those exponential distribution. Then that summation is basically a gamma distribution.

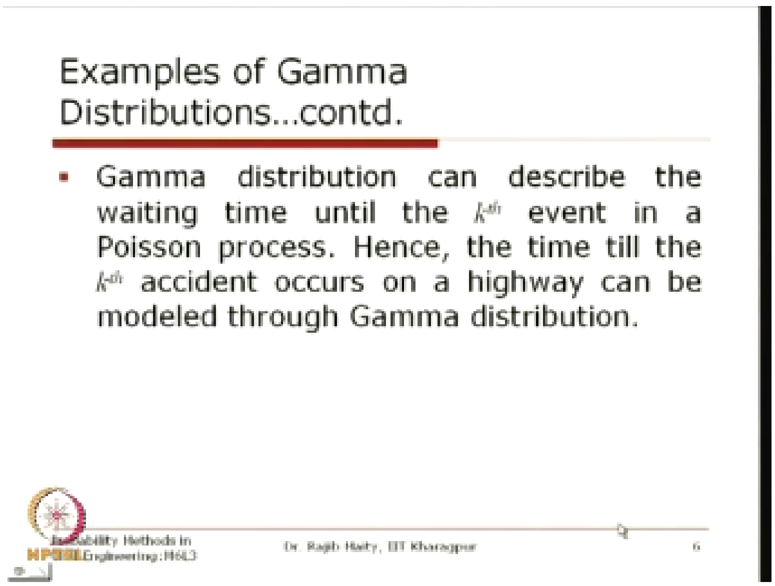
(Refer Slide Time: 09:28)



So this is what using this properties this properties there are many problems can be modeled. So the gamma distribution can be viewed as a sum of a number of exponentially distributed random variables. Hence if a construction equipment consists of a number of components whose individual life span is described by the exponential distribution then the life of the total system may be modeled using the gamma distribution.

So there are different components and each component is having a life span of some x1, x2, x 3 and all these if all these variables generally can be modeled you know that this time to the first occurrence of one event can be modeled through this exponential distribution. So now the full system can consist of this summation of this distribution, so the total system the life of the total system can be modeled using this gamma distribution.

(Refer Slide Time: 10:28)

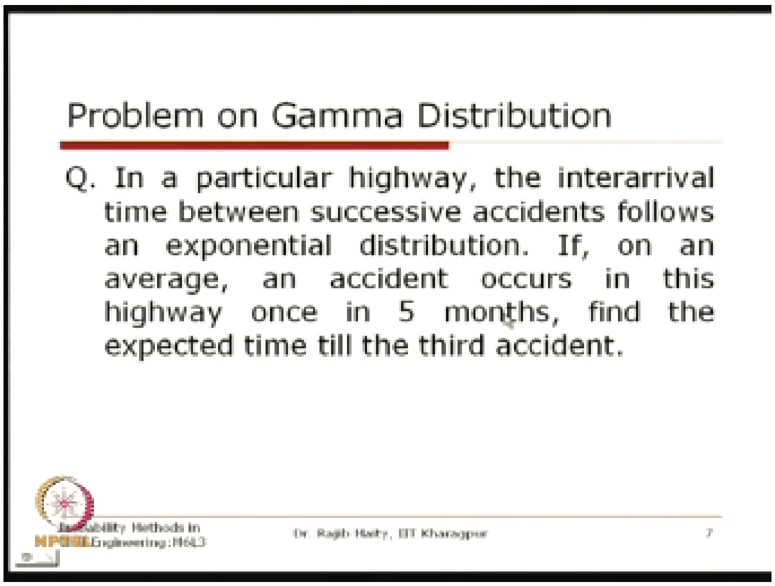


Another thing is that this is this is mostly used to address different problems of this transportation engineering that this gamma distribution can describe the waiting time until the k^{th} event in a poisson process. Hence the time till the k^{th} accident for example one example is taken from this transportation engineering that there are the accidents take place either on the railway or on the highway.

So that you we can we can take that what is that what is the waiting time till that k^{th} accident can take place. So that k^{th} accident take place means there are basically the summation of this k inter arrival times of the of the accidents. So that it is a summation of those that exponential distribution which is the inter arrival time is modeled through. So this one so this time till the k^{th} accident.

Which is summation of the k individual exponentially distributed random variables. So this can be modeled through the gamma distribution as it is explained in the last point. So we will take a similar problem here.

(Refer Slide Time: 11:41)



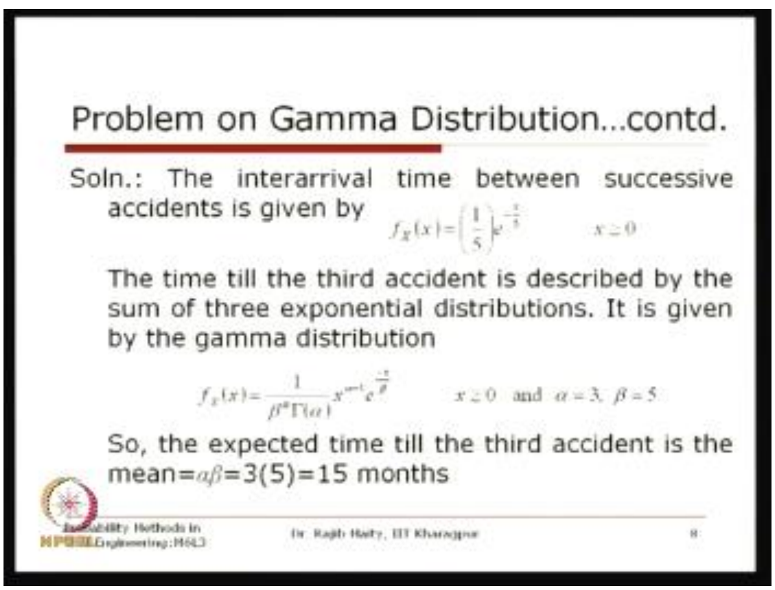
The problem is easy but it is very interesting to understand that how it is linked to this gamma distribution. So in a particular highway the inter arrival time between the successive accidents follows an exponential distribution. If on an average an accident occurs in this highway once in five months find the expected time till the third accident. Now thing is that so this is known that we know that the exponential distribution that we have describe earlier that and it is given that the average time between the accident is five months.

So that the parameter of this exponential distribution which is lambda is equals to 1 by x-bar. So this 1/5 months. So this is the parameter for this exponential distribution. Now we have to find out what is the what are the parameter associated parameter for the gamma distribution and we know that the expected time is you looked for. So just the multiplication of that two parameters that is α and β.

So that will give you the expected time for till the third accident. Now the straight forward thing here is that one accident in the five months and we are assuming that these are the dependent events that is the inter arrival time between the first and between the between now and the first accident and then from the first accident to the second and then from the second to the third. So all these three inter arrival times are independent to each other.

So that if for the one inter arrival time the expected time is five months then the expected time for the three such events can be easily multiplied by this 3 and by this five months we can get the answer that is the time is the fifteen months. Now if we want to convert it from the exponential distribution to the gamma distribution.

(Refer Slide Time: 13:30)

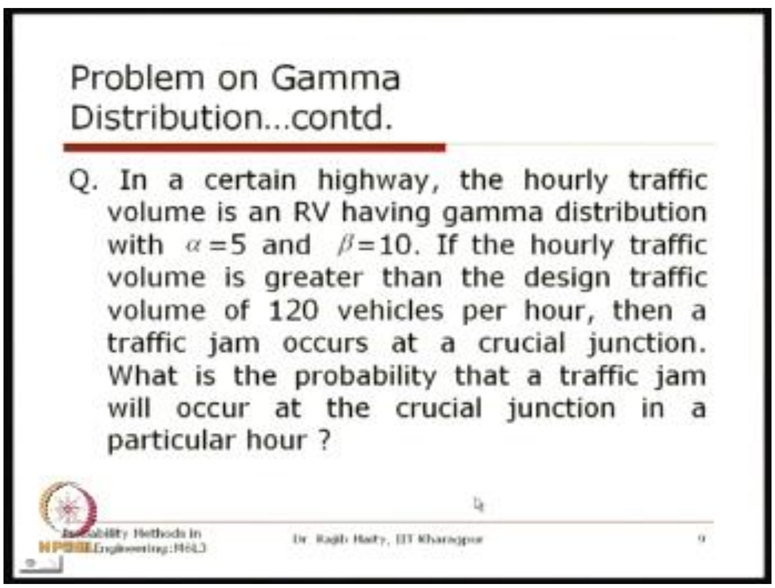


Then first of all the inter arrival time between the successive accidents is given by that distribution of this exponential distribution as it is mentioned in the problem that is fx(x) = 1/ $\lambda e^{\lambda x}$ and this λ = 1/ 5. So 1/5 $e^{-x/\lambda}$ obviously x > = 0.So if you get this one then the time till the third accident is described by the sum of the three exponential distribution. So thus 3 means from today.

So from today I am starting what is the time till the first accident this is a one inter arrival. So one the first incident occurs then from the first incident to the second incident then second to the third. So the three exponential distributions are there so the so the gamma distribution will have the parameters x>= 0 and the α should be equals to 3 and β is equals to is 5. So β is nothing but you know that β is the your that 1/β is here that that lambda.

So this β is 5 and there are three inter arrival times. So this one so the and so we have seen that parameters are α =3 and β = 5.So the expected time till the third accident is the mean that is α x β. So 3 x 5 so fifteen months as we have just expecting just by looking the problem itself. So this is how as how many exponential distribution we are adding and what is the parameter for the exponential distribution. So that which come from this lambda and that number of those distributions are being added from which we are getting this gamma distribution.

(Refer Slide Time: 15:18)

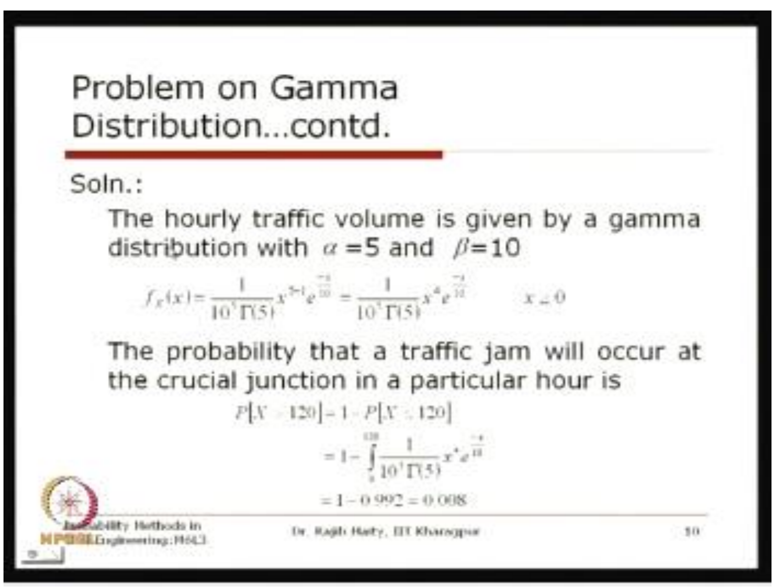


So we will take another one another problem this is also related to the transportation engineering and sometimes we design the road network for certain traffic volume and then we can see we can estimate that how this traffic volume can be modeled through and this traffic volume also can be modeled through this gamma distribution and from there we can answer some of the real life problem.

That what is the probability of the traffic jam and all so this type of problem is taken here once. So in a certain highway the hourly traffic volume is an random variable is a random variable having gamma distribution with $\alpha = 5$ and $\beta = 10$.If the hourly traffic volume is greater than the design traffic volume of this 120 vehicles per hour then a traffic jam is possibly can occur at a at a crucial junction of the road network.

What is the probability that a traffic jam will occur at the crucial junction in a particular hour. So first of all so it means traffic jam will occur that is the random variable that I have to define that is what is the hourly traffic volume that is the random variable here which is a gamma distribution parameters are given. So now we have to find out what is the probability that particular random variable will be greater than equal to 120. So that is the overall problem is given here.

(Refer Slide Time: 16:42)

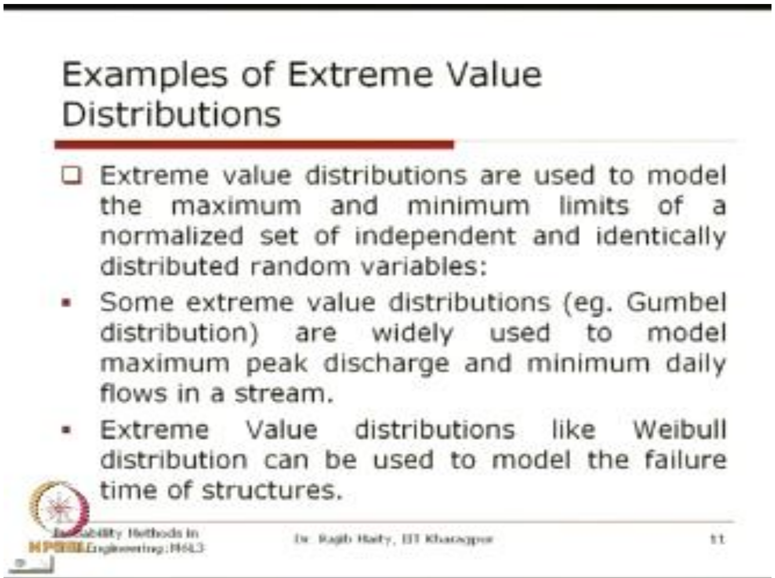


So the hourly traffic volume denoted by this random variable x is given by the gamma distribution with $\alpha = 5$ and $\beta = 10$. So the distribution that we can get from this gamma distribution which is that your fx(x) = $1/10^{5}$ then gamma5 and then x^{5-1} So these are just we are using this those parameters α - 1 and $e^{-x/\beta}$. So we will just put this value here to get that complete form of this gamma distribution.

Now the probability that a traffic jam will occur at the crucial junction is a is a in a particular hour is given by this one that is probability of $x > 120$.Now $x > 120$ obviously you know that by this time that it can be the total probability $-x < 120$ which is 1- 0 to 120 of these integration of this full this distribution and obviously dx. Now if you do this integration you can use some of the software to do this one this numerical integration or you can use some table also to get the value this value will come as 0.992 so 1 - 0.992 gives you the 0.008.

So the probability of this traffic jam is a very less here you can see this also can be inferred from the data that is that is given that is $\alpha = 5$ and $\beta = 10$ that means the expected time expected traffic volume is that 10 multiplied by 5. So which is the 50 vehicles per hour and we are looking for the probability which is exceeding 120 vehicles per hour so the probability that is there so is obviously should be very less which is 0.008.

(Refer Slide Time: 18:55)



Now we will go through some of this extreme value distributions first of all we will just see what is this extreme value distribution and what are this it is possible application this extreme value distributions are used to model the maximum or minimum limits of a normalized set of independent and identically distributed random variables so there are some random variables is available to us and we are interested only for their extremes, extremes means here either we are interested to know the nature of its maximum side or the nature of it is the minimum side.

So this when we are when we are interested for this type of variable that is suppose that one data set is available to me and I am just interested to know its maximum one the maximum value out of this data set and what is the distribution of the maximum so that is the extreme value that we are referring to and its distribution it is associated distribution is refer to as this extreme value distribution similarly for the minimum also we will see and there are different types basically is available that we will discuss through.

So the extreme value distribution when we are when we are talking about we are basically referring to the event where out of the set of from this data or out of the set of this random variable we just want to pick up it is one of the extreme either maximum or the minimum and want to know its probabilistic distribution so if you see from the example from the civil engineering one is that for anything suppose that we are talking about the stream flow then whether the low flow is the one extreme and the maximum flow is the another extreme.

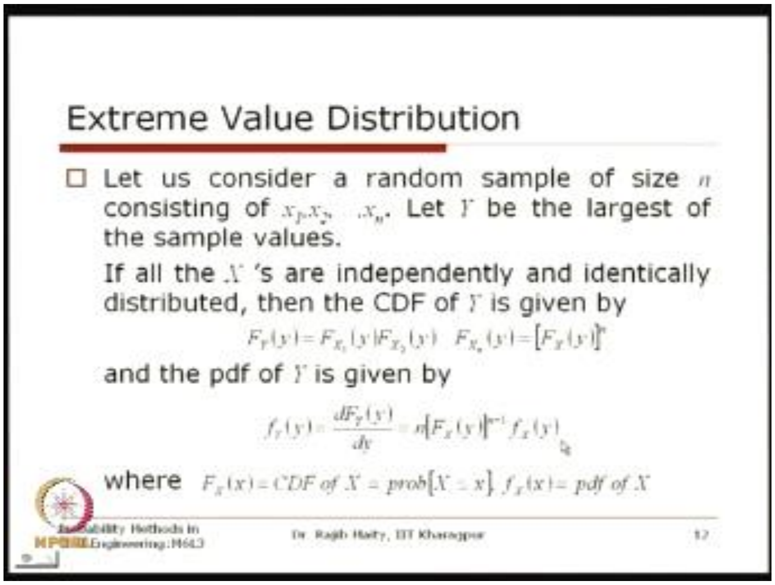
Similarly if we see that the structural components so its failure due to the several reasons are there one of the thing is that fatigue the failure due to fatigue so how many times it can what is the total time it can go through and what is the maximum time that it can pass through that can be also be modeled through this extreme value distribution similarly in the other application also from this environmental engineering or the transportation engineering, whatever the data that we are having if we are interested to model only it is extreme side either maximum or the minimum one.

Then we will refer to this type of distribution so some of the extreme value distribution that is the Gumbel distribution is one that I was also mentioning at the beginning of this lecture that this extreme type distributions are widely used to model the pick discharge and minimum daily flow in a in a stream there are other extreme value distribution like the Weibull distribution which can be used basically this is the reverse Wei bull type distribution which is the type three distribution which can be modeled for this failure of this structure also so here we will we will take that first we will take that generalized extreme value distribution.

Then we will discuss about its different type then we will go through this will take into details as to what is this Gumbel distribution and how to model this peak discharge and this minimum flow in a stream so this is important because this extreme value distribution means so far as this water resource or hydrology is concerned we are always interested to know that either the very low flow or the or the very maximum flow because both are is needed to be explode for its for is the societal impact.

So this we will see so that when we are talking about that there is a maximum or minimum limits of a normalized set of this independent and identically distributed random variables so we will just see the first of all we will see what is the specific distribution.

(Refer Slide Time: 23:05)



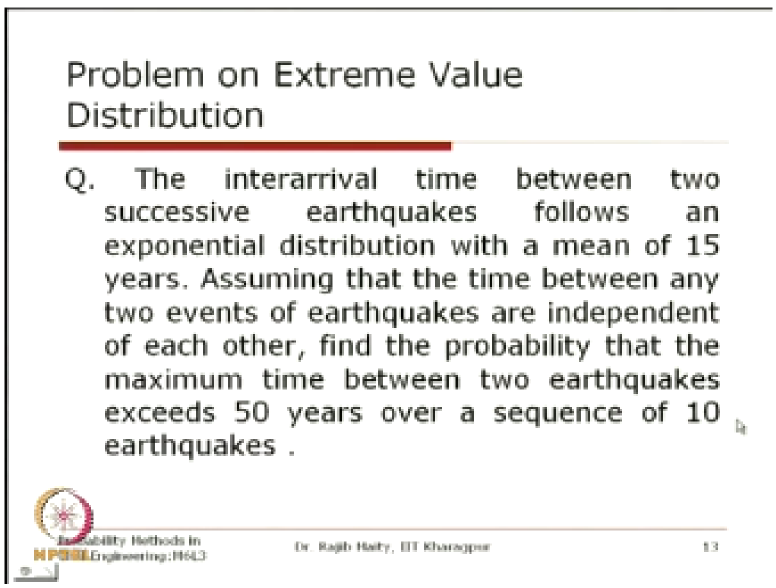
That we can that we can take you through whatever the knowledge that we have seen from this CDF and the pdf of the if the assumption on the assumption that it is independently and identically distributed let us consider a random sample of size n consisting of this x 1 x 2 and up to x n so this is basically the data that is available to us let y be the largest of this sample value so we are interested to know out of this n which one is the maximum so that we are denoting as this y now if all the x's are independently and identically distributed then the CDF of y is given by that we know from our earlier discussion.

Is that the CDF that is the fy(y) should be the product of their individual cumulative distribution now when we are taking that taking that data dataset so each and every observation can be can be treated as one random variable so like that this is the this is the distribution of this the first one that is x 1 x 2 and this is for the x n now as they are they are as they are independent that is why we have we got their product as to get their joint distribution as they are identical that is if we just say that all these are identical and equal to that f x.

Then we can say that f x of now it is expressed through the y^n so that there are n different random variables are there so power n will give you that CDF of that the largest value of the sample and now obviously once we get that CDF we can get its pdf also probability density function that by differentiating that so this Fy (y) is equals to that differentiation with respect to y which is nothing but n the cumulative density power n - 1 multiplied by its pdf of the individual random variables.

So where this fx x is a CDF of x that is probability of x less than a specific value of x and this f small fx x is the pdf of that x now these are all means obviously express through that one variable which is y.

(Refer Slide Time: 25:35)

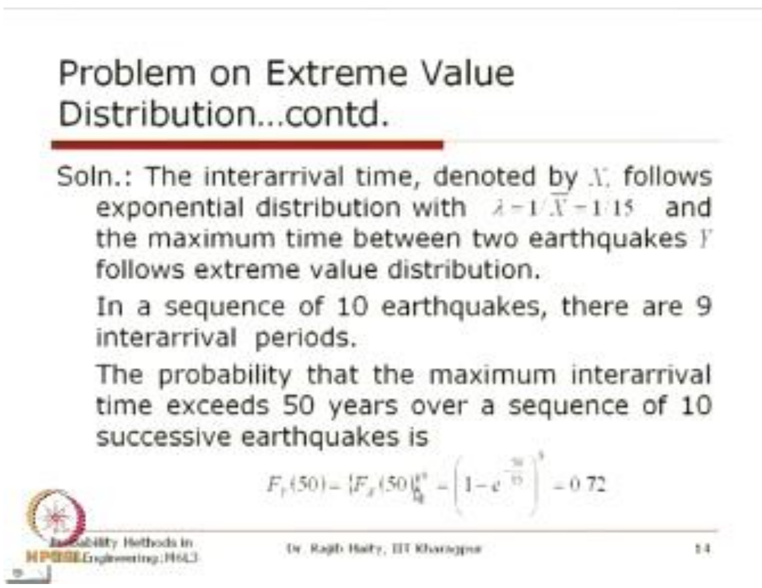


Now suppose we will just see the same thing through 1example that is a inter arrival time between two successive earthquakes so again we know that these are this can be easily treated as this independent and this can also be treated as this identical so both are same distribution if we assume, whether we can answer this one there is a maximum inter arrival time so the problem related to that one so the inter arrival time between two successive earthquakes follows and exponential distribution with a mean of 15 years.

So we know that mean between two successive earthquake is 15 years assuming that the time between any two events of the earthquakes are independent of each other so this is that that first assumption so find the probability that the maximum time between two earthquakes exceeds50 years over a sequence of 10 earthquakes so the inter arrival time having a mean of 15 years we can model it through this exponential distribution.

Their independence we get their joint distribution by their product and then what we are interested to know is that what is the maximum inter arrival time then what is the probability that the maximum time will exceed 50 years over a sequence of 10 successive earthquakes.

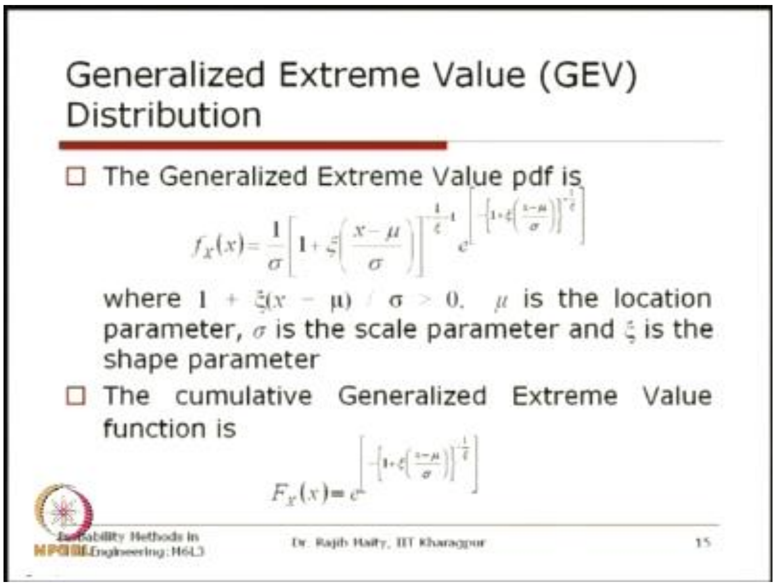
(Refer Slide Time: 27:05)



So the inter arrival time which we can denote by x follows an exponential distribution with $\lambda = 1$ by x bar that is 1 by 15 the maximum time between two earthquakes that is denoted by y, follows an extreme value distribution now in a sequence of 10 earthquakes there are 9inter arrival inter arrival periods now the probability that the maximum inter arrival time exceeds 50 years over a sequence of 10successive earthquakes is that, is this that is f y of 50 is = fx 50^9 so this is basically we are taking it from that this expression so this n is now here is total number of inter arrival times which is 10 -1 obviously that is 9.

So this 9 if we take and this is that this is this is their $1-e^{-50/0.15}$ which is equals to 0.72. One correction, this is not the maximum interarrival time exceeds 50 years. Rather, this 0.72 is the probability of maximum interarrival time be less than 50 years. Now, if you want to know the maximum interarrival time exceeds 50 years, then it should be subtracted from total probability1. So, the maximum interarrival time exceeds 50 years should be 1-0.72 that is 0.28.

(Refer Slide Time: 28:47)



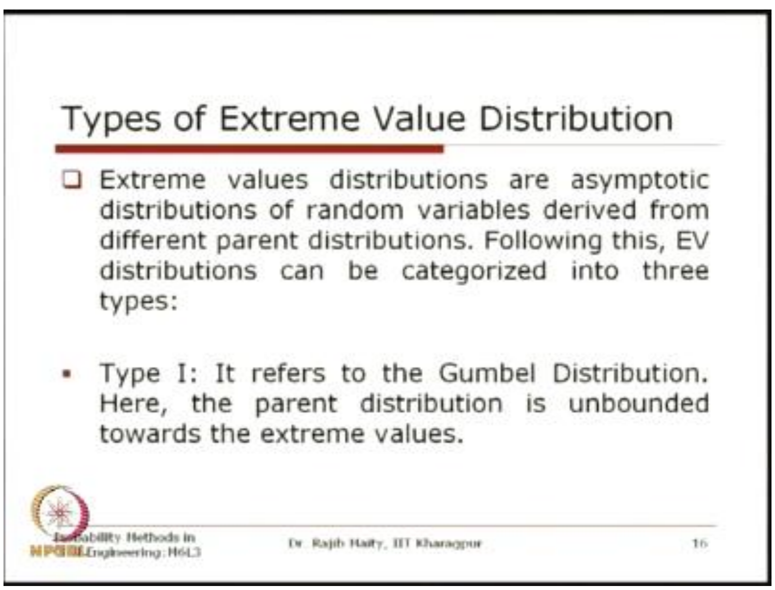
Now, we will see the generalized extreme value distribution that first of all, this is a, this distribution that we are going to discuss it is a generalized frame work. Now, this generalized extreme value pdf is expressed through a little compression expression, where this $f_x=1/\sigma[1+\xi(x-\mu/\sigma)^{1/\xi-1}]$ that exponential power again, the same expression that is $-1+\xi(x-\mu/\sigma)^{-1/\xi}$. Now, if we just see that their individual element then, it will be more meaningful that the first thing is that, this distribution is the, its range, its support should be such that that $1+\xi(x-\mu/\sigma)$ should be greater than greater than 0.

Where this μ and σ that we know already, which is similar to our that other distribution, that is μ is your the location parameter and σ is your scale parameter. Basically, both are location parameters, means where it is, means it is giving the information about its mean and this is giving its information about its variants.

Now this ξ is another parameter which is the shape parameter. Now, this ξ is one of this crucial parameter in this distribution in the sense, that it generally controls the tail behavior. Now depending on this, whether this tail behavior, how it will behave that we can classify into three different types. The first thing is that if this ξ is tending to 0, or whether this is negative or this is positive. So, this three cases can lead to three different types of the of the distribution that we will discuss in a minute.

And for this one, if we just follow the same principal of getting its the cumulative distribution function, then it can be shown that this $F_x(x)=e[-\{1+\xi(x-\mu/\sigma)\}^{1/\xi}]$. Now, as I was telling that this ξ is a shape parameter, depending on which it is either tending to 0 or it is positive or it is negative.

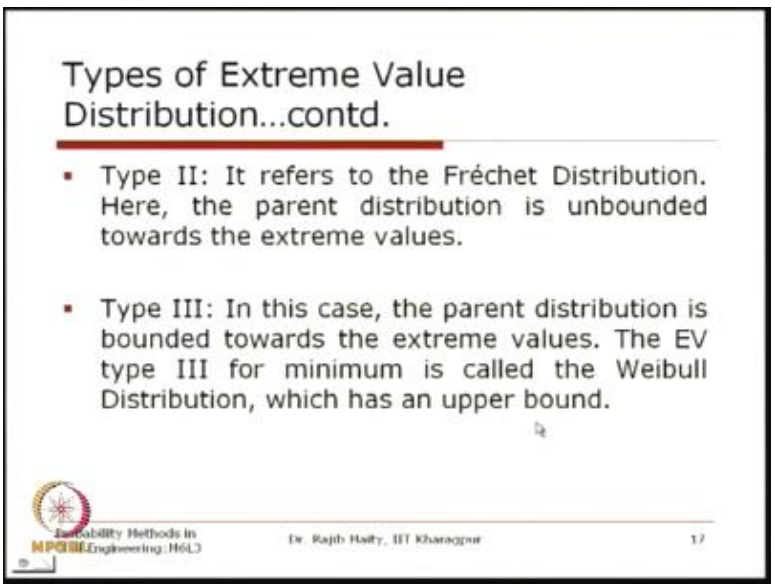
(Refer Slide Time: 31:26)



So, in these three cases, there are three different types are there. So, this extreme value distributions are asymptotic distribution of the random variable derived from the different parent distribution. Following this extreme value distributions, can be categorized into three types. The first type that is type one, it refers to the Gumbel distribution. As I was also mentioning earlier, this type, this Gumbel distribution is also known as these type one distribution. Here, the parent distribution is unbounded towards the extreme values.

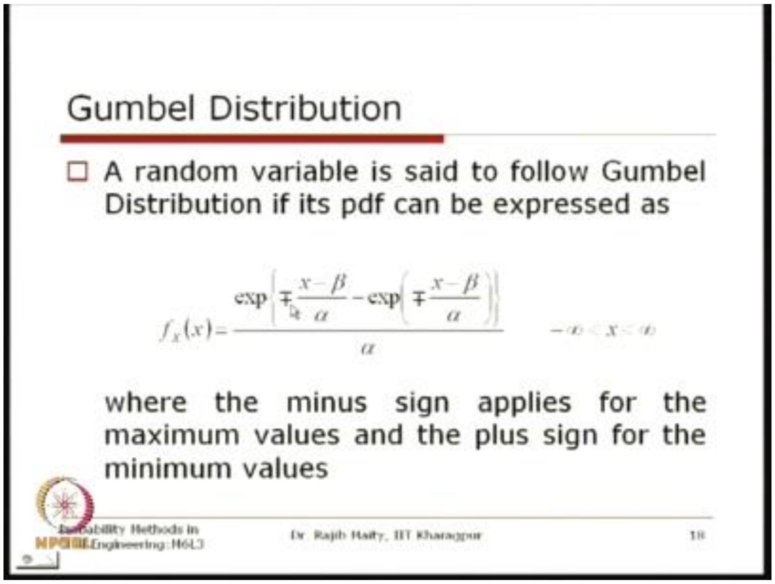
Now, unbounded towards the extreme values means, if you are interested, suppose that the maximum one or the largest one, if we tell, then we can see that. This for example, that this normal distribution towards the positive extreme is unbounded and even the gamma distribution towards the positive boundaries unbounded, like that.

(Refer Slide Time: 32:16)



Similarly type two it refers to the Frechet distribution. Here, the parent distribution is unbounded towards the extreme values and the type three in this case, the parent distribution is bounded towards the extreme values. The extreme value type three of the minimum is called the Weibull distribution, which has an upper bound.

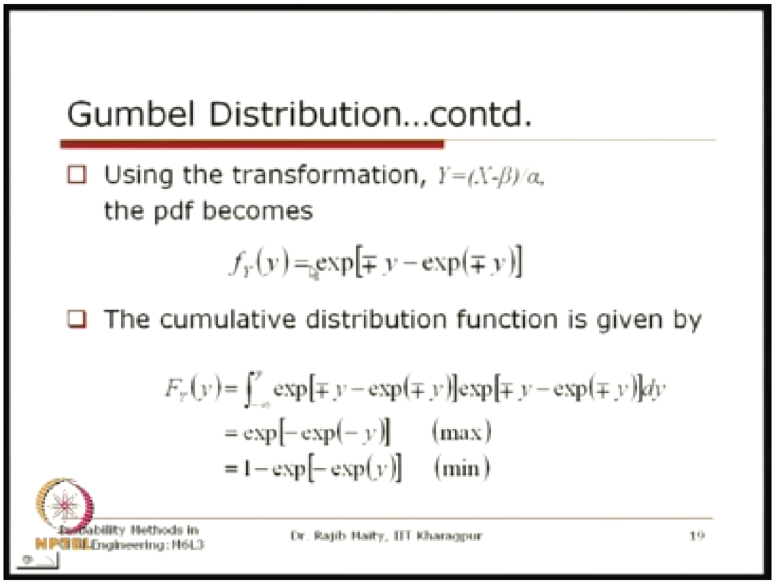
(Refer Slide Time: 32:47)



So, with these three types of this distribution, now we will see, one of this the type one distribution which is the Gumbel distribution and as I told, that this is having a very wide application. So far as the hydrology and water resource is concerned, to analyze the extreme values of the stream flow either it is low flow or it is the high flow. So, a random variable is said to follow the Gumbel distribution if its pdf can be expressed as like this; that is, $f_x(x)=\exp\{\mp x-\beta/\alpha-\exp(\mp x-\beta/\alpha/\alpha)\}$ this x has a limit from this $-\infty$ to $+\infty$.

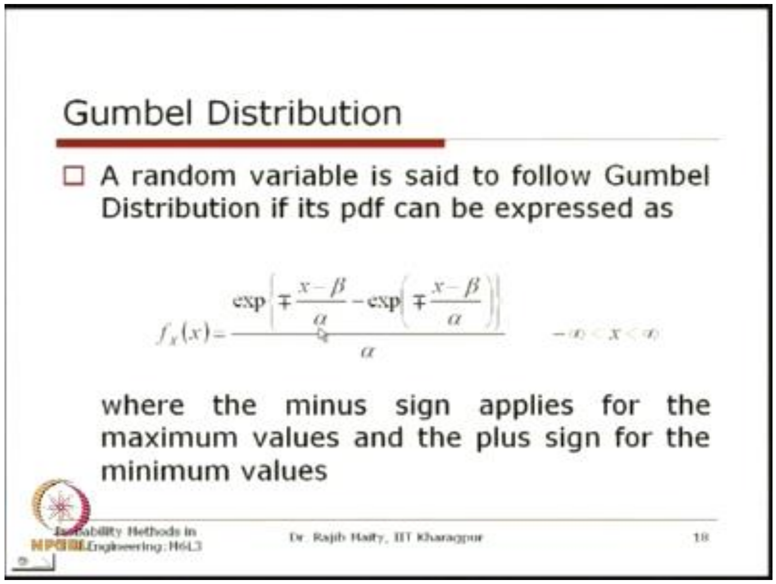
Now this both, this signs have some means, we have to use one of these signs as I have mentioned the - and +, one we have to use for the maximum and other one we have to use for the minimum one. So, here it is mentioned that where the minus sign applies for the maximum values and the plus sign is applicable for the minimum values. So, if you are interested to model the higher side of the extreme, that is the maximum values, then we have use this minus sign and this is, then the, and if it is for the, if it is for the minimum one, then you have to use the plus sign.

(Refer Slide Time: 34:16)



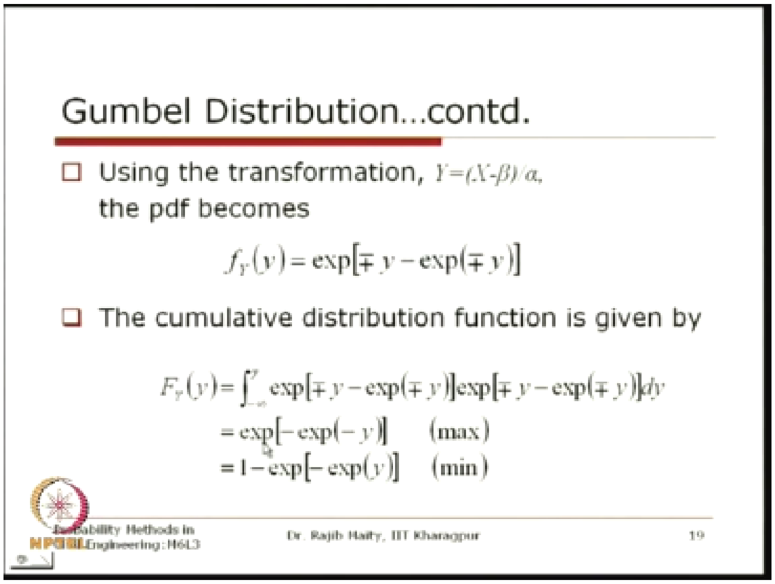
Now, if we use this transformation that is Y=(X-β)/α then the pdf can be again expressed.

(Refer Slide Time: 34:27)



Then this one just by transforming that (X-β)/α is equals to that Y.

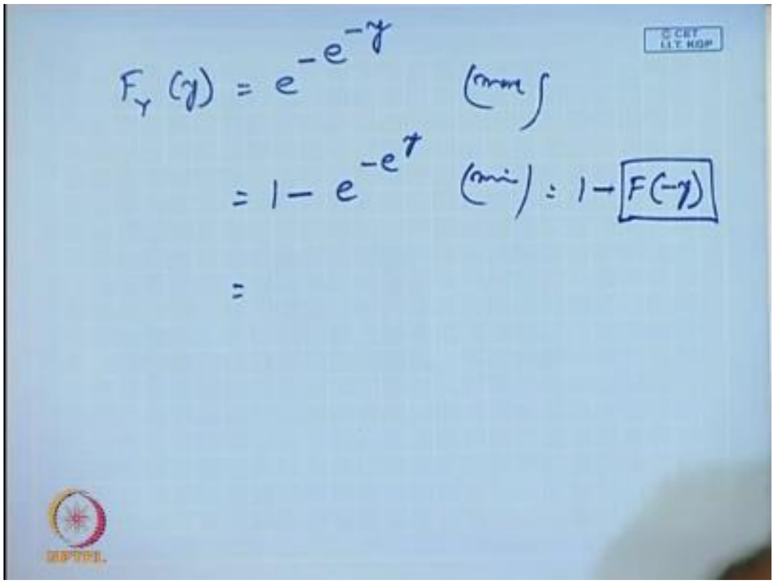
(Refer Slide Time: 34:30)



So, this pdf can be expressed by exponential of this [∓ y- exp(∓y)] again that same, that sign convention, that minus is for the as I told, the minus is for the maximum and the plus is for the minimum. The cumulative distribution function is given by we can do this integration from this -∞ to this specific value of this y and we well get that for the maximum one, it comes at exp[-exp(-y)] and for the minimum one it is 1-exp[-exp(y)]. So, even though so, with the complicate, we started with the complicated or little bit looking cumbersome distribution.

But, whenever we come to this cumulative distribution, it generally becomes simpler and very easy to remember also.

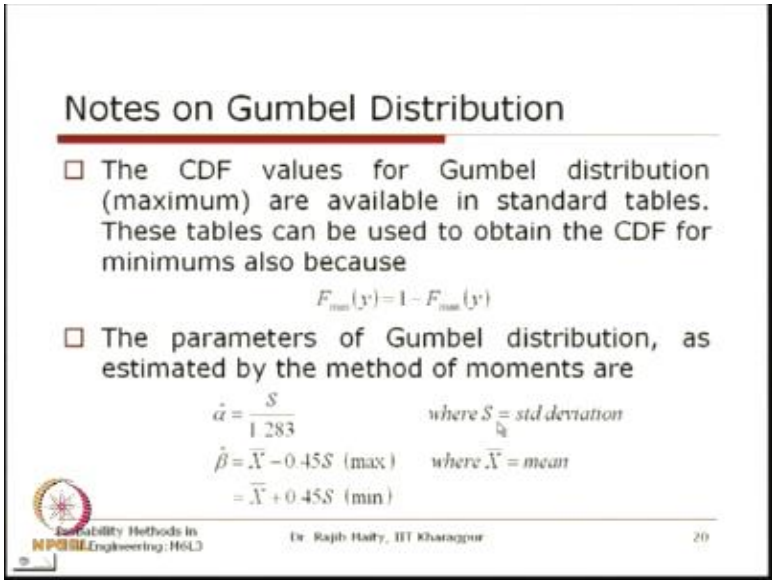
(Refer Slide Time: 35:29)



This means, this is that F_Y of this y is equals to, generally we remember it like this, e^{-e-y} and this is for the maximum. And similarly, $1-e^{-ey}$ this is for the minimum. So, you can see that one, that is, if we just write that this can be for this minimum one, the minimum one can be written as 1-F(-y) so this one is basically for the for the maximum one. So, generally this tables are available for the, for one of this if it is available only for the maximum also, then also you can use the same table for the minimum through this one.

Because this, if I change this attribute to this -y, then that one is subtracted from 1 will give you the required value for this minimum one.

(Refer Slide Time: 36:29)



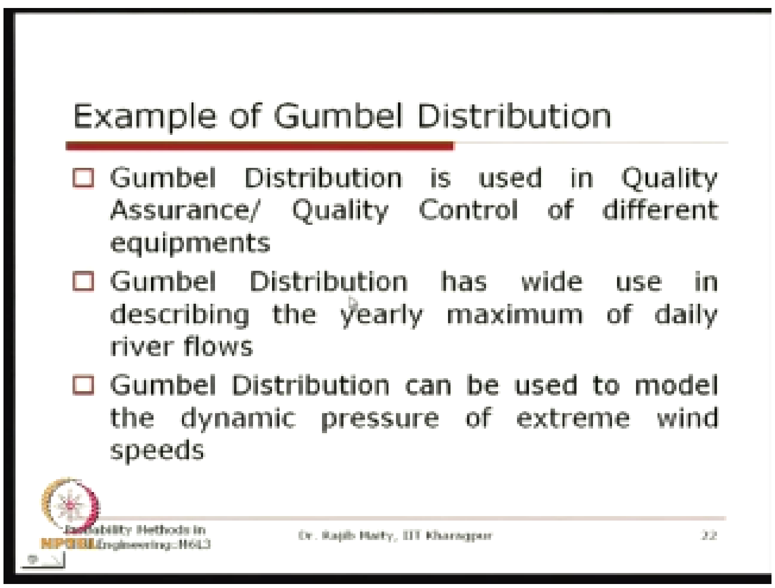
So, we will see one thing that is some notes first that is, if the CDF values for the Gumbel distribution for the maximum are available in the standard tables, this tables can be used to obtain the CDF for the minimum also. Because, as I was telling that F_{min}=F_{max}, this will be -y for this maximum one this will be minus .So, the parameters of the Gumbel distribution as estimated by the method of moments are this αβ. So, from this method of moments how it is so, we have not yet so far in this course, we have not discussed about this parameter estimation is in different methods.

As I was mentioned, that this parameter estimation will be covered in this module 7. So, till that you can just remember that these are one of the method of this parameter estimation which is known as this method of moments. So, we will be discussed these things the different methods of this parameter estimation in the next module. So, using that method of moment method, this α can be that estimated value of these α this is the parameter of this Gumbel distribution is equal to S/1.283 where S is the standard deviation.

This β cap is the $\bar{X}$-0.45S for this maximum and $\bar{X}$+0.45S for the minimum, where this $\bar{X}$is the mean. So, by knowing the, if the data is available, we can calculate what is its sample estimate of this standard deviation, sample estimate of the mean and we can use that some of those s and x bar information to get that parameter α and β. So once we know this parameter α and β, we know what is y, and why because is that x - α / β.

So from there we know what the probabilities are or what are the particular the question that is asked for the mean of this extreme value distribution is given by this β + 0.5772 α for the maximum one and β - 0.5772 α for the minimum one. Variance is given by1.645 α^2 the Coefficient of skewness is a constant value for both this maximum and minimum it is 1.1396. The example of Gumbel distribution, this Gumbel distribution is used in the quality assurance or quality control of different equipments.

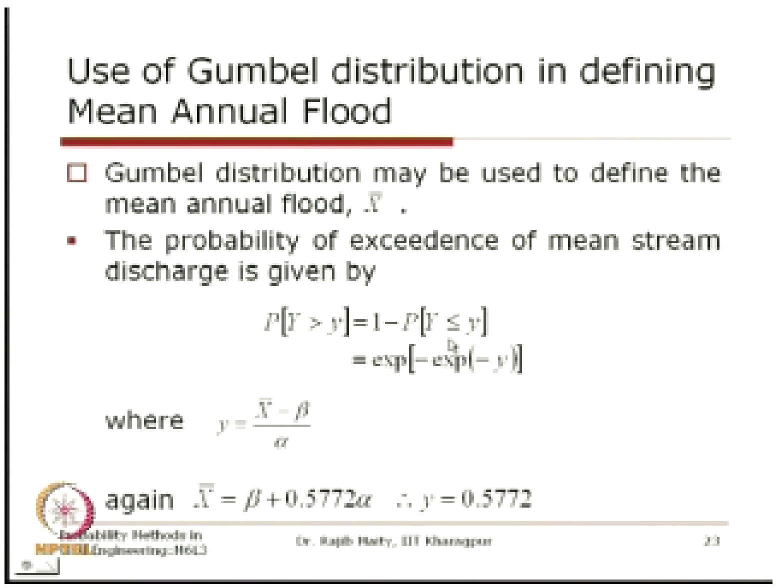
(Refer Slide Time: 39:19)



This Gumbel distribution has a wide use in describing the yearly maximum daily river flows. Gumbel distribution can be used to model the dynamic pressure of the extreme wind speed. So, these things can be always means, whenever as I was mentioning, always whatever the random variable that we are talking about, if we are looking for one of its extreme, then this distribution can be used. These are very; these are the cases where we have seen in this in the civil engineering, where the wide application of this distribution is there.

Now this one example we will see here, that is use of this Gumbel distribution in defining this mean annual flood. So, mean annual flood, the Gumbel distribution we used to define its mean annual flood. So, what is, that is the probability of exceedence of this mean stream discharge is given by this probability of y >y which is basically 1 - probability of y < y.

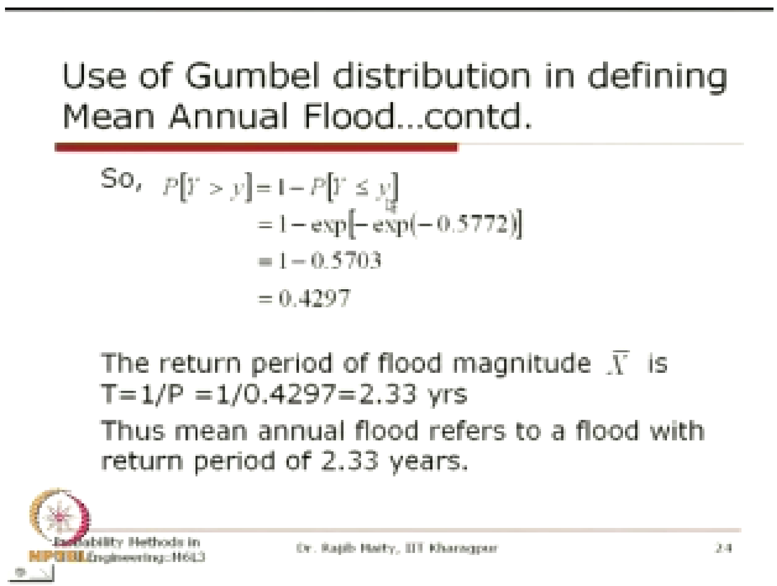
(Refer Slide Time: 40:25)



This one we know that it is one exponential of minus exponential of - y. So this one again we know that this y = x /-β/α. So, basically what we are looking for is that, the mean annual flood. What is it, how can we define that? So from the extreme value distribution so with respect to its return period and all that we will see.

So, this y = x /-β/α. again. So, this we know that x /= β + 0.5772 α. which we have seen from this earlier from the estimate. So, that y = 0.5772. So, whatever this probability, will get the probability of exceedence, that will if you put this value of 0.5772, then we will get what is this that probability. If you put in that expression then, it will come that that probability of y greater than this distribution is equals to 1 - probability of y < y.

(Refer Slide Time: 41:41)

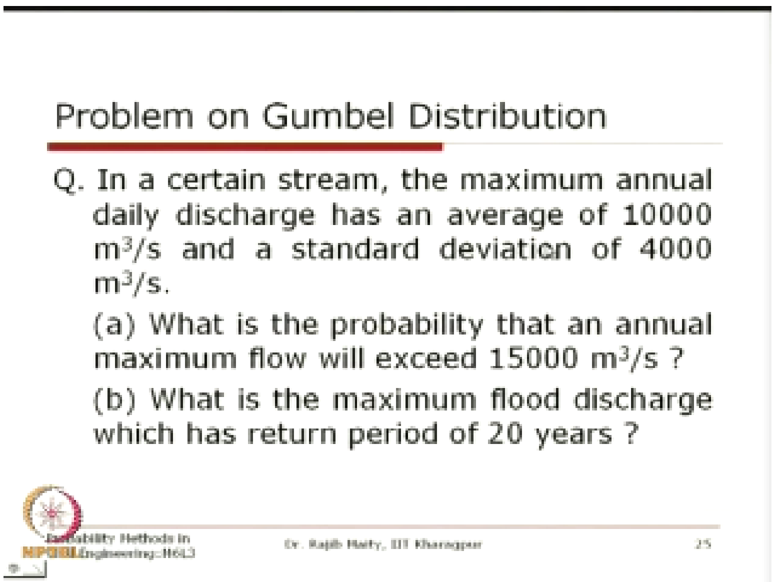


That is the cumulative distribution here. So, that is that exponential of minus exponential of - y and y in that, as we have seen here is the distribution for that0.5772. If you solve this one, it will see that it is 1 -0.5703 which is the 0.4297.Now once you get this probability, what we can define is that, is the is what is its return period. So, the return period we have seen in this in one of this lecture, previous or previous to previous lecture that this is the 1 / p. So 1 by the probability of that particular event.

So if we get this one, we will get the return period of that particular event. That is T = 1 /p. So, 1/0. 4297 is 2.33 years. Thus the mean annual flood refers to a flood with the return period of 2.33 years. So that mean annual flood is having a return period of 2.33 years. So this information is being, is can be used to find out what is the mean annual flood. We will take one more problem on this, because that is just one of these applications that we have shown.

Now, we will take one example on this Gumbel distribution, where we will see some of these answer for the maximum annual daily discharge. So in a certain stream the maximum annual daily discharge has an average of this 10000 meter cube per second and a standard deviation of 4000 meter cube per second. Now, these two information has given. So this can be also, can be estimated from these data also if the data is available.

(Refer Slide Time: 43:53)



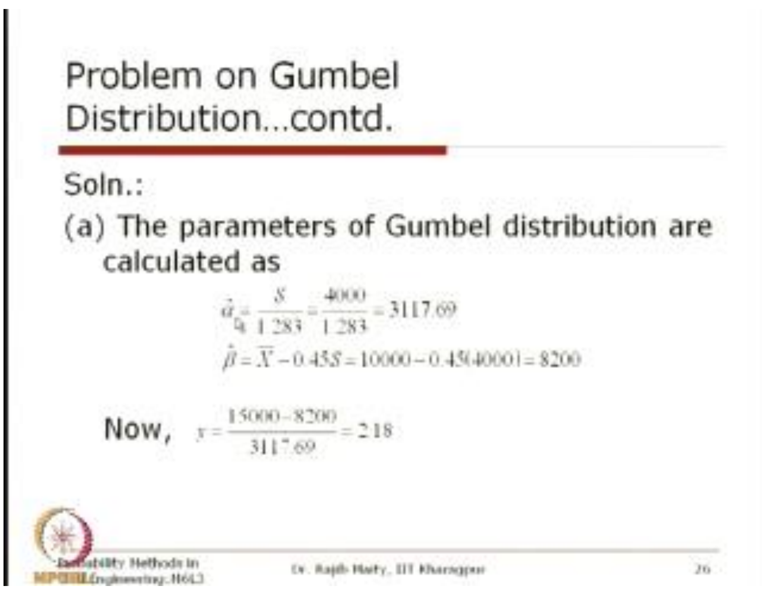
We know this one. Now once we know these two information then basically what we can get is that the parameters of these Gumbel distribution just now we have discussed that. So, those parameters can be estimated. Now we are supposed to answer that what is the probability that an annual maximum flow will exceed that 15000 meter cube per second. Second is the, what is the maximum flood discharge which has a return period of 20 years.

Again that we know that if the return period is given, we can calculate what is this that non exceedence and what is its exceedence probability just by getting is1 / 20 so basically there are structures of these related to the water resource have some specific return period we have to consider and based on that we have to find out what is the magnitude of the flood is coming, and with that value, we have to use that one.

Because, whatever the historical data that are that is available to us, may not reflect that that complete nature. For that reason only, we are, we have to we have to do this exercise to find out what is the maximum possible event or maximum possible magnitude that can occur with that return period. With that value, we have to with this is tremendous useful for this reason purpose and all.

So one such example is shown here, so that two parameter that we are talking about is obtained from the available data that is recorded for that particular sight and from there with through the extreme value distribution and then I will try to find out these answers. So the parameters of these Gumbel distributions can be calculated from this.

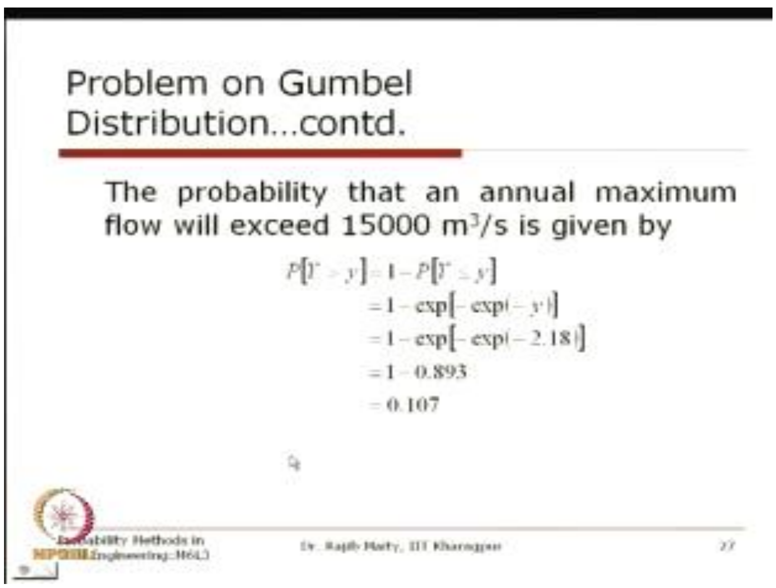
(Refer Slide Time: 45:50)



Their expressions that is α^= s by 1.283, which we have shown just now that is it that method of moments estimated through the method of moments. So, this s is now here shown that 4000 divided by 1.283 which is 3117.69. The β^ is also that a mean minus 0.45 times of this standard deviation s and which is a value of 8200. Now so this, now what we will get we will get this information on the reduced variable. We are supposed to know the question that, what is the probability that it will exceed that15000 meter cube per second?

So, this 15000meter cube per second minus α/β so that should be the reduced variant. That should be the transformation we have to do, before we can get that answers. So from this Gumbel distribution, basically to know, what is that what is the probability of x >15000 what is the probability of y > 2.18. So, this is why we are just transforming this, what is the original data through this x –α/β.

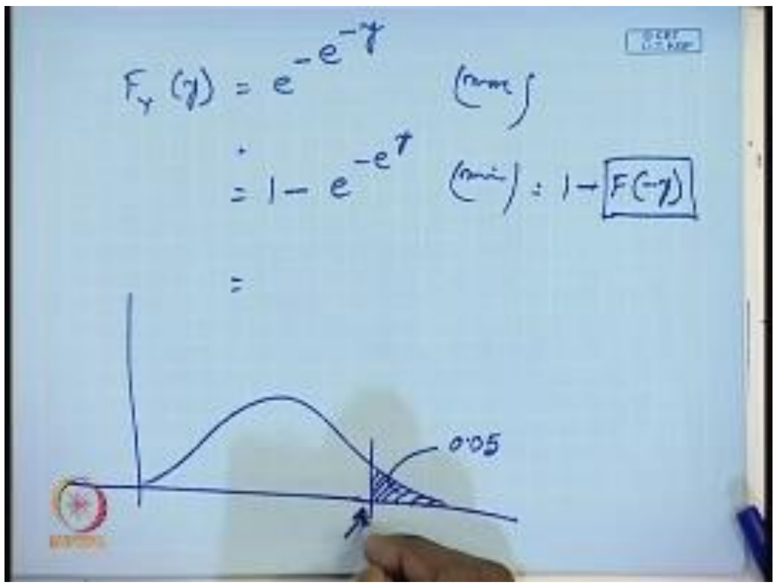
(Refer Slide Time: 47:10)



So the probability that that annual maximum flood will be, will exceed this 15000 meter cube per second is given by the probability that y greater than this value is equal to1 minus this cumulative probability. So this1 minus exponential of minus exponential y and just now we have calculated the y = 2.18. So if we put this one in this expression, then we will get this 1 is 0.893. So, the probability that we will get finally, is 0.107.So, the probability that an annual maximum flow will exceed 15000 meter cube is 10.7percent you can say. The second thing, second question that was looked for is, what is the, that what is the magnitude may here so earlier we have calculated the probability here the return period is given.

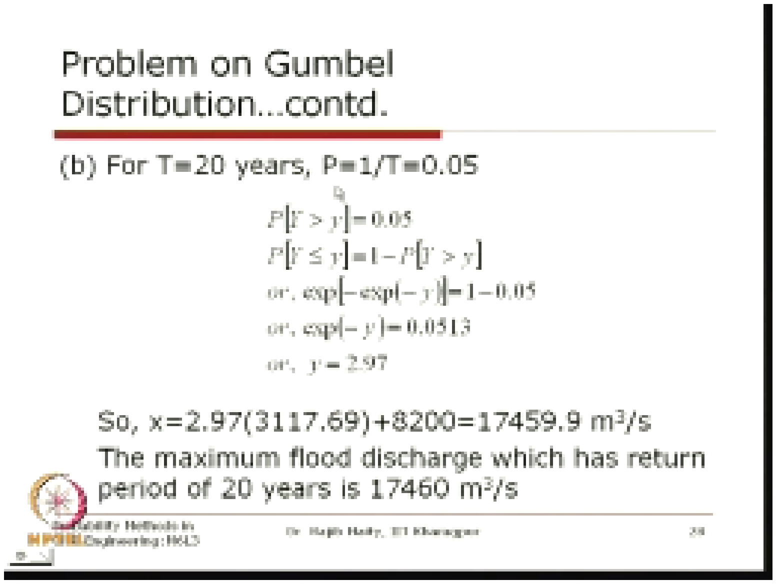
So we are supposed to find out what should be the corresponding magnitude so here the return period 20 years is given first of all we have to find out what is the probability that is so as we have seen in this last problem also that probability is equals to 1/t that is so this 1/t it is the point 0.05 so this is basically that on the right extreme there is if the distribution looks like this so basically we are looking for this particular value higher these one is your point 0.05

(Refer Slide Time: 48:56)



So this particular magnitude want to know from this one so it is called the return period t=20 years now depending on the what is the project. Now depending on this, what is the project that is under consideration, this can change? So, this can change to even 50 years or 100 years or so. So similarly, based on the, what is the return period we are considering, based on that this will this will change. So, once this probability, we get. But, whatever may be this return period, once we get the corresponding probability, then we will be using it from this cumulative distribution and we will get the answer.

(Refer Slide Time: 49:39)



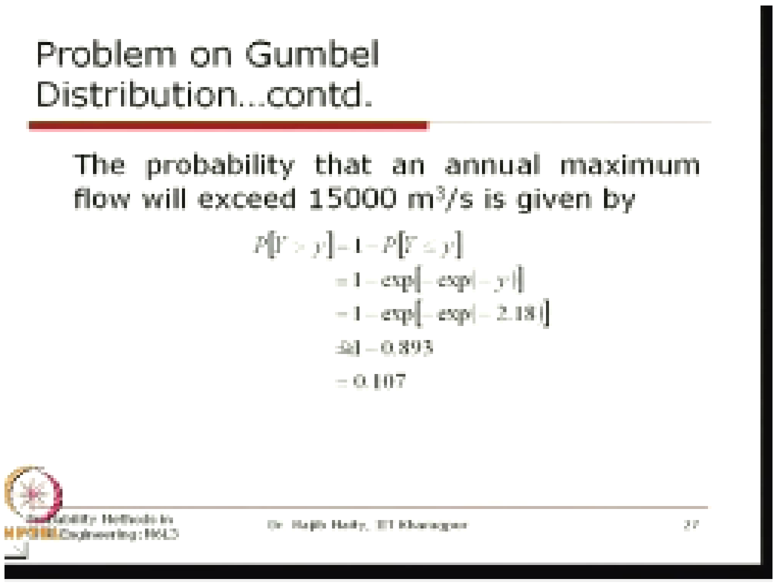
For example, here the probability is your 0.05.So, this probability of y greater than this magnitude that we looking for; obviously, this magnitude first we are looking for in the scale of y. So, that is = to 0.05.So, this probability of y less than = to y is = to 1 - probability of Y greater than this magnitude. So, or what we can write that this exponential of this exponential of exponential of - y which is =to 1 - this 0.05.So, once we do this one and after this, after we solved it for the y finally, we will get the value of this y is = to 2.97.

So, this probability of y greater than 2.97 is basically is this probability 0.5.Now, we have to get it back to the original scale. That is, what is the value of this magnitude of this stream flow, which is now is = 2.97 multiplied by your that 2.97 multiplied by that that value is α or this α plus this β. Basically, this is just inversion of this expression. So this, if we do we will get that 17459.9meter cube per second.

So, we can say that the maximum flood discharge which has a return period of 20 years is = to 17.460M^3 per second. So, similarly, in the previous problem what we have seen is that, that mean flood, that mean annual flood that is having the returned period of 2.33 years.

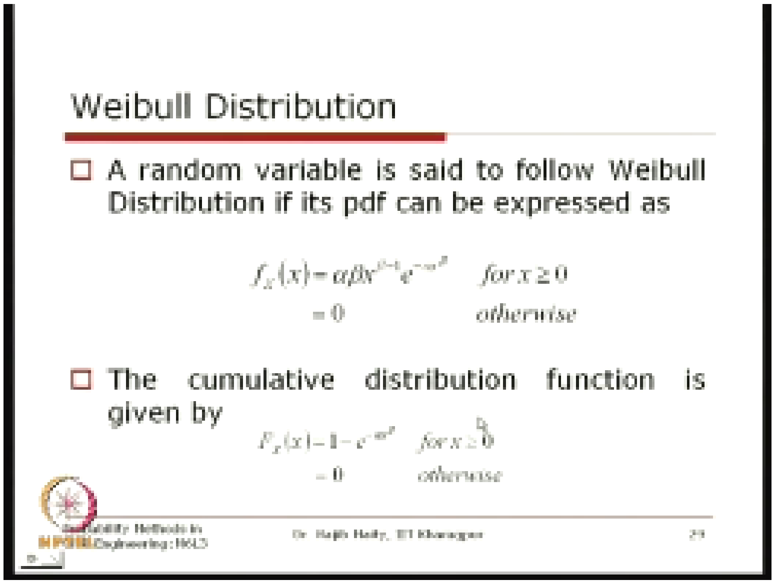
So, if we calculate from this data, if we want to know, if we calculate it from here that, what is the magnitude of the return period? So, magnitude of the return period means you have to follow the first one.

(Refer Slide Time: 51:57)



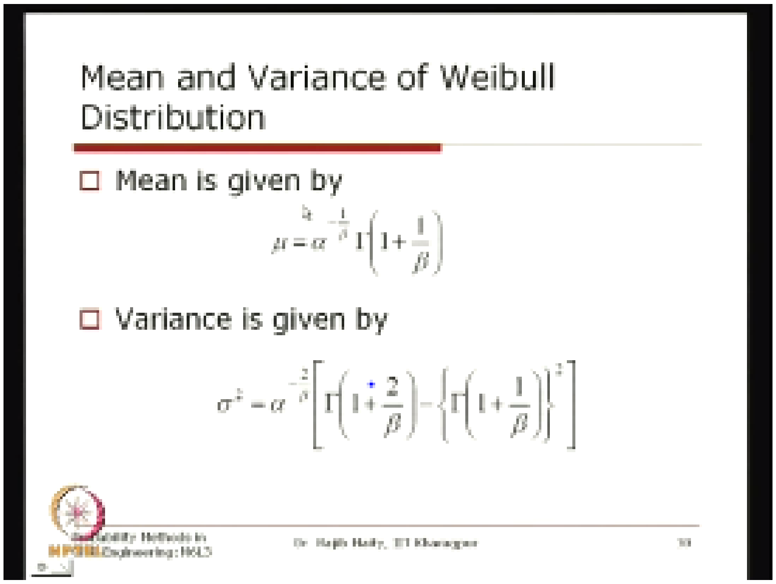
Magnitude of the return period for the for the mean value, then will get that and then will first get the probability, and then we make it inverse and will get that that is 2.33 years.

(Refer Slide Time: 52:13)



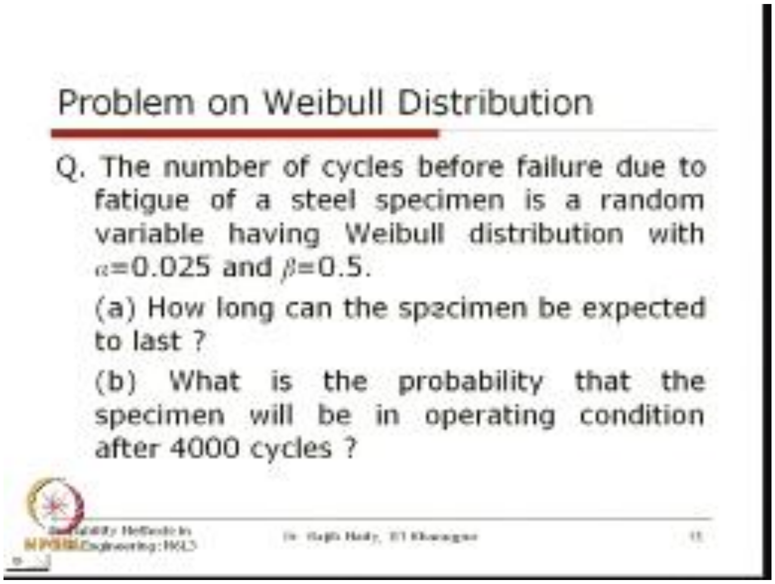
Then, we will take that Weibull distribution. Random variable is said to follow the Weibull distribution, if its pdf can be expressed as this one; that f x x = to α β x power β - 1 and e power - α x power β for the x is greater than = 0. The cumulative distribution function of this Weibull distribution is = distribution to 1 - e power - α x power β. So, we will use this one to get this Weibull distribution and will see some application in this civil engineering.

(Refer Slide Time: 52:56)



So, this mean of this distribution is mean= to, this can be express through its parameter. That is, α power - 1 by β and gamma of this 1 plus 1 by β. Variances also given by this expression sigma square is = to α power - 2 by β gamma of 1 plus 2 by β - gamma 1 plus by β whole square.

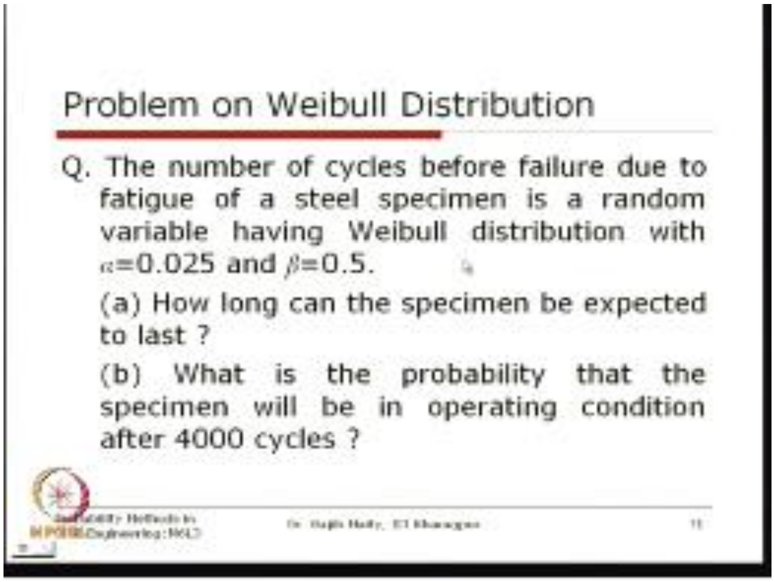
(Refer Slide Time: 53:29)



So, with this one if we just take one example of this of this Weibull distribution, as we was mentioning this is generally used for the failure of this structure. So, one such example is taken here through this Weibull distribution. So, this number of cycles before the failure due to fatigue of a still specimen is random variables having an Weibull distribution with α = to 0.025 and β = to 0.05.Now,I hope that you know this failure due to fatigue means this is there are there are several experiments also can be referred to.

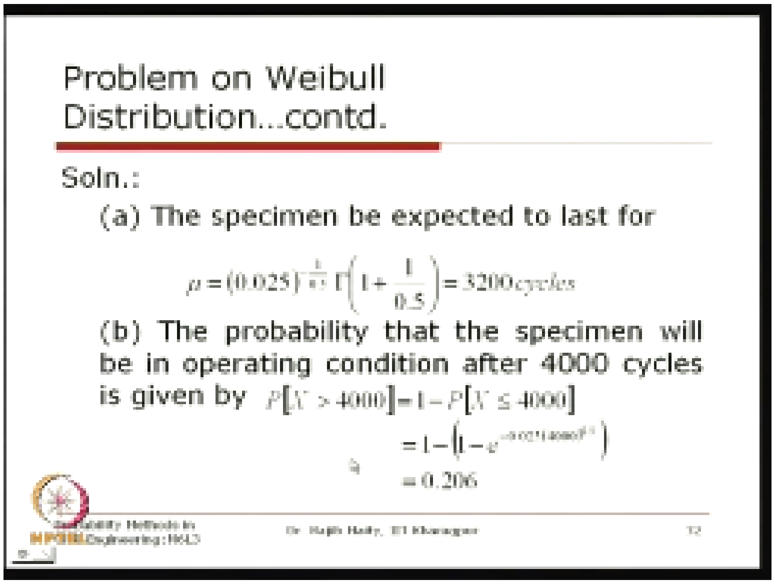
Basically, what does it mean is that it is subject to a reversal of the force? Either it is in the compression, or it is in the it is in the tension with this force pattern is getting reversed and the load is applied which by the by the static analysis, it can be shown that this structure is safe under that load. But, once it is go on for this reversal of force and there is certain cycle. After certain cycle, the structure may fail even though the applied load is within the within the this specified limit. So, like that one. So, how many cycles it can go through that can be modeled through this through this Weibull distribution.

(Refer Slide Time: 54:51)



One such example is shown here, for which the parameter α = to0.25 and β = to 0.05.The question is, how long can the specimen be expected to last and so, the expected value, we have to get and what is the probability that the specimen will be in operating condition after 4000 cycles.

(Refer Slide Time: 55:14)



So, the first question is straight forward which is the mean that we are looking for, and this mean we can expresses through its parameters. So, this mean 0.025 power 1 by0.05 gamma of 1 plus 1 by 0.05.This gamma value we can get from the table and we can calculate this mean, which can which is 300 sorry 3200 cycle. So, any specimen on an average that is expected value of the specimen that can last before failure is that 3200 cycles.

The second question is that probability that the specimen will be in the operating condition after 4000 cycle. Obviously, this probability will be less. So, that can be given by this probability that x greater than 4000, which is the 1 - the probability of x less then= to 4000, which we can get from its cumulative distribution function from this from this 1 - c power - α x power β.

So, if u put that expression here. So, we get that probability is = to 0.0 sorry0.206. So, this is a probability that the specimen will be in operating condition after4000 cycles So, in today's lecture also, we have taken some more probability models gamma distribution, extreme value distribution, Weibull distribution .In that extreme value, we have discussed in detail about this Gumbel distribution and its application to analyze the extreme value, extreme stream for values, the annual maximum and all.

So, we will be taking some more examples because these are all the continuous distribution that we have discussed so far in this module. We will also discuss that discrete random variable, which are also; there are some wide application in this civil engineering difference civil engineering problem and that distribution we will take in the next lectures. Thank you.

Probability methods in civil engineering

End of lecture 33

Next: "Probability Models using Discrete Probability Distributions" in Lecture 34

NPTEL Video Recording Team

NPTEL Web Editing Team

Technical Superintendents

Computer Technicians

A IIT Kharagpur Production

www.nptel.iitm.ac.in