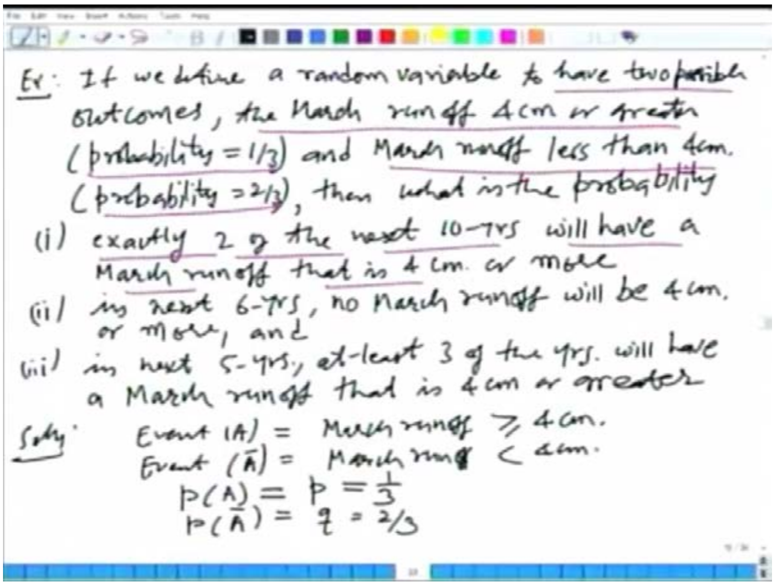
Advanced Hydrology
Prof. Dr. Ashu Jain
Department of Civil Engineering
Indian institute of Technology, Kanpur

Lecture – 34

Good morning and well come to this video course on advanced hydrology, we are into stochastic hydrology module, and we started a new chapter yesterday on frequency analysis. What we did was we looked at some basic concepts about the frequency analysis and that is why it is needed and what are its advantages and limitations? Then we looked at what is called the plotting position? The plotting position method, this is one of the very approximate or simple methods of frequency analysis where in we looked at many formulas you know which can be used to calculate the plotting position which is nothing but the probability of accidence.

Then we moved ahead and define what is called a written period? Once we know the plotting position or the probability of accidence we can take the reciprocal of that that is defined as the written period or the recurrence interval. Then we looked at a very interesting and important discrete probability distribution called the binomial distribution. First we looked at its limitations or the assumptions under which or the criteria under which it can be applied or it can be used. And then we looked at it is basic equation and the couple of coronaries that is where we ended the last lecture. But I am going to be able to do today is get started with the numerical example on the binomial distribution. And then we will look at some other probability distributions which are use for hydrological frequency analysis.

(Refer Slide Time: 01:59)



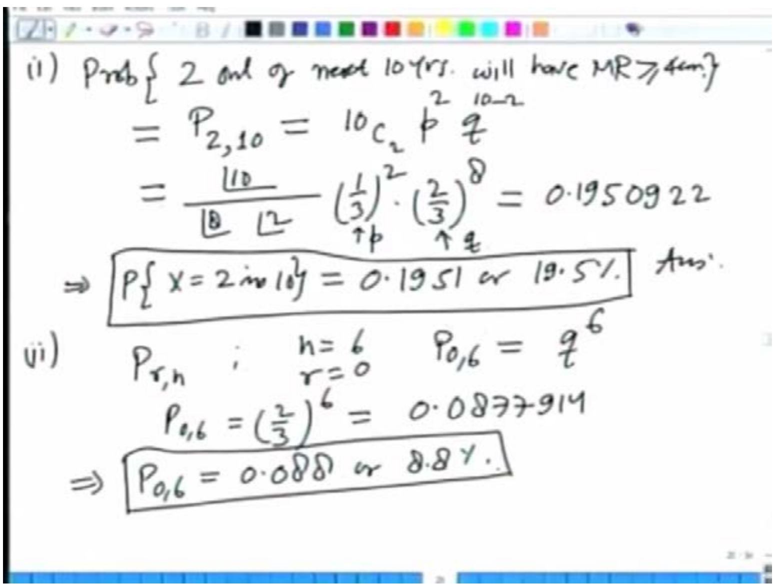
So, the example problem which we are going to look at today goes like this. If we define a random variable to have two possible outcomes, that is the March runoff being 4 centimetre or above or greater with a probability of 1 by 3 and March runoff less than 4 centimetres with a probability what will that be well that will be obviously 2 by 3. Then what we have to answer is what is the probability? That one exactly 2 of the next 2 of the next 10 years will have a March runoff that is 4 centimetre or more that is part one. In the second part you have to find the probability that in next 6 years no March runoff will be, no March runoff will be 4 centimetre or more. And the last one or the third part of this is that in next 5 years, at least 3 of the years will have a March runoff that is 4 centimetre or greater.

In this example, as you see what is given is that the data that is given is that the March runoff is greater than 4 centimetre equal to or more than 4 centimetre. So, that is the event which has been defined, this may be useful for some practical purposes which we are not concerned about right now. But what we are you know going to see is that this is a very special type of situation in which the event has been defined in such a manner that there are only 2 possible outcomes, what are those? The March runoff is greater than or equal to 4 centimetre that is event A let us say or event 1 or one possible outcome. And the other one is what? The March runoff is less than 4 centimetres and the probabilities of those 2 events are given let us say p and q. So, those are given to us. So, we see that the criteria for the application of the binomial distribution are satisfied here.

So, we can apply the binomial distribution if needed. And then we have 1 2 3 different types probabilities in which you see that we have to find the probability of occurrence of this event or times in n trials. All of these can be classified into one of those. So, let us see let us summarise the data which we have as per as the binomial distribution is concerned and then we are going to use it. So, let us look at the solution in which what we have is let us say you have event A as what as the March runoff greater than or equal to 4 centimetre that is event A. And then event B or I would say a bar that is non occurrence of event A is March runoff less than 4. If you see that is how we are defining to have 2 possible outcomes that is the March runoff 4 centimetre of greater with p is equal to this and the March runoff less than this with probability of this.

So, this is given to us. So, what we have is then the probability of event A occurring let us say we call it p this is given to us as what? 1 over 3. Similarly, the probability of event A bar occurring is let us say q which is given to us as 2 by 3. So, this is the data that are given to us. Now, we will look at start looking at what we have to find? In the first one we have to find what exactly 2 out of the next 10 years will have a March runoff that is greater than 4 centimetre, how can we find this? So, what we have to see here is that we have to find a kind of probability in which we have to say that the event A will occur 2 times in next 10 years. So, this is similar to saying an event will occur R times in n trials is not it. So, we can apply the binomial distribution directly here.

(Refer Slide Time: 09:14)



So, let us look here the first one is what we have to find is the probability of 2 out of next 10 years will have I will use you know the short form here the March runoff greater than or equal to 4 centimetres. That is what we have to find is the first part, what will that be? That will be nothing but P 2 comma 10 is not it. So, what is this? This is saying that we have to find the probability of occurrence of the March runoff being greater than equal to 4 centimetres exactly twice in 10 trials or in next 10 years that is what we are saying. That will be as per binomial distribution will be what P subscript 2 comma 10. And as per the formula we can write it as what? It will be 10 C 2 p to the power R, what is R? R is 2 exactly twice and q to the power n minus R n is 10 minus 2. And we can then put the values of p and q and the expand this combination, permutations and combinations.

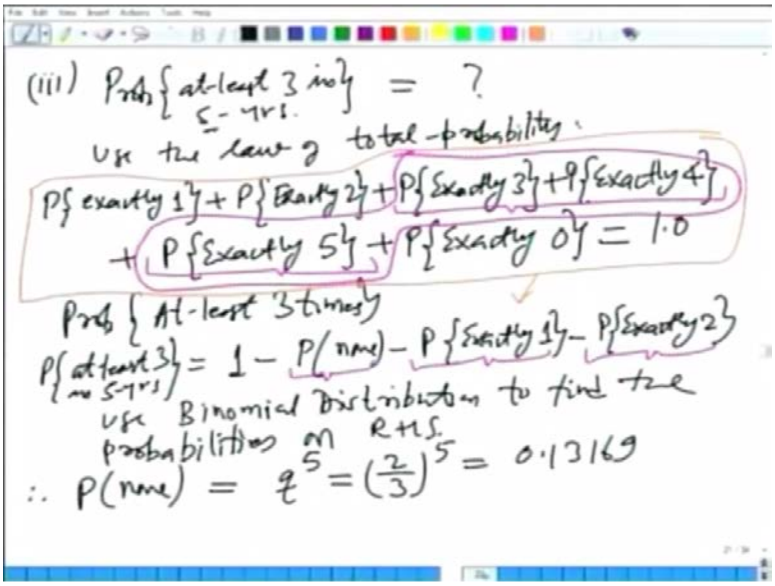
So, this will be n factorial over n minus R factorial into R factorial. So, it will be 10 8 and 2 then 1 by 3 raise to the power what? 2 this is p; this is p. And then multiplied by it will be 2 by 3 raise to the power 8 and this is nothing but your q which we have just written. So, if you simplify that it will be 0.1950922 that is P X is equal to 2 in 10 and ensure that is how I am writing or P 2 comma 10 is equal to what? 0.1951 or approximately 19.5 percent. You see that we are able to find this kind of probability using the binomial distribution which can be extremely useful in the design purposes. Look at the second one, let me go back the second one says what is the probability that in next 6 years? Come here in the next 6 years no March runoff will be 4 centimetre or more, what does that mean in terms of your P r comma n? No March runoff will be greater than or equal to 4 centimetres in n particular years or in 6 years in this case so n is equal to 6.

So, you say that your n is equal to 6. What is r? Then we say no March runoff is greater than or equal to 4 centimetres means r is equal to 0. So, you say no March runoff is more than that particular value that is in 6 years none of the years will have March runoff more than 4 centimetres, this was actually one of the corollaries we have seen yesterday. So, what would that be? P 0 comma 6 will be equal to what will be nothing but your q to the power 6 if you go back and look in your notes that is what it will be. So, you just put the value of q so that your P 0 comma 6 is going to be equal to what? 2 by 3 raise to the power 6 and that will be your answer. And if you simplify it will be 0.0877 and 914 to be more precise.

So, your probability of 0 comma 6 is equal to 0.088 or I would say 8.8 percent it is a very small probability at that event will not occur given the data. So, this is your answer for the second part. Now, let us move on and look at the third one which says what let me go back here this one it says that find the probability that in next 5 years. In next 5 years at least 3 of the years will have a March runoff that is this. So, event A will occur a minimum of 3 times in next 5 years that is what we have to find.

So, this is a slightly different kind of probability in which the at least is involved. And how do we solve that? We have seen that the expression for at least once, what was the expression for at least once, which was one of the corollaries of the binomial distribution. At least one the probability of at least once is equal to what? 1 minus q to the power n, but here we have at least 3 of the next 5 years we will have that particular event A occurring. So, we will do is we will go to the basics and look at our, what is called the law of total probability.

(Refer Slide Time: 15:11)



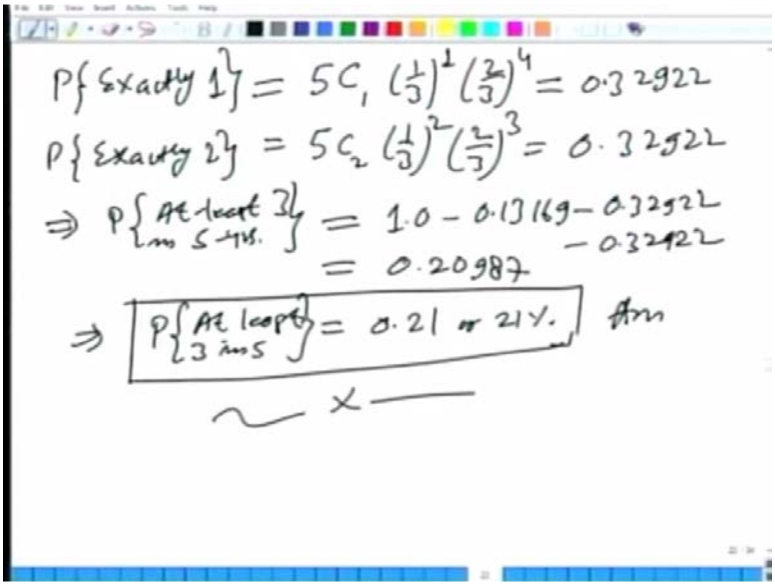
So we have to find the probability of what? At least the event A occurring at least 3 times in 5 years that is equal to what? That is what we have to find out. Now, if we use the law of total probability which we had seen yesterday, how we can write that? We have n is equal to 5 here. So, we can have probability of exactly once the March runoff being grater then equal to 4 centimetre will occur exactly once plus probability of March runoff being exactly twice plus exactly 3 times plus probability of March runoff being equalled or exceeded by 4 centimetre exactly 4 times plus the probability of March runoff being exceeded by 4 centimetre. Exactly 5 times in 5 years anything else yes plus you have probability of non occurrence of that event that is exactly 0 times. These are the all possible combinations or possibilities when you are conducting this trial in 5 successive years.

What can happen is either the event will not occur at all event A that is we are talking event A is what? The March runoff being greater than or equal to 4 centimetre. So, either it will not occur exactly it will be 0 times or it will be exactly once. Once means it occurs ones and in the remaining 4 years it does not occurs. And then there can be many combinations of that, but the binomial distribution will take care of that. Another possibility is twice exactly twice, exactly thrice, exactly 4 times and exactly 5 times. So, all of this, sum of all of this probabilities has to be equal to 1 by the law of total probability. So, if we say that this is all going to be equal to 1.0 then how can you find the probability of this event A occurring at least 3 times? At least 3 times will be equal to what is nothing but the sum of exactly 3 4 and 5 is not it?

So, that is what actually we are talking about this if you sum these things all these this is what? Constitutes at least 3 think about it and see what we are trying to say here. So, that you can say there probability of at least 3 in 5 years is going to be equal to what? 1 minus P of none minus P of exactly 1s minus P of exactly twice by this equation I think you should be able to see this, what I have done is I use this equation, this whole thing to get this. That is at least thrice will be equal to what 1 minus you throw everything else on the other side exactly 0 or none exactly ones and exactly twice.

Now, our only problem that remains is we find these guys what is this? What is this and what is this? So, we do that we have we can us the binomial distribution that is use binomial distribution with the data given to find the probabilities on the right hand side. So that means what will be probability of none? It will be q to the power 5 in 5 years is not it? This is 0 times in 5 successive years or 5 trials which is q to the power 5 or what is q? q is 2 hour 3 to the power of your 5 and that will be 0.13169. Let us move on.

(Refer Slide Time: 20:40)

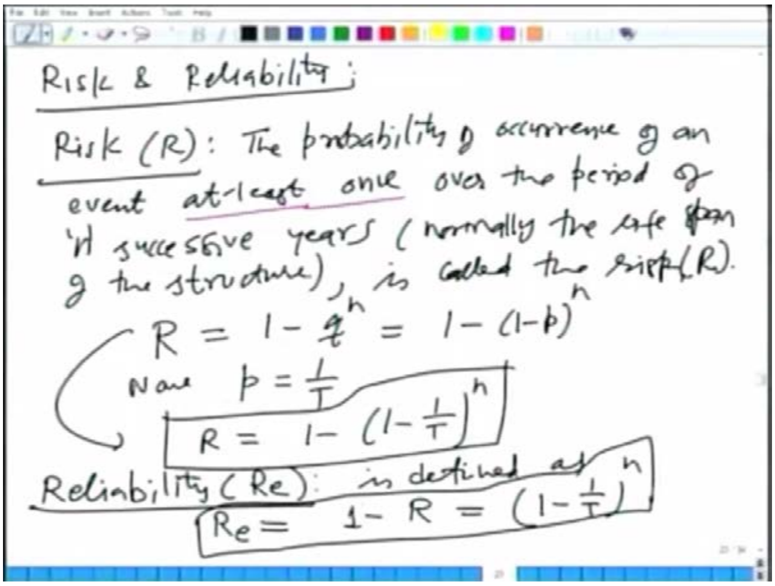


What we have to find next is probability of March runoff being greater than or equal to 4 centimetre. This event occurs exactly once in 5 trials will be equal to what? It will be 5 C 1 then p to the power r, what is r? r is 1 and q to the power n minus r n is 5 r is 1. So, it will be 4. You can simplify that and that should come out 0.32922. And similarly, we can find out exactly twice is going to be equal to what 5 C 2 1 by 3 to the power 2 and 2 by 3 to the power what? 3 and that will be 0.32922. If you see therefore, your probability of at least thrice in 5 trials or 5 years is going to be equal to what? Well 1 minus 0.13169. You just put these values which we have found minus 0.32922 minus 0.32922 that is equal to 0.20987. Therefore, you can say that your required probability of at least 3 times in 5 trials is what? Is 0.21 almost or 21 percent, so that is your answer.

So, you see that this binomial distribution is extremely useful in finding out various kinds of probabilities none once exactly once, exactly twice, exactly you know R times and also at least once, at least twice, at least thrice, those kind of things. And this type of probability is very useful in finding out the risk and reliability and the economic value of a particular project, because during the life span, life span of a hydraulic you know water resources project the risk may be that you do not want the, the design flood to occur. Let us say more than 2 times, the safety of the, a particular structure will get endanger if we know that the flood is going to occur let us say more than twice it can sustain may be 2. So, you want to find out what is that probability?

That will we can translate into the risk associated with that that will translate into the economic values and the funding patterns and planning and everything depends upon all of those things. So with this we will move on and define a few concepts before we move on to the next part in this chapter.

(Refer Slide Time: 24:03)



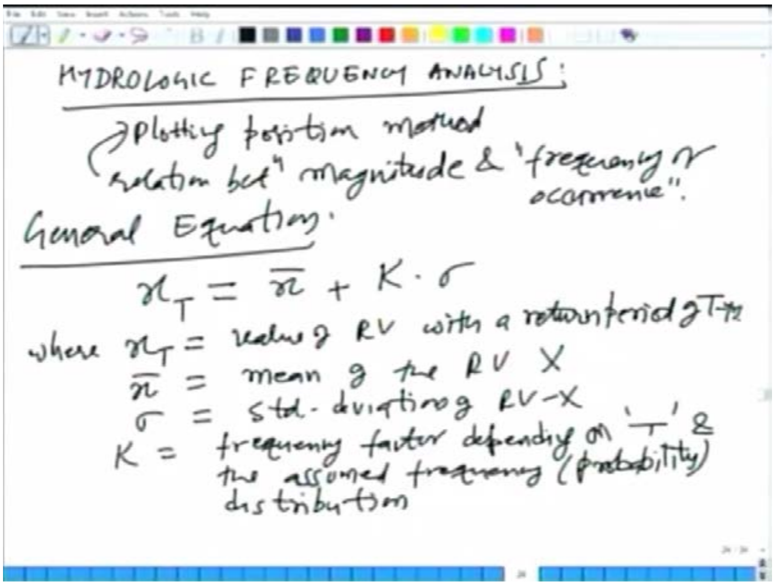
And the next concepts are going to be we define what is the risk and reliability? The risk we are going to denote as let us say R. It is nothing but the probability of occurrence it is nothing but the probability of occurrence of an event at least and I am going to underline this here. We have already seen all these things how we can these probabilities at least how much, how many times? Once, at least once over the period of n successive years over the period of n successive years which is normally the life span of the structure life span of the structure. For example, if you are designing a dam it may be design for a 100 year life or if you are designing a bridge it may be 70 years or so this quantity is called the risk R that is all.

So, what is the risk the way we have defined is that the risk nothing but the probability of occurrence of an event. And that event is your worst case scenario the flood design flood or may be something else at least once during the life span that is the risk associated with the project that even if the flood occurs once or if it is equalled or exceeded once that is the risk we want to find out that probability and we want to minimise that. How can we minimise that risk associated with that project? That is a separate topic we are not going to go into that. But this is how the risk is actually defined the occurrence of the flood or an event at least once. And we know how we can determine it.

So, if we write the expression for R it is nothing but what? It is 1 minus q to the power n, q is what? q is the non occurrence and n is the a life span or the number of trials or the number of years. What is q? Well it is 1 minus P as we have seen, because there are only 2 possible outcomes whether the flood occurs or it does not occur. So, it is 1 minus P to the power n. Now, what is P? Or P is equal to in terms of written period as we have seen the reciprocal of this is nothing but the written period, so P is going to be 1 by T. So, if we use this expression then you will have R is equal to 1 minus 1 minus 1 over T raise to the power what? n. So, what is the risk associated with a project? Well it will be at direct function of the written period for which it is designed. Let us say if you are designing a dam for a 100 year event then you can calculate the risk associated using this formula or similarly, if you are designing a bridge for a 50 year event or 50 year flood then you can find out the discuss associated with that using this.

This is about the risk and then we have another thing which is called the reliability. And this we will denote as R e. If the risk associated with the project is R then the reliability is simply defined as its compliment or you say R e is equal to 1 minus R. So, what will be 1 minus R 1 minus of this whole thing? That will be 1 minus one over T to the power n, so this is your reliability of the project. So, if the risk associated with a project is 10 percent, what is the reliability? Reliability is 90 percent, because the sum has to be equal to 100 percent. So, they complement each other so risk and reliabilities are defined in such a manner then 1 is the other 1is 1 minus of the other.

(Refer Slide Time: 29.10)



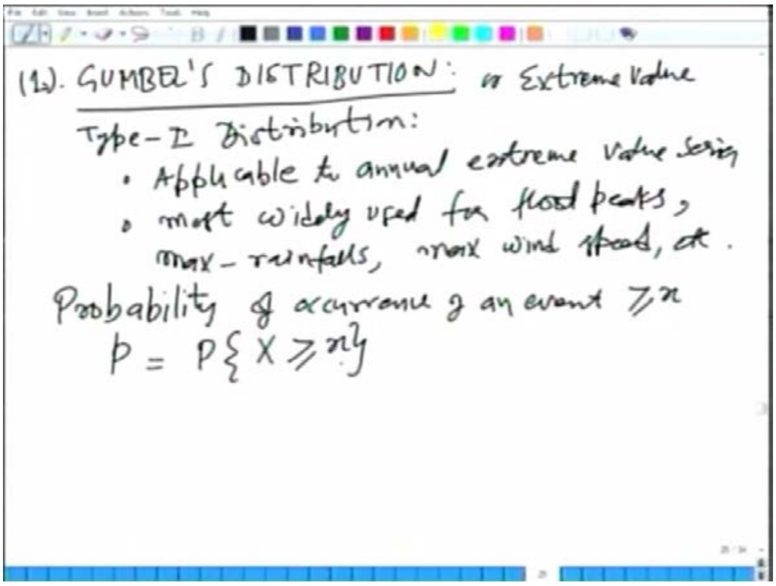
So, with these concepts what we do is let us move on to what is called the hydrologic frequency analysis, hydrologic frequency analysis. We have already seen what is called the plotting position method is not it? So, plotting position method is nothing but a crude way of establishing a relationship between the magnitude and the frequency of analysis, what do we do in frequency analysis? We establish a relation between the magnitude and what? And the frequency of occurrence of a particular event, and plotting position method is one of those which we have already seen.

Now, what we are going to be able to do is look at a general frame work of this frequency analysis. So, we will write down a general equation and then we will look at some specifics it says that your x T is equal to x bar plus K times of your sigma. So, this is the general frequency analysis equation where what are these quantities? Let us define these, what is x T? Well x T is the value of your random variable with or corresponding to a written period of T x. So, the objective as we know is what is to determine the magnitude corresponding to a particular frequency of occurrence and the frequency of occurrence we are representing in terms of written period.

So, the objective is to find out the value of x T, x T is the value of your random variable corresponding to a written period of T s. The next one is x bar which all of you know is the mean or the average of the random variable X, Good morning and welcome to this video course on advanced hydrology. We are into stochastic hydrology module. And we started a new chapter yesterday on frequency analysis. What we did was we looked at some basic concepts about the frequency analysis and that is why it is needed. And what are it is advantage of your random variable X. And K is something important which is defined as the frequency factor, what is K is the? K is the frequency factor depending on frequency factor depending on the T and assumed frequency or probability distribution. So, what is K here is that? K is a factor or it is called the frequency factor which will depend upon what? It will it will be function of a minimum of two things.

What are those? One is the written period T itself and the other one is the assumed probability distribution or the probability distribution that particular hydrologic random variable is suppose to follow. So, whenever we are carrying all this frequency analysis it is irrespective to a, a specific probability distribution, we have seen the example of binomial distribution earlier. But now, we are what we are going to do is we will look at you know 2 3 different methods of fitting this frequency analysis or determining the value of K depending upon the different probability distribution. So, this K as we said is the frequency factor that depends upon T and what and the assumed frequency or the probability distribution. So, let us look at a few probably distributions which are very popular in hydrology.

(Refer Slide Time: 34:13)



And the first one of them we will see a, what is called the Gambel's distribution. I am sure you may have seen this earlier in your UG classless Gambel's distribution. This is also called the extreme value, extreme value type 1 distribution EVT 1 famously known as distribution. As the name suggests it is applicable or useful for the extreme values it is applicable to annual extreme value series well I am writing here annual, but it does not have to be. But mostly these are the annual extreme events, we apply it to or the other thing about this distribution is that it is most widely used, most widely used for flood peaks or the peak discharges and maximum amount of rainfall for a particular duration maximum wind speed, etcetera.

What this distribution gives us is the probability of occurrence of an event, such that it is magnitude is greater than or equal to x that is your p is equal to how it is define? As we have seen X random variable is X such that it is magnitude is equal or exceeded greater than equal to x. So, small x here is just a number and capital X here represents what? A capital X represents the random variable. And this is as per the Gambel's distribution given by 1 minus e to the power minus e to the power minus of y subscript T where y subscript T is what? It is a dimension less reduced variant.

So, an intermediate variable actually which is defined like this, this is given as 1.2825 times x T minus x bar divided by sigma plus 0.577 this is your y T or the reduced variant. And this equation actually is applicable for population where the sample data set is very large or infinite where this number actually is under the limits we will define little later that this is nothing but we will define as as subscript n and this number you see here this actually we will define later as y n bar. But these are the numbers for the case when the number of data points is a very large where x T is what? x T is the flood magnitude with a probability of...