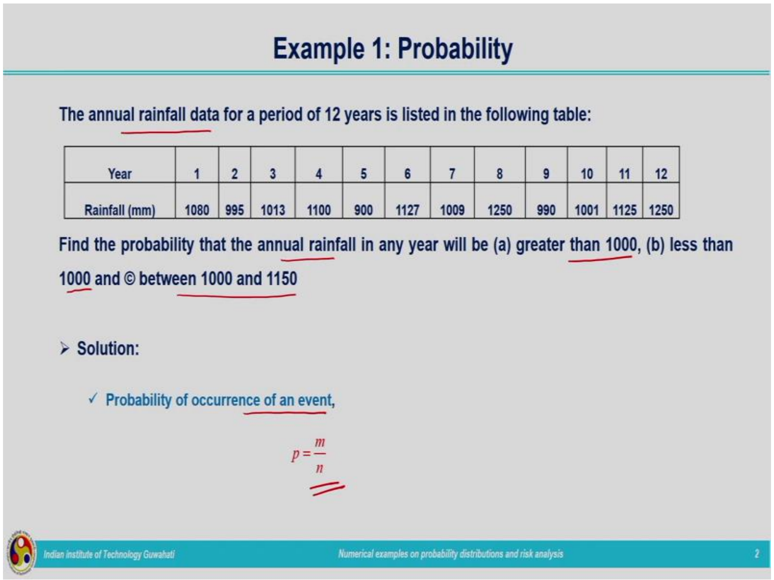
Engineering Hydrology
Dr. Sreeja Pekkat
Department of Civil Engineering
Indian Institute of Technology, Guwahati
Lecture 79
Numerical Examples

Hello, all. Welcome back. In the previous lecture, we were discussing about hydrologic design. We have stopped with the topic of risk analysis. Today, we need to solve some numerical examples on risk analysis. In addition to that, we have to solve some of the examples on probability distributions also.

These different types of distributions we have studied in the previous module. That time I did not work out the examples related to different probability distributions, because I wanted to give you some idea about the return period, probability of exceedance all those things. So, today we will combine the numerical examples on probability distribution and also risk analysis.

(Refer Slide Time: 1:30)



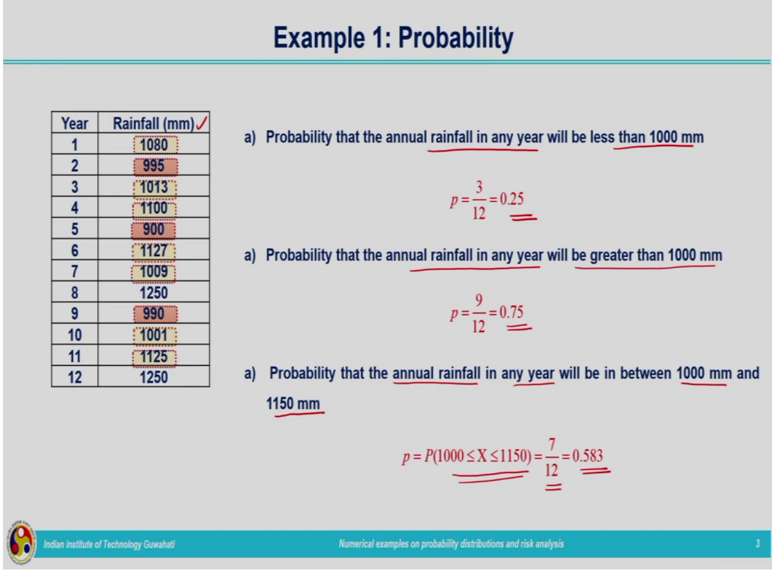
Let us start with solving the examples. First example is on probability. The annual rainfall data for a period of 12 years is listed in the following table. 12 years annual rainfall data is given to you. Find the probability that the annual rainfall in any year will be greater than 1000 millimeter, less than 1000 millimeter, and between 1000 and 1150 millimeters.

We need to estimate the probability related to different conditions. One is related to annual precipitation or the annual rainfall less than 1000 millimeter, second one is greater than 1000 millimeter and the third one is the rainfall value between 1000 and 1150 millimeters, here we do not have any distribution and any other details. So, we will make use of the probability expression by number of events divided by total number of data in the sample. The probability of occurrence of an event is nothing but p is equal to

$$p = \frac{m}{n}$$

m is the number of events which are satisfying the conditions and small n is the total number of sample data, here in this case it is 12.

(Refer Slide Time: 2:57)



These are the data corresponding to annual rainfall, annual rainfall data in millimeters for 12 successive years. We need to find out probability that the annual rainfall in any year will be less than 1000 millimeters. So, here we can count how many events are there representing the value less than 1000 millimeters. So, here we are having corresponding to second year it is 995, fifth year it is 900 and ninth year 990. These are the three events which are less than the given condition. So, small m is 3 and small n is 12. So, we can calculate the probability by using the formula

$$p = \frac{m}{n} = \frac{3}{12} = 0.25$$

Now, the second part is to find out the probability for the events corresponding to the value greater than 1000 millimeters, probability that the annual rainfall in any year will be greater than 1000 millimeter. So, here three events are less than 1000 millimeter and we are having total events. So, remaining events will be greater than 1000 millimeters. Here in the data points, we do not have any data corresponding to rainfall is equal to 1000 millimeter, if that case is there that we have to exclude if we are calculating the probability of occurrence of events greater than 1000 millimeters. So, here in this case that is not there. So, remaining events which we have calculated i.e., three events with rainfall less than 1000 millimeter and remaining 9 events will be representing a value greater than 1000 millimeters. So, we can calculate the probability as

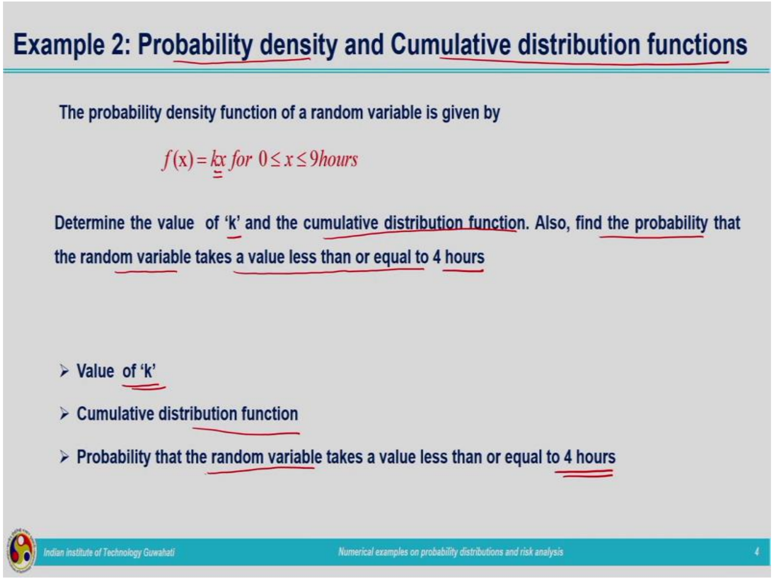
$$p = \frac{9}{12} = 0.75$$

So, if you add these two probabilities, you can and get equal to 1. Now, the third question is that probability that the annual rainfall in any year will be in between 1000 millimeters and 1150 millimeters. We need to count the number of events coming within this range and then we need to calculate the probability. So, if we are checking the values, we can understand that the annual rainfall between 1000 millimeter and 1150 millimeters we are having corresponding to first year, third year, fourth year then come sixth year, seventh year, tenth year and eleventh year. So, in total we are having seven events which are coming within this range given that is 1000 millimeters and 1150 millimeters. So, the corresponding probability is given by

$$p = P(1000 \le \mathrm{X} \le 1150) = \frac{7}{12} = 0.583$$

This is very simple problem to calculate the probability corresponding to different conditions given.

(Refer Slide Time: 6:21)



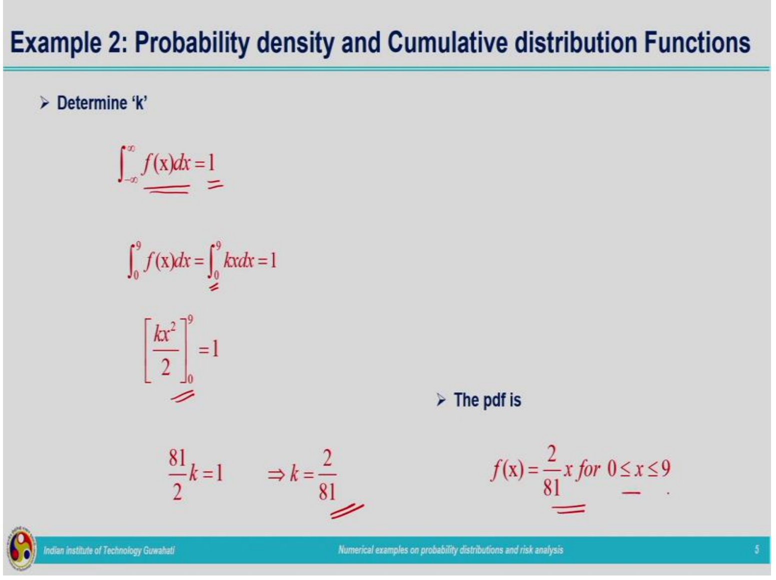
Now, we will move on to second example. Second example is related to probability density and cumulative distribution functions. The probability density function of a random variable is given by

$$f(x) = kx \ for \ 0 \le x \le 9 hours$$

Determine the value of *k* and the cumulative distribution function. Also find the probability that the random variable takes the value less than or equal to 4 hours.

So, here you can understand that we have been given with the probability density function, it is representing the continuous random variable. *f(x)* that is PDF is given by *kx* and the range is also given to us. We need to determine the value corresponding to *k*, cumulative distribution function and also probability corresponding to a particular condition given in the question that is *x* less than or equal to 4 hours. We need to calculate the value corresponding to *k*, we need to find out the cumulative distribution function and third part of the question is to find out the probability that the random variable takes a value less than or equal to 4 hours. So, we will start solving the example.

(Refer Slide Time: 7:42)



First, we need to determine the *k* value, probability density function is

$$f(x) = kx$$

How can we find out the value corresponding to *k*? From the theory of probability, we know that

$$\int_{-\infty}^{\infty} f(x)dx = 1$$

This is corresponding to continuous random variable, if we integrate the probability density function within the limits $(-\infty)$ to $(+\infty)$ it should be equal to 1 that is representing the law of total probability. So, here we can substitute $f(x) = kx$ i.e.,

$$\int_0^9 f(x)dx = \int_0^9 kxdx = 1$$

and we can integrate within the range 0 to 9. It is given in the question that the PDF which is given that is $f(x) = kx$ is valid for a value corresponding to *x* within the range of 0 to 9. So, this is a simple function, we can integrate within the limit 0 to 9 it will be taking the

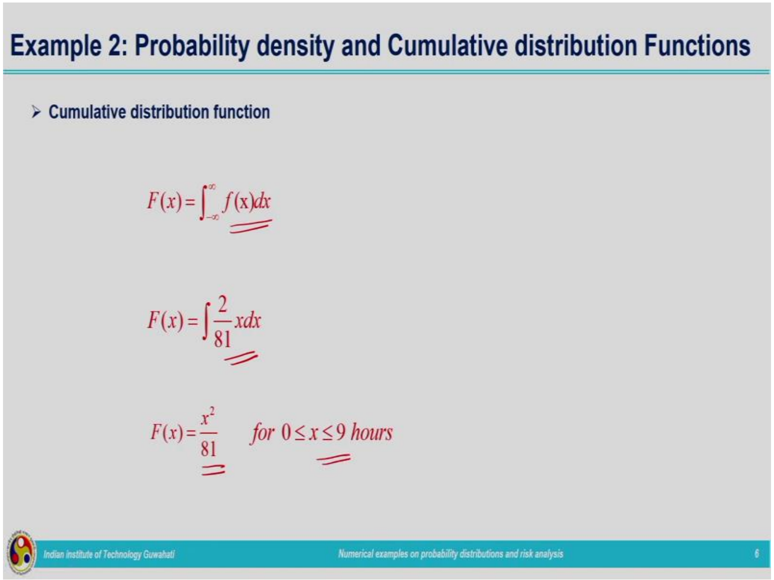
$$\left[\frac{kx^2}{2}\right]_0^9 = 1$$

Then we can substitute the limits 9 and 0, it will be giving us the value of $k = \frac{2}{81}$. The value of k is determined from the probability density function. So, our *f(x)* will be taking the form,

$$f(x) = \frac{2}{81} x \; for \; 0 \le x \le 9$$

This is our probability density function. Once the probability density function is there with us, we can find out the cumulative distribution function that is the cumulative distribution function we can obtain by integrating the probability density function within the respective range and if the cumulative distribution function is given to you by differentiating that will get the PDF. So, here PDF we have found out we can find out this cumulative distribution function.

(Refer Slide Time: 9:44)



F(x) is given by

$$F(x) = \int_{-\infty}^{\infty} f(x)dx$$

Now, we can make use of our PDF here, that is capital *F(x)* CDF is given by

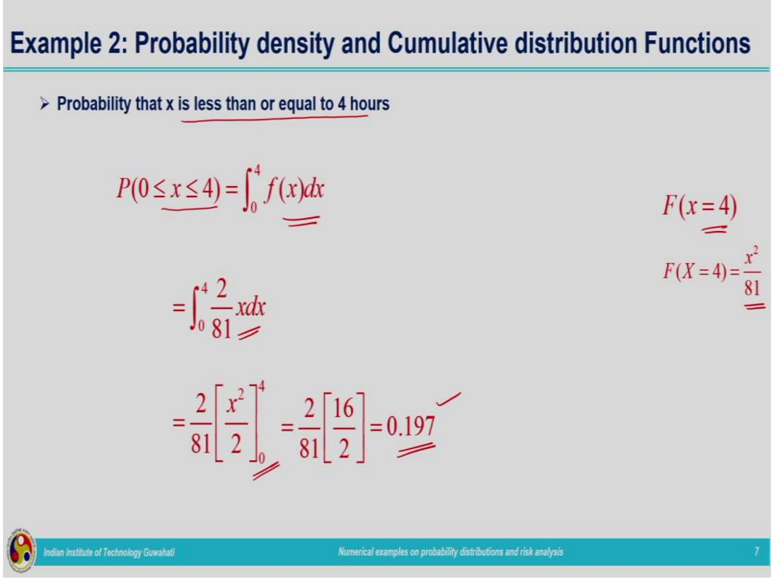
$$F(x) = \int \frac{2}{81} xdx$$

When we integrate this, we will get

$$F(x) = \frac{x^2}{81} \; for \; 0 \le x \le 9 \; hours$$

You can substitute the values 0 and 9 here and get the value corresponding to that it will be coming out to be capital *F(x)* equal to 1. Corresponding to in between values, that is *F(x)* less than or equal to 4, less than or equal to 5, that way if you want to calculate, you keep the function as such, corresponding values of *x* can be substituted in this, and its answer can be found out. So, here I am keeping the function as such, CDF $F(x) = \frac{x^2}{81}$. This is the way in which we will be finding out the cumulative distribution function from probability density function.

(Refer Slide Time: 10:53)



Now, one more part is left that is the probability corresponding to an *x* value less than or equal to 4 hours. So, that is given by

$$P(0 \le x \le 4) = \int_0^4 f(x)dx$$

So, we can get the value corresponding to it by $\int_0^4 f(x)dx$. This is by making use of the probability density function, we can compute it by making use of cumulative distribution function also. So, when we substitute for $f(x) = \frac{2}{81}x$, integrated within the limit 0 and 4, i.e.,

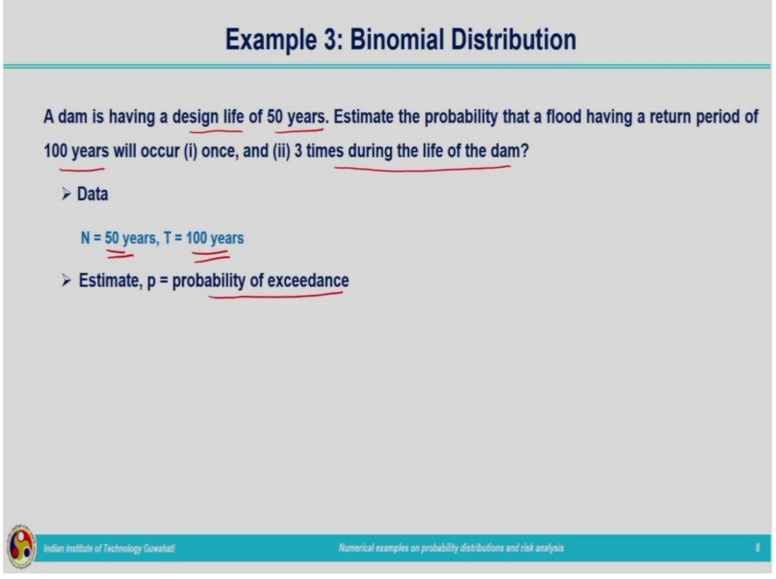
$$= \int_0^4 \frac{2}{81} x dx$$

So,

$$= \frac{2}{81}\left[\frac{x^2}{2}\right]_0^4 = \frac{2}{81}\left[\frac{16}{2}\right] = 0.197$$

This is calculated by making use of the probability density function. Instead of PDF, we can make use of CDF also we can find out the value that is $F(X = 4)$. So, if you are substituting the CDF, CDF is given by $F(X = 4) = \frac{x^2}{81}$, x is 4, $\left[\frac{4^2}{2}\right]$, that is $\frac{16}{81}$. So, we will get the same answer as that of 0.197. So, this is a simple example related to PDF and CDF.

(Refer Slide Time: 12:26)

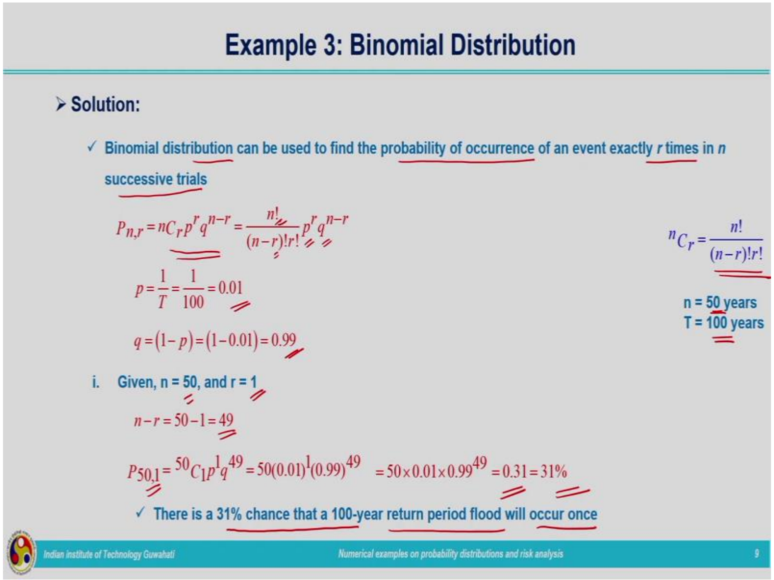


Now, let us move on to an example related to binomial distribution. We have seen different types of distribution corresponding to continuous random variable and discrete random variables. When we were discussing about discrete random variables, we have looked into two types of probability mass function. In the case of discrete random variables, we were using the terminology PMF, it was not PDF. So, here we are going to solve one example related to binomial distribution.

A dam is having a design life of 50 years. Estimate the probability that a flood having a return period of 100 years will occur once and 3 times during the life of the dam. So, here we are having a dam which is having a design life for 50 years, we need to estimate the probability that a flood having a return period of 100 years, which may occur once in this design period, and second part is which may occur thrice in this design period. This problem can be solved by means of binomial distribution, because this is related to discrete random variable. I did not solve this at the time when I was discussing about binomial distribution, because, you were not familiar with the term of design period of a structure that we have completed in the previous lecture only, that is why I have kept this question to be solved after completing that particular topic.

So, a dam is having a design life for 50 years, you know, what is meant by design life and we need to compute the probability that a flood having return period of 100 years will occur once and three times during the life of the dam. So, here we are having capital N is equal to 50 years and return period is equal to 100 years. 100-year return period flood we are considering and the probability corresponding to that will be $p = \frac{1}{T}$, these two are inversely related. So, p can be calculated by making use of $\frac{1}{T}$. So, N is given that is design life is 50 years and T is equal to 100 years. Estimate probability of exceedance? We need to calculate small p that is the probability of exceedance.

(Refer Slide Time: 15:02)



We can make use of binomial distribution to find the probability of occurrence because the event exactly occurring r times in n successive trials. The number of trials we can consider is 50, because the design life period is 50. So, if we are considering every year so, we are having 50 number of times and out of that *r* times the flood will be occurring. So, in such cases, we will be making use of binomial distribution. Binomial distribution is given by

$$P_{n,r} = nC_r p^r q^{n-r} = \frac{n!}{(n-r)!r!} p^r q^{n-r}$$

Where,

$$nC_r = \frac{n!}{(n-r)!r!}$$

So, here we need to identify different terms, *n* is given as 50 years and *T* is given as 100 years, i.e., design period is 50 years, return period is 100 years. So, here we can substitute the values corresponding to *n* and *r* after finding out the values corresponding to *p* and *q*. So,

$$p = \frac{1}{T} = \frac{1}{100} = 0.01$$

We know probability of exceedance is nothing but 1 by return period, return period is 100. So, p can be calculated as 0.01 and what is q?

$$q = (1-p) = (1-0.01) = 0.99$$

Now, we can substitute in the binomial distribution for the first condition *n* is 50 and *r* is equal to 1 that is the flood is occurring once, now, we can calculate

$$n - r = 50 - 1 = 49$$

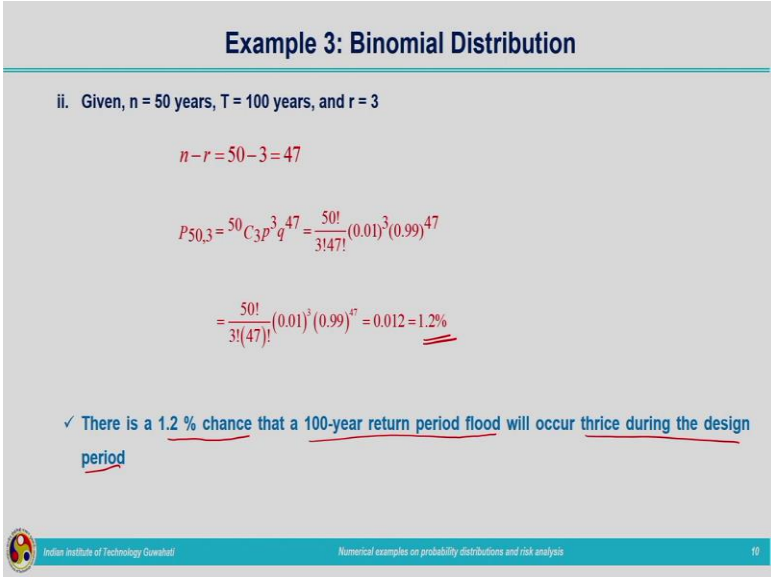
Now, you can calculate the value corresponding to

$$P_{50,1} = 50C_1 p^1 q^{49} = 50 \times (0.01)^1 \times (0.99)^{49} = 50 \times 0.01 \times (0.99)^{49} = 0.31 = 31\%$$

So, the probability of occurrence of the flood exactly once in the design life period of the dam is calculated by means of binomial distribution that was found out to be 31% i.e., there is a 31% chance that 100-year return period flood will occur once in the lifespan of the dam.

Now, we can move on to the second part of the question that is the probability of occurrence of a flood thrice in the lifespan of the dam. So, here in the second part *r* is equal to 3, there is no change as far as *n* and *T* are concerned. Design period *n* is equal to 50 and the return period T is equal to 100. So, there will not be any change taking place in the values corresponding to p and q, but only difference is there in the value of r, r is equal to 3.

(Refer Slide Time: 17:53)



So, n is equal to 50 years, T is equal to 100 years and r is equal to 3 times,

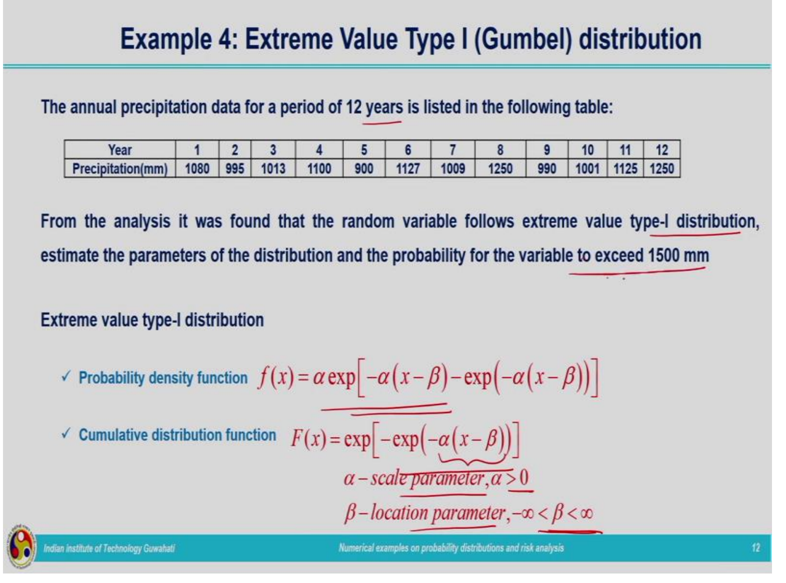
$$n - r = 50 - 3 = 47$$

So, we can substitute in the binomial distribution expression

$$P_{50,3} = 50C_3 p^3 q^{47} = \frac{50!}{3!(47)!} \times (0.01)^3 \times (0.99)^{47} = 0.012 = 12\%$$

We can calculate the value as 1.2% i.e., occurrence of the flood thrice within that design period is lesser that is 1.2% e. There is a 1.2% chance that 100-year return period flood will occur thrice during the design period of the dam. If it is once the probability of occurrences chance is more and in the case of thrice occurring case it is 1.2%. So, this is related to binomial distribution. Similar to the binomial distribution, we can solve other distributions such as Poisson distribution, gamma distribution, etcetera.

(Refer Slide Time: 18:59)



Now, let me move on to the extreme value distribution that is Gumbel's distribution. Extreme value type I is termed as Gumbel distribution. The annual precipitation data for a period of 12 years is listed in the following table. This is the same data which we have used in the first problem. From the analysis it was found that the random variable follows extreme value type I distribution. Estimate the parameters of the distribution and the probability for the variable to exceed 1500 millimeters.

So, here the data corresponding to annual precipitation is given to us. We are not going to check whether it is following extreme value distribution or not, but it is given in the question that it is following extreme value type I distribution that is the Gumbel's distribution. We need to find out the parameters of the distribution and also probability for the variable to exceed 1500 millimeters.

So, extreme value type I distribution we need to have idea about the PDF and CDF. We have already made use of these distribution for solving one numerical example on frequency distribution. So, here let us re-look into the distribution function, probability density function is given by this particular expression

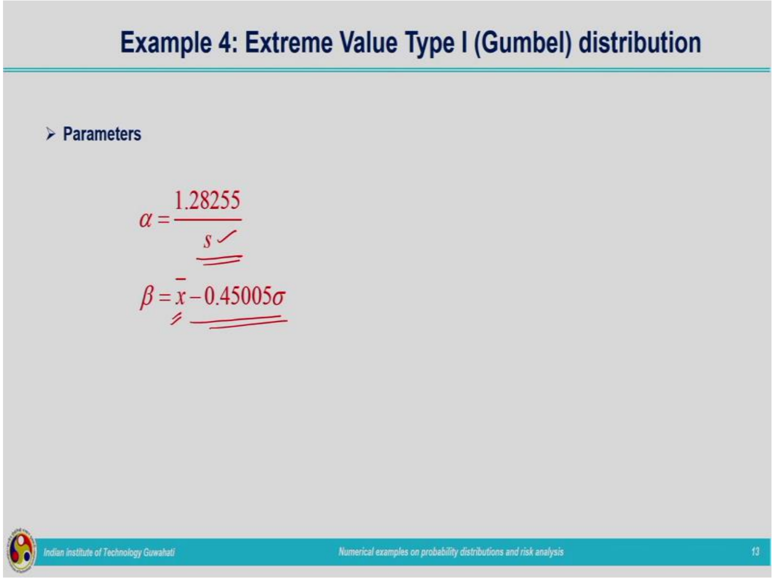
$$f(x)=\alpha \exp\left[-\alpha\left(x-\beta\right)-\exp\left(-\alpha\left(x-\beta\right)\right)\right]$$

and cumulative distribution function is given by

$$F(x)=\exp\left[-\exp\left(-\alpha\left(x-\beta\right)\right)\right]$$

So, here we have substituted for $(-\alpha(x-\beta))$ as y that is the Gumbel's reduced variate and α is the scale parameter which is greater than 0 and β is the location parameter which lies between $-\infty < \beta < \infty$. So, we need to find out the distribution parameters that is α and β and we need to find out the probability for the variable to exceed 1500 millimeter.

(Refer Slide Time: 21:06)



Let us start solving the example, first we need to find out the parameters. Parameters α and β. We are having the expression corresponding to that

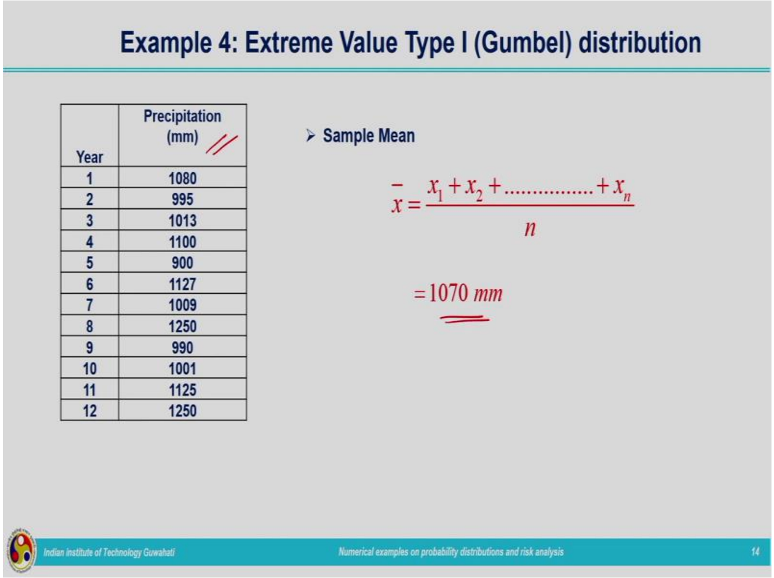
$$\alpha = \frac{1.28255}{s}$$

and

$$\beta = \bar{x} - 0.45005\sigma$$

Here you can see one term corresponding to s and another term corresponding to $\bar{x}$ is present. So, s is representing the standard deviation of the sample data and $\bar{x}$ is representing the sample mean. So, first we need to calculate the standard deviation and the mean corresponding to the sample data, then only we can go for the computation of α and β.

(Refer Slide Time: 21:56)

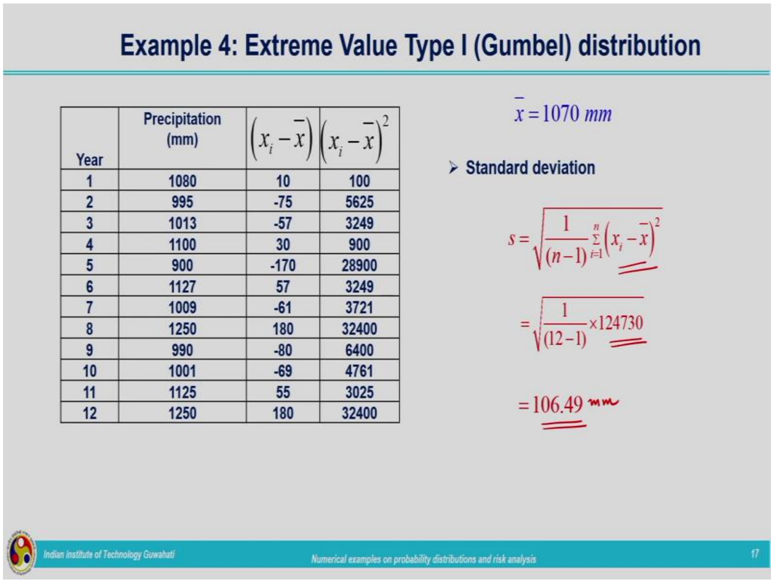


So, these are the data given to us. Now, we can compute the mean value by taking the average of these values given for precipitation. So, sample mean is given by

$$\bar{x} = \frac{x_1 + x_2 + \dots\dots\dots\dots + x_n}{n}$$

Here we are having data up to x_{12}. So, we will be dividing it by 12. So, this can be calculated as 1070 millimeters.

(Refer Slide Time: 22:28)



Now, next step is to compute the standard deviation sample standard deviation. $\bar{x}$ is 1070-millimeter, standard deviation is given by the formula,

$$s = \sqrt{\frac{1}{(n-1)} \sum_{i=1}^{n} \left(x_i - \bar{x} \right)^2}$$

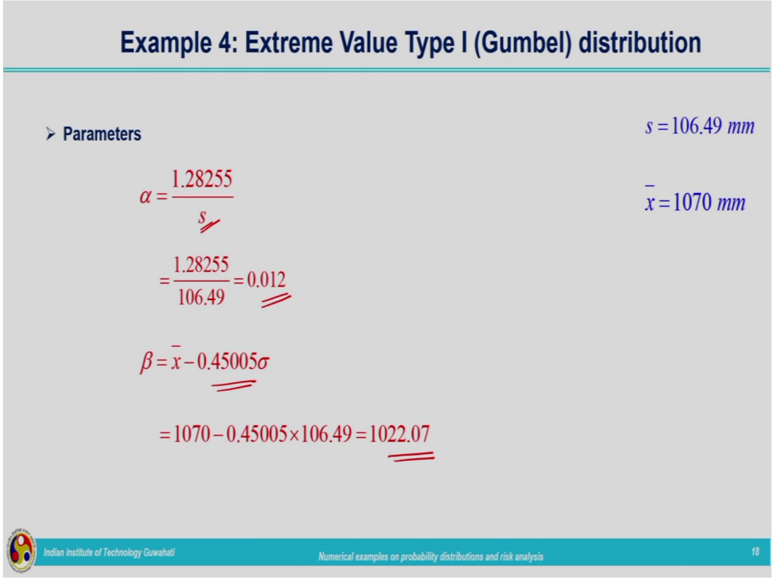
So, here in the denominator we are having $(n-1)$ because it is corresponding to the sample, if it is related to population we will be using *n*. So, *n* is 12 here and we need to calculate $\left(x_i - \bar{x} \right)^2$.

So, x_i is given by the precipitation corresponding to each year, *i* varies from 1 to 12, we can compute $\left(x_i - \bar{x} \right)$. It is calculated over here in this column and next we can calculate $\left(x_i - \bar{x} \right)^2$ that also calculated. And after that we can get the summation of $\left(x_i - \bar{x} \right)^2$ and that can be substituted here in this expression for *s*. The $\sum_{i=1}^{n} \left(x_i - \bar{x} \right)^2$ is coming out to be 124730. So, once we substitute for that value we can calculate s as

$$s = \sqrt{\frac{1}{(12-1)} \times 124730} = 106.49 mm$$

So, this is the standard deviation. Once we have computed sample mean and the corresponding standard deviation, we can move ahead for the computation of the parameters of the Gumbel’s distribution.

(Refer Slide Time: 23:57)



So, the parameters α and β we need to compute, α is consisting of standard deviation and

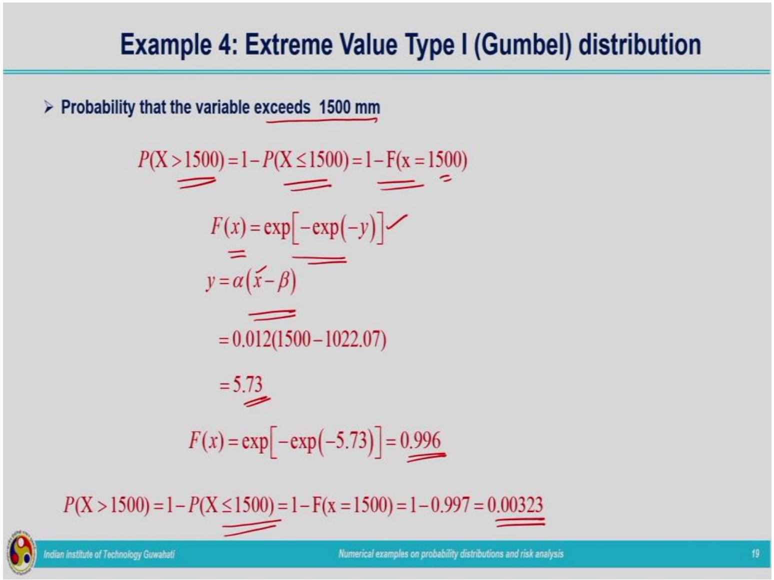
$$\alpha = \frac{1.28255}{s} = \frac{1.28255}{106.49} = 0.012$$

and

$$\beta = \overline{x} - 0.45005\sigma = 1070 - 0.45005 \times 106.49 = 1022.07$$

So, we got the values corresponding to different parameters of Gumbel's distribution.

(Refer Slide Time: 24:39)



Now, the second part of the question is to calculate the probability. Probability that the variable, random variable exceeds 1500 millimeters that is our rainfall is exceeding 1500 millimeters. What is the corresponding probability? That can be calculated by using the formula $P(\mathrm{X} > 1500)$. So, that is obtained by making use of the cumulative distribution function, this is equivalent to

$$P(\mathrm{X} > 1500) = 1 - P(\mathrm{X} \le 1500) = 1 - \mathrm{F}(\mathrm{x} = 1500)$$

This will be giving us the value corresponding to the probability that the variable exceeds 1500 millimeters. So, here we need to compute the $\mathrm{F}(\mathrm{x} = 1500)$.

We know the cumulative distribution function F(x) is given by

$$F(x) = \exp\left[-\exp\left(-y\right)\right]$$

What is *y*? *y* is given by

$$y = \alpha\left(x - \beta\right)$$

Here *x* is nothing but 1500, α and β we have already calculated. Now, we can substitute in the expression for *y*, the values of α, β and *x* and

$$y = 0.012(1500 - 1022.07) = 5.73$$

So, *y* is obtained as 5.73. Now, we can substitute the value of *y* as 5.73 in the cumulative distribution function *F(x)* and it is calculated

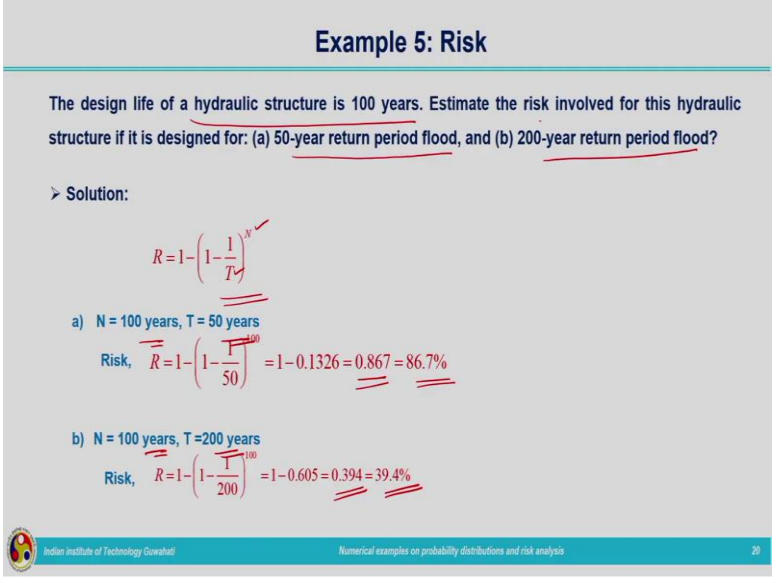
$$F(x) = \exp\left[-\exp\left(-5.73\right)\right] = 0.996$$

So, the probability can be calculated as

$$P(\mathrm{X} > 1500) = 1 - P(\mathrm{X} \le 1500) = 1 - \mathrm{F}(\mathrm{x} = 1500) = 1 - 0.997 = 0.00323$$

The probability that the variable exceeds 1500 millimeters very, very small. You can observe the data series itself, in that there is no data corresponding to 1500 millimeters. So, based on the probabilistic distribution, when we calculate the values coming out to be very, very small. So, this way we can make use of Gumbel's distribution to calculate the probability for that first we need to find out the parameters of the distribution. Similar way, we need to calculate the case with the other distributions also. I am not solving the examples related to that, the methodology followed will be the same.

(Refer Slide Time: 27:13)



Now, let us move on to solve some examples related to risk. Let me first read out the question. The design life of a hydraulic structure is 100 years. Estimate the risk involved for this hydraulic structure if it is designed for 50-year return period flood and 200-year return period flood. Two cases are there, the design period of the structure design life of the structure is 100 years and we need to compute the risk in two cases, one for the case of 50-year return period flood and the second one is for the case of 200-year return period flood. How can we compute the risk the formula which can be utilized for finding out risks we have discussed in the previous lecture that is given by

$$R=1-\left(1-\frac{1}{T}\right)^N$$

Where, T is nothing but our return period, capital N is the design life of the hydraulic structure.

The design life of the structure is 100 years, only the return period is changing in both the cases. N is equal to 100 years and T is equal to 50 years for case A. We can substitute that in the expression for risk

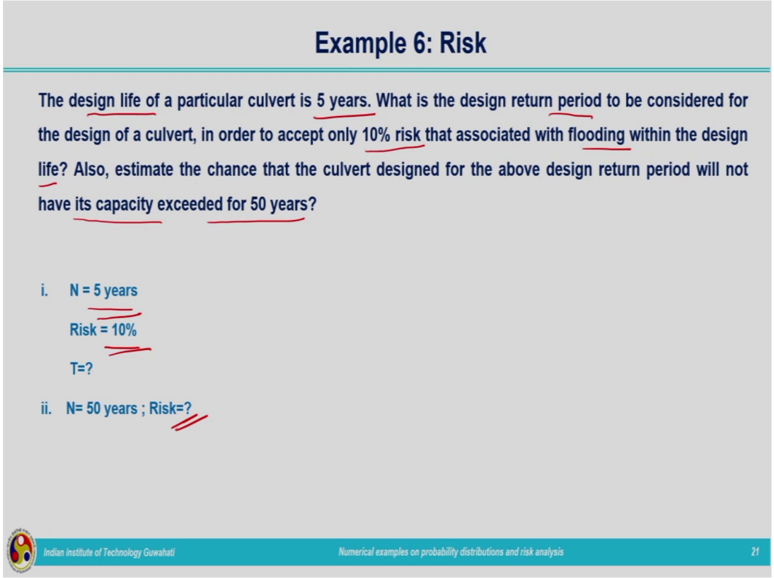
$$R=1-\left(1-\frac{1}{50}\right)^{100}=1-0.1326=0.867=86.7\%$$

It can be calculated as 0.867, it is around 86.7%. So, 50-year return period, if we are talking there is a chance of risk of about 86.7%. Then coming to the second case that is 200-year return period flood. While discussing about the risk factor I was explaining that we need to choose the return period according to the importance of the structure. We will not be using the same return period for all the types of structures. So, in the case of dams and other major structures, the return period which will be considered will be very high. So, definitely you can understand that when the return period is increased to 200 years, the risk associated with that will be less compared to a return period of 50 years. So, we can solve the second part, N is equal to 100 years and T is equal to 200 years. So, when you substitute in the expression for R, you can calculate the risk associated with that as

$$R = 1 - \left(1 - \frac{1}{200}\right)^{100} = 1 - 0.605 = 0.394 = 39.4\%$$

So, the risk is reduced. Here we need to have two data, one is the return period of the extreme event and also the design period corresponding to the particular hydraulic structure.

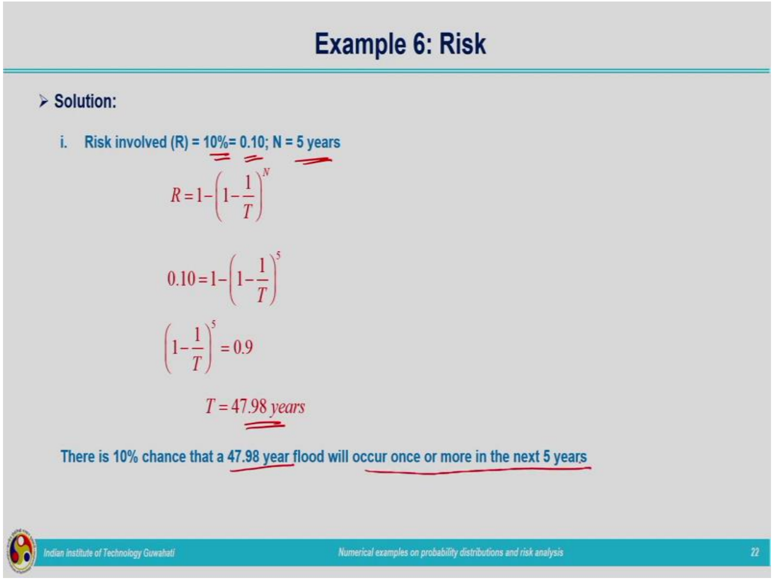
(Refer Slide Time: 29:53)



Now, let us solve one more example related to risk. The design life of a particular culvert is 5 years. What is the design return period to be considered for the design of the culvert in order to accept only 10 percentage risk that can be associated with flooding within the design life? Also estimate the chance that the culvert design for the above design return period will not have its capacity exceeded for 50 years. Two parts are there, slightly difficult to understand the question, read it carefully. We need to find out the return period corresponding to the design of a culvert which is having a design period of 5 years and we can allow a risk of 10 percentage, that is the first part. By allowing a risk of 10 percentage we need to calculate the return period and then the second part is to find out the design period. Second part is to find out the chance that the culvert designed with the return period which we have found out in the first part will not have its capacity exceeded for 50 years.

Let us start solving the problem then you will be able to understand. In the question it is given that the design life of the culvert is 5 years. So, 5 years design life is there for the culvert, those types of structures will be given less importance that is why the design life is given 5 years and the return period considered also will be less in such cases. Here we need to compute the risk associated with the structure. *N* is equal to 5 years and risk allowed is 10 percentage. We need to compute the return period corresponding to this allowed risk. We are having the formula for risk, we need to compute the return period *T* and second part is to compute the risk. We are having the design life of 5 years and also for which return period the risk has to be computed, the return period which we are computing in the first part of the question. Two parts are there in this question, first part is to find out the return period corresponding to a risk of 10 percentage, allowed risk is 10 percentage and the design life is there 5 years and second part is that if you are considering the return period of the one which we are calculating in part one, how much will be the risk associated with that?

(Refer Slide Time: 32:38)



We can proceed solving the question. The risk involved *R* is 10 percentage, that is 0.1 and *N* is equal to 5 years. We will substitute in the formula *R* is equal to

$$R = 1 - \left(1 - \frac{1}{T}\right)^N$$

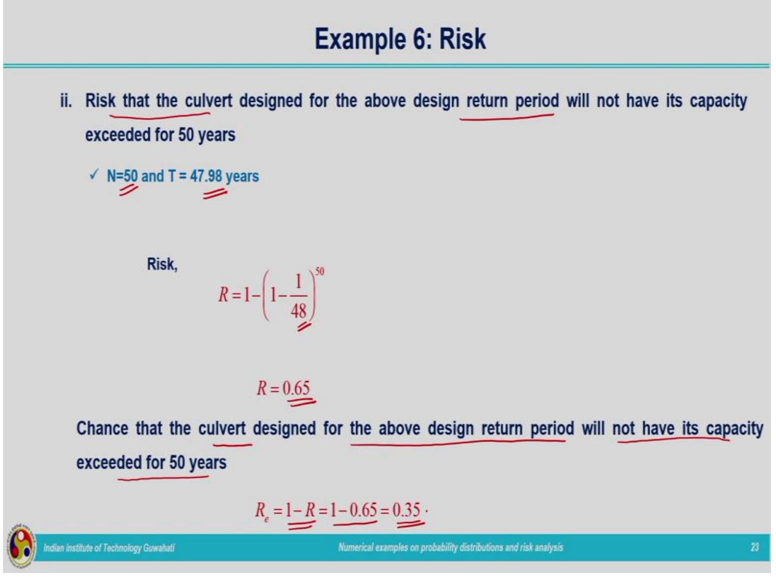
$$0.10 = 1 - \left(1 - \frac{1}{T}\right)^5$$

$$\left(1 - \frac{1}{T}\right)^5 = 0.9$$

$$T = 47.98 \; years$$

As far as a culvert is concerned, this return period is very high, 47.98 years and in majority of the cases we will not be considering 47.98 years, we will be approximating it to 50 years. So, this is the design return period, which can be calculated from the question for a risk to happen around 10%. There is a chance that 47.98-year flood will occur once or more in the next 5 years.

(Refer Slide Time: 33:34)



Second part of the problem is to calculate the reliability. For that first we will compute the risk associated with that. So, we are going to make use of new design return period and in that case, N will be 50 and T will be 47.98 which we have computed in the first part. 47.98 can be approximated as 48 years and we can compute the risk associated with that R is equal to

$$R = 1 - \left(1 - \frac{1}{48}\right)^{50}$$

So, T we have substituted as 48 and risk can be calculated as 0.65.

$$R = 0.65$$

65% of risk associated with this. But we have been asked to find out the reliability that is the chance that the culvert design for the above design return period will not have its capacity exceeded for 50 years. How can we calculate reliability? Reliability is nothing but it is 1 minus R, i.e.,

$$R_e = 1 - R = 1 - 0.65 = 0.35$$

It can be calculated as 0.35 or 35%.

(Refer Slide Time: 35:00)

So, these are the different types of numerical examples, which can come under these topics. You need to try with different example problems and exercise problems. You will be getting different numerical exercise and example problems from these reference textbooks. Try to solve maximum number of problems related to this topic. So, here I am winding up the problem-solving session on probability distribution and risk analysis. Thank you.