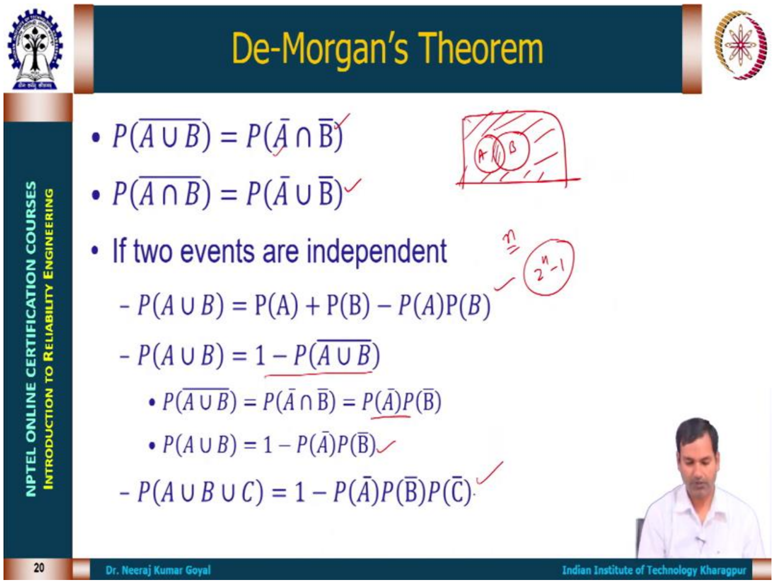
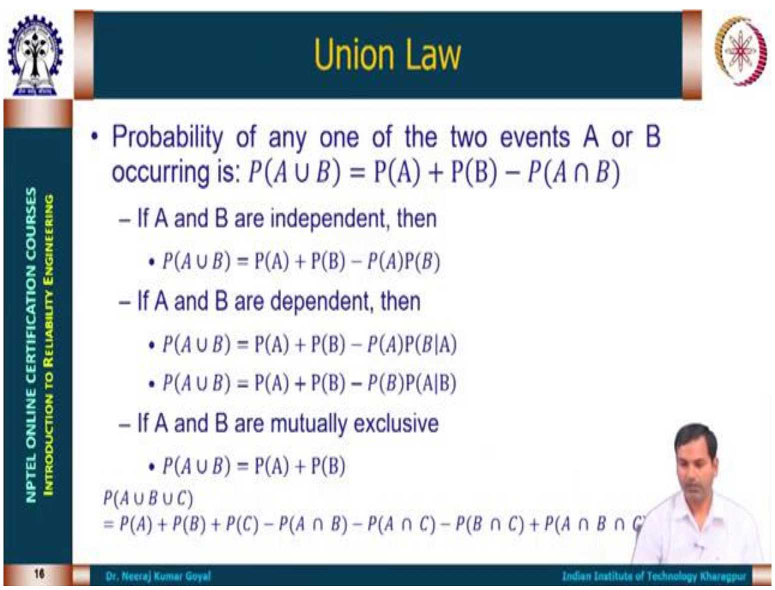
Introduction to Reliability Engineering
Professor Neeraj Kumar Goyal
Subir Chowdhury School of Quality and Reliability
Indian Institute of Technology, Kharagpur
Lecture 05
Probability Basics (Contd.)

(Refer Slide Time: 00:36)





Hello everyone, we have now reached lecture number 5. This is a continuation of our previous lecture, lecture number 4 on probability basics. We will first discuss De Morgan's theorem. De Morgan's theorem is related to the complement of events.

If we have two events, A and B, and we are interested in their union, taking the complement of it means we are talking about the area outside A union B. This can be obtained by finding the intersection of what is outside A and what is outside B. Therefore, anything outside A but not part of B gives us the complement of A union B.

Similarly, if we are interested in the intersection of A and the complement of A intersection B, we need to find the union of what is outside A and what is outside B. This is because everything outside B and everything outside A has to be covered.

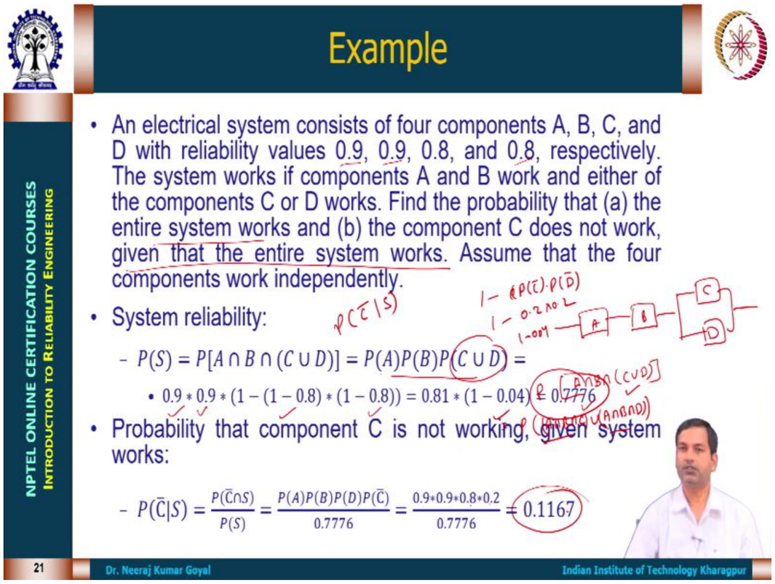
So, in a simple way, we can see that when we take the complement of a group, the individual elements will be complemented and the signs will be reversed. If it is a union, it will become an intersection, and if it is an intersection, it will become a union.

We can use this property to simplify cases where two events are independent. For example, if we have n individual events that are a union, we would have had 2 to the power n minus 1 terms to solve and evaluate. However, using De Morgan's theorem, we can simplify this as the probability of A union B is equal to the product of 1 minus the probability of the complement of A union B.

The complement of A union B is the intersection of A complement and B complement, which can be multiplied to get their probability since they are independent. So, the probability of A union B can be expressed as 1 minus the probability of A union B complement, which is equal to 1 minus P A complement P B complement.

Similarly, for three events, we can simplify it as 1 minus P A complement P B complement P C complement, which is much simpler than solving it the other way. If there were four events, with D also involved, we can express it as the probability of D complement multiplied by the previous expression.

(Refer Slide Time: 03:26)



Let's take an example where an electrical system consists of four components: A, B, C, and D. Their reliability values are given as 0.9, 0.9, 0.8, and 0.8. The system will work if A and B work, and if either C or D works. Generally, if we represent the system diagrammatically, it will look like A-B, and then either C or D can work. So, the system will work if A works, B works, and either C or D works. Now, we want to know the probability of the entire system working, which means A works, B works, and either C or D works. This can be represented as the probability of (A intersection B intersection C) union with (A intersection B intersection D), or we can evaluate it directly.

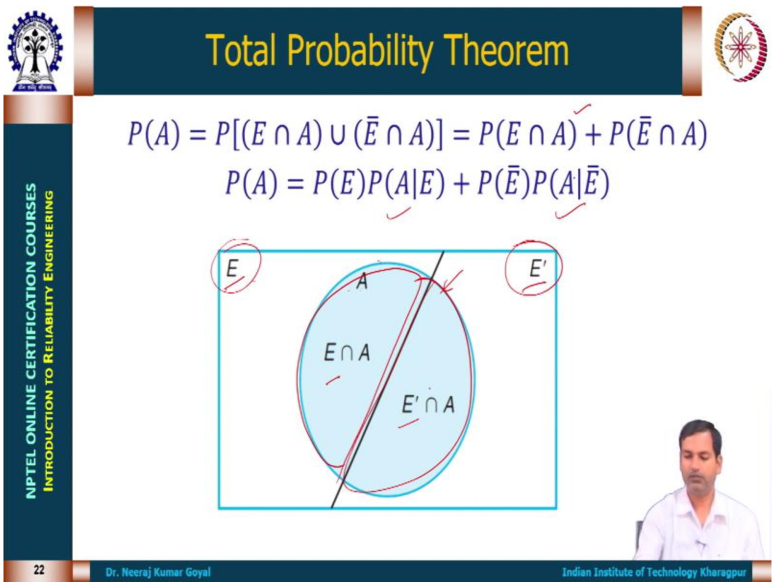
If we calculate it directly, then we can solve it in this way: the probability of A is 0.9, the probability of B is 0.9, and we want to calculate the probability of (C union D) working. We can calculate this by subtracting the probability of (C not working) and (D not working) from 1, i.e., 1 - probability of (C bar) times the probability of (D bar). What is the probability of (C bar)? That is 1 minus 0.8, and the probability of (D bar) is also 1 minus 0.8. So, 0.9 times 0.9 times 1 minus 0.2 times 0.2 gives us the probability of 0.7776. Therefore, we are able to calculate this.

$$P(S) = P[A \cap B \cap (C \cup D)] = P(A)P(B)P(C \cup D) =$$
$$\cdot 0.9 * 0.9 * (1-(1-0.8)*(1-0.8)) = 0.81*(1-0.04) = 0.7776$$

Now, our second question is: what is the probability that component C does not work given that the entire system works? That means, in the case where A, B, and D work, but C does not work, the system will still work. We can represent this as the intersection of (C bar) with S, where S represents the event that the entire system is working. So, if we take the intersection of (C bar) with S, then C dot (C bar) will become 0 and will be removed. The only thing remaining will be PB PA PB PD and PC bar PA PB PD and PC bar divided by the probability that the system is working. The probability of the system working is already available to us, and we can calculate it as 0.9 times 0.9 times 0.8, and the probability of (C bar) is 0.2. This gives us a result of 0.1167, which means that there is only an 11.67% chance that component C is not working when the system is working.

$$P(\bar{C}|\ S) = \frac{P(\bar{C} \cap S)}{P(S)} = \frac{P(A)P(B)P(D)P(\bar{C})}{0.7776} = \frac{0.9 * 0.9 * 0.8 * 0.2}{0.7776} = 0.1167$$

(Refer Slide Time: 06:31)

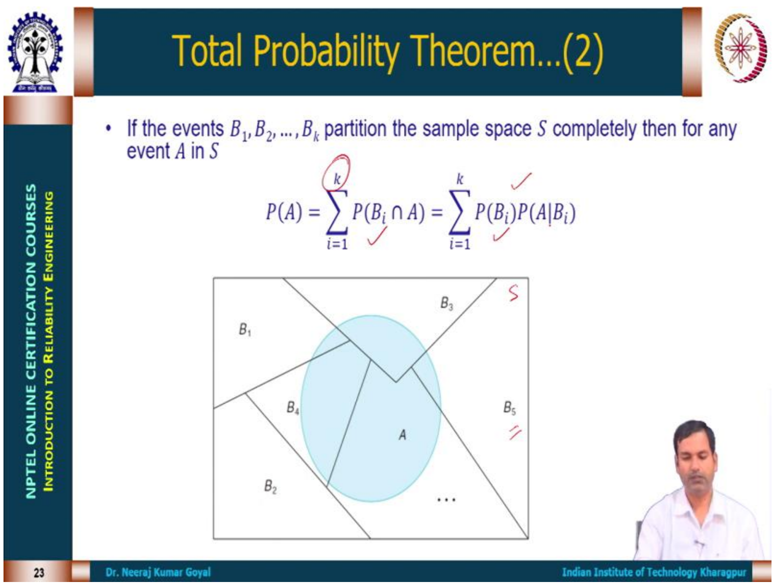


There is another important term called the Total Probability Theorem. The Total Probability Theorem is useful when we have mutually exclusive events. For example, we may have two events, event A and event A bar, where A bar is the complement of event A. We divide the entire sample space into two parts, E and E bar, which represent these mutually exclusive events. Now, we want to know the probability of A, but we only know the probability of event A occurring with E or without E.

To calculate the probability of A, we need to take the intersection of A with both E and E bar, since they represent the total sample space. Therefore, the probability of A is the sum of the intersection of A with E and the intersection of A with E bar. If the events are dependent, we can calculate this probability by multiplying the probabilities. However, if the events are independent, we can directly calculate the probability without using the Total Probability Theorem. The Total Probability Theorem is mostly used when there is a dependency between events.

$$P(A) = P[(E \cap A) \cup (\bar{E} \cap A)] = P(E \cap A) + P(\bar{E} \cap A)$$
$$P(A) = P(E)P(A|\ E) + P(\bar{E})P(A|\ \bar{E})$$

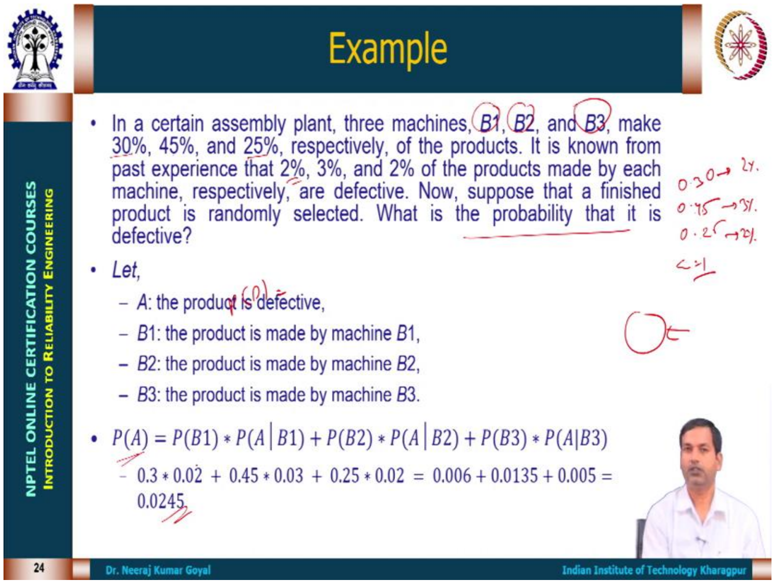
(Refer Slide Time: 07:51)



This total probability theorem if let us say we divide the whole sample space into here like it is shown in 5 events. So, k number of events, if k number of events mutually exclusive they do not share any common space. In that case, I want to calculate then so again, summing up the intersection probabilities from i equals 1 to k will give me the probability of A or conditional probability taking condition into account probability of Bi given multiple availability of A given Bi has already occurred.

$$P(A) = \sum_{i=1}^{k} P(B_i \cap A) = \sum_{i=1}^{k} P(B_i)P(A \mid B_i)$$

(Refer Slide Time: 08:27)



So, this case, how is it useful, let us take an example that if we have an assembly plant where 3 machines are there, B1 B2 B3 and they are making 30 percent 45 percent and 25 percent of the production, so 30 percent production 0.3 is covered by B1, 0.45 is covered by B2 and 0.25 is covered by total will be as we see total production is divided into a total sample space is divided into 3 parts total will always be 1, 0.3 0.45 0.25 will be sum will be equal to 1.

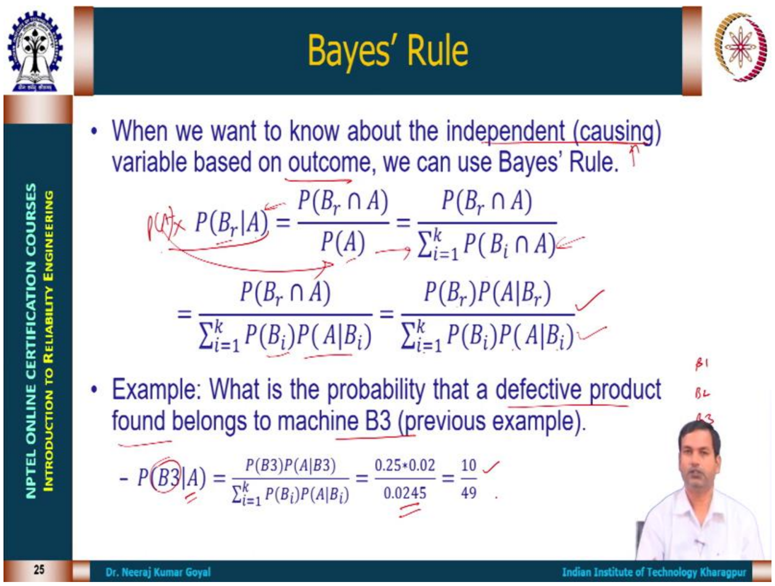
Now, here, we know that this machine produces 2 percent defective. This machine produce 3 percent defective and this machine produce 2 percent effective now, what will be the probability if I am having a product which I have selected from the output which is randomly selected, now, what is the probability that it is defective the probability that it is defective is equal to probability that it is coming from machine 1 and it is defective probability that it is coming from second machinery it is defective and third machine it is defective so we can get it.

So, the probability of A is equal to product B1 into probability of A given B1 probability B2, A given B2 and probability of B3, A given B3. So, 0.3 into 0.02 like this, if you sum up we are able to get the probability of defective.

$$P(A) = P(B1) * P(A \mid B1) + P(B2) * P(A \mid B2) + P(B3) * P(A \mid B3)$$
$$= 0.3 * 0.02 + 0.45 * 0.03 + 0.25 * 0.02 = 0.006 + 0.0135 + 0.005 =$$
$$0.0245$$

There is a Bayes rule, what happens generally here in this example, as we see, the response variable is defective and the independent variables are the machine output, and they are defective, they are not changed, they are independent, but outcome which is the product defective is dependent on the machines. So, machines are independent variables and outcome is the dependent variable.

(Refer Slide Time: 10:29)



But in case of Bayes we can do the interchange, we can assess the probability for independent variables given the dependent variable is known so, here what we want to know independent that is the cause.

So, we want to know the probability of cause based on the outcome, so, we want to check our cause based on the evidence or based on the observations which you are having. So, we want to know what is the probability that the outcome is coming from a section Br given that outcome A is observed, so, this is equal to we know that this is simple reverse formula, we know probability of Br intersection A is probability of A into Br into this but if you take probability of A here, this becomes probability of Br A is equal to this.

So, this probability of A if you use total probability theorem, it can be expressed like this. So, probability of Br given A can be calculated using this formula, this can be also calculated in terms of conditional probabilities like this.

So, we are able to calculate probability of Br given A which is equal to probability of Br multiplied with probability of A given Br divided by probability of B for all cases, probability of B into product A given Bi i equal to 1 to k like the example which are taken earlier we are 3 machines were there B1 B2, B3.

$$P\left(B_r \mid A\right) = \frac{P\left(B_r \cap A\right)}{P(A)} = \frac{P\left(B_r \cap A\right)}{\sum_{i=1}^{k} P\left(B_i \cap A\right)}$$
$$= \frac{P\left(B_r \cap A\right)}{\sum_{i=1}^{k} P\left(B_i\right) P\left(A \mid B_i\right)} = \frac{P\left(B_r\right) P\left(A \mid B_r\right)}{\sum_{i=1}^{k} P\left(B_i\right) P\left(A \mid B_i\right)}$$

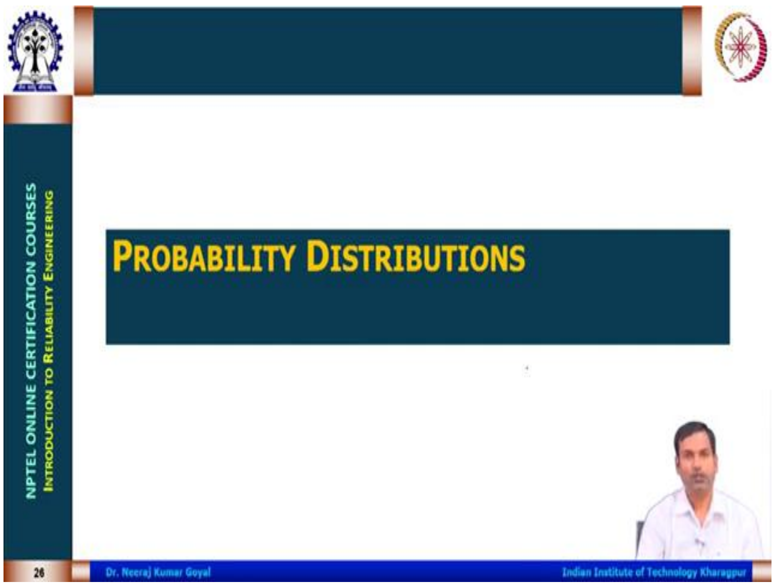
Now, we want to know opposite, we want to know what is the probability that our defective product which we have found belonging to machine B3 that means, given that product is defective, what is the probability that it has come from machine number B3 that means probability of B3 given A so B3 given A means, we can use this formula probability of B3 multiply the probability of A given B 3 divided by for all cases for the Bi or we can directly have PA, which we have already calculated 0.0245.

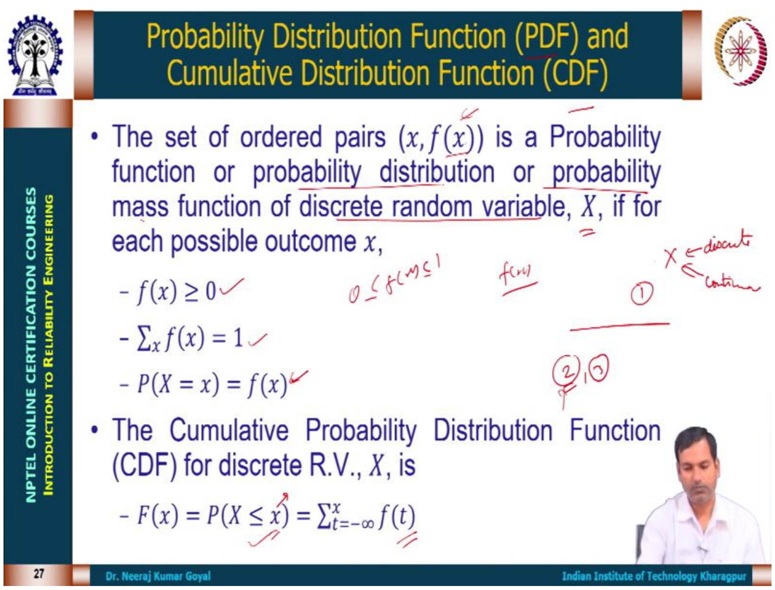
$$P(B3 \mid A) = \frac{P(B3) P(A \mid B3)}{\sum_{i=1}^{k} P\left(B_i\right) P\left(A \mid B_i\right)} = \frac{0.25 * 0.02}{0.0245} = \frac{10}{49}$$

Now, what is the probability of B3 probability of B3 is 0.25 and probability of defective if this it is B3 is 0.02. So, this if we see solve then this comes out to be 10 divided by 49. So, that means 10 divided 49 and is the probability that a defective product which are found belongs to machine number B3, similarly we can get for B2 B1 like that.

So, it helps us to investigate the matters and see that what are the more important cases from where it might have come so, we can do cause analysis also or we can take the do the reverse analysis also using the Bayes rule. There is lot of lot of the things about Bayes which you can read more about it.

(Refer Slide Time: 13:21)





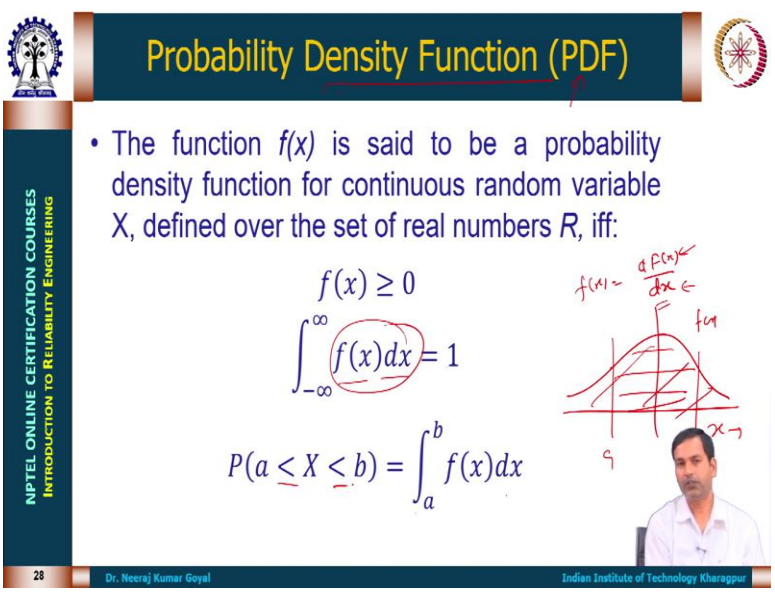
Next, we can move on to probability distributions. Probability distributions are common terms used like PDF and CDF. The set of ordered pairs x, f(x) is the probability function or the probability distribution or the probability mass function. We call this the probability distribution function or probability mass function for discrete random variables. There are generally two types of random variables - discrete and continuous. For a discrete random variable, the pair x and f(x) is called the probability distribution function or probability mass function. Here, fx represents the probability because the total probability of 1 is divided into various small probabilities for all outcomes. f(x) is always greater than or equal to 0, and the summation of f(x) will give you 1. f(x) is also a value between 0 and 1, representing the probability that the random variable x has a value equal to x.

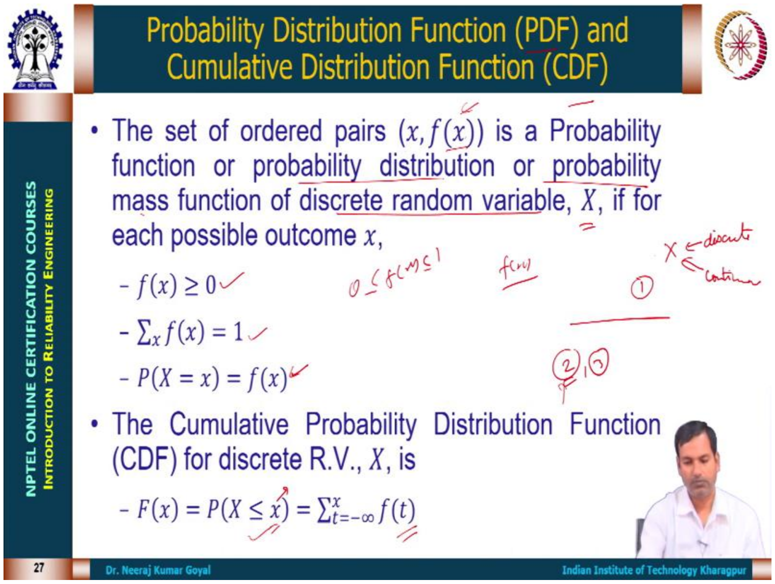
$$f(x) \geq 0$$
$$\Sigma_x f(x) = 1$$
$$P(X = x) = f(x)$$

The cumulative probability distribution function is the summation of probabilities. If we want to know the cumulative probability up to event x, we can sum all the probabilities from minus infinity to x. For example, if x is 2, then the cumulative probability includes all probabilities up to x=2. As x increases, the cumulative probability also increases. If we want to calculate the probability for x=3, the cumulative probability for 3 includes the probability up to 2 plus the probability of 3. Therefore, the cumulative distribution is always increasing and never decreases.

$$F(x) = P(X \leq x) = \sum_{t=-\infty}^{x} f(t)$$

(Refer Slide Time: 15:18)





Then comes probability density function same both for discrete and for continuous both we are saying PDF, but there is a little difference in when we talk about the continuous random variable for continuous random variable this is called density function not the distribution function. So, it is called density function because here f(x) is equal to df(x)/dx it is the slope.

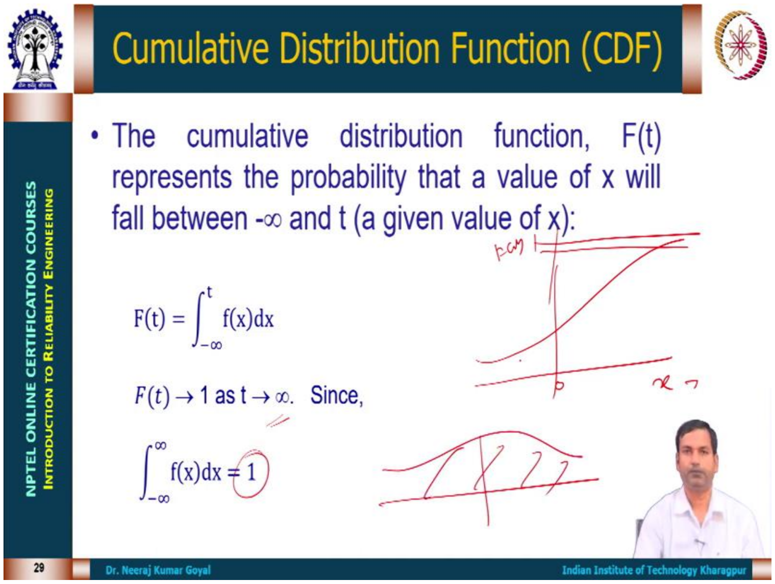
So, fx is the unitless quantity x is not the unitless quantities, it is divided. So, this is the density function. So, it is not unitless while f(x) is unitless in case of discrete random variable as we see when we multiply dx then only it becomes unitless. So, when we multiply with dx and integrate now, integration of small f(x) versus all value for all values of x gives you so, area under this curve is always 1.

Now, if I want to calculate the probability from A to B that is area under the curve from A to B that is A to B integration of small fx Fx dx this gives me the probability that it is lying between. Here in case of discrete we should be careful with the size whether equal and here, in case of continuous whether we put equal or do not put equal that does not make much difference because point probabilities are equal to 0, point does not have any dimension so, it will not have any area under this. So, because of that area under this point probabilities are 0.

$$f(x) \geq 0$$
$$\int_{-\infty}^{\infty} f(x)dx = 1$$
$$P(a < X < b) = \int_a^b f(x)dx$$

So, because of that equal sign here, where we whether we put or do not put does not make much difference, we could only we should only be careful because, so, that meaning is correct. But for discrete distribution, it matters if you see because value of x it will be counted or not it depends on the equal sign or not if I say x is less than x, then well probability of x will not be counted here. So, that makes a difference here.

(Refer Slide Time: 17:24)

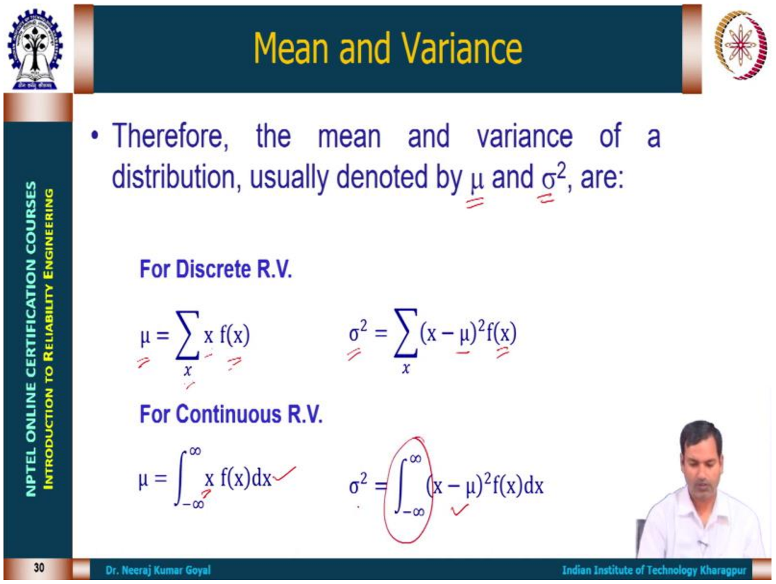


For cumulative distribution functions for same continuous random variable I can get it by integration. So, that is from minus infinity to t*f(x)dx will give me the cumulative distribution for this we know that if t is infinity then f t will become 1.

So, f t will look like something like this it will always reach to 1 finally, and will always start from 0 so, and if I take area under the curve for f(x) then it is always giving me 1 this we have seen already.

$$F(t) = \int_{-\infty}^{t} f(x)dx$$
$$F(t) \to 1 \text{ as } t \to \infty.$$
$$\int_{-\infty}^{\infty} f(x)dx = 1$$

(Refer Slide Time: 18:07)

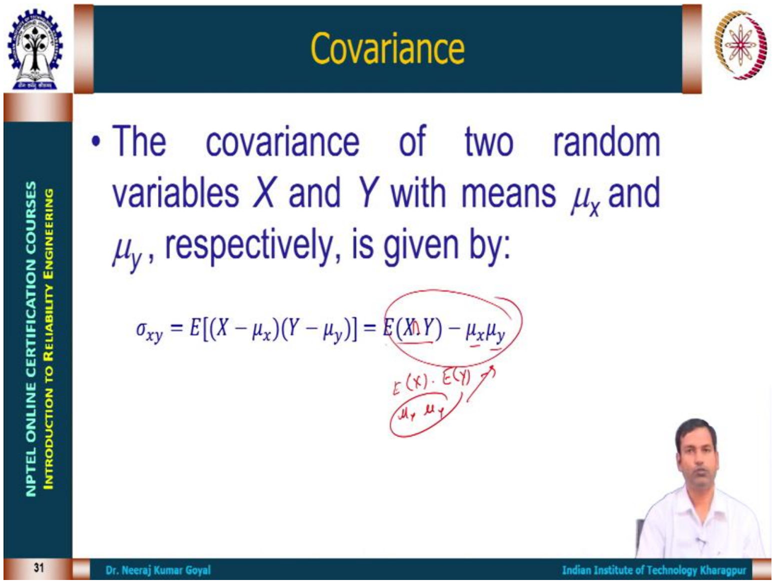


Now, there are two important terms which we use in engineering decisions, one is mean and another is variance means tells us where the centralization of the values is there and variance tells us how much spread is there how much is the variability is there in the values.

So, for discrete random variable if I want to calculate mean it is nothing but summation of x multiplied by f(x), f(x) tells the wait that or how much proportion of time x is supposed to occur. So, if we multiply this and sum it up, we get the value of mu and variance has given us variances generally the second moment around mean.

So, from mean how far is it so, from mean whether it is negative side or positive side whatever is the distance that is x minus mu since we are taking a square this will always give the positive quantity. So, this is whatever distance we have that is the square of that multiplied by the f(x). So, that gives me the variance for continuous random variables same thing is done f(x) is taken f(x) is taken it is integrated over f(x)dx. Similarly, x minus mu whole square f(x)dx, if you see integrated over full range, full range of full values of x.

(Refer Slide Time: 19:33)

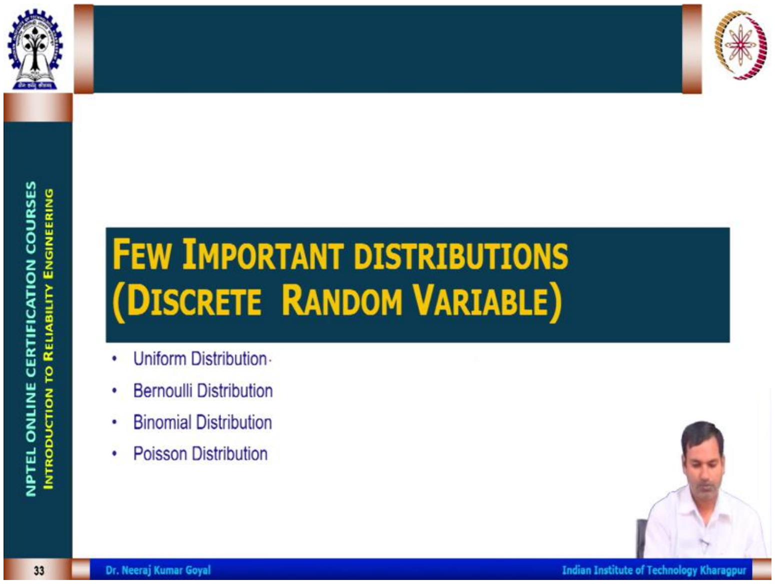


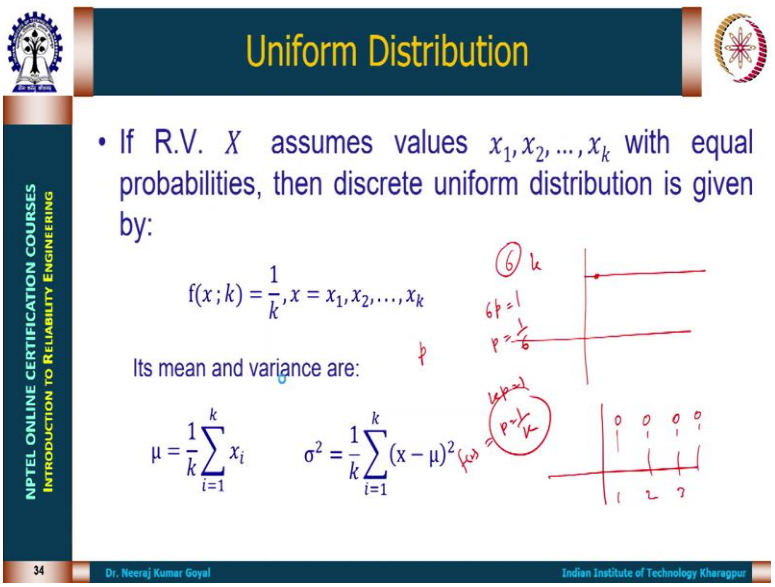
Covariance is used for when there is a dependency among X and Y if there is if these are independent then covariance will be equal to 0. So, covariance of x and y is equal to the expected value of x minus mu x and y minus mu y that is if you saw that becomes the expected value of x into y or we can say intersection y and mu x into mu y.

So, if they are independent then we know this is the value of expected value of x multiplied with the expected value of Y, which is mu x, mu y so that will be 0, but if they are dependent, it will not be equal to mu x mu y so, in that case only, we will have the covariance value.

$$\sigma_{xy} = E\left[\left(X-\mu_x\right)\left(Y-\mu_y\right)\right] = E(X.Y) - \mu_x\mu_y$$

(Refer Slide Time: 20:16)





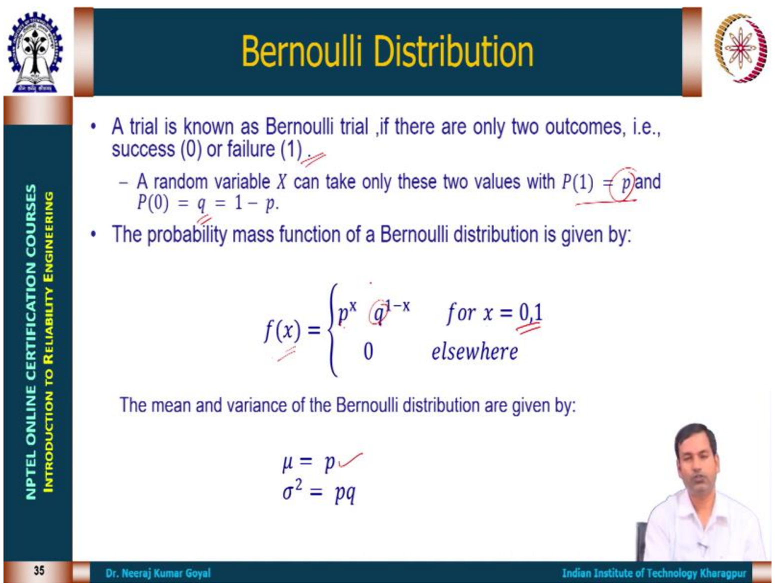
We will discuss some important distributions. So, we will discuss some discrete random variable distributions, first this distribution is uniform, then Bernoulli, then binomial, then Poisson. So, uniform distribution as you know uniform is equally likely. So, if we say since we are talking about discrete, so, we have let us say, if there are k possible outcomes 1 2 3 three possible outcomes are there for like when we say tossing a coin, then we have two possible outcomes.

If we say when we are rolling a dice, then we have 6 possible outcome. So, this gives us the probability now, all outcomes are equal probable, since all outcomes are equal probable, let us say probability is p. And if I am talking about let us say 6 outcomes, then all are P so, this will become 6 p.

So, summation we know should be equal to 1, so, 1 p will be equal to 1 upon 6 if I say K outcome, then k p is equal to 1 so, p will be equal to 1 upon k. So, the probability this p is nothing but f(x) is constant all outcomes are equally likely with the probability 1 upon K. The mean and variance for this distribution are given mean will be nothing but summation of this multiplied with 1 upon K because all f(x) is same so, xi into 1 upon k.

So, summation of x divided by k will give the mu and sigma squared will be 1 upon k summation of x minus mu whole square.

(Refer Slide Time: 22:00)



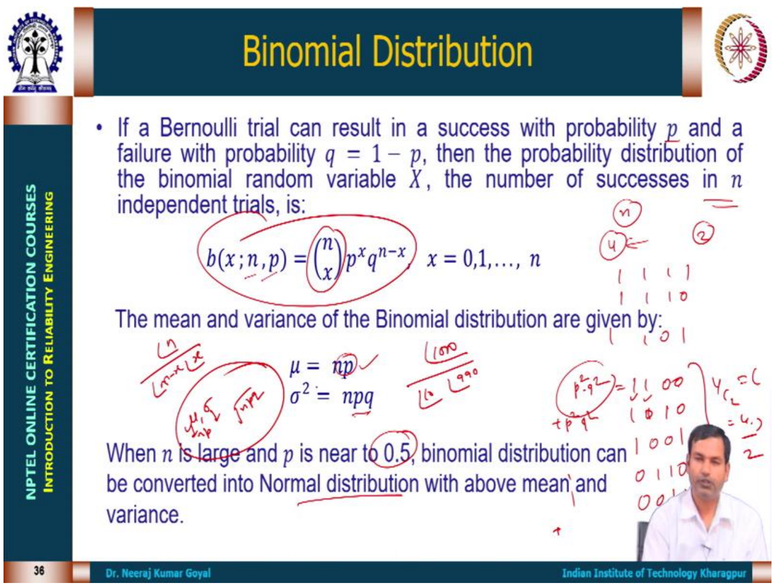
Then there is another distribution which is used for Bernoulli like for tossing a coin, tossing a coin is actually equal probable so, we use we can use uniform. But if the two events are not equally likely, then we can use a Bernoulli for equally likely also, we can use Bernoulli but, there are two possibilities, let us say in case of reliability, we talk about two outcomes either success or failure.

So, here the random variable represent whether the system is state is success or failure. So, success state's probability is p if success state's probability is p then by default failure probability will be 1 minus p that is we are representing as q now, Bernoulli distribution can have two outcomes either 0 or 1, 0 means unsuccess 1 means success.

So, fx can be given as p to the power x q to the power 1 minus x. So, when I put x equal to 0 p to the power 0 will become 1 q to the power will become 1. So, that will become q if I put x equal to one p to the power 1 will become p, q to the power 1 minus one 0 will become 1.

So, this becomes p, mean and variance for this is given as mean of this f(x) is p and variance has given us pq but generally this is for single item or single coin toss or single item failure, but many most of the time we are having multiple events happening together.

(Refer Slide Time: 23:28)



So, if a Bernoulli trial can result that means, if you do multiple Bernoulli trials, and each trial, the probability is same, probability of success or failure is same, probability of success is p and probability of failure is 1 minus p, then, this will become a binomial distribution. In binomial distribution, the Bernoulli trial is done n number of times.

So, here we have two parameters, number of trials and probability of success p. So, binomial distribution probability can be calculated as nCx, because, for each outcome, let us say n outcomes are there, I am, let us say 4, I am doing the 4 trials, and so, 4 trials can be like all are success or one of them is failed. One of them fail can be in various ways like this.

$$b(x;n,p) = \binom{n}{x} p^x q^{n-x}, \quad x = 0,1,\dots,n$$

So, when I am talking about X number of successes out of it, let us say I am talking about 2 success out of 4. So, in that case, the possible cases of 2 success out of 4 will be 1100 1110 1010 01. There will be 6 possible states 01100011 and 0101 4C2, 4 into 3 divided by 2. So, that becomes 6 states.

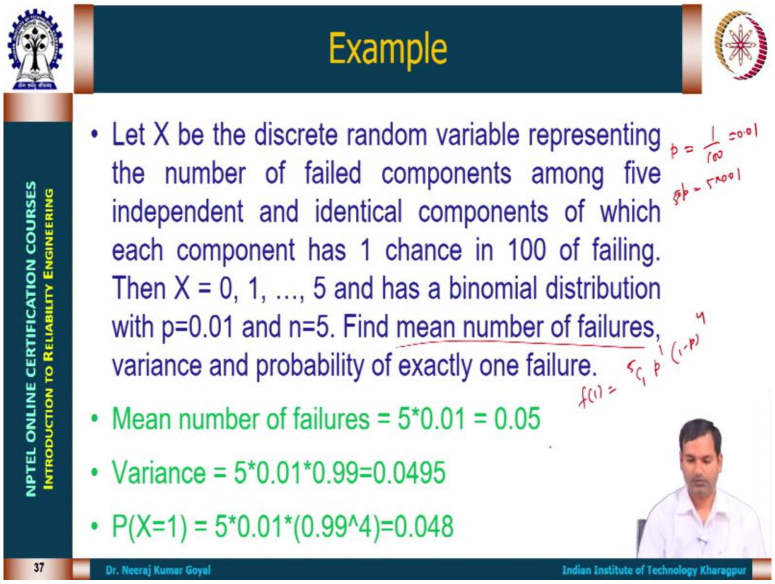
So, have what happens here in all state what is the probability, probability is same, this is success, this is success. So, p into p, p squared and q into q, q squared. So, the total probability will become this plus again p squared q squared like this.

So, if I sum it up that becomes ncx p to the power x, q to the power n minus x, x is number of successes and n minus x will become number of failures, this gives me the total probability binomial probability for observing x number of successes out of x trials, where p is the probability of success, mean of this is given as like for Bernoulli it was p now, for e, n trials are done so, mean becomes n into p.

And sigma squared is also summation of all sigma that is pq some n number of times that becomes npq as we see that calculation of this ncx factorial n divided by factorial n minus x n factorial x if I say n is very very large if n is 1000, if I want to know 10 out of 1000 then this probability calculation would be difficult, the value of factorial 1000 will become very high.

So, in that case, this binomial distribution is many times approximated as either normal distribution or Poisson distribution. So, when value of P is somewhere around middle 0.5 0.4 0.5 0.6 or around that somewhere in middle then in that case, this binomial distributions follows the normal distribution for normal distributions, we know there are two parameters mu and sigma. So, mu will be equal to NP here and sigma will be equal to square root of NPQ these two parameters if you use then for normal tables we can use to evaluate the probabilities.

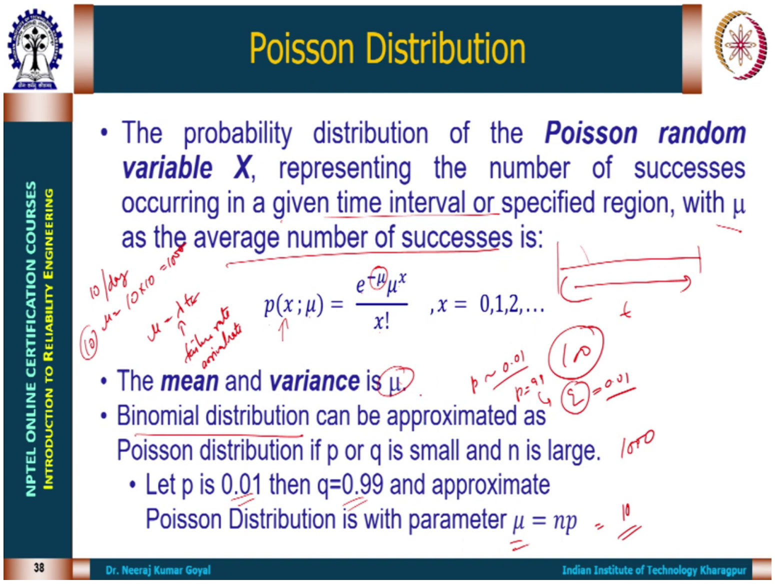
(Refer Slide Time: 27:05)



Let us take an example, if x a random variable representing the number of failed components among 5 independent identical components, each component has one chance in 100 of failing. So, probability of failure here is p equal to 1 out of 100 that is 0.01.

Now, I want to know x equal to 0 1 2 has a binomial distribution, we have already found that that is probability is 0.01 and n is equal to 5 now, I want to know the mean number of failure, so, the mean number of failure will be np, so, 5 into p so, 5 into 0.01 that will be 0.05 if I want to know the variance that is npq 5 into 0.01 0.99 0.0495 if I want to know the probability of exactly 1 failure that means f of 1 that is our 5 c 1 and p to the power 1 and 1 minus p to the power 4. So, this we calculate we get the 0.048.

(Refer Slide Time: 28:26)



Then comes Poisson distribution, Poisson distribution is probability distribution where number of successes but here like here we were talking, in binomial we were talking about our samples all possibilities, sample space were the countable like 5 10.

So, that was also discrete, but in case of Poisson we take the possibility in continuous space that is time. So, let us say we want to know how many failures in 100 hours in that case, we will be using Poisson distribution. So, here number of success occurring in a given time interval we want to know or any specific reason that means, let us say 100 kilometers someone has table what is the number of failures a person can observe or how many breaks a person will take.

So, let us say mu is the average number of successes, mu is only parameter for Poisson distribution, that is giving the mean number of successes. So, probability of x can be obtained as probability of x number of successes I can get with the same parameter mu that is given as e to the power minus mu multiplied mu to power x divided by factorial x where x can be 0 to infinity and mean and variance both are equal to mu here.

$$p(x;\mu) = \frac{e^{-\mu}\mu^x}{x!} \quad , x = 0,1,2,\ldots$$

So, the Poisson distribution has single parameter mu, and the parameter value is, here in reliability this mu is considered as lambda t where lambda is the failure rate or it is also called as arrival rate and t is the time. So, multiplying constant failure rate arrival rate multiplied by time will give you how many.

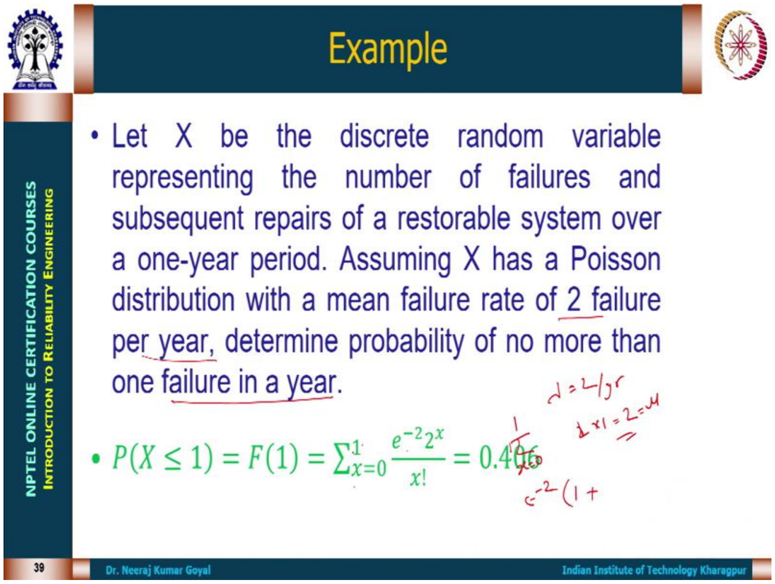
So, let us say arrival rate is 10 per day, let us say our production rate is 10 per day. So, if I want to know how much is the production in 10 days that will be equal to 10 into 10. So, mu will be equal to 100 that is 10 is the arrival rate and t is the 10 days, so, that will become 100, same way I can this mu will be calculated.

So, thus as we discussed earlier binomial distribution if the value of P is somewhere around 0.01 0.1 means it is on lesser side or higher side let us say 0.99, in the case of 0.99 I can consider it equivalent to q, I will use q as p and rather than np as q in that case q will be equal to 0.01.

So, I will use q as p in that case what will happen the probability will become smaller. So, when probabilities are smaller in that case I can use, convert the binomial distribution as the Poisson distribution.

So, if p is let us say 0.01 and then what will happen q will be 0.99 so, mean will become n into p. So, if let us say n is 1000. So, 1000 into 0.01 will give me 10 so, my mu will become 10 in that case, I will be able to use the poison distribution with mu equal to 10. So, as we see here that binomial distributions many times is difficult to evaluate. So, we may be able to use the Poisson distribution as the.

(Refer Slide Time: 31:53)



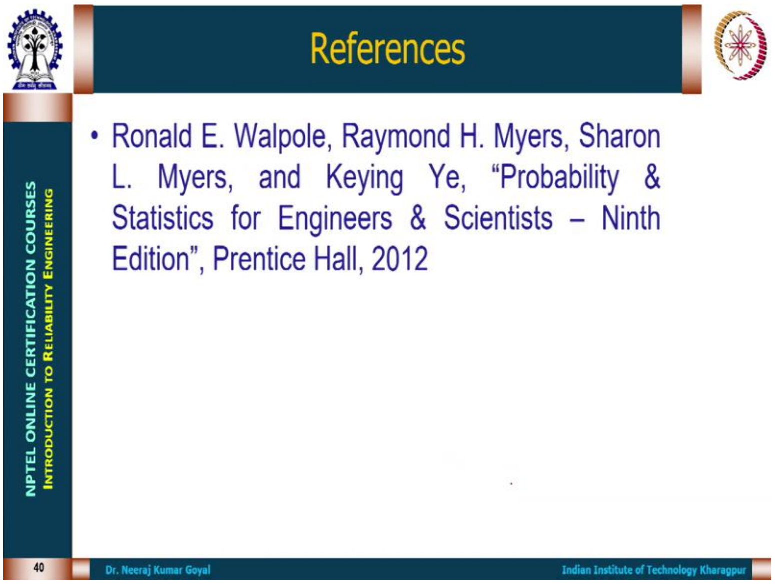
Let x be the random variable which is representing number of failures and subsequent repairs of restorable over a 1 year period assuming x as a Poisson distribution with mean failure rate of 2 failures per year.

So, lambda is 2 per year. And we want to know the probability that there is no more than one failure no more than one failure means, we want to know the probability of x equal to 0 and 1. So, probability that x is less than equal to 1. So, I want to know the probability of cumulative probability up to 1, F 0 or 1.

So, x equal to 0 to 1 e to the power minus, now 2 per year failure and I want to know for 1 year so, that will be 2 into 1 that will be mu will be equal to 2. So e to the power minus 2, 2 to the power x divided by factorial x, now, if I put x equal to 0 this will become e to the power minus 2 into 2 to the power 01 so 1 plus if I put x equal to 1 this will become 2 divided by factorial 1, 1 plus 2. So that will become 3 into e to the power minus 2 which if we saw we get 0.406.

$$P(X \leq 1) = F(1) = \sum_{x=0}^{1} \frac{e^{-2}2^x}{x!} = 0.406$$

(Refer Slide Time: 33:05)



"So, that's all for now. For more details, you can refer to the Walpole book which can help you understand probability and the importance of continuous distributions in reliability theory. We will discuss them one by one in our upcoming lectures. Thank you."