

Probability Methods in Civil Engineering
Prof. Rajib Maity
Department of Civil Engineering
Indian Institute of Technology, Kharagpur

Lecture No. # 22
Conditional Probability Distribution (Contd.)

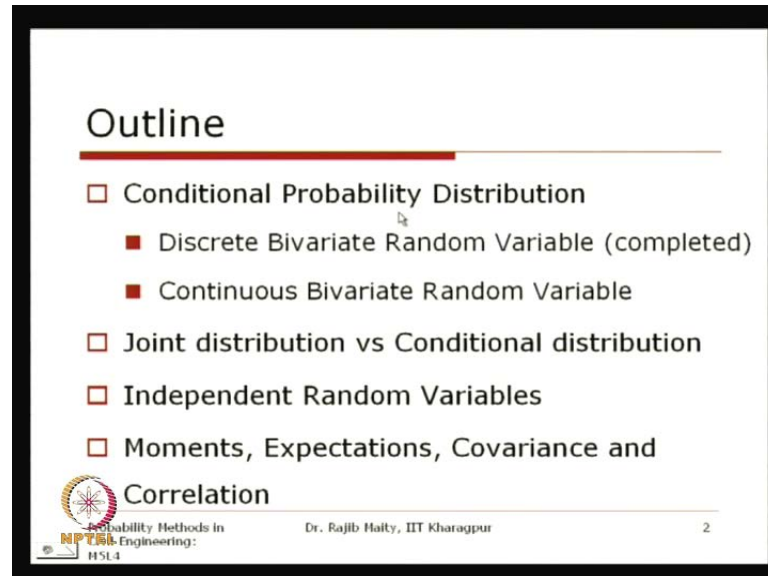
Hello, and welcome to this lecture of this module. We are discussing multiple random variable in this module, and so far what we have covered that their joint distribution, marginal distribution. And from last lecture, we started that conditional probability distribution for the multiple random variables. And in last lecture, what we have seen we have understood the concept, and particularly with respect to the discrete random variable. And how we can get their conditional probabilities, we have discussed through a problem which is where the random variables are in discrete nature. We will continue the same concepts, same conditional probability discussion on this conditional probability density, and this we will continue for the continuous random variable from this lecture.

So, basically we are continuing that the conditional probability distribution which we will basically cover in this class for the continuous random variable. And we will discuss some problems, where we need to deal with the continuous random variable that we will see. And after that, we will go gradually to the discussion and different measures of association of different random variables; this is very important, because when we are talking about that different random variables, and we are considering their joint occurrence, their joint probabilistic behavior. So, basically we want to know their kind of association with each other or sometimes if the number of random variables is more than two, then we may be interested to know that, what association is two random variables when the other random variables are either partial doubt or taken care of.

But basically as you know from this last couple of lecture that we are basically starting with the two random variables which is the special case, which is known as the bivariate random variable. We will also discuss those measures in terms of the two random variables, and the similar concept can be taken forward for the more than two random

variables. So, that will be our basically I should not say that main goal, but one of the most important goal to understand the concept of multiple random variables.

(Refer Slide Time: 03:04)



So, to start that we will first see that continuous random variable in case of the continuous random variable, how we can get this conditional probability distribution, and as I told that this discrete bivariate random variable case means, in case of this conditional probability distribution, that we have completed in this last lecture. And today, we will continue with the same concept of conditional probability distribution, and we will know how we can determine this in case of the continuous random variable.

Then also we will go with the joint distribution versus conditional distribution; this is in case of the both the cases means discrete as well as continuous, then the concept of independent random variable we will see in the light of this conditional probability distribution and the sum of probabilistic condition. We will see, to declare two random variables to be independent and after that gradually we will move to different measures of association among the random variables that we are dealing with so far, in this module through, their moments expectation covariance correlation may be all this property will not be covered in this lecture itself but gradually this will be the overall outline, how we will proceed one after another, so we will continue now with this continuous bivariate random variable for the conditional probability distribution.

(Refer Slide Time: 04:41)

Conditional Probability Distribution of Continuous Bivariate RV

□ Conditional Probability Density Function

- For continuous variables, the conditional density function of Y given X is:

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$$

- Hence:

$$f_{X,Y}(x,y) = f_{X|Y}(x|y)f_Y(y) = f_{Y|X}(y|x)f_X(x)$$

 Probability Methods in
Electrical Engineering:
HSE-4

Dr. Rajib Halty, IIT Kharagpur

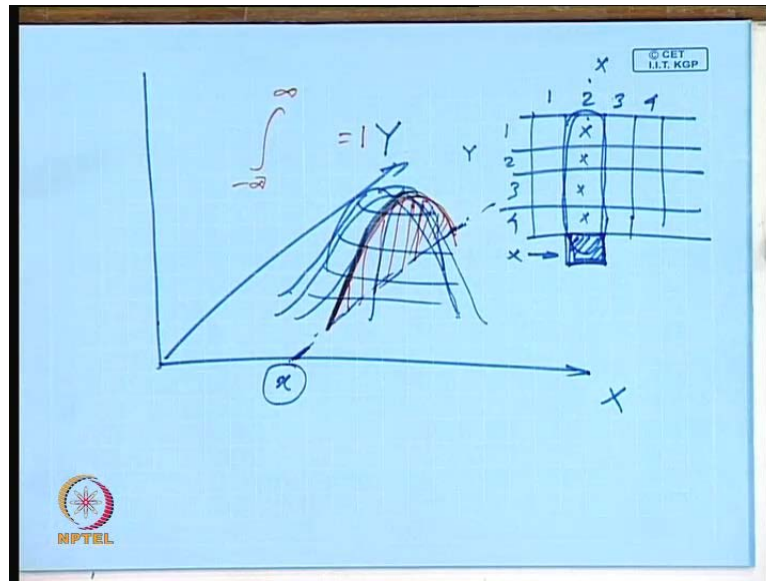
3

So, if we recall from our last lecture, that this conditional probability distribution is what we call it, when the information of the one random variable is known to us. For example, this is that distribution of Y given X and we have seen in the last lecture, in case of the discrete random variable that; this is the joint distribution in case of discrete; it is pmf, that is joint pmf we have seen, and that divided by the marginal of the one random variable on which the distribution is conditioned on. So, here as it is conditioned on the X we are using this $f_X(x)$; which is nothing but the marginal probability distribution for X.

So, in the last lecture, we have seen that, why we need this **need this** means; this form actually mixes; that mixes it suitably, first of all this conditional density is itself, and is a probability density, so the properties that the common density function follows; it should be followed by this one also. So, the first one you know that this should be non-negative.

And second one you now that, this should be that summation over the entire support should be equal to 1. So, here the summation, if you want to know, then that you know that, we have to do the integration over the entire range of the random **random** variable concerned and that should be equal to 1, so basically that the pictorial representation that we give in the **the** last lecture.

(Refer Slide Time: 06:39)



If you refer to that once again here, on this writing pad that; this one suppose that, if you take that; this is that **this is that** two axes or the two random variable axes representing two random variables, and if we are having some kind of surface here.

So, if this is the joint **joint** distribution of this two random **random** variable; that means, when we are talking about this is condition on the X so; that means, this one this particular value of this X, I am just talking about. So, I am looking through if this is; this has occurred. So, what is the distribution of Y, now when we are talking that on condition of this; basically you are taking, a taking you are looking at the cross section; that cross-section is passing through this value X. Now, that in case of the discrete; also we have seen in the last class; that there are say depending on how many specific outcome that the X and Y can have and one example, we have shown that the X can take the 4 specific outcome 1, 2, 3, 4 and Y also can take 1, 2, 3 and 4.

So, then for the condition on that X is equals to 2 then we have considered only in this column. So, this is for the discrete case and similarly, if I just take the same concept to the continuous random variable; that means, I am considering a section of this joint pdf through this line that, where this X is given to you which is equivalent to considering the one column. Now, you have also seen that, if this is the marginal distribution of this of the X then you know this is the summation.

So, first entry, second, third and fourth, so this probability that summed up to get the marginal distribution at x equals to 2, so and that one was taken and divided by this two so that the summation of this all probability is equals to 1, so equivalent in this continuous case, what we will take that so this is the marginal **marginal** distribution that at this section.

So, whatever the section that, I see here in this section; through this section suppose that, this is a section that, I see through this one, through this X . Now, the total this one this whatever the total means; total that whatever the **whatever the** distribution through this x , if that is divided by this joint distribution **distribution** then the new, then the final distribution that we will get that will follow that the second property of a probability distribution function is that; its integration over this range, over this entire range of this y should be equals to **should be equals to** 1.

So, this is the fact that, we are **we are** taking that this $f(x)$ there is a marginal distribution of this of the random variable on which the distribution is conditioned on. So, the expression for **the expression for** the conditional distribution says that **that** if $f(x, y)$; this is the joint distribution that divided by its marginal of the random variable on which it is conditioned will get the conditional distribution.

So, this also we can represent if you want to know that, just after some algebraic just take this one, this side then we get that the joint distribution of x and y is equals to probability of I think from this equation we will get this expression directly that is f , the distribution of y on condition x ; that is y given x **y given x** multiplied with the marginal of x .

Similarly, you can also if you **if you if you** see that, if you just change this one; that is $f(x)$ **$f(x)$** given y then this will be the joint distribution; this will have no change because this is the joint distribution between X and Y , but this marginal will change of instead of x it will be y the marginal of y , you have to consider. So, there also if we take the joint distribution will equate then it will come that $f(x)$ given y x given y multiplied by the marginal of the random variable Y . So, these two expressions are equal because both are equal to their joint distribution.

(Refer Slide Time: 11:40)

Conditional Probability Distribution of Continuous Bivariate RV...Contd.

□ The cumulative distribution function of the conditional variable is:

$$F_{Y|X}(y|x) = \int_{-\infty}^y \frac{f_{X,Y}(x,y)}{f_X(x)} dy$$

 Probability Methods in Engineering: HSE-4 Dr. Rajib Maity, IIT Kharagpur 4

Now, suppose that we are considering that the Y given X. Now, if I want to know the cumulative distribution of that **distribution of that** conditional distribution Y given X, then we have to do the integration of, because this is now the this full expression is now the distribution for that conditional probability density. And so we have to integrate it from the left extreme to a specific value Y, with respect to y only, because we are talking about the Y random variable Y here to get the cumulative distribution function.

(Refer Slide Time: 12:21)

Example on conditional pdf

Q. A storm event occurring at a point is characterized by two variables, namely, the duration X of the storm, and its intensity Y, which is defined as the average rainfall rate. The joint distributed variates of X and Y is assumed to follow exponential bivariate distribution given by:

$$f_{X,Y}(x,y) = [(a+cy)(b+cy) - c] e^{-ax-by-cxy}$$

where $a=0.06$, $b=1.3$ and $c=0.09$. Find the conditional probability that a storm lasting $X=4$ hours will exceed an average intensity of $Y=3$ mm/h.

 Probability Methods in Engineering: HSE-4 Dr. Rajib Maity, IIT Kharagpur 5

We will take of **take of** one problem here, which is on this storm event the similar problem was taken earlier; but here we will specifically discuss about the conditional distribution and we will see that, how we can get that conditional and how it is, how we can apply that whatever the conditional distribution we have **we have** obtained from this just now what we discuss.

So, this problem is on a storm event; which is occurring at a point and it is characterized by two variables and these two variables are; one is that the duration of the storm event and second one is the intensity of the storm event or we can say that; this is the maximum intensity. So, what happens in general that if we see the behavior of the **of the** storm event then generally, if the duration increases the maximum intensity decreases so and vice versa so that there could be some probabilistic relationship between this two; this two variables basically this storm events are characterized **characterized** by 3 variables along with this duration maximum intensity as well as the total depth of rainfall.

So, for example, if we see that once storm event which is **which is** lasting for 1 hour then the maximum intensity, and on the other hand, if we take the duration which is; which for 6 hours then the maximum intensity; this two for the 6 hours maximum intensity will be lesser than the storm event of this 1 hour **one hour** duration. As, I told that this storm events are generally characterized by with also another random variable; which is the total depth of rainfall, but here as we are discussing the bivariate case now, we are considering the two random variables, so for a storm event we have considered only two random variables to attributes; one is that its duration another one is maximum intensity.

So, there are two random variables now; one is X and other one is Y, and their joint distribution of this X and Y is assumed to follow an exponential bivariate distribution, which is given by this expression $f(x, y)$ this is joint distribution; which is equal to $a + cy$ multiplied by $b + cy$ minus c exponential minus a times x minus c times y ; this a , b , c are different constants and that generally depends on the specific location and that and this has to be determined.

Where it is supplied that; this a equals to 0.06, b equals to 1.3 and c equals to 0.09, so with these parameters with the values of this constant. So, our joint distribution is completely defined, now which is look for is to find the conditional probability that a storm event lasting X equals to 4 hours like so the duration is 4 hours will exceed an average intensity

of 3 millimeter per hour. So, here the unit of the intensity is millimeter per hour and that X is your in hours. So, which is actually this constants are **are** of course, the related to the units of this random variable; that is being **being** considered. So, now, for the as I just told that if the duration increases then the **then the** intensity will decrease, so here the question is that.

After giving this joint distribution question is so far, the storm duration is X. So, now, the conditional probability we should obtained on condition X; that is distribution of Y on condition X. Now, condition of X is that; X is equals to 4 hours. Now then we have to determine, what is the average intensity? So, the **the** probability that the average intensity will exceed 3 millimeter per hour; this is **this is** what we have to determine.

(Refer Slide Time: 16:56)

Example...Contd.

Sol.:

Since the conditional pdf of the storm intensity for a given duration is

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{[(a+cy)(b+cy)-c]e^{-ax-by-cy}}{ae^{-ax}}$$

$$= a^{-1}[(a+cy)(b+cy)-c]e^{-by-cy}$$



Probability Methods in
 Engineering: HSL4

Dr. Rajib Maity, IIT Kharagpur

6

To solve this one; first of all we have to **we have to** obtain this conditional distribution and you know that for to obtain this conditional distribution, we have to know its joint distribution as well as the marginal distribution of the random variable; on which this conditioned on, as we have seen that; this is conditioned on the x, here x is the duration of this storm, so we need to know the marginal density of x, and in the earlier lecture, if you refer to them from there you can see that; for this type of distribution the marginal distribution of the x is on **is on** exponential distribution and it with the **with the** parameter a.

So, this f_X is the **the** marginal distribution of x is, a exponential minus $a \cdot x$ so this is the exponential distribution for the x . Now, once we know this **this** marginal distribution; that means, it is straightforward to get that their conditional distribution it is just the division of this joint distribution which we know, and for its marginal distribution which is a e power minus x . So, we get the form of this **form of this** distribution that is conditional distribution and of course, here what it is not mention here is that, the both x and y you know that for the exponential distribution support from 0 to infinity. So, here as we are this both; this x and y are there are. So, this support for this x and y both are from 0 to infinity. So, this is the complete description of this joint distribution of y given x .

(Refer Slide Time: 18:42)

Example...Contd.

The conditional cdf is:

$$F_{Y|X}(y|x) = \int_0^y a^{-1} [(a + cy)(b + cy) - c] e^{-by - cxy} dy$$

$$= 1 - \frac{a + cy}{a} e^{-(b+cx)y}$$

which yields,

$$P[Y > 3 | X = 4] = 1 - F_{Y|X}(3|4) = 1 - 1 + \frac{0.06 + 0.09 \times 3}{0.06} e^{-(1.3 + 0.09 \times 4)3}$$

$$= 0.038$$



Probability Methods in
Engineering:
H5L4

Dr. Rajib Halty, IIT Kharagpur

7

Now, we have to **we have to** get its **its** probability so we can directly do that integration or what we can do, we can also get that cumulative density first and cumulative density, we have seen just from previous slide that; as it is that on the y on condition x . So, we have to do the integration from the left extreme to the specific value of y to get its **get its** cdf; that is the conditional cdf, conditional cumulative distribution function. So, at the left extreme is here 0, so 0 to y and this is the conditional distribution function; this we have to integrate with respect to y and after this integration and we have to follow some integration by parts also and you will see that, it comes to the form of 1 minus a plus cy divided by a exponential minus b plus cxy , so this is the cumulative; this is the conditional cumulative distribution function for y given x .

So, now, I think just we have to put those values here we want to know that what is the exceedance probability of this Y greater than 3 given; that X is equal to 4 and you know that, as it is given that Y greater than 3. So, if we just detect from the total probability which is equal to 1, then it will be that Y less than or equal to 3, given X equal to 4 so that means, that is the $f(y, x)$ given 4 means Y less than or equal to 3, given X equal to 4.

So, we just put this value of Y equal to 3, and X equal to 4 here with the other parameter that was defined and we get the probability that; it will exceed at 3 millimeter per hour for a 4 hours duration storm is equal to 0.038.

So, you see from this example that, if we know the joint distribution and the corresponding marginal distribution then this type of any **any** such question can be answered. So, one question we have seen that what is the probability that it will exceed three millimeter per hour for a 4 hour storms and we have seen that the probability is very less. So, it will help to take a decision find as the probability is very less; that is 0.038 then we can say that getting more than 3 millimeter per hour may be very less probable. So, like **like** this any of this type of answer we can get.

But one thing that, I was mentioned earlier also here is that to, if we know the marginal distribution for two random variables then it is not that easy to know, what is that joint distribution, only one thing we can say, from based on the theory is that, if the joint distribution is yours is joint normal distribution or joint Gaussian distribution then, what we can say that their marginal also will be a normal distribution, but again that **that** the reverse that is, it we cannot say it is vice versa.

So that if I say that two random variables X and Y ; both are having the marginal density of normal distribution X and Y we cannot say that the joint distribution, what is that joint distribution so even though we are starting this type of problem that this joint distribution is given to us. So, getting the joint distribution from the marginal's need not be very easy task and in this module towards the end we will discuss about a recently or relatively newer concept in the civil engineering; which is known as copula and this copula is generally one of the latest thing that is there to get their joint distribution that, we will discuss later for the time being we are starting of whatever the problems we are

describing **we are describing** that their joint distribution is known. So, this is the question ask for and we got this probability for this particular condition.

(Refer Slide Time: 23:25)

Independent Random Variables

□ Two random variables X and Y are statistically independent if and only if their joint density (pdf) is the product of their marginal densities, that is,

$$p_{X,Y}(x,y) = p_X(x) p_Y(y) \quad \text{for discrete RV's}$$

$$f_{X,Y}(x,y) = f_X(x) f_Y(y) \quad \text{for continuous RV's}$$



Probability Methods in
Engineering:
H51.4

Dr. Rajib Hait, IIT Kharagpur

8

Now, we will discuss about another important concept is called the independence and in the light of this conditional distribution. So, first of all if we just directly state, what is how we can decide that whether two random variables are independent or not mathematically. So, there are two straightforward expressions are there; one for discrete and other one for the continuous, but before I got to that **that** mathematical expression. If I just want to know, what does this independence means simply for two random variables, if we **if we** declare that the outcome of one random experiment does not have any influence on the outcome of the other one then we can say that this two are independent.

Now, to test this one mathematically we have to **we have to** take the help of that joint distribution as well as their marginal distribution that is two random variables. If it is available it is not necessary that two random variables only we can say even more than two; that is, if it is more than two or n numbers or random variables are available then, if you know their **their** marginal distribution of this n random variables. And if we know their joint distribution as well then the check is that, whether the product of the marginal distribution is equal to their joint distribution if this is true; then we can declare that this is independent.

So, the statement states that two random variables X and Y are statistically independent, if and only if their joint density pmf of course, pmf for the discrete their joint density is the product of their marginal densities, that is $p_{X,Y}(x,y)$ joint pmf between x and y is equals to that $p_X(x)$ marginal for x multiplied by $p_Y(y)$ marginal for the y . And this is for the discrete random variable and for the continuous joint the density pdf joint pdf between x and y is equals to marginal density of x multiplied by marginal density of y ; this is for the continuous random variable.

Now, in the light of their that conditional distribution there we will just see in a minute, how we can get this expression but in this light, what I like to stress is that, this is to declare it is statistically independent; this is the condition should be if and only if. So, this one this price I want to stress that is; this is the condition if it is satisfied then only we can say that X and Y are independent.

(Refer Slide Time: 26:31)

Independent discrete random variables

- If the events $(X=x)$ and $(Y=y)$ are statistically independent,

$$p_{X|Y}(x|y) = p_X(x)$$

and $p_{Y|X}(y|x) = p_Y(y)$
- Also

$$p_{X,Y}(x,y) = p_{X|Y}(x|y)p_Y(y) = p_{Y|X}(y|x)p_X(x)$$
- The above equation can be written as:

$$p_{X,Y}(x,y) = p_X(x)p_Y(y)$$



Probability Methods in
Engineering:
H5L4

Dr. Rajib Halty, IIT Kharagpur

9

Now, in the light of the conditional distribution that we discussed, so far if you see that is we have told that the outcome of one random variable does not have any influence on the other. So that means, here that there are suppose this is for the discrete case there are some two events; one is that X equals to x specific outcome, and Y equals to y ; this two are statistical independent, then obviously that probability that x given y should be equals to probability of that x ; that is marginal of the

x, because just now we know that this the outcome of one had no influence on the outcome of the other.

So, the whenever we are telling that this x given y, so this given y have no influence on this joint distribution, that means; that the probability of this x given y is equals to the probability of x, because this is have no influence of the outcome of other and vice versa. The probability of y given x is equals to the probability of y; that is the marginal of marginal density for y, and also from the conditional distribution, what we have seen that just few minutes back that this joint density is equal to that conditional multiplied by the multiplied by its marginal so this will be p, not f_p means we are using the notion p for the discrete random variable.

So, this condition multiplied by its marginal, so it is condition on y so multiplied by the marginal of y or the conditional density condition of x multiplied by its marginal density. Now, if you just replace this conditional probability with this what we **what we** have discuss just now that is p_x on condition y is equals to p_x or the p_y on condition of x which is equals to p_y, then all this expressions leads to that this joint density; that is p(x, y) is equals to their product of their marginal; so, this is in case of the discrete random variable.

(Refer Slide Time: 28:58)

Independent continuous random variables

- If the events $X=x$ and $Y=y$ are stochastically independent, then

$$f_{X|Y}(x|y) = f_X(x)$$

and $f_{Y|X}(y|x) = f_Y(y)$
- Again, $f_{X,Y}(x,y) = f_{X|Y}(x|y)f_Y(y) = f_{Y|X}(y|x)f_X(x)$
- Therefore $f_{X,Y}(x,y) = f_X(x)f_Y(y)$



Probability Methods in
Engineering:
H5L4

Dr. Rajib Halty, IIT Kharagpur

10

And similarly, if we see that for the continuous case also the similar thing; only in case here we have to consider the pdf not the pmf, so that pdf of this; that is x on condition y it

should be equal to its marginal, that is $f_X(x)$ and similarly $f_Y(y)$ given x that is pdf of y given x should be equal to pdf of y alone, because the outcome of this y does not have any influence from the outcome of x . So, following again the similar concept here, also we can just write that; this joint distribution we know that is equal to the conditional distribution multiplied by its marginal and it is conditional on y . So, marginal is $f_Y(y)$ here and conditional on x if it is conditional on x that is given x then it should be multiplied with the marginal of x .

Now, this conditional probability, if we replace from this their marginal itself then both this expression will lead to that joint density joint pdf is equal to the product of their marginal distribution; that is $f_X(x)$ multiplied by $f_Y(y)$. So, these are the two conditions; one for the continuous another for the discrete one which should be satisfied and if this is satisfied then only we can say that this X and Y are statistically independent.

(Refer Slide Time: 30:29)

Example

Q. The amount of water X that can be supplied through two different routes of water distribution system. First (usual) route supplies at a rate x_1 per hour. Second, less amount of supply, $x_2 (< x_1)$ is also provided if usual route is needed to be closed for maintenance and the probability of facing such situation is π . If the hourly demand, denoted by, Y , is met, system is considered to be working fine, otherwise failed.



Probability Methods in
Engineering:

Dr. Rajib Haity, IIT Kharagpur

11

Now, we will take another problem on this continuous random variable, where we will **we will** say discuss again the similar thing just the previous problem was based on this discrete random variable. And here we will take another problem on this water distribution system, and where that when we will see that **that** is the probability of its failure.

Just little background to this problem is that some time for this water distribution system we use alternate path and from the if everything is usual, then we determine that **that**

peak demand, and just on that this supply is determined and of course, the supply is determined with the various factors; that is its socioeconomic factor water availability and sort of things is there. So, we determine this much water is the available and also that demand is also variable you know that in during the peak **peak** hours the demand is more, and during the lean hours the demand is less. Now, when we can say that this system is working fine is that whenever the supply of the water is sufficient or sufficient in the sense; that it is the more than, what is demand Then we say that system is working fine, but if the supply is less than whatever **whatever** the demand then the system is declared as failure.

Now, the failure of a system is, what additional distribution system is the also can have manyfold reasons; one reason here that we have taken in this problem is that the sudden requirement of this closing of the path. So, sometimes there are some alternate routes are available, but this alternate routes generally having due to the **due to the** various factor, if it is possible to supply only very less amount of water through the other **other other** path.

So, this is also now that the less amount of water, if it is during the peak hours there is chance that the **the** required demand may not be supplied during the peak hours so that time the system generally fails. So, based on that of this now, there are two different supplies where we can say that whether, which supply is available now and again depending on the hourly demand also varies. So, for example, that from the peak hours also says that morning 7 am to 8 am or 8 am to 9 am.

So, what is that, what is the demand and in this demand times there are we can now imagine that this, if we consider these are the random variables then the this random variables are also related to each other and dependent on each other before we can determine the probability of failure. So, one such example problem is taken here in the context of that on the conditional probability for the continuous random variable.

So, the problem states the amount of water that can be supplied through two different routes of the water distribution system. So, the among this two different routes the first one is the usual route and this through this first route the water can be supplied at a rate of say X_1 unit per hour, and second through the second route the less amount of supply is possible; which is denoted as X_2 and X_2 is less than **less than** X_1 ; which is also **also**

provided if that is the second route is used if the usual route is needed to be closed for the maintenance or any other reason.

And the probability of facing such situation; such situation means here, that I am using the alternate path not the usual one. So, this say situation the probability of this situation is p_i , so whether the supply of the water is supplied to the community through the usual path or in or through some alternate path. So, the probability of supplying the water through the alternate path is p_i and obviously, this will be less. Now, if the hourly demand denoted by Y is met then the system is considered to be working fine otherwise it is failed, now this now once we are declaring that is the hourly demand, if we consider that two different hours.

(Refer Slide Time: 35:24)

Example...contd.

The distribution of Y is

$$F_Y(y) = \lambda e^{-\lambda y} \quad \text{for } y \geq 0$$

Let the random variables Y_1 and Y_2 represent the demand per hour during the peak hours commencing at 7 am and 8 am respectively. Let A represent the failure of the system in these peak hours, i.e., required amount is not supplied to meet the demands. What is the probability of the event A ?



Probability Methods in
Engineering:

Dr. Rajib Maity, IIT Kharagpur

12

For example, here that two different times; one is the 7 am and 8 am which are considered to be the peak hours is **isis** considered for two different random variables Y_1 and Y_2 , before that this demand can follow a specific distribution and it is declared that this distribution of this y ; that is the demand is hourly **hourly** demand, it follows a distribution that is a lambda exponential $\lambda e^{-\lambda y}$, for **for** y greater than equal to 0, you know that, this support for this exponential distribution is 0 to infinity.

Now, if we consider the two random variables here that is Y_1 and Y_2 , which **which** represents the demand per hour during the peak hours commencing at 8 am and

commencing at 7 am and 8 am; that means, are 7 to 8 and 8 to 9 respectively this two are the requirement that **that** demand is also **also** vary.

But any way we have considered that both the things are following for the simplicity sake we have to assume that both the demands following the same distributional form; one is that $\lambda e^{-\lambda y}$ which is an exponential we could have even consider that two variables here is following two different type of distribution; means the type of distribution can be kept same that is the exponential distribution, but this parameter can be changed that is for the Y_1 it is λ_1 and for the Y_2 it could have been λ_2 , but just for the demonstration purpose here it is considered that both are following the same distribution.

Now, let A represents that the failure of the system in this peak hours, that is a required amount of water is not supplied to meet the demands. Now, the question is what is the probability of the event A event at this is, this is a spelling mistake event A so that during this peak hours; that means 7 to 8 and 8 to 9 this two time whatever the demand is that is Y_1 and Y_2 this required supply is not there.; that means, that Y_1 is greater than **greater than** either x_1 or x_2 whatever is being supplied or that Y_2 is greater than that x_1 or x_2 whatever the supply during that time, so, have to determine the failure of the system through this.

(Refer Slide Time: 38:08)

Example...Contd.

Sol.:
The failure of the system can be written as:

$$P[A] = 1 - P[Y_1 \leq X, Y_2 \leq X]$$

i.e taking the complementary probability

Also, it is known that

$$p_X(x) = 1 - \pi \quad \text{for } x = x_1$$

$$= \pi \quad \text{for } x = x_2$$



Reliability Methods in
Engineering:
H5L4

Dr. Rajib Maity, IIT Kharagpur

13

To solve this one the failure of the system can be written as like this, so this is the a that event we have declare and so that means, as I told that Y_1 should be greater than that X that is the supply which is you know that this supply here is this type of random variable is call that dichotomous **dichotomous** means it can take only two possible values. So, here we have seen that it can be it can take only the two values; one is x_1 other one is x_2 . So, x_1 is the usual supply and X_2 is the supply through the alternate route so the failure means Y_1 should be greater than, what is the supply that, Y_1 should be greater than X or that Y_2 is greater than greater than X .

Now, the failure is also here it is represented as A_1 minus this Y_1 less than X and Y_2 less than equals to **less than equals to** X , because of this the help from the cumulative density the concept of the cumulative density you know that we need to know that less than equal to. So, we have just considered that; this is 1 minus of this either that Y_1 less than equal to X and Y_2 less than equals to that supply X .

Now, it also known that so far, as the distribution of this x is concerned that is that I told that it is a dichotomous variable, so it can take only two possible value **value**; one is that x_1 and other one is the x_2 and in the problem it is declared that probability of facing such situation that; such situation means, when we supplying the water through the alternate path that is x_2 is the x_2 is the supply so that probability is your p_1 and. So, the probability of supplying that x_1 amount; that is when x equals to x_1 ; obviously, this should be 1 minus p_1 , because this two are the mutually exclusive events. So, either of this two should occur, so if the probability of one event is p_1 it should be that for the probability of the other one should be 1 minus p_1 .

So, now this is the marginal; this is basically the marginal distribution of that **of that** x and which you are can say that for this two different case.

(Refer Slide Time: 40:42)

Example...Contd.

- Using the conditional probability and considering the possible closure of one route between 7 and 8 am, the above equation is expanded as:

$$P[A] = 1 - \left\{ P[Y_1 \leq x_1, Y_2 \leq x_1 | X = x_1]P[X = x_1] + P[Y_1 \leq x_2, Y_2 \leq x_2 | X = x_2]P[X = x_2] \right\}$$

$$= 1 - \left[F_{Y_1, Y_2 | X}(x_1, x_1, x_1)p_X(x_1) + F_{Y_1, Y_2 | X}(x_2, x_2, x_2)p_X(x_2) \right]$$


Probability Methods in Engineering: HSI 4

Dr. Rajib Maity, IIT Kharagpur

14

Now, using the conditional probability and considering the possible closure of one route between that 7 and 8 am, the above equation is can be expanded that is the probability of failure this thing can be expanded that. So, 1 minus as I just let once again refer to this one, so 1 minus $Y_1 \leq x_1$ and $Y_2 \leq x_1$. So, we are now expanding this **this** term that is $Y_1 \leq x_1$ and we are also expanding that $Y_2 \leq x_1$.

How we are doing that is that, the probability of Y_1 that is between 7 to 8 is less than x_1 comma probability of $Y_2 \leq x_1$, because we are assuming both the times supply is usual on condition that X is equals to x_1 which is multiplied by its marginal. So, this is the conditional multiplied by its marginal that is X is equals to x_1 this and the other situation is the supply is x_2 ; that is probability that $Y_1 \leq x_2$ and $Y_2 \leq x_2$ on condition that supply is x_2 this multiplied this full quantity multiplied by the what is the probability of x_2 ; that means, this is the cumulative conditional distribution for this Y_1, Y_2 when this variable when the x , now in the supply is x_1 and this is the cumulative distribution when the supply is x_2 .

So, which is represented through this expression though this cumulative distribution multiplied by the marginal of that X for the value x_1 and this cumulative multiplied by the marginal for the **for the** value X equals to x_2 .

(Refer Slide Time: 42:39)

Example...Contd.

□ The available supply X is assumed to be independent of the Y_1 and Y_2 s, thus

$$P[A] = 1 - \left\{ P[Y_1 \leq x_1, Y_2 \leq x_1]P[X = x_1] + P[Y_1 \leq x_2, Y_2 \leq x_2]P[X = x_2] \right\}$$

$$= 1 - \left[F_{Y_1, Y_2}(x_1, x_1)P_X(x_1) + F_{Y_1, Y_2}(x_2, x_2)P_X(x_2) \right]$$

□ Assuming that Y_1 and Y_2 are independent

$$F_{Y_1, Y_2}(x_1, x_1) = P[Y_1 \leq x_1, Y_2 \leq x_1] = P[Y_1 \leq x_1]P[Y_2 \leq x_1]$$

$$= F_{Y_1}(x_1)F_{Y_2}(x_1)$$

 Reliability Methods in
Engineering:
H51.4

Dr. Rajib Maity, IIT Kharagpur

15

So, this two are obtained to and after that, if the available supply x is assumed to be independent of the Y_1 and Y_2 here, one thing I want to mention that even though this is written here it is not for the practical when we decide that, what should be the supply for this one there lot of analysis is been is done to determine that what should be the supply, but once we have decided that, now the what is meant here for the independent; the independent means the demand and supply are independent what we mean that.

Now, the variable demand and the supply these two variables are assumed to be independent; that means, I have decided from some from the analysis that the this time the supply will be this much amount. Now, this one whether the distribution system needs the maintenance or not so that a due to the sudden maintenance and this type of reason we need to change the supply, but the demand is always, whatever the requirement of the whatever the requirement of the community so that is why it is declare that if we assume that this two event are independent to each other than how this expression is for the simplified.

So, that is why it is explain that, if the available supply X , either x_1 or x_2 is assumed to be independent of that Y_1 and Y_2 thus it can be written that; this Y_1 less than equals to x_1 , Y_2 less than equals to x_2 that condition that x is equals to x_1 is now, omitted which is multiplied by their marginal probability that is probability x equals to x_1 .

So, this is the cumulative distribution for this Y_1, Y_2 which is in case of this x_1 ; it is multiplied by the marginal probability of x_1 and other one, if it is x_2 then it is multiplied by its marginal probability $\times 2$.

Now, if we assume that this Y_1 and Y_2 are also independent; Y_1 and Y_2 independent means whatever what is the demand during the 7 am to 8 am and these two demand between the 8 am to 9 am, if this two demand are independent then we can write that this cumulative distribution that is for anything, that is for, if it is the now it is the explained here for the case of x_1 is equals to probability of y_1 less than x_1 and probability Y_2 less than x_2 and we know from the independence just now we discuss that. So, this is the joint distribution which should be equal to the their marginal; that is probability of Y_1 less than x_1 multiplied by probability of Y_2 less than x_2 .

So that; that means, these are that cumulative distribution for the Y_1 and Y_2 , so this cumulative distribution these multiplications should give that that cumulative distribution for the joint if this Y_1 and Y_2 are independent.

(Refer Slide Time: 46:05)

Example...Contd.

- Considering again that Y's are identically distributed

$$F_{Y_1, Y_2}(x_1, x_1) = [F_Y(x_1)]^2$$

$$\text{and } F_{Y_1, Y_2}(x_2, x_2) = [F_Y(x_2)]^2$$
- Then,

$$P[A] = 1 - \{ [F_Y(x_1)]^2 p_X(x_1) + [F_Y(x_2)]^2 p_X(x_2) \}$$
- Since $p_X(x_2) = \pi$ and $p_X(x_1) = (1 - \pi)$, we obtain

$$P[A] = 1 - \{ (1 - \pi) [F_Y(x_1)]^2 + \pi [F_Y(x_2)]^2 \}$$



Probability Methods in
Engineering:
H51.4

Dr. Rajib Maiti, IIT Kharagpur

16

So, this expression is now replaced in this form and what we will get the probability of failure is that, this now again; this cumulative distribution when we say that probability of y (x_1) multiplied by Y_1 and Y_2 and in the problem also for the simplicity sake we assume this two are following the same distribution same exponential distribution.

So, with the **with the** identical parameters well so that is why here we are just expressing this is to be square. So, multiplied instead of multiplying that $f(y_1, x_1)$ and $f(y_2, x_2)$ it is $f(y, x_1)$ square and this the other one that is $f(y, x_2)$ square.

Now, these two expressions are replacing that expression for the probability of failure which is equal to $1 - (f(y, x_1))^2$ multiplied by the probability of supplying x_1 amount plus probability of $f(y, x_2)$ that is a cumulative distribution of y for x_2 square multiplied by the probability of supplying amount x_2 .

Now, these two are known that probability of x_2 equals to π and probability of x_1 equals to $1 - \pi$. So, putting this expression we obtained the probability of A equals to $1 - (1 - \pi) (f(y))^2$ for x_1 whole square plus this π multiplied by $f(y)$ for x_2 square.

(Refer Slide Time: 47:38)

Example...Contd.

- We have distribution of Y as:
$$F_Y(y) = \lambda e^{-\lambda y} \quad \text{for } y \geq 0$$
- The failure probability of system can be written as:
$$P[A] = 1 - (1 - \pi) (\lambda e^{-\lambda x_1})^2 - \pi (\lambda e^{-\lambda x_2})^2$$



Reliability Methods in
Engineering:
H5L4

Dr. Rajib Maity, IIT Kharagpur

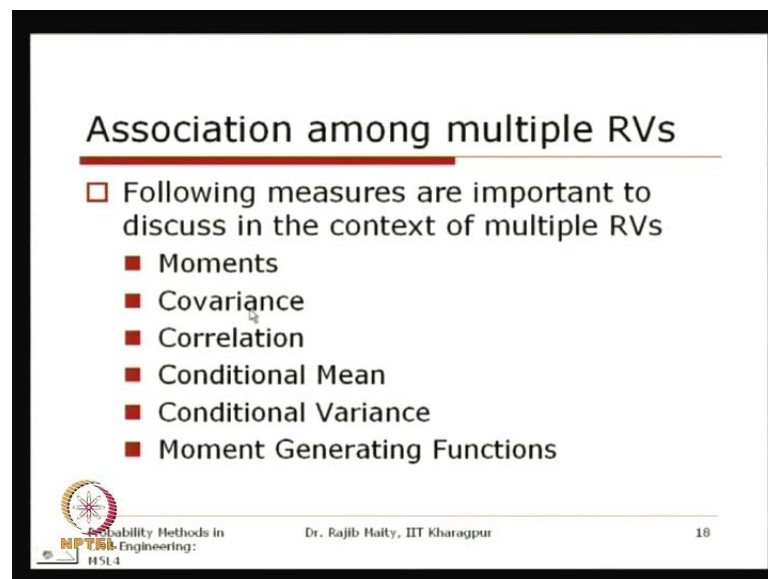
17

Now, you know these cumulative distributions are for the **for the** exponential distribution is also known that is $1 - e^{-\lambda x}$, this will be you know that **1** minus e power minus λx , if this is the exponential density then their cumulative density you know this will be $1 - e^{-\lambda x}$. So, this is a mistake it should be instead of $\lambda e^{-\lambda x}$ it should be $1 - e^{-\lambda x}$. So, this should be the cumulative distribution and the support is same y greater than 0.

So, this cumulative distribution should be placed here, that is $1 - e^{-\lambda x_1}$ and $1 - e^{-\lambda x_2}$ for the x_1 value square and $1 - e^{-\lambda x_2}$ power square. So, this whole quantity now these parameters we know as we have declared that, we know that distribution; so that means, we know these parameters also; that means, λ we know and this π we know, if we know this parameter we know that this what is the amount of that is x_1 and x_2 their magnitude. So, with this magnitude we can calculate that what the probability of this failure is as a system designer.

So, this type of probability we should know that this type of probability should be minimized. So, once it is satisfied by the designer the system designer then that the design can be accepted. So, the basic idea is that after getting this probability; this probability should be minimized to the extent possible so we have discussed.

(Refer Slide Time: 49:31)



Association among multiple RVs

- Following measures are important to discuss in the context of multiple RVs
 - Moments
 - Covariance
 - Correlation
 - Conditional Mean
 - Conditional Variance
 - Moment Generating Functions

Probability Methods in
Engineering: HST-4

Dr. Rajib Haity, IIT Kharagpur

18

So far, the conditional probability and the practical application in some of the civil engineering problem we have seen, and also now what the next thing that we just was referring to in this at the starting of this lecture is that, this whenever we are talking about this multiple random variable we are interested to know their joint association.

Suppose that, if we just start our discussion with the bivariate random variable only that is the X and Y there are two random variables are there. Now, how they are associated with I cannot say they are related with, because that is I that may not be the right word,

but the association when we are referring to that can be the linear association or the non-linear association or their **their** variability with each other their covariability between two random variables all these things are important and these things are also generally obtained from their probabilistic that their joint probabilistic behavior.

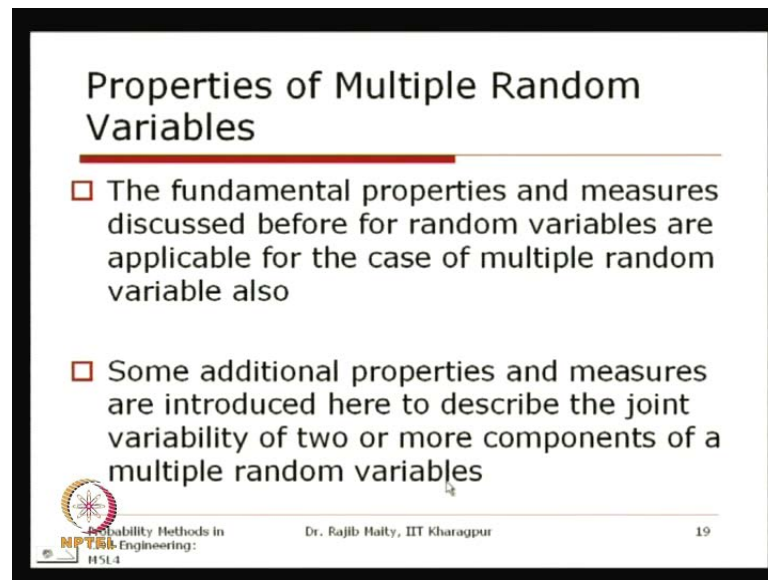
So, this will discuss now, one after another and we will go through this concept that is the one; is that moment then the covariance correlation conditional mean, conditional variance moment generating functions maybe all these things will not be covered in today's lecture, but we will just start with one after another. And we will know that, what the important information that we can obtain from this **from this from this** measures for the multiple random variables.

So, you also know that these moments and the expectations we have discussed in case of the single random variable and thus; obviously, here also the similar concept will be used, but it has to be extended to the higher dimension for the single random variable, what we have seen that, it is just one **one** excess if you just go for the pictorial representation on based on the one axis we have discussed their area concept, that is the moment when we are discussing we have discussed in the concept of the moment.

Now, if we just extend it for the bivariate case; that means, it is now a three dimensional space. So, two axes represent two random variable and other one is the probability density. So, now, if we suppose that the expectation if we just extend to the two dimensional case. Now, whatever we have considered for the area in the one dimensional that the single random variable, now the same thing for the two dimensional case it will be the volume.

Now, you can imagine for, if it is more than two then it is something going to be three dimensional for three random variables and one dimension for their probability density, so it can be extended according to it.

(Refer Slide Time: 52:21)



Properties of Multiple Random Variables

- The fundamental properties and measures discussed before for random variables are applicable for the case of multiple random variable also
- Some additional properties and measures are introduced here to describe the joint variability of two or more components of a multiple random variables

 Probability Methods in
Engineering: HSL4

Dr. Rajib Maity, IIT Kharagpur

19

So, any way to start with the bivariate case we can just go for the first is that the fundamental properties then the measures, that we have discussed for the **for the** random variables these are applicable for the case of the multiple random variable also. And some additional properties and the measures are introduced here to discuss their joint variability of the two or more components of the random variables.

So, as I was just telling that the **the** fundamental concept that we have used earlier for the single random variable will also be **will also be** applicable, and the same concept should be extended. First we understand for the two-dimensional, case and then we will go for we will take that for the one higher dimensional; at least that is for the mathematical expression purpose.

So, this detail discussion on this moment, and its expectation then one of the most important things is covariance, and from the covariance we will know that the correlation; correlation is the measure of the linear association. This linear is important this, then gradually we will go on the other properties as well for the multiple random variables. And this will be started discussing in the next lecture. **Thank you.**