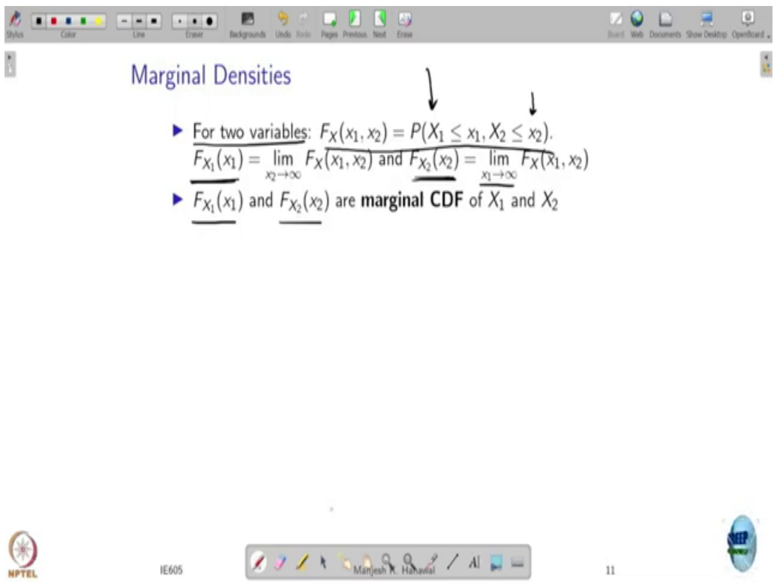
Engineering Statistics
Professor Manjesh Hanawal
Industrial Engineering and Operations Research
Indian Institute Of Technology, Bombay
Week: 03
Lecture: 12
Joint Distribution of RVs and Marginal Densities

(Refer Slide Time: 02:43)



Now based on this joint distribution we can recover the marginal. So, far we are talking about one-one random variables. So, if I am going to give you joint behavior of 2 random variables, you should be able to recover behavior of individual random variables. So, let us say I have 2 random variables and their joint probability distribution is given to me CDF is given to me. Now I am only interested in the one random variable X1 not about X2.

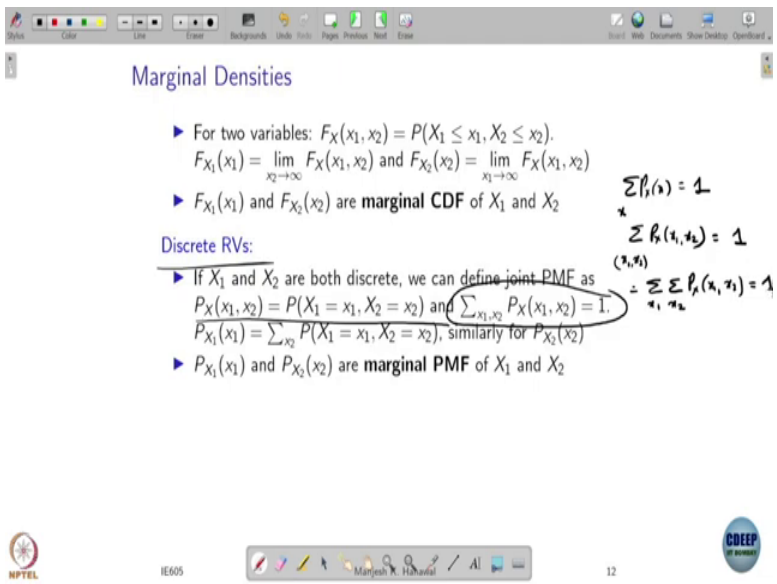
Now how to do its behavior, now one possibility to do this you let X2 go to Infinity in this function. And what does X2 letting go to Infinity means you are letting this go to Infinity that means you are allowing X2 ≤ Infinity that means you are allowing all possible values of X2, then any effect has to come from only X1, and that is what now we are able to recover the effect of random variable X1 and that is why we are just calling it a F(x1).

And similarly if your focus is only on X2 you similarly let X1 go to Infinity in this function you let X1 take all possible value whatever the effect remains it has to come from only X2. And in such case this F(x1) F(x2) called marginal CDFs.

Now which one will have if I am talking about 2 random variables, which one will have more information. Case 1, I will provide you joint CDF. Case 2, I will provide you individual CDF. Which one will have more information? Joint because if I provide you joint you can always recover the densities individual marginal but it is not always the case that if I provide individual behavior you may not be able to get the joint behavior.

So, joint you can get the marginal but from the marginal you cannot get the joint behavior that is why providing a joint information is always going to be difficult like I mean you need to have good amount of information.

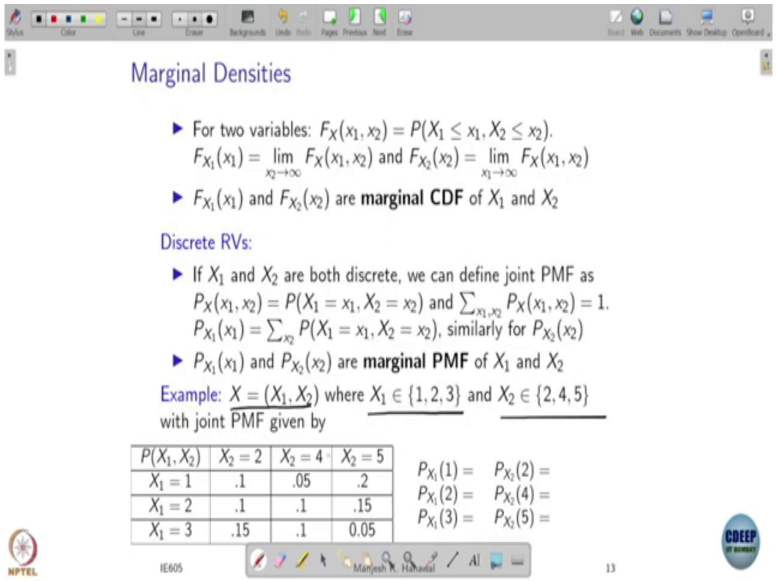
(Refer Slide Time: 03:00)



And this is something simple I will just skip this part I mean whenever you are going to deal with the discrete random variables you will just join probabilities is just asking X1 is taking x1 and X2 is taking x2 and obviously if I have this joint probability if I add over X1 and X2 this joint probability it has to add up to 1.

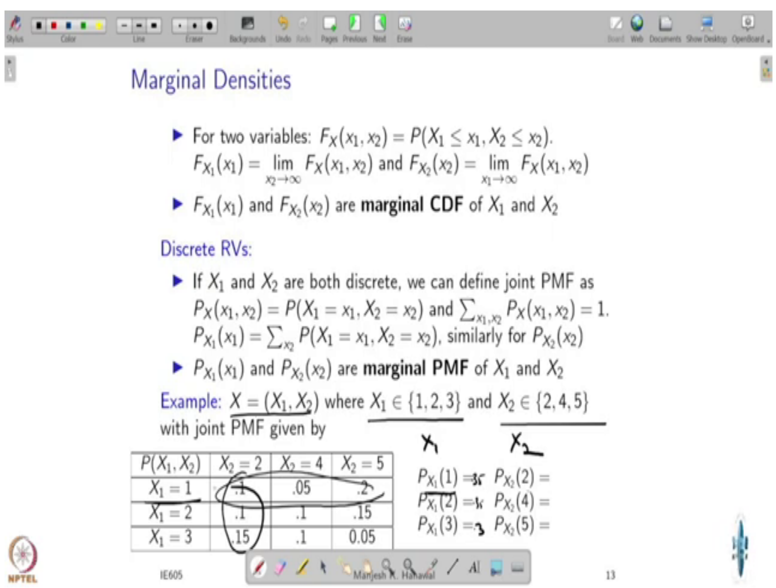
Earlier I have this like X if I add over all possible values of X it should one. Now it is a joint probability like X1 and X2 now X1 and I have to add over all possible values of X1 and X2 and then you should add up to 1. Or this is same as in double summation X1 and X2 and then P(x1, x2) this should be equals to 1. You understand this?

(Refer Slide Time: 04:00)



Now let us quickly do this exercise. Let us say I have 2 random variables X1 and X2 X1 takes 3 values 1, 2, 3 and X2 takes values 2, 4, 5 and I have given that joint probability mass function in this table. So, if I add it over all possible values in this matrix it should add up to 1 then only it is a valid joint probability mass function.

(Refer Slide Time: 04:41)



Now from this let us try to recover the marginals. First let us focus on random variable X1 and this is random variable X2, what is this probability? This is asking what is the probability that my random variable X1 takes value 1 I am interested in this value and in that case I do not care about what is the possible values of X2. So, I will just take all the values and you will get, what is the next one?

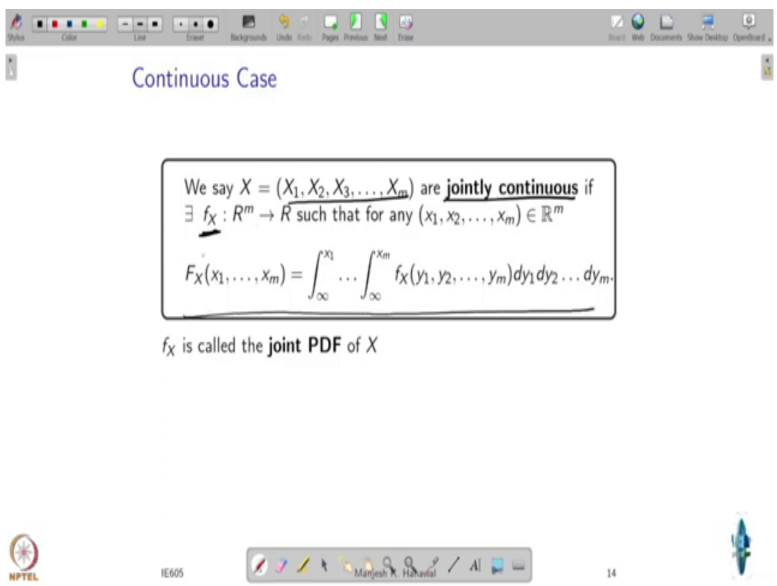
Student: .35

Professor: What is the last one?

Student: .3.

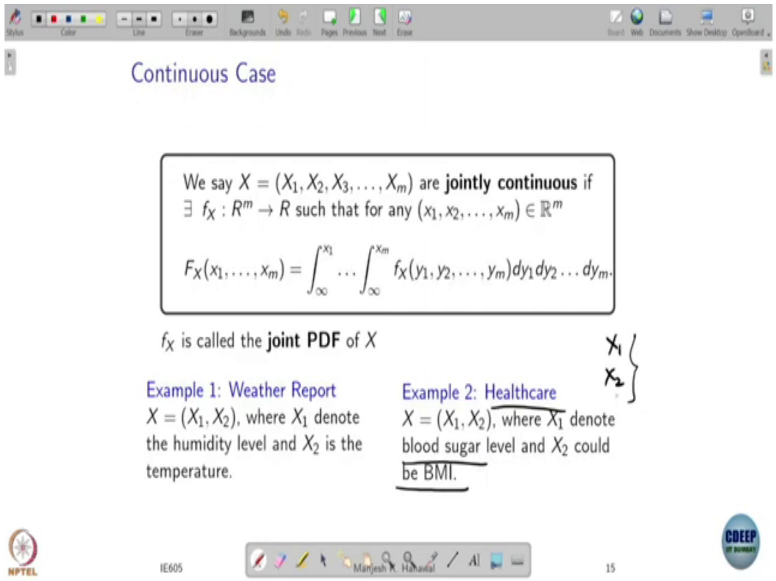
Professor: Point 3 and like that if you want to compute this you have to sum up the elements in each of the columns.

(Refer Slide Time: 05:31)



Now this is for the discrete case, what about the continuous case? Continuous we were going to extend the definition analogously the jointly we were going to say that a given bunch of random variables are jointly continuous if there exist a function multivariable function f such that this relation holds, that is CDF can be given as integration of this function over all possible random variables. So, there should be minus minus here and you see that this is a simple straightforward extension of what we have for a single random variable.

(Refer Slide Time: 06:27)

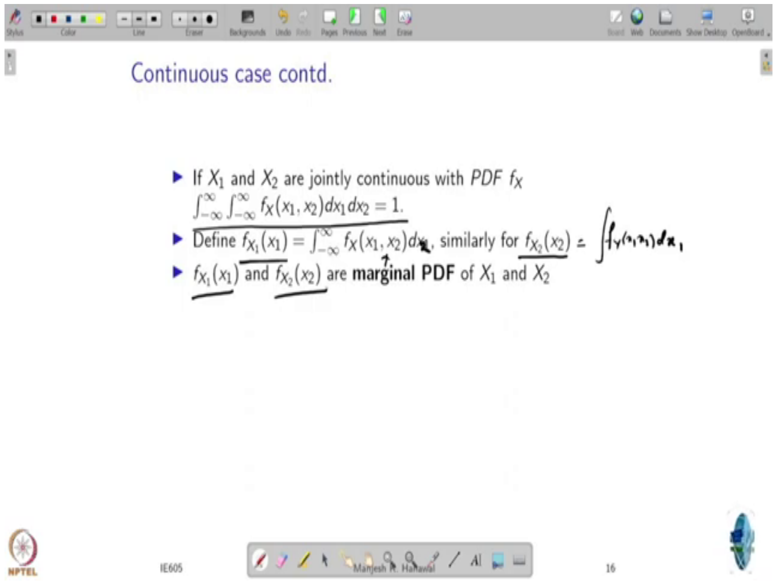


That was like in the simple case we had this is like minus infinity to x, f of let us call this y dy and now we are just like making this x to be x1 up to xn and now putting integration which each one of them and such whenever such an f exists we are going to call it as joint PDF. And examples again pretty obvious I will just look into one example let us take Healthcare example, suppose certain health condition depends on 2 values one is your blood sugar level and another is your body mass index.

Doctor is there if you go to him he is going to ask you to perform 2 tests, he will ask you to get and test the test your blood sugar level what are the outcome let us call that a random variable X1 and whatever your body mass index let us call that X2. Maybe there they are dependent and one has influence over other like if your body mass index is high maybe you will also have sugar who knows.

So, then there is some dependency there and we have to look into them together. Doctor he has asked doctor has asked you both the reports because he want to make a decision looking them together not just like a blood sugar level or not just like your BMI. He wants to look them together and then make a decision because they together influence something and that is where this one has to deal with this joint PDF.

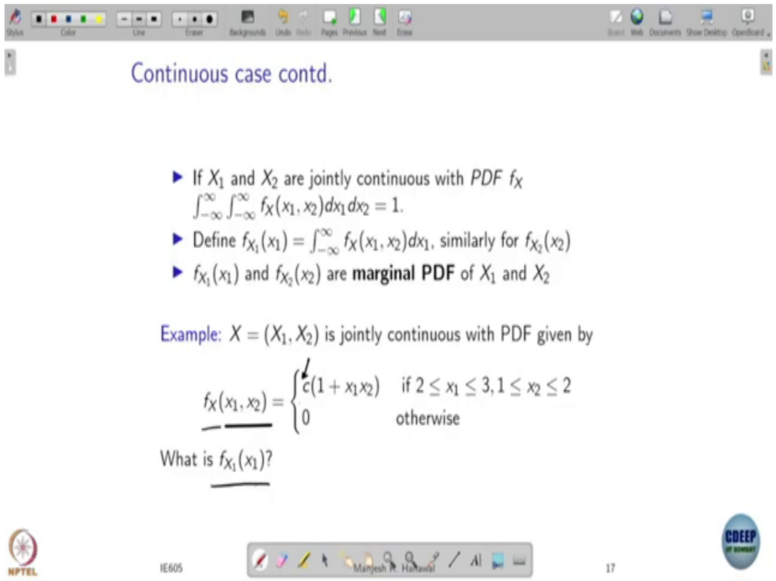
(Refer Slide Time: 08:25)



Like the way we did it for discrete random variables there is a notion of marginal PDF also and that we can get first basic properties our joint PDF should integrate to 1 over all the variables or which we are interested. If f(x) and x1 x2 this x1 and x2 variables it should integrate to 1. And now if you want, you are only interested in random variable X1 and do not care about the influence of X2 what you need to do is you need to let maybe there should be d2 here.

What you need to do is you let integrate this over all possible value of X2 then what remains is only the influence of X1 and you will get the effect of only random variable X1. And similarly you can integrate this function over dx1 and then you will get only the influence of x2 and this f of x1 and f of x2 they are called marginal PDFs of random variable X1 and X2. So, again from joint PDF we are able to recover the marginal PDFs.

(Refer Slide Time: 10:01)



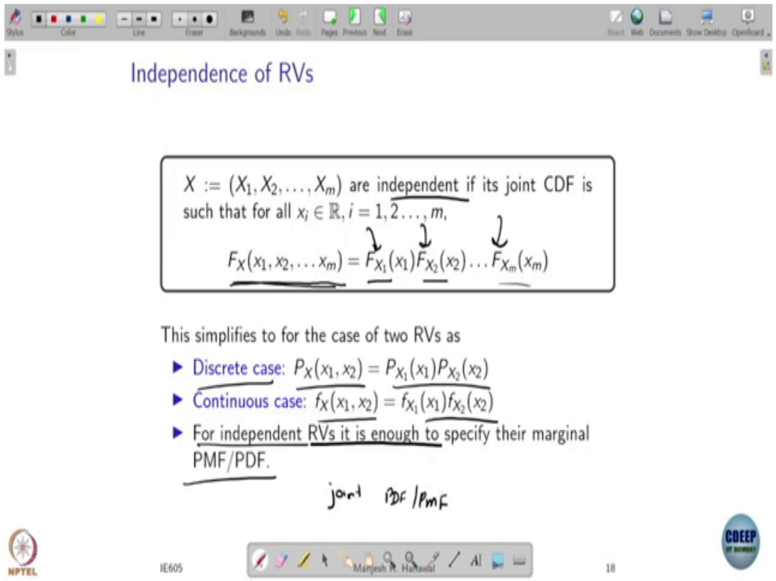
I will leavethis example you can work out like there is a one joint PDF we have given, so there is some constant c. Can this c can take any value here?

Student: No.

Professor: No, right, how we are going to find the value of c? You will integrate this function between 2 to 3 on x1 and between 1 and 2 on random variable x2 and equate it to 1 you will get one equation where in terms of c and then you solve that equation it will give you the value of c.

And that will completely define your f(x1, x2) even though I did not specify c here, c cannot be arbitrary here c has to be some particular value for which only this is a valid joint PDF from that joint PDF you can go back and recover your marginal PDFs.

(Refer Slide Time: 11:08)

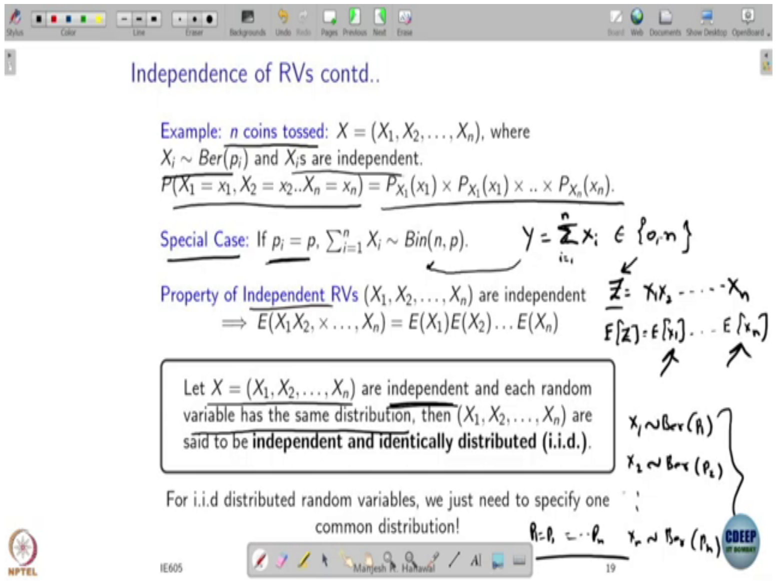


Now let me quickly look into this one aspects called independence of random variables. Now the way we introduce independence of events we are going to define a bunch of random variables are independent if their joint CDF can be expressed as product of their marginals, once you have this joint CDF I can get this marginal, this marginal, this marginal and this marginal.

So, if this can be written as the product of this marginal then it is called independent these random variables are called independent. So, notice that here in whenever it is independent it is just enough to provide CDF of each of the random variables and from that I already have a joint CDF. So, earlier I said that if I give you marginal CDF I may not able to recover the Joint CDF but however if I say additionally they are independent then providing marginal is enough you already readily get your joint PDF or joint CDF here.

So, this simply translates to in the discrete case if I have 2 random variables x1 and x2 if they are independent then their probabilities can be expressed as a product of their marginal probabilities and similarly in the continuous case if I have that joint PDF it could be expressed as product of their marginal if they are independent if they are not independent this property may not hold. So, if random variables are independent it is enough to specify their Probability Mass Functions or Probability Density Functions we do not need joint PDF or PMF because that can be recovered from their marginals.

(Refer Slide Time: 13:28)



Let just me talk this and then we will take a break. Let us say there are n coins each one of them is a Bernoulli I hope all of you recall, what is a Bernoulli random variable? Bernoulli we said comes with a parameter p and I am denoting the ith random variable with parameter pi here. And now I am saying this Xis are independent. Suppose now I want to compute this probability what is the probability that X1 is first random variable is going to take value x1 second random variable is going to take value x2 like that this because of their independence I can write like this because once they are independent this is it.

Now as a special case let us say all these pis are some same value and what I am going to do is? I am going to add all these random variables Xis. Now if all these Xis are Bernoulli. What is the possible values of Bernoulli random variable can take?

Student: 0 and 1.

Professor: Then what is Y can take?

Student: 0 to n.

Professor: 0 to n, so this is going to take 0 to n. I mean zero to n now you can verify this this Y is nothing but a Binomial random variable with parameters n and p. Now if you go back and compute all the probabilities what is the probability that Y equals to 0 what is the probabilities at Y equals to 1 what is the probability that Y equals to n you will exactly get what is we have defined for binomial random variable. So, what is the relation then between binomial and Bernoulli random variables?

Student: Sum of Bernoulli random variables.

Professor: Just sum of Bernoulli random variables?

Student: Constant probability.

Professor: That is from constant probability of success and?

Student: Have to be independent.

Professor: They have to be Independent. So, binomial random variable with parameter n and comma p is nothing but it is the sum of n independent Bernoulli random variables with the same parameter p. So, that is where the independence often simplifies the things and you will see connection between different distributions.

A consequence of this Independence is if you want to take product of these random variables like let us say I have defined a new random variable let us call Z and I am going to take product of this and now if I want to take the expectation of this Z, I do not need to go and find out its PMF or PDF altogether, I can directly compute this as expectation of X1 all the way up to expectation of Xn.

That is I can just need to find out the expectation of individual random variables and if I take the product that will give me the product of this that will give me the value of expectation of set, this is true only they are independent if they are not independent this simplification does not work. Now one last definition we have is if you have this bunch of random variables if all of them are independent they are independent and also each of them have the same distribution.

For example, in the case of binomial we said, there are n Bernoulli random variables each of them has the same parameter that means all of them have the same distribution. Like that if we have in independent random variables and all of them have the same distributions then this bunch of random variables are called independent and identically distributed.

So, one quick example is suppose let us take I will take X1 to be Bernoulli with parameter p1, X2 is going to be Bernoulli with parameter to p2, like that xn is going to be Bernoulli parameter pn I will tell you that they are independent. But can I say they are IID?

Student: No.

Professor: No, why?

Students: Because parameter is not same.

Professor: Yeah, they are they need not be same distribution even though they are Bernoulli but the parameters are different. If I say that independent and further say that all this p1 equals to p2 all the way up to pn are equal then they are IID. Let us stop here.