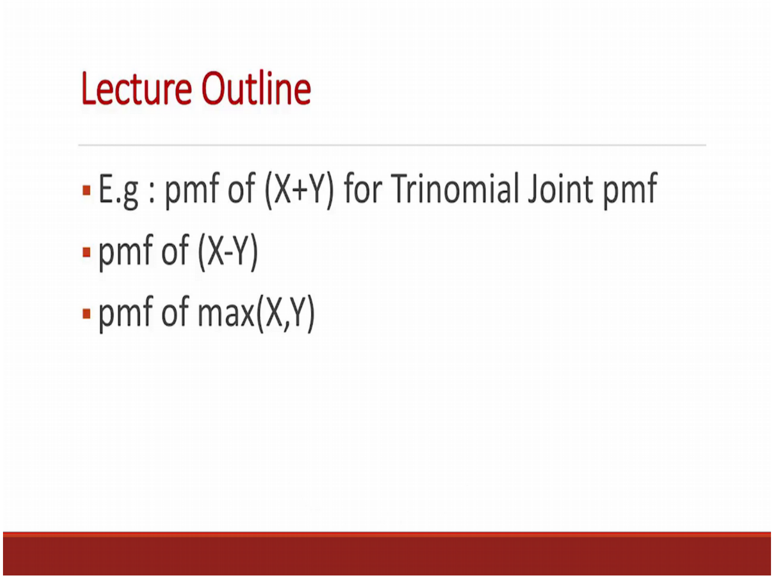
Probability Foundations for Electrical Engineers

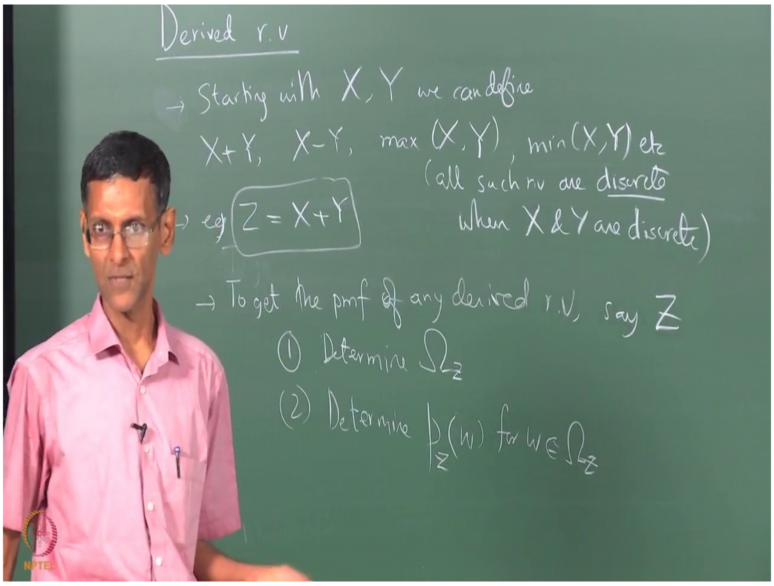
Prof. Aravind R
Department Of Electrical Engineering
Indian Institute of Technology, Madras

Lecture 15 - Part 1

(Refer Slide Time: 00:13)



(Refer Slide Time: 00:21)



Last class we looked at this particular case of X plus Y in some detail, right. But of course, you can define starting with X and Y which are rare for the time being they are discrete. You can define any kind of combination which makes sense or which is important in some situation or the other. I have given a few examples of there X minus Y is as sometimes as usually the X plus Y or max min and so on, right. Remember max and min slightly this would makes in Y in that it is you have to some you may have to work a little harder to understand it or just a larger of the 2, right? And when is this moral of the 2 that is all. And they are all discrete random variables when X and Y are discrete. We are giving a name only as if now we are giving a name only to X plus Y which is which we which we calling Z.

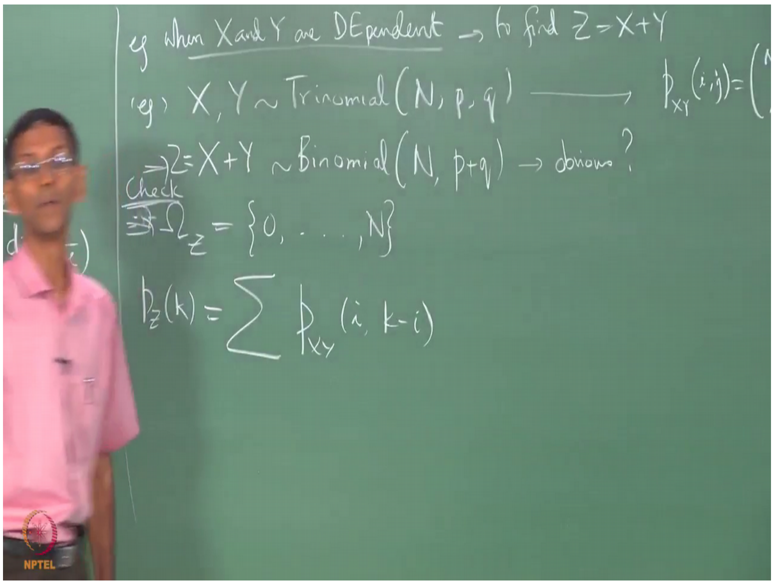
Later on, may be during this lecture, I will work with some one of one of 2 the others in we give some names at that point right. So, as of now we have Z equal to X plus Y. And we pointed out that we could get the to get the pmf of Z or anyone of them of any of the derived random variable what is what are the 2 steps? The first step is to determine omega of the for that random variable, who Z let say. So, you must determine the range in which or the set of numbers in which Z takes values. Then the second step of course, is to determine the probability that Z takes some value in that set. So, you have to do this for all possible values in omega Z. Only then the pmf is complete. And how do you do this? You just look at what combinations of X and Y give you that particular value of W. And the joint distribution of X and Y will automatically tell you what the joint probability is. And that joint probability is inherent inherited by Z.

The summation is if any occur only when you have multiple points in omega X Y which map to that particular value W. So, why did we get a summation in and the correlation case? The convolution summation came about because you have when X and Y integer valued, you have many combinations of X and Y which give you a particular integer value of Z. For example, if you though throw 2 dice and say Z equal to X plus Y, you can get for example, omega Z will be 2 to 12, that is obvious then, but any number like 6 or 7 or 8 will have lots of combinations of X and Y which give you that particular Z right. 3 and 3 will give the 6 4 and 2 also give the 6 2 and 4 will give 6 etcetera.

So, all of those probabilities where would we added to get the probability of Z being 6 or 7 whatever. And then you can easily verify the convolution right result even without going into convolution. It is just a nice triangular or psuedo pseudo trapy, I mean trapezoidal in this I know it is triangular, right. P Z in this dye roll case 7 has a maximum probability, right the sum anyway. So, I think I mean, I am not going to repeat what I said last week beyond this.

We have to move on to something different. So, today what we will do is to take the case of dependent X and Y right.

(Refer Slide Time: 04:11)



Student: (Refer Time: 04:14).

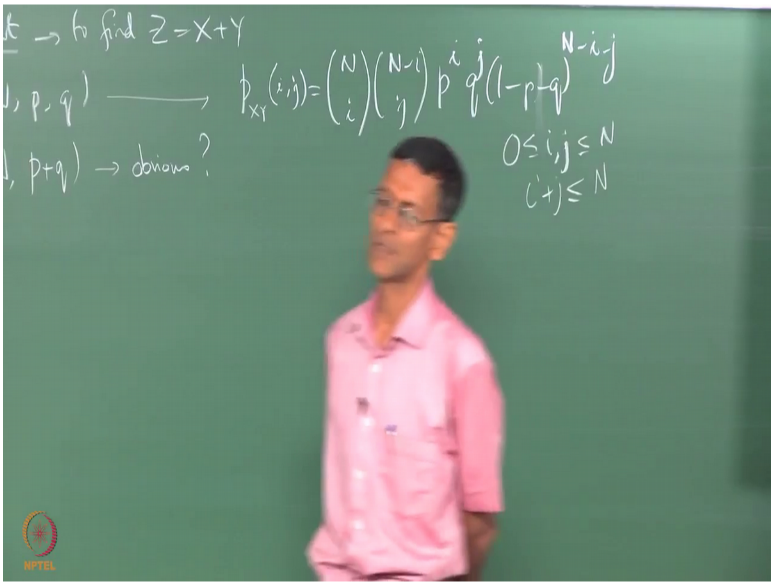
Not independent, and the simplest example would be X and Y are jointly trinomial with parameters capital N P q. This is a I think the notation's value obvious, right. P is the count even count for a X and q is the probability of the even count for Y. Exactly like we have done earlier, right. No there is no change in notation capital N is what we used earlier.

So, we already know the answer to this right. So, if I say if I will look at, then if I look at X plus Y it is basically the sum of 2 counts. So, it has to be what? You are basically putting both in a single box. So, do we know the answer to this or not. What is, right? What is X plus Y going to represent in the trinomial case? X plus Y is going to represent the count of either the event one or the event 2. Remember, we said those events have to be exclusive otherwise it is not trinomial, is not it? So, X plus Y must be binomial with parameters N and P plus q. This we know is k N can never increase, because it you have a total of N trials or N times you are throwing N balls throwing into bins, right. You do not have more than that.

So, you cannot get 2 N or any such things. So, that it has to be limited by this N, but the probability of that box would become bigger, and it will become P plus q. Because you are looking at either the first event or the second event not you not really distinguish between the 2, you know it's so, this is the count that you get one, I mean either event one or event 2 right. So, how do we derived this?

So, this is what can I say this. So, this sake at the camera I have to say that it is try to put some text down here on the board. So, that I put obvious put question mark, anyway. But let us just derive it, and see for answers what, what it you know that we do in fact, get it. Now I think what I have to do here is I have to go here and write the expression this P X Y of let me see what notation i comma j, right I used?

(Refer Slide Time: 06:51)



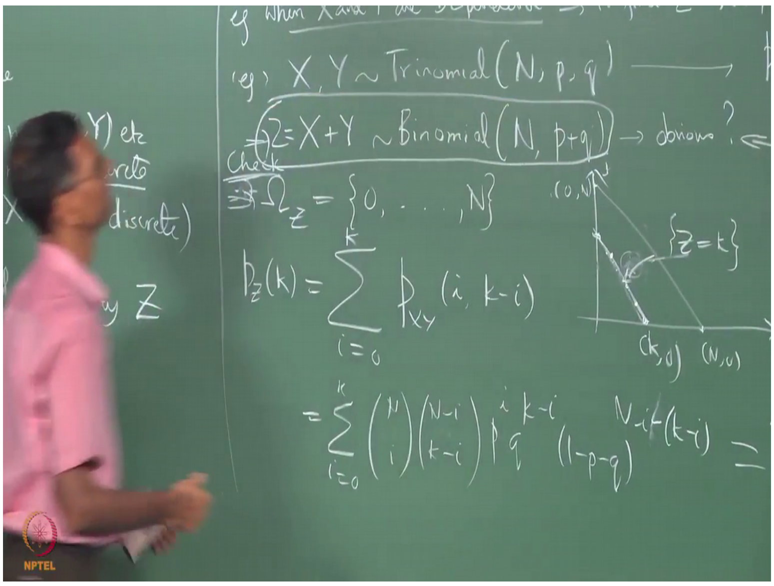
i comma j is N choose i N minus i choose j P to the power of i q to the power of j 1 minus P to the power 1 minus P minus q to the power of N minus i minus j. So, this for 0 less than i comma j less than N. i plus j is also strictly is unequal to n.

So, this is basically or starting pmf. So, here X and Y are both integer valued. So, what is, right what is omega Z? So, I am going to check this result. Omega Z is what? Both X and Y need not happen at all. So, obviously, Z cannot start from 0 or both the X and Y can be 0. So, omega Z starts from 0, and goes all the way up to N. It cannot be more than N, right. What we said, right. The X plus I am sorry, you are repeating the experiment N times. And so, the count of the 2 events the for which X and Y are the counts like they some cannot be more than N anyway.

So, therefore, Z is also integer valued it takes values only in this set. So, what is P Z of k now? So, what I have to do is I will write the expression down P X Y of i comma k minus i, this anyway I have, right. This is what I have to start with. So, for a particular value of k, right. I have I can X can be i and Y it can be k minus i, what will be the range of I is the summation of what we done this for the poisson case right. So, this exactly the same thing we the limits are very similar or a right. So, again I going from 0 to k. Remember, we are we are looking at the probability the is the Z equal to k.

So, if Z is k X cannot be more than k. X can be at most k, isn't it? But here because a dependent this cannot and will not factorize into P X of I into P Y of k minus i. That is a more important difference, between that and this. So, let me just put on some key just the key manipulations here just to make sure I do not waste a time. So, let me also draw this diagram out here.

(Refer Slide Time: 09:51)

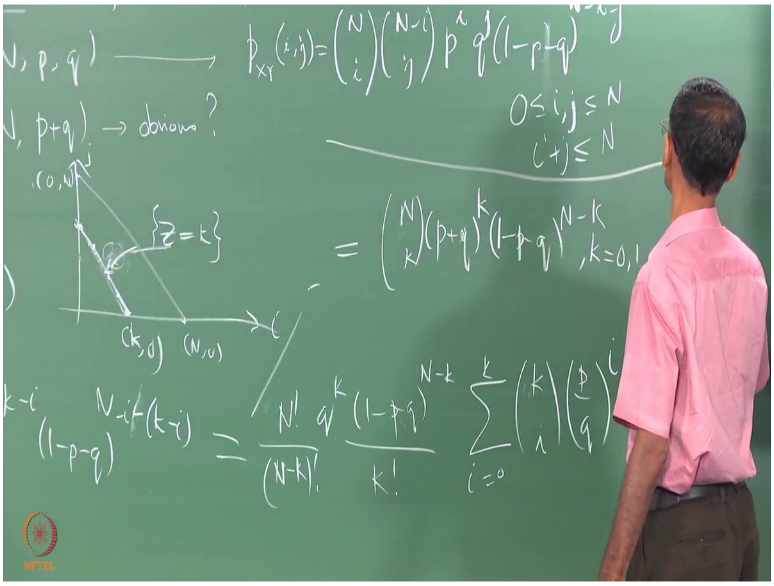


So, what I am doing is; this is the N if I draw this as i and j, this is the point N 0 and this is the point 0 comma n. So, the k will be some line parallel to this.

So, this will be the point k comma 0. And this is 0 comma k. And you going up and down this line. So, this line corresponds to let me say Z equal to k. Z equal to k is a same as X plus Y equal to k right. So, you are basically taking some point on this line and adding up all the sorry, you are taking this line for that by you can adding up all the points they fall on that line. The probabilities of all the points at line, on this line.

So, you can easily do this, I am not going to spend too much time on this because manipulation is not very convenient to the board, but let me just write some key steps. So, if I do this i equal to 0 to k just simply. So, I have to substitute this out there, N choose i N minus i choose k minus i P to the power of i q to the power k minus i 1 minus P minus q to the power of N minus i minus k minus i. Though all I am doing this substituting k minus i for j.

(Refer Slide Time: 11:47)



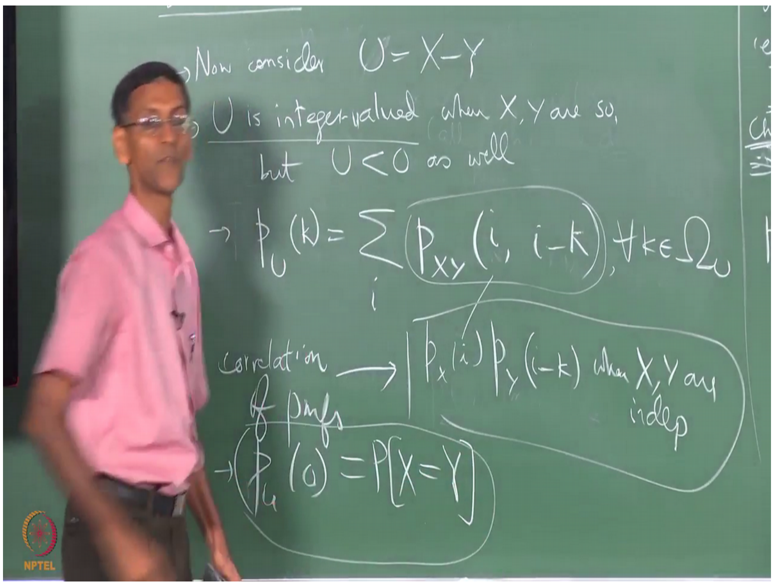
And then, so, P by so, I have here P by q power i, this q power k can it can come out N factorial. So, essentially what I will get will be let me just write some skip some step and write this. So, N factorial divided by N minus k factorial q power k 1 minus P minus q to the power of N minus k, by k factorial times summation i equal to 0 to k this will be just k choose i P by q whole power i. I want you to work this out. So, not I do not think is too much to ask. Just expand out the this factorial. So, expand out both factorials and I mean both comma this thing binomial coefficients and write it like this. And you can clear what is that summation we have seen this type of thing earlier, it is just 1 plus P by q whole power what is it?

Student: (Refer Time: 12:58)

Whole power k. So, I think you will you can easily show from all of this this will become equal to what? This becomes equal to N choose k and draw a line here, N choose k P plus q to the power of k 1 minus P minus q to the power of N minus k. The point is; this 1 minus P plus N minus k N minus q this thing is already come here itself. So, this just case repeated here. So, what is this? This is valid for all k, you are not said you are not put any condition on case. So, this is true for k equal to 0 up to capital N. That is as you move this line all the way from this point to here. What is that pmf? Binomial with parameters capital N and P plus q.

So, this is an example of a dependent sum calculated I hope calculated pmf from first principles. So, the same thing you should be ready to do of course, we do not have any I mean other as I said the all other examples that I have seen with joint pmfs are more contrive, right. Only this trinomial thing seems to be the more in a sense most acceptable thing you know all others look like they more textbook is (Refer Time: 14:23) than anything else further unlike this anyhow. So, this example I think should illustrate the point let me not spend too much time on this. Is there any question here? So, I put a tick mark here to say the this is what we wanted this is exactly matches this.

(Refer Slide Time: 15:09)



What about X minus Y? Remember, right now we will give some name to X minus Y. Now consider I am not going to spend too much time on this. X minus Y, U when X and Y are integer valued U is also integer valued, but you can also take on negative values as suppose to only positive values when X and Y, even if X and Y are both positive, right. U is are so, but U can be less the 0 as well. So, for example, if we take the difference of 2 dye rolls, how will you go what finding the probability that U equals minus 1? It is a same principle. You look at combinations of X and Y which give you minus 1 and all those things right.

So, it turns out that actually. So, let us a take this case of U integer valued, right. It turns out that P U of k is the double summation over I of interest P X Y if I put I here I must put i minus k here. Why i minus k? It is this difference is what? Is k, but k can be positive or negative? Sometime we have a symmetry this in this situation and you may find that it is enough to calculate positive k then for example, in dye rolls the difference P U of plus 1 will be probably equal be equal to P of minus 1. May not be any no I do not know I take that back I have not checked it out thoroughly right, but this in general again uses the point is that it uses there is geo discrete it uses this joint pmf which is the key without the joint pmf in general what I want to say is you cannot proceed of course, in the independent case what happens? It will factor into P U of I into P Y of sorry P X of I into P Y of i minus k.

So, then this factorization happens for independent, right. Then you get what is that operation called on the 2 sequences P X and P Y have you seen this in your dsp class.

Student: (Refer Time: 17:51) correlation.

Its correlation. So, only if you can get this you get the correlation of the pmf. So, for integer valued sum independent sum you get convolution for integer value is different you get correlation. So, let me not I am not going to work I mean this this kind of manipulation is best left homework exercise I am not going to do any example here. Just want to point out one thing though, what is P U of 0? P U of 0 by has some significance. P U of 0 is basically identical to the probability that X equals.

Student: Y.

Y, right. These 2-events U equal to 0 is identical to the probability that X will equal Y. So, there is no different. Now the next thing I will quickly look at is the max and the min, but I am not going to look at the dependent case. Because the dependent case is like trinomial is very messy to handle. We look at the independent case.

Student: Sir, can you (Refer Time: 19:22).

Pardon me.

Student: Can you repeat this part.

Which part?

Student: (Refer Time: 19:26) X minus Y equals to does it denote number of time this events one happen?

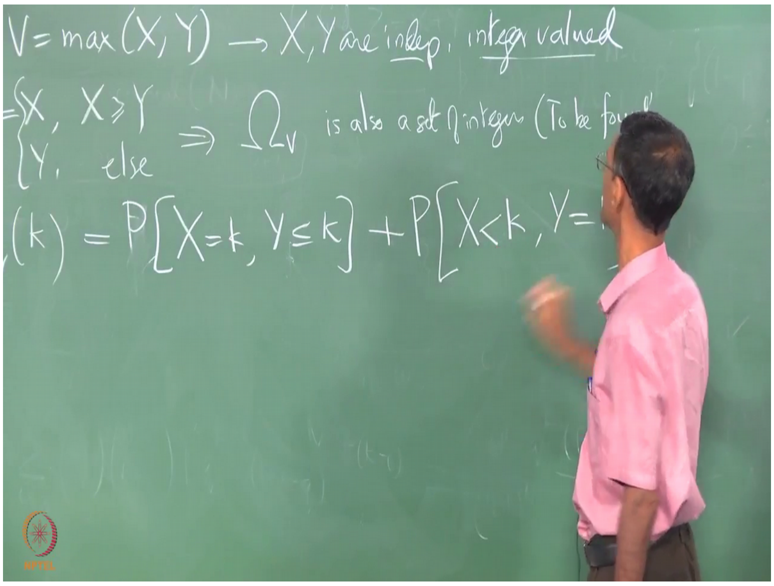
U equals X minus Y is just the difference between the count X all, right. They the value taken by X and the value take by Y. So, on a trial if you find their own 2 dyes for example, right. You get 3 on 1 3 for X 4 for Y you will take the value minus 1, that is all. In the trinomial case, right, if you are looking at X and Y as a count of the one and count of 5 on N rolls of dye. Again, you can look at the difference. So, the number of times you get 1 minus a number of times you get 5 is U. This is okay?

Please ask, right. Do not hesitate to ask, you have any question is, right. And this is and this is clear that P U of 0 must be X equal to Y, when does you take when you can U equal to 0 only when X equals Y if and only if. So, whenever X and Y takes on 2 values you will take a value a value just uniquely determine by the values of X and value of Y. So, that is that concept is so important that I have to repeat this several times; that is, X and Y jointly determine everything about derived random variables, right. Everything from the they set in which (Refer Time: 20:55) derived random variables can take values to exact values on trials at the whole thing right.

So, derive is completely something which is the which is based only on you know the existence of U Z whatever as you might wanted to define like max min, and so on all of them are dependent, right. Define based on them underline X and Y. It is a kind of processing you do, right. Like you receive X and Y you process X and Y in the in the way in which you want to do it. And then you derive whatever this is.

Now, to (Refer Time: 21:31) that point even further let us look at max and min, I just now said, but I am not going to do it for the dependent case because the dependent case is messy and I would not try I tried riding it here.

(Refer Slide Time: 21:49)



And I mean, I would done it in the past, but it is not been very nice to work with it. So, let me let me take this as a, right as a test case.

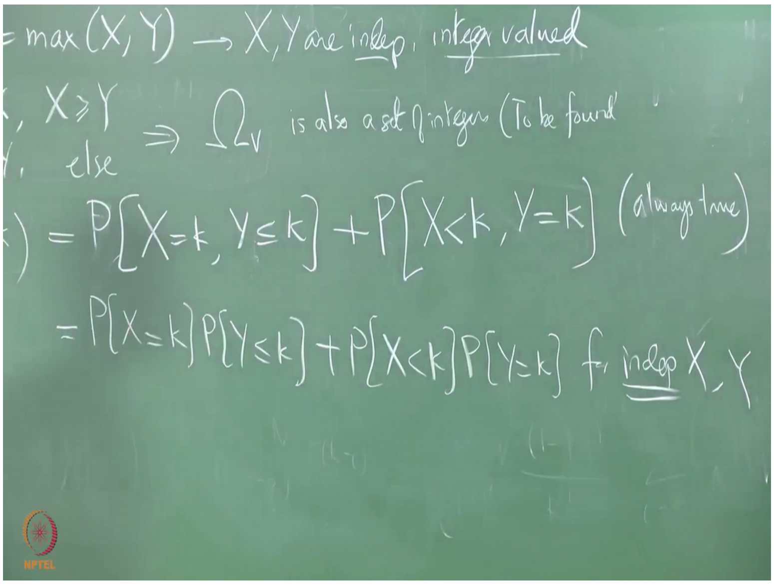
So now I am going to take X and Y are independent integer valued. So, first of all what you understand by max X Y. So, how do I determine v from X Y? So, v equals X I when you say it is a X? When X is greater than Y, otherwise you said Y. This is the meaning of max exactly like you have for 2 any 2 numbers. You can always take the maximum of 2 numbers. So, you can take the maximum of 2 random variables also.

So, outch this is so, it does not matter. So, how we are also assuming that they are independent and integer valued. So, if they are integer valued you; obviously, v also has to be integer valued now let me take the what example when I come prepared to discuss just it matter let it be just any integer value does not matter.

So, what is right? So, omega v is also a set of integers. Now not the full set or sub. So, the value this you know. So, what the minimum of value of X is a minimum value of either X or, right. What would you say it is whatever I mean if both X and Y start from 0 then the minimum can be 0. If X is a starts from 0 and Y starts from 2; obviously, the minimum cannot be sorry the maximum cannot be 0, it has to 2. So, you have to determine this omega v appropriately right.

So, let me not write any notation on this. So, I want to look at just the probability. So, this has to be done.

(Refer Slide Time: 24:09)



So, this needs to be done to be found, right. On a case by case basis now, I just want to write an expression for P v of k, right. For that any integer valued case. So, what is P v of k? You get k when, what you get when X let say X takes the value k. Then Y must be maximum k, right. Or Y can be k and X can be smaller than k, but you should not count anything twice. So, it turns out that P v of k is basically the probability that X equals k Y is less than equal to k fine this is one part of it which is fine. So, note that whenever this happens the maximum will be k, plus P the probability that X is only smaller than k because I have already look at the equals case here. So, to avoid double counting I have to insist that X is smaller than k when Y equals k. I do not want to count X be X equal to k Y equal to k twice.

So, this is true even when X and Y are dependent, right. This does not be co independence. This is always true. In the independent case, what can you how it is simplified. These are all joint probabilities involving X and Y, but when X and Y independent then this event X equal to k becomes independent of the event Y less than k. So, this becomes so, this manipulation I am just simply writing it ah I am not going to go you know go further there is I think is best left to homework exercise this point. Because, right doing extensive manipulations to the board is not very conducive to an interesting class I think.

So, we leave at this be once you get any problem down to this level it is easy to do. Because specially if you have eventually involving one random variable you can easily find this, right. I hope this is not a problem of anybody P of Y less than less than equal to k we in fact, when this kind of thing will use a lot in the continuous random variable case, right.