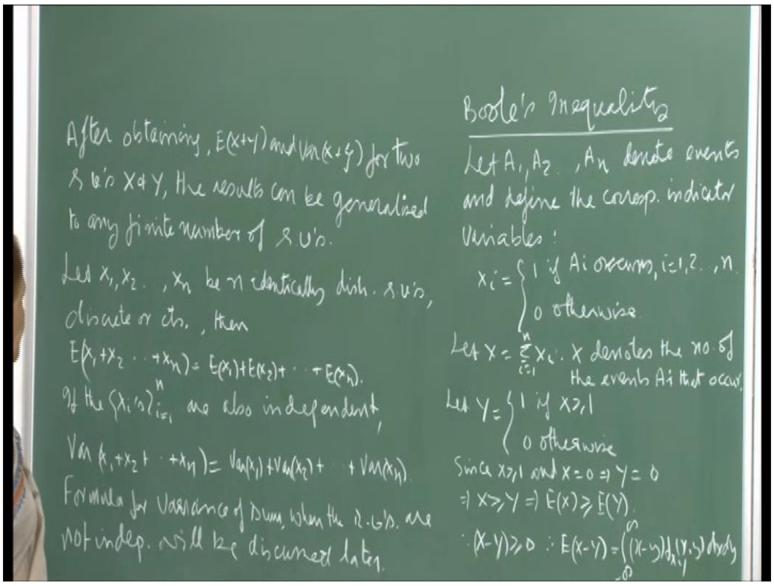
Introduction to Probability Theory and its Applications
Prof. Prabha Sharma
Department of Mathematics and Statistics
Indian Institute of Technology, Kanpur

Lecture - 14
Chi-square R.V. Sums of Independent Normal R.V. Conditional distr

(Refer Slide Time: 00:16)



So, after obtaining expression for E x plus y, in fact we showed that E f x plus y is E x plus E y and under independence of x and y, we showed that variance of x plus y is equal to variance of x plus variance of y. Then we can generalize these results for any finite number of random variables. So, therefore if you have x 1, x 2, x n as a n identically distributed random variables, discrete or continuous, then expectation of x 1 plus x 2 plus x 10 will be equal to expectation of x 1 plus exception of x 2 plus expectation of x n. So, want to show that say that this is you can always take the some of the expectation. Sorry, I mean expectation of the sum as sum of the expectations. So, first these have to be identically distributed.

Now, if the x size are also independent, then we can also extend this result at variance of x 1 plus x 2 plus x n is equal to variance of x 1 plus variance of x 2 plus variance of x n, because as we saw if the case 2 variables that the product term vanishes. So, here also because independence, the product term because you will have two products of two things at a time. That means of the kind you will have an expression like expectation of x i minus expectation of x i into x j minus expectation of x j. So, under independence this will go inside and therefore each of the term will vanish. So, you will only get, this square terms will be left. When you square up this minus the expectation of the sum and therefore, you will get the sum of the variances. So, under independence, you get that result. Now, formula for the variance of sum when the random variables are not independent will be discussed later.

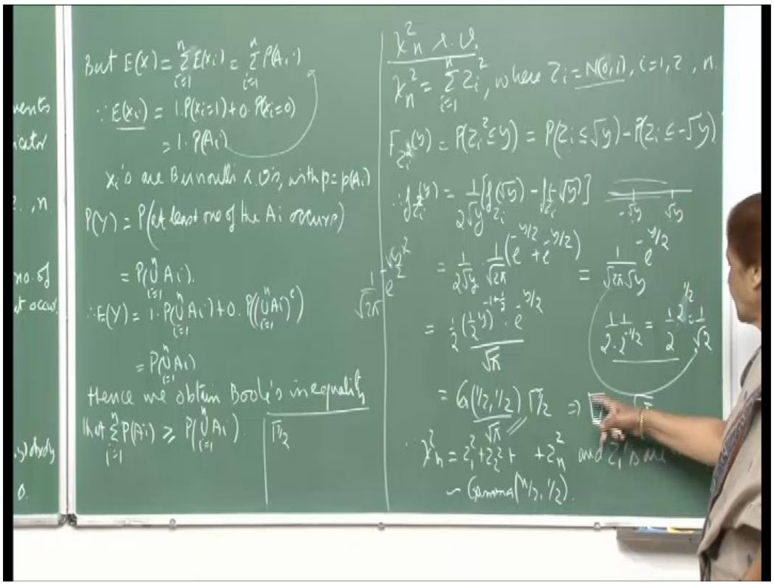
So, we will give you a general formula when the variables are not independent. So, this is one thing and so, we can use it for computing various expectations. So, now let me discuss interesting result which is called Boole's inequality, and let me first just describe what we want to say here. This is that if you have A 1, A 2, A n as n events and when corresponding to these n events, I define the corresponding indicator variable. That means, x i is 1. If a i occurs i varying from 1 to n and its 0 otherwise. So, therefore, then I will let x be the sum of these indicator variables, these n indicator variables. So, in other words, x denotes the number of the events a i that occur fine because x i is 1 if a i occurs. So, if x for example, if capital x is 5, then that means, 5 of a i's have occurred. This will add up to 1 plus 1 for 5 events which have occurred, right and define another variable y which is equal to 1. If x is greater than or equal to 1 and 0, so otherwise means that you see if x is 1 or greater than 1, then y is 1, but if x is 0, then y will be 0.

So, from definition it follows that x equal to 0 implies y equal to 0. Since, your x is otherwise greater than or equal to 1. See either an event occurs or it does not occur. So, if anyone of the events occurs, then x will be at least 1 and if none of the events occur, then it will be 0. So, therefore, x is always greater than or equal to 1 or it is 0. I mean if 1 of the not always. What I mean is that if at least one event occurs, the next will be always greater than or equal to 1. Otherwise, if none of the events occur, then x will be 0 and in that case, y will also be 0.

So, therefore, it implies that x is greater than or equal to y, right because x will take value 1 or 2 or 3 your y is 1 and whenever x is 0, your y is 0. So, it is clear that x is always greater than or equal to y, and this implies that your expected value of x will also be greater than or equal to expected value of y because this being x minus y is non-negative. Therefore, expectation of x minus y. Now, I am just writing the general expression here. So, that means, for example a general expression for e x minus y would be minus infinity to infinity x minus y f x y x y d x d y. So, this integrant is non-negative and therefore, the integral will be non-negative. So, it follows that your E x must be greater than or equal to

E y, ok. So, that is the important result. That is how through this, we will derive the Boole’s inequality finally.

(Refer Slide Time: 05:17)



Now, look at E x. So, E x is expectation of i take it inside being linear function. So, I can exchange the expectation and summation sign. So, sigma i varying from 1 to n expected x i, but expected x i is what expected x i 1 into probability x i equal to 1 plus 0 into probability x i equal to 0. Since, x i take values only 1 and 0. So, this is 1 into probability x i is equal to 1 plus 0 into probability x i equal to 0. So, this is 0, and this probability x i equal to 1 is probability of occurrence of a i and therefore, this is 1 into P A i which is P A i, right. So, you are some of the expectations of x i is equal to sigma P A i.

Now, x i’s are, you can see that x i’s are Bernoulli random variables, right because x i’s take value 1 or 0 and the probability of success as you can call it is P A i for x i, right. Now, what is probability y? So, probability y is probability that at least one of the A i occurs. So, this is union A i varying 1 to n. So, this is at least one of the A i occurs. So, well which way would we are saying y is equal to 1 if x is greater than or equal to 1 and this translates to x is greater than or equal to 1 if at least one of the events A 1, A 2, A n occurs, right. So, this is probability union i varying from 1 to n A i. So, expectation E y will be 1 into because y is 1 if x is greater than or equal to 1. So, this is union i varying from 1 to n A i plus 0 into probability union A i complement, right. Probability of this complement of this event and therefore, this is also equal to E y is equal to simply probability of union A i varying from 1 to n.

Hence, we obtain Boole's inequality which says that sigma i varying from 1 to n P A i is greater than or equal to probability union A i varying from 1 to n. So, in other words, it says that probability of occurrence of at least one event out of given n events. So, probability occurrence of at least one event is no greater than the probability or some of the probabilities of occurrence of individual events. So, this in other words is your Boole's inequality which you might say is simple to accept, or this sounds very reasonable, but we had to go through this process to be able to derive this inequality, right. So, this says at least one of the events must occur. So, probability of that and that cannot be greater than some of the probabilities of the individual events, ok.

Now, let us further go on. I want to again continue with using whatever we have developed about adding up random variables and then, computing their expectations and other results through summing up random variables. So, now, look at the chi square and random variable. So, this random variable is defined as summation i varying from 1 to n z i square, where each z i is standard, normal i varying from 1 to n, right. Thus, the expectation of z i is 0, and the variance of z i is 1. So, each of the z i is standard, normal, then look at these. So, we want to compute first of all, we want to compute the c d f of z i square. So, this will be let me write, make it clear.

So, F z i square y will be probability z i square less than or equal to y, and in one of the earlier lectures, I have already discussed. So, what you are saying is that your z i square should be less than or equal to y, which means that your z i should be less than root y. So, it should lie between minus root y and root y, your z i, right, so that the square does not exceed y. So, this probability can be written as because you will be taking if this probability z i less than root y, then you want to subtract z i less than minus root y. So, this is the probability, right and therefore, when we differentiate this side, I will get the P d f of z i square, this is z i square which will be equal to (( )). So, from here we will differentiate. So, see derivate of a root y will be 1 by 2 root y in to f of z i root y minus f of z i minus root y and for standard normal c.

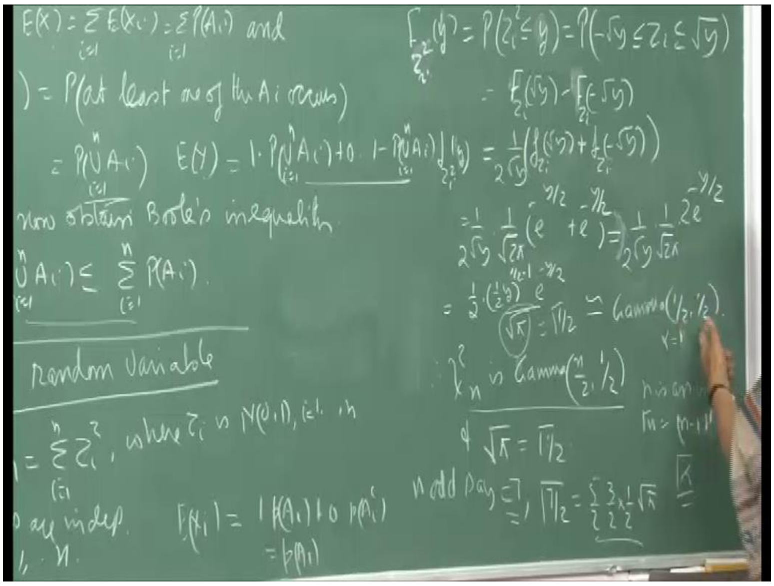
You see what is your this thing is standard normal 1 upon root 2 pi because sigma is 1 is equal to E raise to minus y square. So, it is root y. So, it will be root y square on 2 because sigma square is 1. So, therefore, when you square up the minus root and root y, they are both give you E raise to minus y by 2. So, this becomes, so this is twice E minus y by 2 and these two, so that cancels out. So, I am left with 1 upon root y 1 upon root y E raise to minus y by 2. Now, this I can rewrite because you have 1 upon root 2 and 1 upon root 2, I am writing as 1 upon 2 into 1 upon 2 raise to minus 1 by 2, right. I am doing this, so it is 1 by 2 is going, yeah.

So, 1 by 2 into 1 upon 2 minus 1 by 2, you write this way and this is left by root 2. Root 2 cancels out and you are left with 1 by root 2, right. So, this whole thing I am writing as 1 by 2, 1 by 2 y raise to minus 1 plus half root pi because this is 1 upon 2 raise to minus half. So, 1 upon 2 raise to minus half into 1 by 2, this whole thing is actually equal to 1 by root 2 which appears here 1 by. So, this 1 by root 2, I am writing in this way, right and now, you see if you can remember your gamma distribution, then my lambda is half and my alpha is half because this is alpha minus 1 and then, this is lambda y raise to alpha minus 1 e raise to minus lambda y and then, lambda. So, this is my gamma. Of course, when you look at the p d f gamma p d f, then it has to be divided by gamma alpha.

So, what I am doing is I am writing this as gamma p d f and then, multiplying by 1 by 2, gamma 1 by 2 because I have divided here this expression, this numerator I have divided and multiplied by gamma 1 by 2. So, when I divide this by gamma 1 by gamma of 1 by 2, the whole thing becomes gamma p d f with parameters half and half, and I have gamma 1 by 2 here and thus, a root pi. Now, since this is p d f on the left hand side, this should also be p d f and therefore, you see that these two must be equal. So, this implies that earlier lecture when we were talking, discussing the, when I introduced the gamma distribution m, I told you that it take that gamma of half is root pi, but now you have an immediate justification that gamma of half must be equal to root pi. And of course, as I said that for other fractional values of gamma f, this gamma function you can tables are there and for integer values we had already seen for positive integers. We also saw that gamma alpha will be alpha minus 1 factorial and so on. So, now we continue with this discussion. So, therefore, each z i square has a gamma half, half distribution, now gamma square n. Sorry, chi square n is z 1 square plus z 2 square plus z n square and z i's are independent.

So, then applying the m g f results, we see that you can add up the p d f's here. The parameters and you will again because each is gamma, each z i square is gamma distribution half and they are n of them, they are independent. So, therefore, the sum will be gamma and by two half. That means there will parameter lambda will be half and this will be n by 2. So, you see how I mean using all the results that we have. So, I thought this was good way to show you how we use these tools that we are generating and then, you can see the breakup of. So, once you have a gamma distribution, then you can see that by adding up these independent gamma distributions a h gamma random variables, you get a chi square n and of course, in a special way.

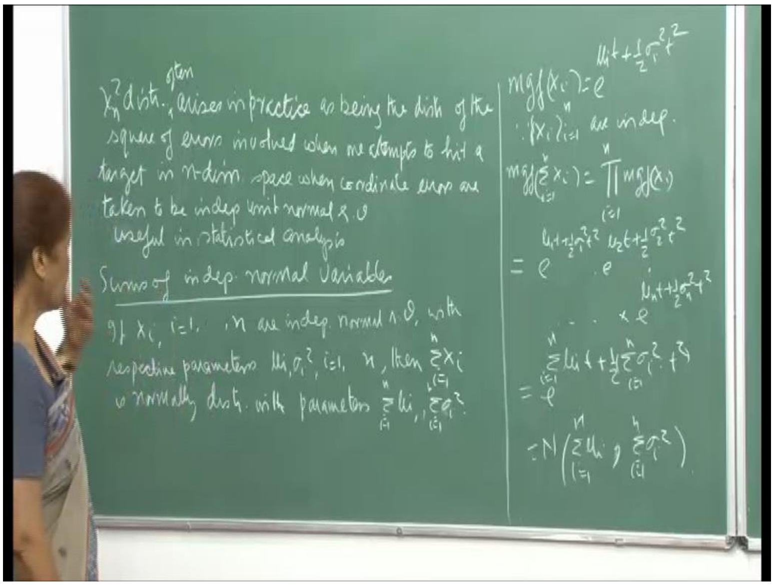
(Refer Slide Time: 15:12)



So, this is another interesting result. So, when you have seen that if n is an integer, then gamma n is simply n minus 1 factorial. So, if n is even, then your this thing will be factorial of n by 2 because n is even then this is an integer. So, gamma of n by 2 will be n by 2 minus 1 factorial, but if n is odd then you will be left with gamma half. So, for example, if you say n is 7, then gamma of 7 by 2 will be 5 by 2 into gamma 3 by 2, then 3 by 2. Sorry 7 by 2 gamma 7 by 2 is 5 by 2 gamma 5 by 2 which will be 3 by 2 gamma 3 by 2 and then, that will be half gamma half which is root pi here. So, this you can compute. So, therefore, now you know you can compute this for all values of alpha integer, non-integer you can find out, right.

So, this was an application of, therefore you see first you square up independent, normal, standard normal variants. Sum them up, you get a chi square distribution and you are showing is that for n, this is chi square with n degrees of freedom. So, then chi square n is actually obtained by adding up gamma half.

(Refer Slide Time: 16:34)



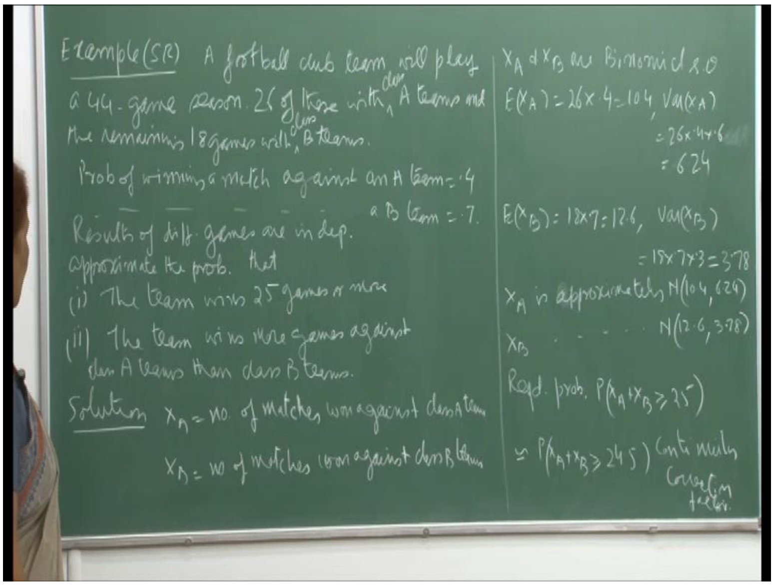
So, I should also mention the importance of chi square distribution. So, if you are talking of chi square n distribution and we said that the n stands for the number of standard normal variable that you are squaring and adding up. So, you can think of because you see this is x i minus mu i upon sigma i whole square. So, this would be your z i, alright and your z i square.

So, you are summing up and so this can be treated as this. This can be looked up on as you know to an attempt to estimate the errors involved and one attempt to hit a target in n dimensional space when coordinate errors are taken to be independent unit probable random variables, right. So, this is you know you can difference from the mean or whatever kind of error you want to, you know talk about and talk about their distribution and so on. Then, chi square random variables come in very handy in that and in fact, it is most widely used distribution in statistical analysis. So, chi square distribution has lot of importance and very often used for your statistical analysis. So, now let me get back to sums of independent normal variables and this is if x i’s are independent normal random variables, x i’s varying from 1 to n or independent normal variables with respective parameters mu y sigma i square, right. That means the x i th normal random variable mean is mu i, and the variance is sigma i square i varying from 1 to n, then sigma x i is again normally distributed with parameters sigma mu i, and that mean as sigma mu y and the variance as sum of the individual variances, because the random variables are independent. So, anyway that we know already that the variance for this would be sigma i square and we also know that the mean for sigma x i will be sigma mu i. These results we have already done, but to show that these sums will again be normally distributed, that is the important thing.

Now, here again I am going to see. The thing is that I have been using m g f's moment generating functions to talk about summation of independent random variables, but it takes sometimes. They introduce the concept of moment generating function much later. So, they actually do it through you know writing the joint density function because these are independent random variables. So, the joint density function will be a product of the individuals and then, they manipulate that term and actually come to the result. So, maybe you should also do that to get to know better feeling, but I find that the treatment through m g f is very convenient. So, the m g f of x i is E raise to mu i t plus half sigma i square t square. This is for the normal mu i sigma i square and since, x i are independent i variant from 1 to n, m g f of sigma x i is the product of the individual x moment generating functions, right.

So, this we have done already. So, the moment generating function of sigma x i would be E raise to mu 1 t plus half sigma 1 square t square and then, E raise to mu t mu to t plus half sigma 2 square t square and so on up to n, and therefore you can add up because these are the powers. So, E raise to sigma i varying from 1 to n mu i t plus half sigma i varying from 1 to n sigma i square t square. So, this is again as I said that the uniqueness of that means given this m g f, I can immediately conclude that the corresponding distribution is normally distributed with mean sigma i varying from 1 to n mu i and this is variance sigma i vary from 1 to n sigma i square. So, using the m g f, you can get these results much quicker, otherwise you have to though the other root is also not difficult one. It is just that you have to write out these long expressions and then, show that this sum of independent normal random variables will be again a normal random variable and these will be the parameters.

(Refer Slide Time: 21:30)



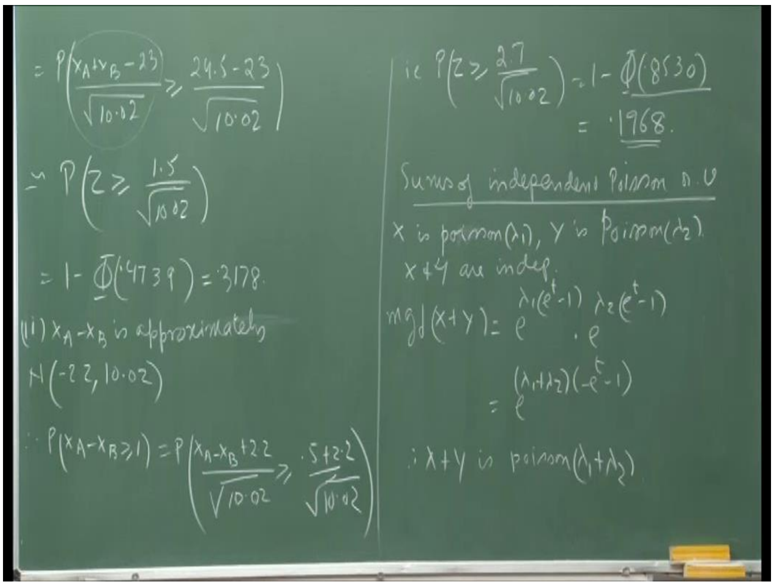
So, to give you an example about how to make use of the fact that sum of independent normal random variables would also be normally distributed, let us look at this example. This football club team will play 44 game seasons. So, you know with during summer, they all play game. So, different teams, so there are 44 games that particular teams will be playing. So, 26 of these with will be with class A team and the remaining 18 games will be with class B teams. So, probability of winning A match against A team is 0.4 because A teams are better than B teams and probability of winning A match against the B team is 0.7. So, results of different games are independent. We are not assuming that there will be any sort of dependence in winning of a game with one team and the other. So, we want the ideas to approximate the probability that the team wins 25 games out of those 444 games played, and second probability that you have to compute is that the team wins more games against class A team's than class B teams.

So, let us start by defining the random variable XA as the number of matches one against class A teams and XB is the number of matches won against class B teams, fine. Now, of course XA and XB are binomial random variables because win is the success and the probability of success is point 4. So, you can find out that out of 26 games played by this team with class A team, then the number of that means, if XA is equal to R, then you will find out it will be a binomial probability and similarly, XB is also a binomial random variable, right. Now, expectation of XA will be 26, right. The formula is NP. So, 26 games played and probability of winning match is 0.4. So, this is NP. So, that comes out to be 10.4 and the variance is NPQ, right which will be 26 into 0.4 into 0.6 is 6.24, then similarly XP being binomial with parameters 18 and 0.7.

So, the expectation of XB will be 12.6 and variance XB equal to NPQ which will be 3.78. Now, the idea is that we start approximating or remember when I told you about approximation of binomial distribution by the normal distribution. The condition was that I mean it is said that if NPQ is greater than or equal to 10, then the approximation is considered good, but here of course that condition is not being satisfied because NPQ in these cases 6.24 and this case, it is 3.78. But still we are going ahead with the approximation just to get an idea because I want to show you the application of adding up normal, where required probability is that XA plus XB together that number of matches one against class A teams, and the number of matches one against class B teams, they must add up to more than 25 or more than 25.

Now, even though I may approximate XN XP by normal, but they are discrete random variables. They are binomial. So, the continuity correction factor must be used here. So, this will be since this is greater than or equal to 25, this will be 24.5 because remember you have this on 25. So, your bar is like this. So, the bar starts from 24.5 and you want to approximate this. You want to include this area because the probability here is greater than or equal to 25. So, it will be 24.5. So, probability XA plus XB is greater than or equal to 24.5.

(Refer Slide Time: 25:50)



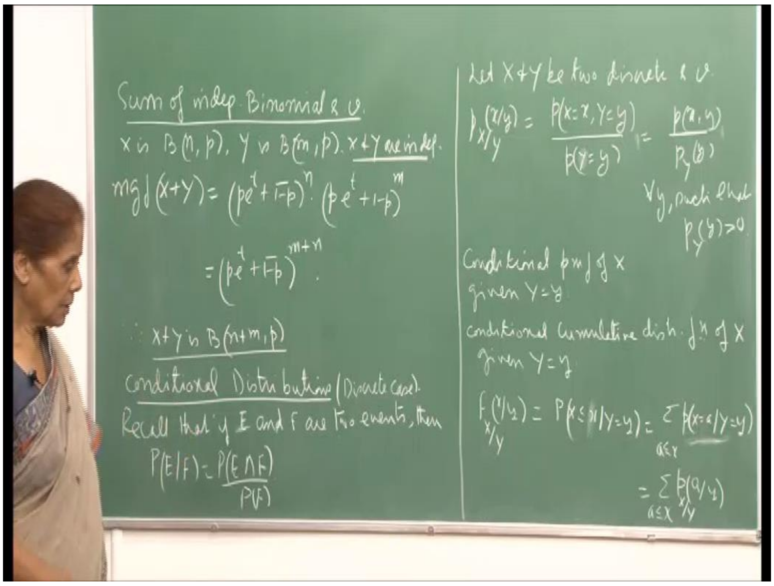
Now, I standardize this probability by subtracting the mean, which has 23. 12.6 and 10.4 adds up to 23 and the two variances add up to 3.7 and 6.24 is 10.02. So, this is what you have. This becomes standard normal variant and therefore, this probability z greater than 1.5 divided by square root of 10.02. So, this will be 1 minus pi of this number comes out to be 0.4739. So, the required probability is 1 minus the normal table, the standard normal probability of this number. So, this is 0.3178.

Now, XA minus XB is also approximately normal minus 2.2 and 10.302 as the variance, right. Therefore, probability XA minus XB greater or equal to 1 because you want the probability that matches against class A team's matches one against class A team's is more than the matches one against class B team. So, therefore, the difference must be greater than or equal to 1 can be more and so, here again we standardize. So, this is XA minus XB minus 2.2 which comes plus and this is under root 10.02 which is greater than or equal to 5.5.

So, here again these term continuity correction factor is used. See a subtract 0.5 from here, so that becomes 0.5 plus 2.2 up on this under root of 10.02. So, that is Z greater than or equal to 2.7 up on under root of 10.02 which comes out to be this. From tables you look up the value which is 0.1 minus of that will be 0.1968. So, this is the probability. So, in fact the probability is low of winning more matches because obviously, this probability is much lower compared to the probability of winning a match against B team. So, as we go on more and more examples of all these concepts that we are talking about. Now, let us come back to sums of independent Poisson random variables. So, X is Poisson lambda 1, Y is Poisson lambda 2 when and X and Y are given to be independent random variables.

So, let us look at the distribution of X plus Y. So, now, since X and Y are independent, m g f of X plus Y will be the product of the m g f's of X and Y. So, m g f of a Poisson lambda 1 is E raise to lambda 1 into E raise to t minus 1 and m g f for Y is E raise to lambda 2 E raise to t minus 1. So, therefore, this adds up to E raise to lambda 1 plus lambda 2 E raise to t minus 1 and this is Poisson lambda 1 plus lambda 2. So, therefore, you immediately get the result and as I told you earlier that you might try to do it directly, right. That means, you may obtain the cumulative density function for X plus Y, distribution function for X plus Y and then from there you can compute.

(Refer Slide Time: 29:11)



So, having learnt the trick to use m g f for finding out distributions of sums of random variables, I will still write it down for binomial and Poisson and so on. I think Poisson we have already done. Now, let us look at the sum of independent binomial random variables. So, here again the x is binomial n, p and y is binomial m, p, then we want to look at the sum and x and y are independent. So, then again m g f of x plus y will be the product of the individual m g f’s. So, here it is p e raise to t plus 1 minus p raise to n, and for the random variable y, the m g f is p into e raise to t plus 1 minus p raise to m.

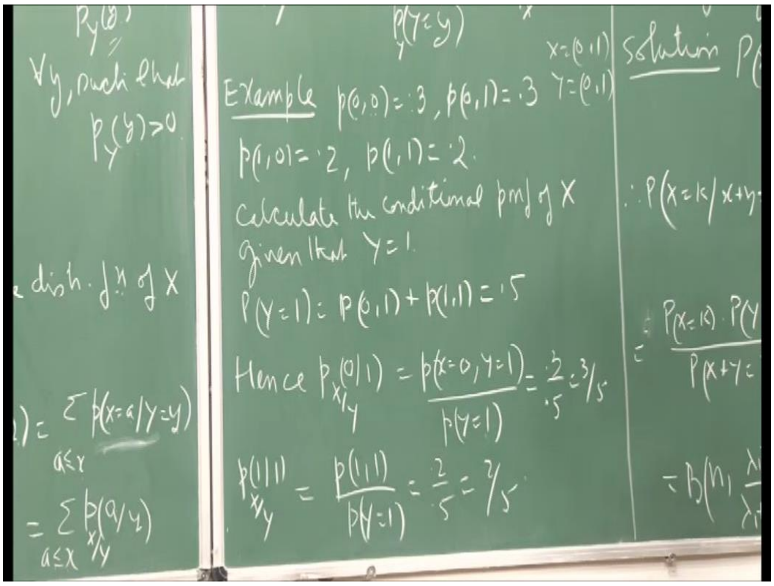
So, when you multiply, the powers get added up. So, this is p e raise to t plus 1 minus p m plus n and therefore, it immediately follows that x plus y is binomial and plus m, p. So, if the probability of success is the same, then if you are looking at two random variables in one case, the number of trials is n. In the other case, the number of trials is m, then the sum will again represent the binomial random variable, either number of trials gets added up. So, probability of success remains the same. So, now, I have seen this thing for sum of these distributions. So, whenever you come across something new, you can read it up and understand.

What is going on? Now, let us look at the conditional distributions also. We have looked at conditional probabilities and we have looked at base conditional probability and so on. So, now, let us look at conditional distributions. So, remember that when E and F for two events, then we define the conditional probability of event E given that event F has occurred and this was defined as probability of E intersection F. That means, both the events must occur divided by the probability of occurrence of F. So, these were the events. Now, when you come to x and y are two discrete random variables, and you want to write down the conditional probability of x given y, where capital Y is let us say small y and x, small x. See if you want to compute this, then it will be again just borrowing it from here. It will be probability x equal to x y equal to small y divided by probability y equal to y which you can write in your notation as p x, y upon probability y equal to small y.

Now, sometimes I may write this suffix, sometimes I may not. So, it does not matter, but you understand from the context that this is for a single variable and this is for a joint p m f, right. So, this is for all y. That means, this conditional probability is defined for all y, such that probability y is greater than 0. I am dividing by number, which I must ensure is positive which is non-zero and since probabilities cannot be negative, so the number must be positive. So, for all possible values of y for which there is a positive probability, I define it this way, right, so conditional p m f of x given this.

Therefore, this defines the conditional p m f of x, given that y is equal to y. Now, conditional cumulative distribution function of x given y is equal to y would be you know f x given y. So, that will be probability x less than or equal to small x, given that y is equal to y which is then probability x equal to a, given that y is equal to y and you are summing up over all a for which is less than or equal to x. So, therefore, the conditional notation is the conditional probability of a given y, where a is less. So, you are summing up over all a less than or equal to x.

(Refer Slide Time: 33:28)

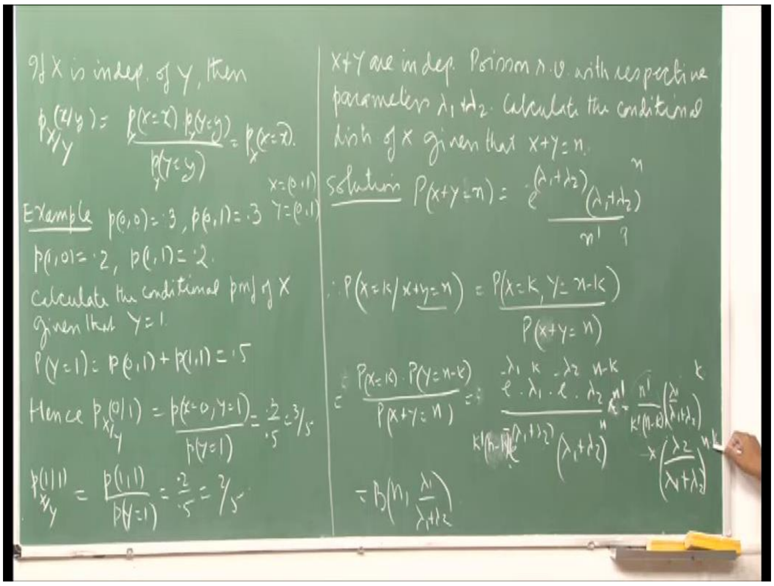


Now, if x is independent of y, then we know that this probability, the conditional probability will be written as because here this will be the product of 2 and divided by probability y equal to y. So, therefore, it will reduce to probability of x equal to x. So, therefore, we are just trying to show you that whatever we did for the events, the same thing goes over for the random variables and there corresponding distributions. Now, just look at this example. If you given that probability of 0, 0 is 0.3. That means, x taking the value 0, y taking the value 0.

So, therefore, your x takes the value 0 1 and your y takes the value 0 1. So, therefore, the four probabilities p of 0 1 is 0.3, p of 1 0 is 0.2 and p of 1 1 is 0.2. They all must add up to 1. So, you want to calculate the conditional p m f of x, given that y is equal to 1. So, first of all you need probability of y equal to 1, right in the denominator. So, therefore, this is equal to p of 0 1 plus p of 1 1, right. That gives you the marginal of y which is a probability of y equal to 1. So, that is 0.5 and hence, probable conditional probability of x given y equal to 1. So, if you want to find out, then you see the possible value of x has 0 and 1. So, you will find out both the probabilities, a conditional 0 given 1 y equal to 1. So, that will be 0 1 divided by y equal to 1, probability of y equal to 1.

So, 0 1 from here is 0.3 divided by 0.5 and that is equal to 3 by 5. Similarly, probability x equal to 1 when y is given to be 1, so that will be p of probability 1, 1 divided by probability y equal to 1. So, it will be 0.2 divided by 0.5 and so this is 2 by 5. So, similarly you can compute the p m f well, ok.

(Refer Slide Time: 35:42)



So, this conditional, you can fix a value of x and then, compute the conditional p m f of y. Yes. Now, this is another interesting example. This says that if x and y are independent Poisson random variables with respective parameters lambda 1 and lambda 2, calculate the conditional distribution of x given that x plus y is n. So, now, the condition is on the sum x plus y equal to n and you want to find out the conditional distribution of x. So, first let us compute probability x plus y, and as I just showed you just a few minutes ago that if x and y both are independent Poisson random variables, their sum will be also Poisson and the parameters will get added up.

So, this is e raise to minus half lambda 1 plus lambda 2 lambda 1 plus lambda 2 raise to n divided by n factorial. So, this is easy because we have already of course, seen the distribution for x plus y then yeah you now want to compute the probability for example, x equal to k when x plus y is n. So, for finding the conditional probability of x given that x plus y is n. So, now, if x is k, then this says that your y must be n minus k, right. See I mean here you will be writing probability, yeah. So, x equal to k and x plus y equal to n that will be product and then, divide it by probability x plus y equal to n. So, intersection of x equal to k and x plus y equal to n is equivalent to the event that x is k and y is n minus k divided by probability x plus y equal to n, and since again x and y are independent, this probability I can write as the product of individual probabilities. So, this will be a probability x equal to k into probability y equal to n minus k divided by probability of x plus y equal to n, and this both being x and y, both being Poisson, this is e raise to minus lambda 1 raise to k divided by k factorial. Then, the other probabilities e raise to minus lambda 2 raise to n minus k divided by n minus k factorial. This may not look very, let me rewrite n minus k factorial, but it is there and then, e raise to lambda 1 plus lambda 2 probability of x plus y equal to n which we wrote down here.

So, this is lambda 1 plus lambda 2 raise 1. Then, there should have been e raise to, yeah lambda 1 plus minus e raise to lambda 1 plus lambda 2 and n factorial goes to the numerator. So, therefore, collect these terms n factorial divided by k factorial and n minus k factorial that comes here, right and then, you see here this is lambda 1 plus lambda 2 raise to n and you have lambda 1 raise to k and lambda 2 raise to n minus k. So, I break up into lambda 1 plus lambda 2 raise to k into lambda 1 plus lambda 2 raise to n minus k. So, then I get that terms lambda 1 upon lambda 1 plus lambda 2 raise to k, and lambda 2 upon lambda 1 plus lambda 2 raise to n minus k and you see these two numbers. That means, lambda 1 upon lambda 1 plus lambda 2 plus lambda 2 upon lambda 1 plus lambda 2. This adds up to 1. So, if I denote this by p and this number is 1 minus p, right.

So, in that case, then this looks like that means a conditional probability x equal to k, given that x plus y is n is binomial. So, therefore, different values of k you will get these probabilities which are exactly the binomial probabilities for the parameters n, and your probability of success is lambda 1 upon lambda 1 plus lambda 2. So, I think through this course I have been trying to show you that even though you have these different random variables, you how you can get through process of addition, conditional , and so on. You can see the connections between the various distributions here. And therefore, you know that makes those things more interesting and of course very useful also.