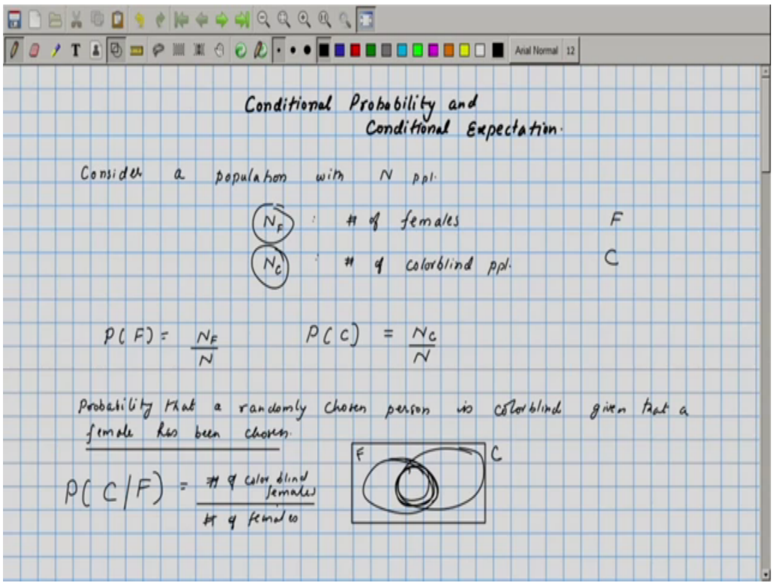
Randomized Algorithms

Benny George Kenkireth
Department of Computer Science & Engineering
Indian Institute of Technology, Guwahati

Lecture – 06
Conditional Probability and Conditional Expectation

(Refer Slide Time: 00:24)



In this lecture, we will learn about Conditional Probability and Conditional Expectation ok. These are two important tools, which we will require for the analysis of many Randomized Agorithms. So, we will start with an example, let us say we consider a population with N people out of which let us say N F are females, this is the number of women. And let us say N C is a number of color blind people ok.

So, we have two events, females and color blind people. We are randomly choosing some person from this group of N people. And when we say randomly each person is equally likely to be chosen. So, when we do this the events that we were interested in are N F and N C. N F the events are F and C; F denotes the event that the randomly chosen person as a female and C denotes the event that the randomly chosen person as color blind.

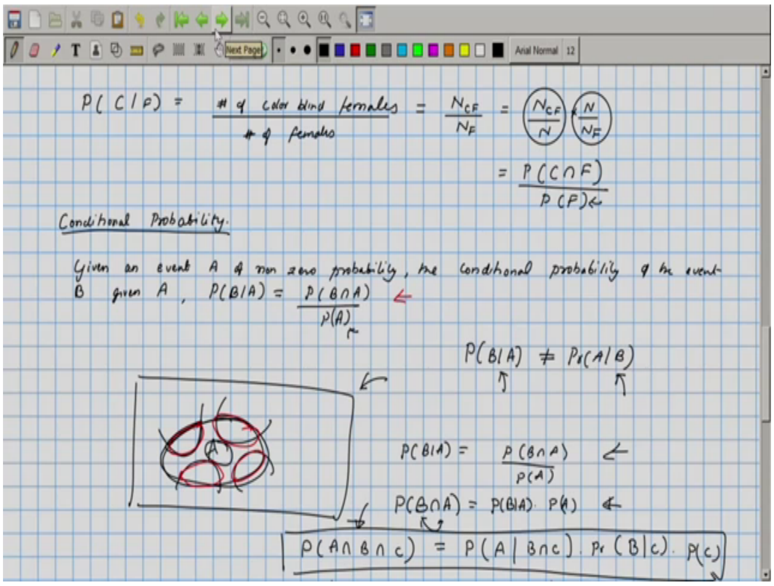
Clearly, the probability of F is equal to N F divided by the total number of people. And the probability of C is going to be equal to N C by N. Now, when we talk about conditional probability, what we are interested in is the event that the person is color blind, there is a conditioning on a particular event. So, let us say suppose we are told that the randomly chosen person is a female, now what is the probability that the person was color blind.

So, we are interested in probability that a randomly chosen person is color blind given that a female has been chosen ok. So, someone has made a choice and it is being told you that the person who is picked was a female. Now, you are asked what is the probability that the person is color blind.

If you look at our sample space, this is the set of females and this is a set of color blind people, they could have an intersection, they may not have an intersection, but this is the general diagram. Now, we are told that the person is a female, then the probability that they are color blind is essentially the size of this region divided by the region of females.

Now, the words that going to be so this probability, we denote will give an notation for that they will denote it by probability C given F ok. We have introduced a notation, but the notion behind is essentially the same. We are interested in computing a probability, but we are restricting the sample space. Instead of considering the entire sample space, now we are restricting to the sample space which consist of only females. So, this probability should essentially be number of color blind females divided by number of females. So, fraction of color blind peoples amongst women that will be this probability.

(Refer Slide Time: 05:06)



So, the probability of C given F is nothing but number of color blind females by number of females. So, we can say that this is number C F divided by number F, which we can rewrite by N C F divided by into N by N F and that is just the probability of. So, N C F divided by N is just the probability of C intersection F divided by probability of F ok, so that would be the probability that a randomly chosen person has color blind given that a female has been chosen.

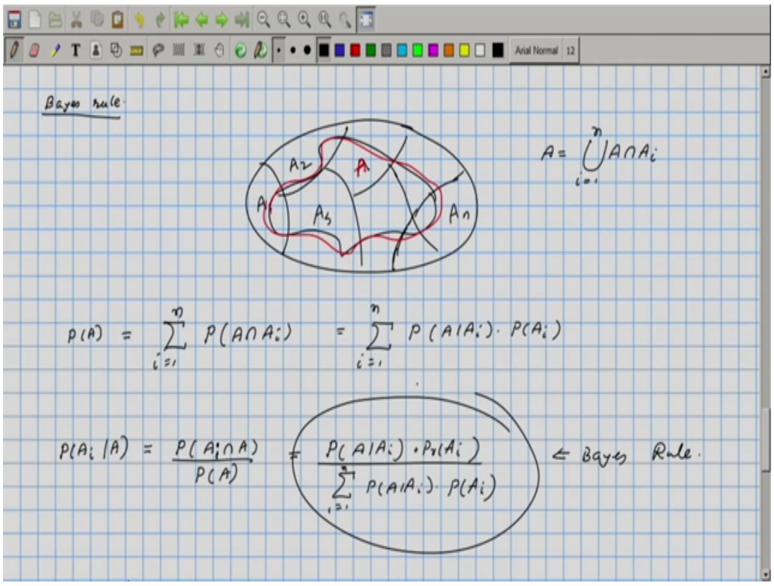
So, the general definition would be something like this. So, we could condition on any possible event, but the event on which we are conditioning in this case the event that the randomly chosen person is a female that should be an event of non-zero probability. So, given an event A of non-zero probability, the conditional probability of the event B given A written as probability of B given A is equal to this is defined as probability of B intersection A by probability of A ok.

So, we can take a sample space a probability space and based on a particular event say A we can define conditional probabilities ok. This is your complete sample space and A was a particular event in it. Now, we are having a new probability space where universe is just A and all the events that are there in this is essentially obtained by taking its intersection with A. So, these would be the new events. And each of those events they will have a conditional probability as given by this particular formula.

Note that probability of B given A may not be equal to A given B ok, because here the conditioning is based on A, here the conditioning is based on B. So, we can write this formula probability B given A is equal to probability B intersection A by probability A. As probability of B intersection A is equal to probability B given A into probability A. So, this can be thought of as a multiplicative form of this particular formula ok.

So, this is a very useful formula, if we are to compute the probability of an intersection ok. So, the probability of A intersection B can be computed as probability of B given A multiplied by probability of A. And this can be generalized as if you had A intersection B intersection C, this probability is equal to probability of say A given B intersection C into probability of B intersection C, which can again be written as B given C into probability of C. So, can be thought of as a chain rule ok. So, if you had intersection of many events conditional probability can be used to determine the probability of the intersection, conditional probability is there often easier to find, if you have if you set up the sample space appropriately.

(Refer Slide Time: 10:09)



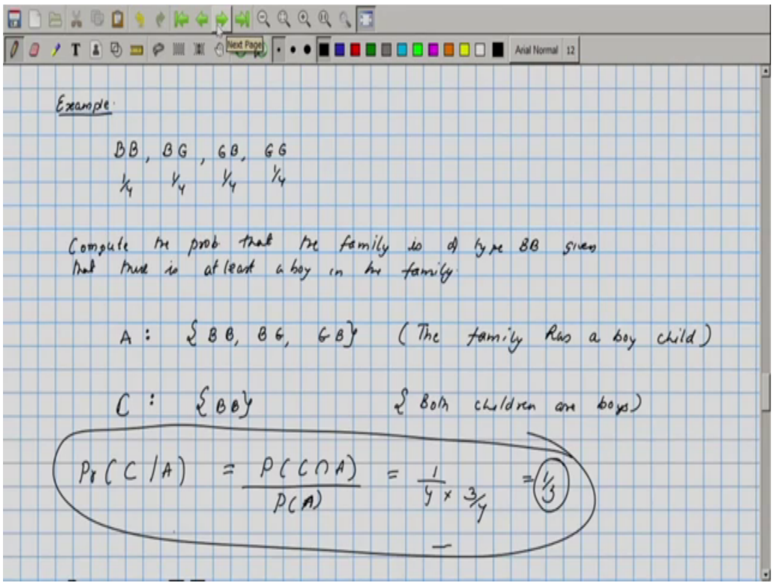
We will see one famous rule called Bayes rule. Now, let us describe a sample space this is our sample space and it is formed by many different events ok. So, these events form a partition of the sample space, let us call that as A 1, A 2, A 3 up to A n. So, we have a sample space which is been partitioned by events A 1 to A n, this partition these events do not have any intersection between them A i intersection a j is empty for each these events.

Now, if you were given a particular event A, event drawn in red color, now that events probability how do we determine? Clearly, the event A is equal to the union of A intersection A i's. So, A intersection A 1 union A intersection A 2 and so on that union would give us A. So, probability of A the rules of the axioms of probability, we can say that is just the summation of probabilities of A intersection A i, i going from 1 to n.

And to compute the probability of this intersection, we can use conditional probability, and that would be probability of A given A i times probability of A i. Now, what Bayes rule tells us is suppose we are told that the event A has occurred ok, what is the probability that A i has occurred. So, we want to compute probability of A i given A ok. So, this can be thought of as probability of A i intersection A divided by probability of A, Probability of A we can use this formula.

And for probability of A i intersection A, we can write it as of A given A i into probability of A i and the denominator is summation i going form 1 to n probability of A given A i into probability of A i. So, this rule is called as the Bayes rule. So, we are interested in probability of one of these A i's happening given that A has happened that can be expressed in this particular format.

(Refer Slide Time: 13:33)



Let us do an example involving conditional probabilities. So, let us look at a family two children ok, and we are interested in the gender of the children. So, there are four possibilities. The first child could be a boy and the second child also could be a boy. We are looking at the case where there are precisely two genders basically a boys and girls. So, BB is a possibility, BG is the possibility which means the first child is the boy and the second child is a girl and GB is the possibility and GG is a possibility ok.

So, our sample space would consist of these four outcomes. If you look at one particular family and they will be in one of these four classes. And we can assume that all of them will occur with one-fourth probability. Assuming that boys and girls are equally likely in any pregnancy, we could say that boy boy happens with one-fourth probability, boy girl happens with one-fourth, girl boy happens with one-fourth, and girl girl happens with one-fourth.

Now, here is a question suppose somebody tells that a particular family that are at randomly chosen family has a boy, we are interested in computing the probability that it is a family with two boys ok. So, we are not told whether it is a first child which is a boy or second child which is a boy, we are just told that the family has a boy. It could be that there is precisely one boy, it could be both children are boys ok.

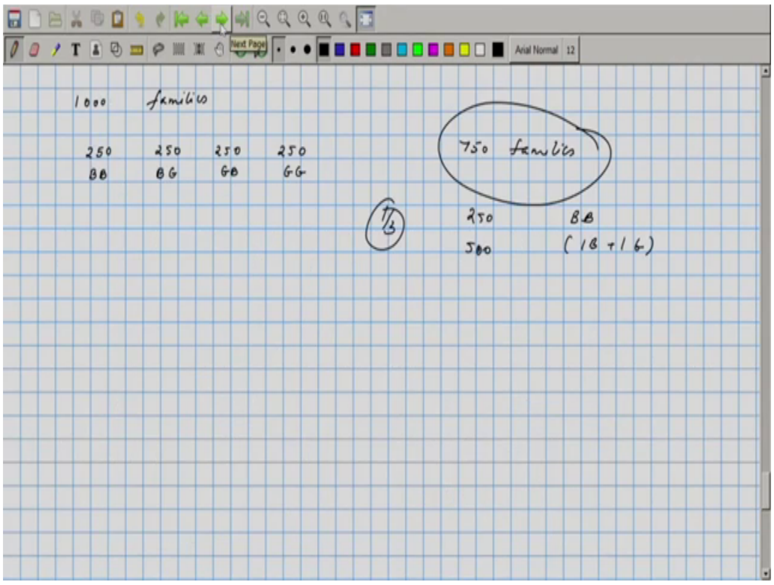
And we are interested in computing the probability that the family is of type BB, which means both children are boys given that there is at least a boy in the family. Think about it for a minute I will pause the video and think it, then will work out the answer and see what it is.

So, let us identify our events ok. Our event A is the event on which we are conditioning is the event consisting of BB, BG, GB. In other words this means, the family has a boy child. The event we are interested in is let us call it as C that would be just BB ok, both children are boys.

So, we are interested in probability of C given A by definition this is equal to probability of C intersection A divided by probability of A. C intersection A happens with one-fourth probability and A happens with three-fourth probability, because all of these three outcomes are disjoined outcomes they happen with one-fourth probability each add up to three-fourth so, this is one-third ok.

So, if you are told that the family has a boy, the other possibility although there are only two possibilities namely the other child is either a boy or a girl still the answer is not 1 by 2 as some of you might have thought, it is 1 by 3. Now, what is the intuitive explanation for this?

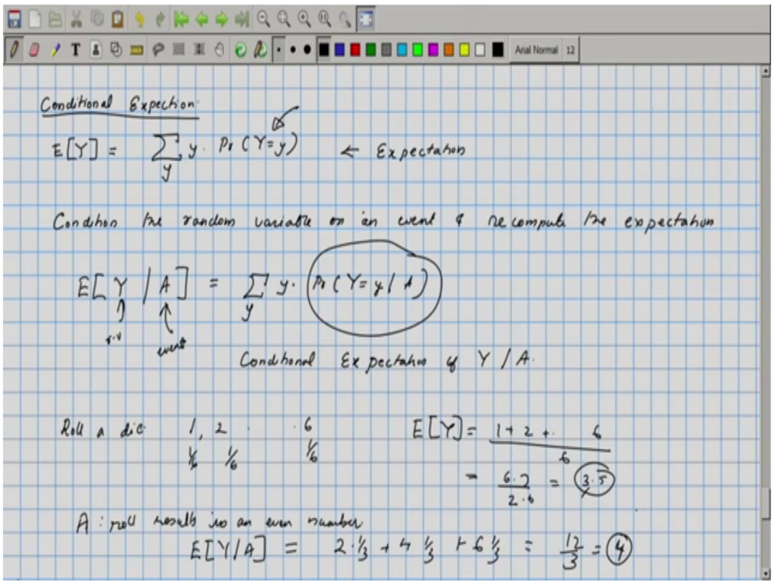
(Refer Slide Time: 18:19)



So, suppose we had 1000 families ok. So, these 1000 families where chosen from our sample space and each these family could either be a BB family or a BG family or a GB family or a GG family. So, we can assume that 250 of them be boys-boys type, 250 would be boy-girl type and 250 would be girl-boy type and 250 remaining 250 would be girl-girl type.

When you focus your attention on families, which contains a boy and at least a boy, those are essentially 750 families. Out of these 750 families 250 of them are boy-boy type and 500 consist of 1 boy plus 1 girl ok. And so you would expect the ratio to be 1 by 3. Now, that is what precisely this calculation shows us ok. So, this is a brief introduction to conditional probability.

(Refer Slide Time: 19:37)



The next thing we will see is something called as conditional expectation. So, just like conditional probability, which is probability based on a condition conditional expectation as essentially expectation based on a condition. There are two distinct notions, which goes by the name conditional expectation will see both of them. And different areas when people say conditional expectation, they might mean either of these notion, they are related but not the same ok.

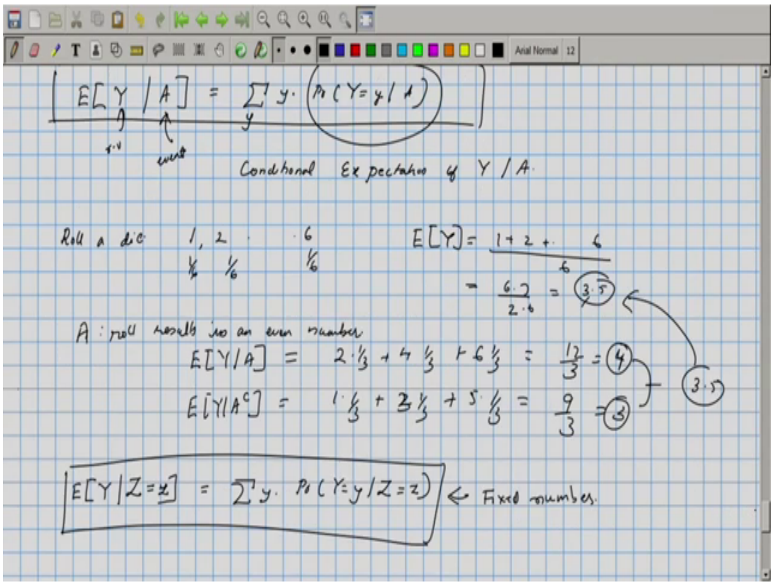
So, let us see the first notion of conditional expectation. So, expectation is always with respect to a random variable. So, let us say will have random variable and it has a certain underlying distribution, then we have it is expectation of Y is just nothing but summation overall the values that Y can take Y into probability that the random variable takes the value Y ok, this is the expectation. We change this probability to a conditional probability and what we get is called as conditional expectation.

So, we have an event so condition the random variable on an event and re-compute the expectation that the informal definition, this would be the conditional expectation. So, more firmly expectation of Y by its random variable given a event A. So, here A is an event, this is equal to sum over all y's, y into probability that the random variable Y takes the value Y given the event A has occurred ok. So, this is the conditional expectation of Y given A.

So, let us look at an example. If you roll a die ok, let say the values that can come are 1, 2, 3, up to 6 all with probability 1 by 6. The expectation, so let us look at the outcome will call it as Y. So, Y is the outcome and Y can take all these values 1 to 6 with equal probability. The expectation of Y would be 1 plus 2 plus up to 6 by 6 that is going to be 6 into 7 by 2 into 6 that is 3.5. This is the expectation of the random variable.

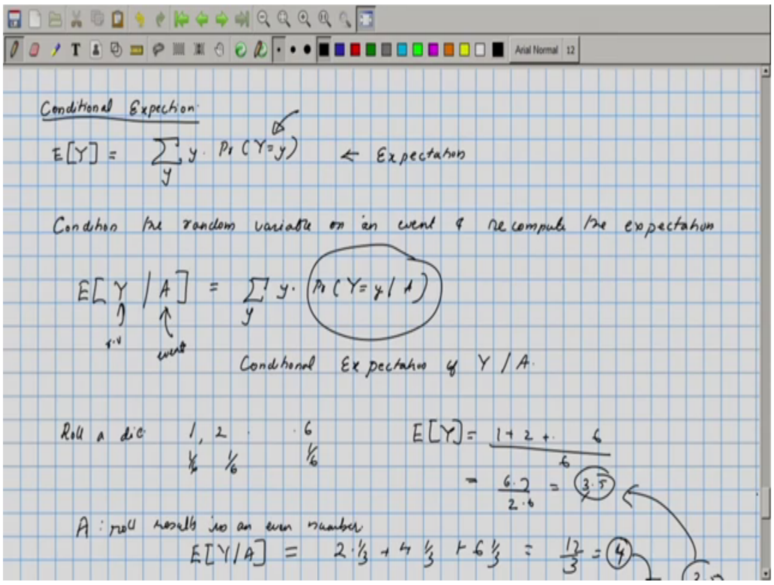
Now, if I say the event A has occurred and A being that the roll results in an even number ok. So, we have told that the result is an even number, now what is a probability. you have to re-compute the probabilities and weigh it accordingly when you do the summation. So, the conditional expectation of Y given this event A is equal to will there are three even numbers that could appear 2, 4 and 6 and all of them being equally likely ok. If you add them up, you will get 12 by 3 that is going to be equal to 4. Of course, if you condition on a complement, you will get say 3 ok.

(Refer Slide Time: 23:37)



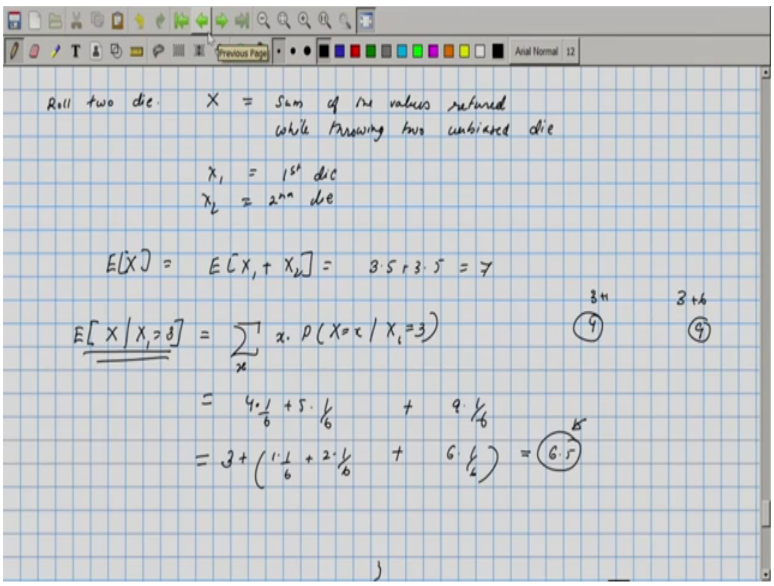
Expectation of Y given another event A complement that is going to be 1 into 1 by 3 plus 3 into 1 by 3 plus 5 into 1 by 3, it is 9 by 3, it is going to be equal to 3. You can see that even and odd are equally likely and these conditional expectations when averaged out over the probability of the events will get 3.5, which is the expectation of the unconditioned random variable.

(Refer Slide Time: 24:13)



When we talk about conditional expectation, the event is often described in terms of a random variable ok. So, the form that we will be most commonly using is the conditional expectation of Y given the random variable Z takes a value small z ok. So, instead of talking about a arbitrary event, we are looking at events which are described using say random variables ok. So, this is equal to summation over y into probability that Y equal to y given random variable Z take a value z ok. So, this conditional expectation, this is going to be a fixed number both of them. If you condition on an event based on a random variable, they are all going to be fixed numbers.

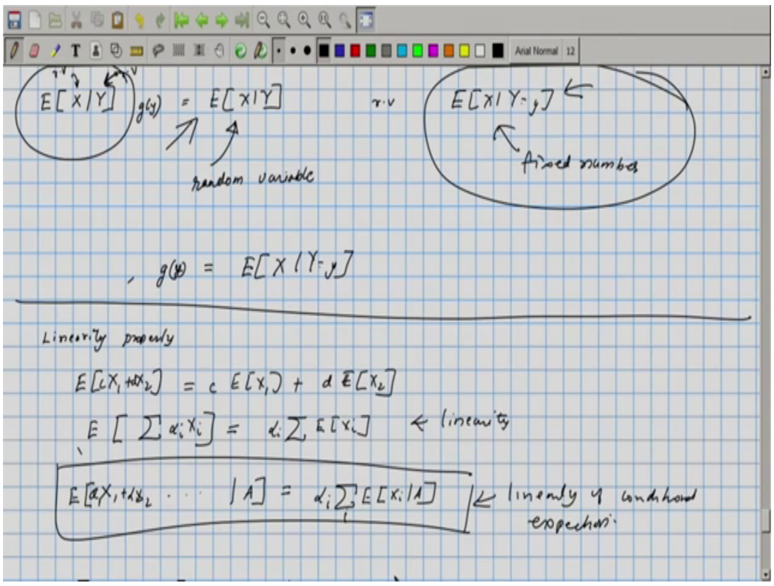
(Refer Slide Time: 25:19)



Let us do one more example. So, suppose the roll two die, so let X denote the sum of the values returned while throwing two unbiased die ok. X 1 let us say that this is the value on the 1st die and X 2 is the value on the second die. The random variable X it is expectation is clearly going to be expectation of X 1 plus expectation of X 2, linearity says this is going to be 3.5 plus 3.5 that is going to be 7.

We could look at say conditional expectation X given X 1 equals 3 by definition this is going to be sum over all values that capital X can take let us denote it by x times the probability the conditional probability that X is equal to x given X 1 equals 3. If X 1 equals 3, then the sum can range from let us say 4. So, 3 plus 1 to 3 plus 6 from 4 to 9, these are the likely value. All the other values happen with 0 probabilities. So, this is going to be equal to 4 into 1 by 6 plus 5 into 1 by 6 plus 9 into 1 by 6 that is going to be 3 plus 1 into 1 by 6 plus 2 into 1 by 6 plus 6 into 1 by 6 that is going to be 6.5. So, the conditional expectation of X given X 1 equals 3 will be equal to 6.5.

(Refer Slide Time: 27:38)



The second notion of conditional expectation, although it is called as conditional expectation, the notion is slightly different. So, you will write a expectation of X given Y. Here X is going to be a random variable and Y is also going to be your random variable.

In the earlier case X was a random variable, whereas Y the second parameter was an event that event could have been described either has an event or by means of a random variable taking a particular value, but here it is just random variable ok. And therefore, this value is not in some sense fully determined. So, in this notation when you write conditional expectation of X given Y, what we mean is a random variable whose value depends on the value taken by Y ok. So, this conditional expectation is a random variable, whereas the other conditional expectation E of X given by Y equals y, this was a fixed number.

In all cases context will be make it clear, whether we are using this notion or this notion. In this course, we will essentially be using only the fixed number notion, but it is good to know what is the random variable corresponding to conditional expectation. So, here this is the random variable and the values so we can think of this sample space for this random variable as the values that Y can take ok. So, if let us view this as a function g of small y so g on small y, it is value is equal to expectation of X given Y equals y ok.

So, as the value of the random variable changes, the conditional expectation takes different values. By describing certain properties of conditional expectation with respect to usual expectation, we had linearity property which says expectation of X 1 plus X 2 is c X 1 plus d X 2 is equal to c times expectation of X 1 plus d times expectation of X 2 or more generally expectation of a summation alpha i X i is going to be equal to summation alpha i expectation X i ok. So, this is the linearity property of expectation.

One can show that this holds for conditional expectation as well. And for both the notion whether you treat it as a random variable or as a fixed number. So, the conditional expectation of X 1 plus X 2 so on given an event A is same as summation expectation of over i expectation of X i given A ok. If you had constants say alpha 1, alpha 2, you could have alpha i here as well ok. So, this is basically the linearity of conditional expectations.