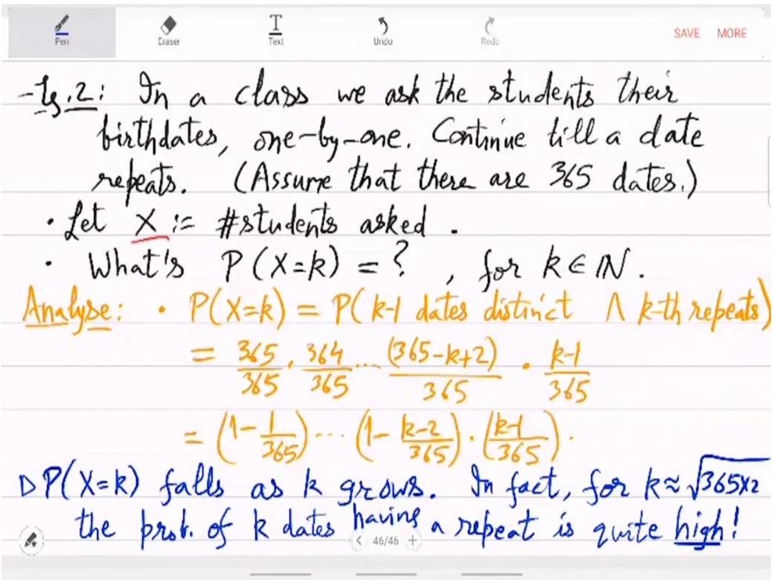
Probability for Computer Science
Prof. Nitin Saxena
Department of Computer Science and Engineering
Indian Institute of Technology - Kanpur

Module - 3
Lecture - 9
Expectation

We started random variables and their expectation. So, random variable is simply a numerical value. It is a variable that takes a numerical value corresponding to every element in the sample space, usually denoted by a big X or big Y. So, the probability of big X equal to small x value is simply the sum of these respective probability of the events. This is called probability mass function for sum of probability omega. And then we saw an example. And let us now see more examples.

(Refer Slide Time: 01:03)



So, second example is, in a class, we ask the students their birthdays or their birthdate one by one. So, we do this questioning one by one, one student at a time, it will not repeat the students, and then we stop whenever there is a birthdate repeated. So, continue till a date repeats, then you stop. And you can assume that there are 365 possibilities. Assume that nobody was born on twenty-ninth Feb. for example. So, there are 365 possibilities.

And, so, every student; if you pick a random student, then probability of some date is 1 over 365. So, let X be the number of students asked. So, that is the random variable of interest in this experiment that how many students were asked before a date repeats. So, what is the probability of X being k? What is the probability that this random variable comes out to be a number k, a natural number fixed in advance? So, let us do this analysis.

So, probability that X = k is basically k - 1 birthday is distinct and kth repeats. So, what is the chance that the first k - 1 are distinct? So, that is, the first one can be anything; so, 365 over 365. Then the next one; the first one is taken, so, next is 364 and so on, k - 1 times. So, 365 minus k plus 2 divided by 365, right? So, first, k - 2 have to be avoided. So, that is the chance of k - 1, first k - 1 being distinct. Now, the kth one can be any of these.

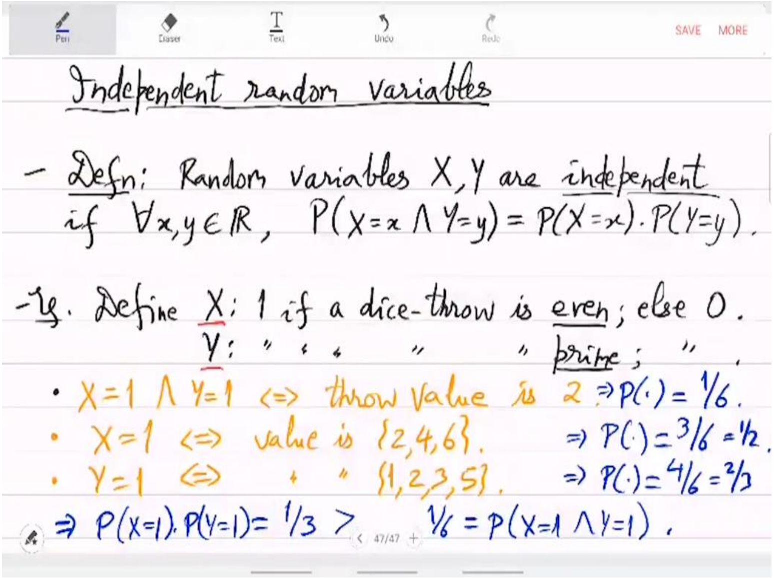
So, k - 1 possibilities over 365. So, simply by the definition of probability, which is favourable events or favourable outcomes divided by total outcomes, you get this expression. So, this expression simplifies as 1 minus 1 over 365 .. 1 minus k minus 2 by 365 times k - 1 by 365. So, that exactly is the probability of X being k. Now, this, for example, when k is 1, then this probability is 0. For k equal to 2, it is still pretty small.

For k equal to 2, it is just 1 over 365. But as k grows, you can show that this probability actually becomes decent. As k grows, this probability actually is good. And no, sorry, as k grows, this is actually decreasing. But anyways, so, actually this probability is falling with k. This decreases with k. So, probability of X equal to k falls as k grows. And already for k equal to 23 or 24, already it is quite small.

So, if you do this experiment, very quickly you get a repeating birthdate. In fact, for k around square root of 365 and maybe double that. So, for k around square root of 365 times 2, the probability of repetition is good. So, the probability of k dates having a repetition is quite high. This is also called the birthday paradox. So, around k 25, if you randomly pick 25 students, then, with a good chance, 2 will have the same birthday, birthdate.

So, it is a paradox because you would not expect that to happen, because it is only 25 people and there are 365 dates. But still, this calculation shows that actually very quickly you start getting same birthdays. That is a side conclusion of this. So, next thing is, what is the meaning of independence for random variables?

(Refer Slide Time: 08:45)



So, recall independent events. How was it defined? It was defined via their intersection and then via conditional probability. So, conditional probability does not change, it remains the same as before, then you want to say that the events are independent. Similarly, here, you will say that the variables are independent. So, random variables X, Y are called independent if for all values x, y, in the range which we are assuming to be real numbers, the probability that X, big X's value takes value small x and big Y takes value small y.

This probability is just the product. So, when this happens for every possible pair of values, then you say, and also it is intuitively clear that these 2 random variables are independent. One does not increase or reduce the other variable. Fixing big X does not say anything about big Y's value. Let us take an interesting example. Define X to be 1, if a dice throw is even, else define it to be 0. So, that is with respect to evenness.

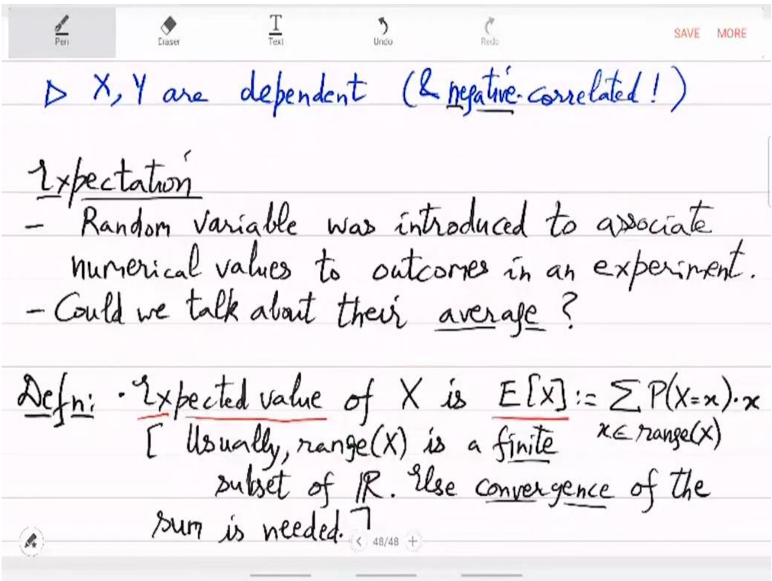
And second random variable, 1 if a dice throw is prime, else call it 0, give the value 0. So, now the question is whether these 2 random variables; these are, the experiment is throwing off a dice with 6 numbers, 1 to 6. So, does this random variable which depends on parity and the other one which depends on primality, are these 2 dependent or independent? So, intuitively, they should be dependent, right?

Because, even means that not prime. In general, you cannot, you do not expect even number to be prime or prime number to be even. So, it should come out to be dependent, X and Y should be dependent, because one implies some information about the other. So, by this definition, what you will do is; let us do the analysis. So, X equal to 1 and Y equal to 1. So, I am already fixing small x and small y to be both 1.

So, what is the probability that both of them are 1? So, that means, the throw value is; so, when are both the events true, even and prime? That is only when 2 is thrown, right? 2 is the only even prime. So, which means that, the probability is 1 over 6. Next is X equal to 1. So, even means 2, 4, 6. And Y equal to 1 means prime. So, the value is 1, 2, 3, 5. So, what are the probabilities? So, probability here means 3 by 6. And this probability is 4 by 6.

So, you can see already, see that these last 2 probabilities, their product is 1 by 3. This product probability or product of the probabilities is 1 by 3 which is greater than 1 by 6 which is the probability of intersection. So, in other words, to put it simply, a intersection has a smaller probability. So, these 2 variables are actually anti-correlated. If one happens, the other is likely not to happen, and in particular they are dependent.

(Refer Slide Time: 15:16)



So, X, Y are dependent and anti-correlated or negatively correlated. So, in particular; so, you now also have a definition of what is positive correlation, what is negative correlation; in terms of numbers you get the definition of these terms. Next is expectation of a random variable. This is a very important quantity. It will tell you; so, essentially, why did you introduce random variables?

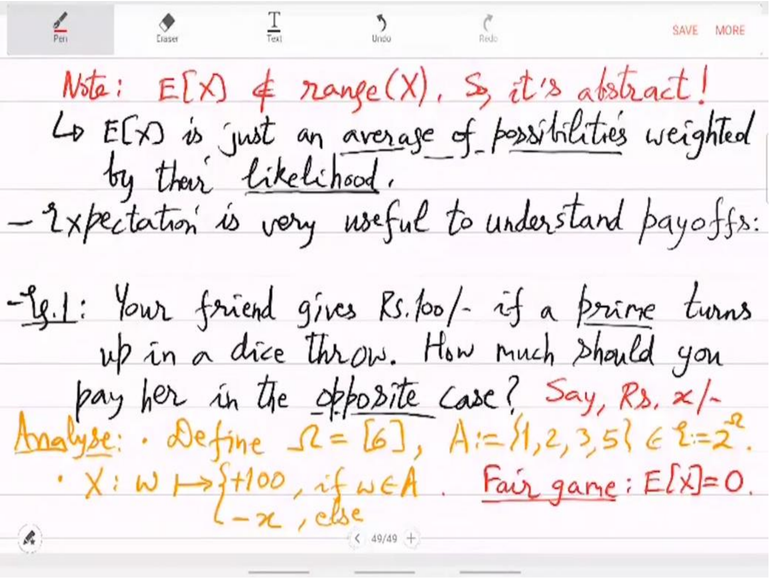
So, random variables were introduced to associate numerical values to outcomes in an experiment. But without doing the experiment, we also want to understand the average of these numerical values. So, could we talk about their average? That is the question at which expectation will answer. Could we talk about an average numerical value, a numerical value which is representative of all possible outcomes in a random experiment?

Without doing the experiment, what can you say is a representative number, a single number. So, that is obviously, in statistics its average, in probability it is expectation or expected value. And the definition will be the natural one which you will expect. So, expected value of X is defined like this. It is, you go over all the values possible in the range of big X, and that value times the probability. So, this probability is basically a weighing X.

So, higher the probability, higher the contribution of X. And obviously, if the value of x is large, then also it contributes more, but it depends on both. So, the total contribution is product of the probability and the value. So, if a value is very large but the probability is very small, then the contribution is around 0 in this sum. So, it is completely intuitive way to define. And obviously, usually the range is just real numbers; is a finite subset of R.

So, usually range is a finite subset of real numbers. If it is infinite, then you have to assume whether the sum converges or not; else you check for convergence, convergence of the sum is needed. Still the definition will be this, but then convergence, it is in sense infinite sum convergence will be required. When it converges, then you see that, that is the expectation.

(Refer Slide Time: 20:47)



One important point is that this expectation, it may not itself be in the range of X, your range maybe a proper subset of real numbers and then the expectation maybe outside that set. What does that mean? It just means that expectation is an abstract value. It may never appear in an experiment as an outcome. So, it is abstract. That you should give some thought about. Expectation is not the outcome which is most possible.

So, it is not the numerical value which appears as an outcome, but it is just an abstract number, like an average. What it says is that, the numerical value in the outcome will generally be around this much, it will not be exactly this much. So, it basically says that expectation is just a weighted average, is an average of possibilities or weighted by their likelihood.

But it may never really appear in the outcome, the outcomes will just be close to this. So, expectation is very helpful, is very useful to understand payoffs in games for example. Payoffs are numerical values; we will say rupees. And understanding payoffs is naturally linked to random variables and then expectation. So, let us see an example. So, your friend gives you rupees 100, if a prime turns up in a dice game, if a prime turns up in a dice throw.

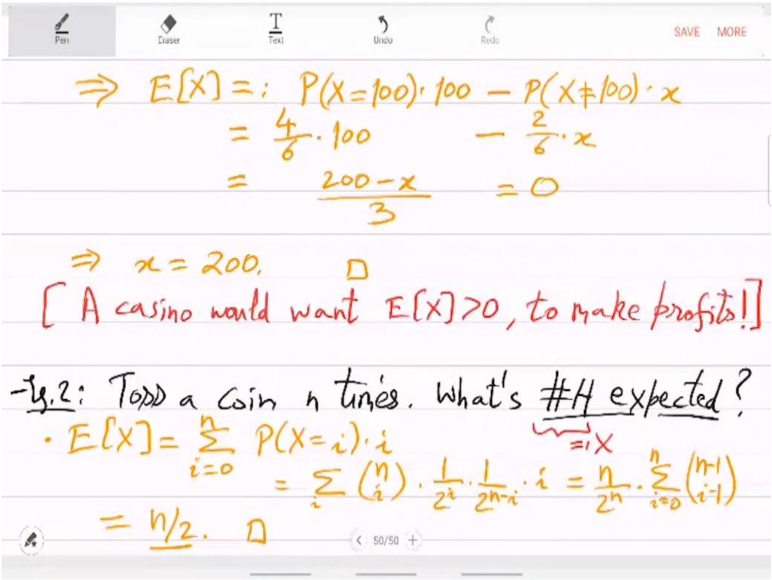
If a prime does not turn up, then you have to pay her something. Now, the question is, how much should you pay in the opposite case, when a composite number turns up. So, if you promise to pay too much, much more than 100, then it might be a bad game. It might be unfair to you. And if you promise to pay much less than 100, then maybe it is unfair to her. So, what should be the number so that both of you play the game with confidence?

So, say it is rupees x, and let us analyse the expectation. So, just to recall; in this experiment, what is the sample space? That is 1 to 6, possible dice throws. What is the event of interest? That is prime numbers. So, it is 1, 2, 3, 5, which is one of the events amongst all the subsets. So, these are the, this gives you the probability space. Probability of course is 1 by 6 for each number. So, this defines a probability space. And then, what is the random variable?

So, random variable is, to a value, dice value omega, it is 100 if omega is a prime, else it is; so, you want to do this analysis from your point of view. So, if omega is an A, you get +100, else you get -x. So, this is your point of view, because you have to decide how much are you willing to give. This is the random variable. And what is the expectation? You do not want the expectation to be negative, you want it to be 0.

And in fact you want it to be positive, but then your opponent will not be interested in that. So, to have a fair game, you have to set the expectation to 0. So, fair game is expectation of X, 0.

(Refer Slide Time: 26:55)



So, what is expectation of X? Expectation of X is; in fact, by definition, it is probability that X is 100 times 100 minus probability it is not times minus x, which is; so, what is the probability that x is 100? So, that is a winning situation for you; that happens with probability 4 by 6. Other case is 2 by 6, composite case, when you have to shell out rupees x. So, which is, so, that is 200 minus x by 3. Correct?

And if you set it to 0, then it means that, to have a fair game, so that both of them agree to play this game, you will be willing to pay 200 for composites and your opponent is willing to pay 100 for primes, because, in this dice throw, primes are more, they are doubly more. So, that is the intuition. And for a game where you win, you should ensure expectation to be positive.

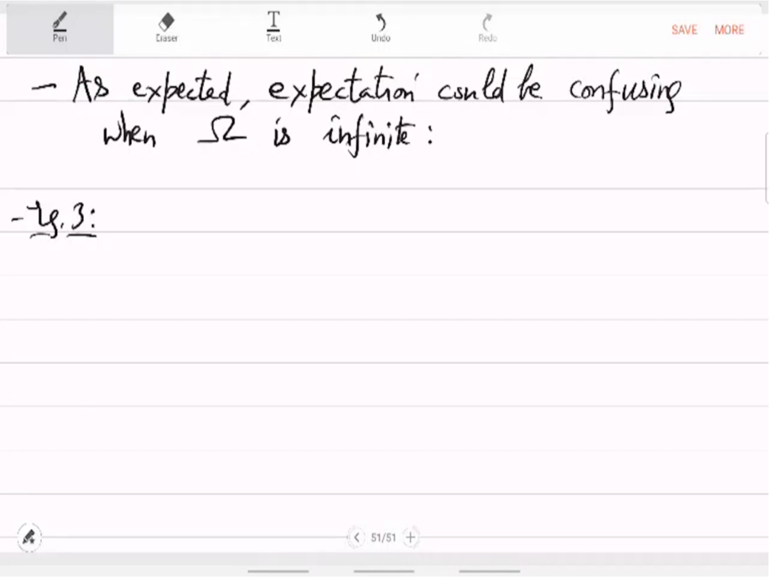
So, for example, if you were designing a casino, then you would want the expectation for you to be positive so that, in a casino, the customers will always lose. So, a casino would want expectation to be positive, to make profits. So, in a casino, games will always be unfair to you. So, those analysis you can do. We will see more examples. So, let us see this other example which is about tossing a coin n times.

What is the number of heads that you get? How many heads do you expect when you toss a coin n times? So, in this analysis, we can do it from the definition. So, you can get i heads, 0 to n. So, probability that this random variable that you are interested in, right; we are calling it X, random variable. So, this is i. What is the chance that there are i heads? So, that is, out of n; so, this is an ordered sequence.

In this ordered sequence of head tails, n many head tails, i places you will choose; and then, there you want head. So, that is 1 over 2 to the i, and the remaining is n - i; and then, the value is i. So, what you get here is, you can simplify it as n by 2 raised to n times i 0 to n and it becomes n - 1 choose i - 1. So, I have simplified n choose i times i, like this. Now, it is a sum of all the binomials, so, you get 2 raised to n - 1.

So, you get n by 2. So, that is it. If you toss a coin n times, then half of the time, you will get heads. This is what it is saying, right? This half is expected. There is no surprise here, no fallacy here. It is for straightforward, and you can do the calculation also quite easily.

(Refer Slide Time: 31:41)



So, as expected, for an infinite sample space; so, till now, the examples we saw, numbers were actually finite, it was very discrete. So, there was no problem in the definition of expectation, no convergence, divergence issues, but expectation could be confusing when omega is, sample space is infinite. It is countable, but still it is infinite; so, you have to do this sum, infinite sum, sum over infinitely many values, say infinitely many real numbers or integers. So, one example of such problems is the problems that you face. Okay, we will do it in a short while.