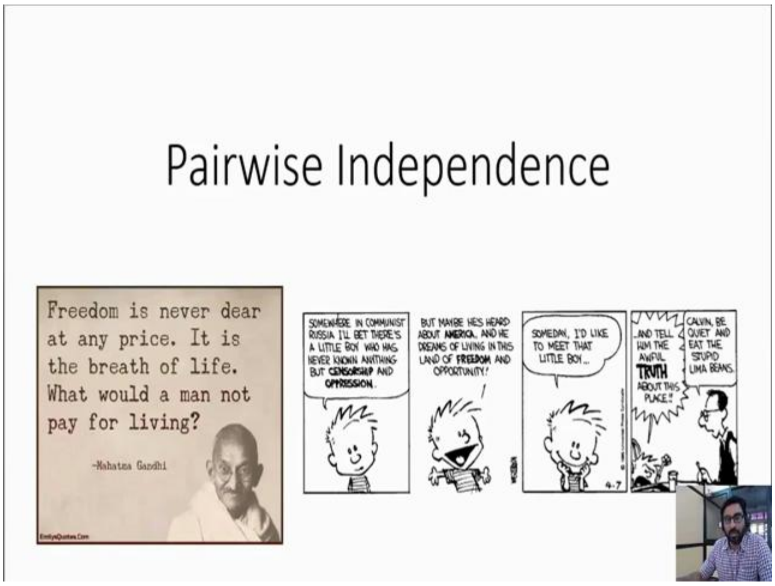
Algorithms for Big Data
Prof. John Ebenezer Augustine
Department of Computer Science and Engineering
Indian Institute of Technology, Madras

Lecture - 27
Pairwise Independence

(Refer Slide Time: 00:14)



Pairwise Independence

Freedom is never dear at any price. It is the breath of life. What would a man not pay for living?

-Mahatma Gandhi

EmilysQuotes.Com
SOMEWHERE IN COMMUNIST RUSSIA I'LL BET THERE'S A LITTLE BOY WHO HAS NEVER KNOWN ANYTHING BUT CENSORSHIP AND OPPRESSION.
BUT MAYBE HE'S HEARD ABOUT AMERICA, AND HE DREAMS OF LIVING IN THIS LAND OF FREEDOM AND OPPORTUNITY!
SOMEDAY, I'D LIKE TO MEET THAT LITTLE BOY...
...AND TELL HIM THE AWFUL TRUTH ABOUT THIS PLACE!!
CALVIN, BE QUIET AND EAT THE STUPID LIMA BEANS.
4-7

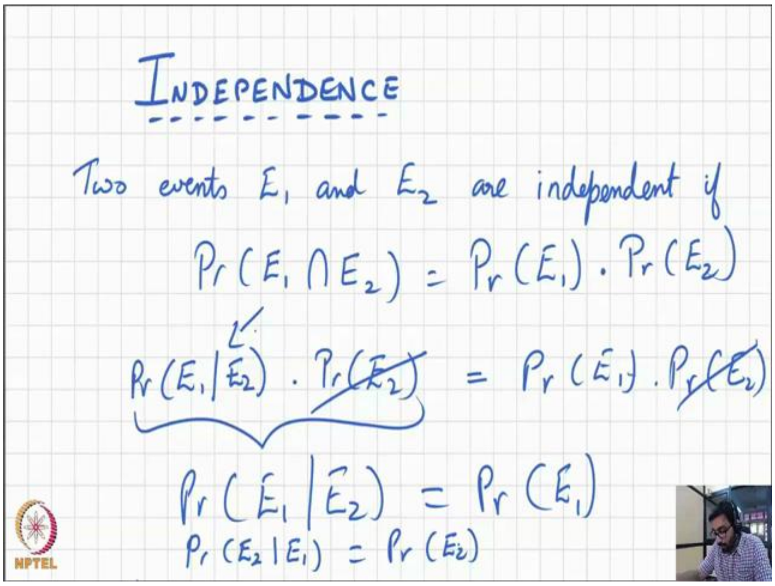
Today’s lecture is about Pairwise Independence. And before we get into the topic of pairwise independence I am like to remind you that the independence that we enjoy is expensive; it is both expensive for us in society; it is also expensive in computer science. Well let us start with society basics a little more interesting.

And so here is a quote from Mahatma Gandhi “Freedom is never dear at any price. It is the breath of life, what would a man not pay for living?” And in the sense he is saying the price of freedom is your life, I mean you should be willing to give your life for the independence that we enjoy, so that is tells you something independence, freedom these are expensive things. And the other aspect in a (Refer Time: 01:13) is that you may think you have independence expect you do not sometimes, and this again is true in computer science we may think we have independent random bits, but maybe we do not. But here is a little cartoon, my favorite Calvin and (Refer Time: 01:33), where you can pause and you know enjoy the cartoon, it talks about the precede independence that Calvin and enjoys in America, but is that true, it is fair interpret the cartoon.

So, in some sense you can think of a place as the variances of independence and still question whether there it is truly independence and that is again true about random bits. Because our time we make analysis of life easy by assuming independence, but is there a good source of independent random bits that is a tough question.

In this lecture, we will try to address this issue and the way we are going to address it, we are going to allow random bits to be less than pairwise independence. And but still maintain some formal understanding of what we mean by that, and that way we can still analyze algorithms or sampling things like that with some mathematical figure, so that is the goal of today’s lecture. So, in some sense we are going to look at random bits without the level of independence that we are used to. So, just let us first remind for ourselves what we mean by that independence.

(Refer Slide Time: 03:13)



INDEPENDENCE

Two events E_1 and E_2 are independent if

$$Pr(E_1 \cap E_2) = Pr(E_1) \cdot Pr(E_2)$$

$$\underbrace{Pr(E_1 | E_2) \cdot \cancel{Pr(E_2)}} = Pr(E_1) \cdot \cancel{Pr(E_2)}$$

$$Pr(E_1 | E_2) = Pr(E_1)$$

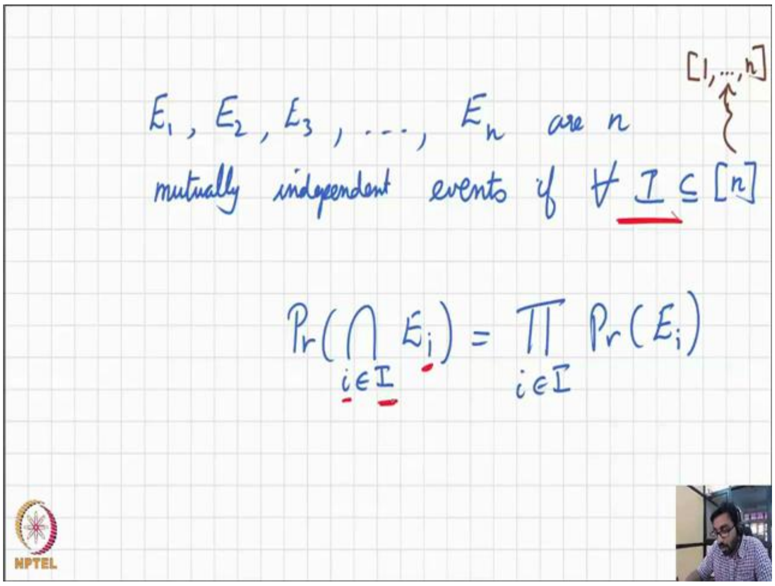
$$Pr(E_2 | E_1) = Pr(E_2)$$

Recall that given to an events E 1 and E 2, we say that they are independent if the probability of this event E 1 intersected with E 2 equals the probability of E 1 times E 2. Another way of stating this precall is probability of E 1 given E 2 times the probability of

E 2. Now that if you recall is nothing but probability of E 1 intersected with E 2; and if that equals probability of this whole left hand side equals probability of E 1 times probability of E 2.

Now, here probability of E 2 can be canceled out, so essentially what we have is that the probability of E 1 given E 2 equals the probability of just the event E 1. So, in other words, the probability the fact that E 2 happen or did not happen has no bearing on the probability of event E 1. And by symmetry, we also have that the probability of E 2 given E 1 equals probability of E 2 when events E 1 and E 2 are independent. So, the independent means here that the rather E 1 occurred or not has no bearing on the probability of event E 2; things become a bit more interesting when you have more than just two events to talk about.

(Refer Slide Time: 04:53)

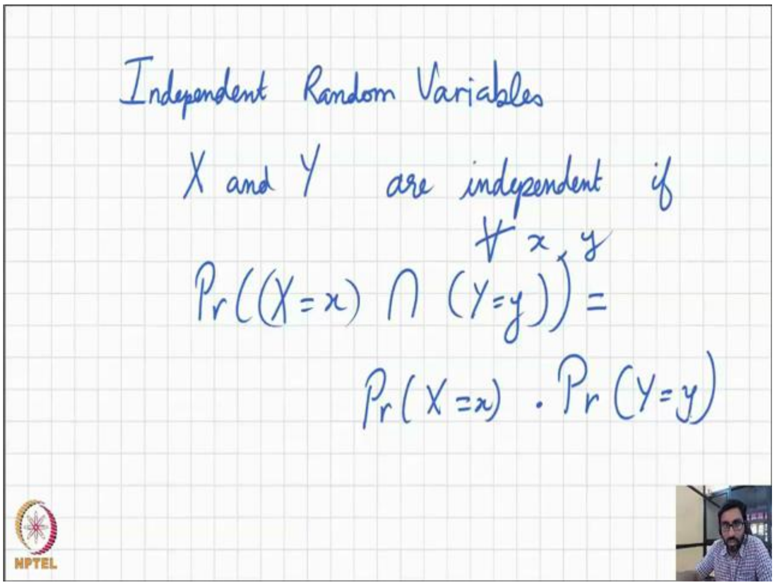


$E_1, E_2, E_3, \ldots, E_n$ are n mutually independent events if $\forall\, \underline{I} \subseteq [n]$ $\xleftarrow{} [1, \ldots, n]$

$$\Pr\Big(\bigcap_{i \in I} E_i\Big) = \prod_{i \in I} \Pr(E_i)$$

So, let us consider some n events E 1 through E n. And we say that these n events are mutually independent, if for every subset of 1 through n basically subset of indices used to denote these events for every such subset I, the probability of the intersection of all events E little i, where little i belongs to the set I equals the product of the individual probabilities. And let me emphasize that this has to be true for all subsets i of the index set 1 through n.

(Refer Slide Time: 05:53)



Independent Random Variables

X and Y are independent if

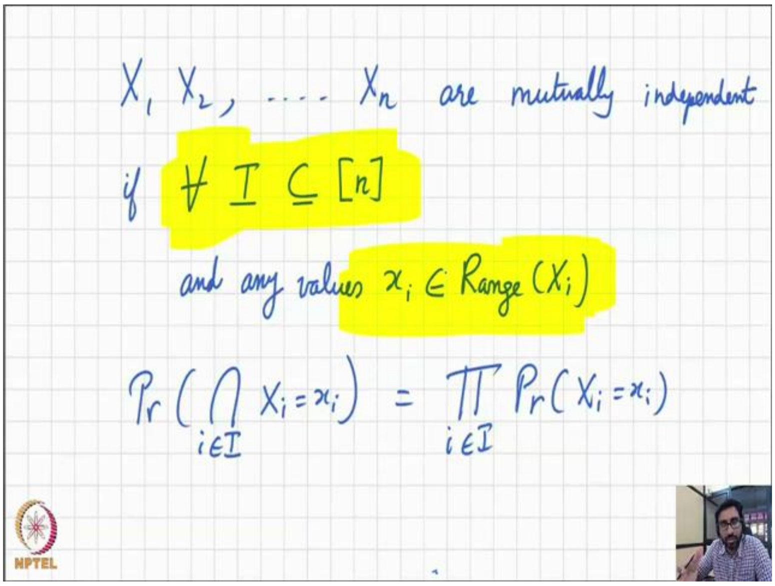
$\forall x, y$

$$\Pr((X=x) \cap (Y=y)) = \Pr(X=x) \cdot \Pr(Y=y)$$

And of course, the notion of independence extends to random variables as well. So, if you have random variables x and y, these two random variables x and y are independent, if the probability that x equals to the particular value x. And the probability that y equals to a particular value y, importantly for all x comma y in the ranges of values that x and y can take the random variables x and y can take.

If this probability on the left hand side equals, the product of the individual probability, there is a probability that the random variable x takes the value little x times the probability that the random variable y takes the value to del y then we say that x and y are independent. And as one might expect this notion extends to multiple random variables.

(Refer Slide Time: 07:05)



$X_1, X_2, \ldots X_n$ are mutually independent

if $\forall I \subseteq [n]$

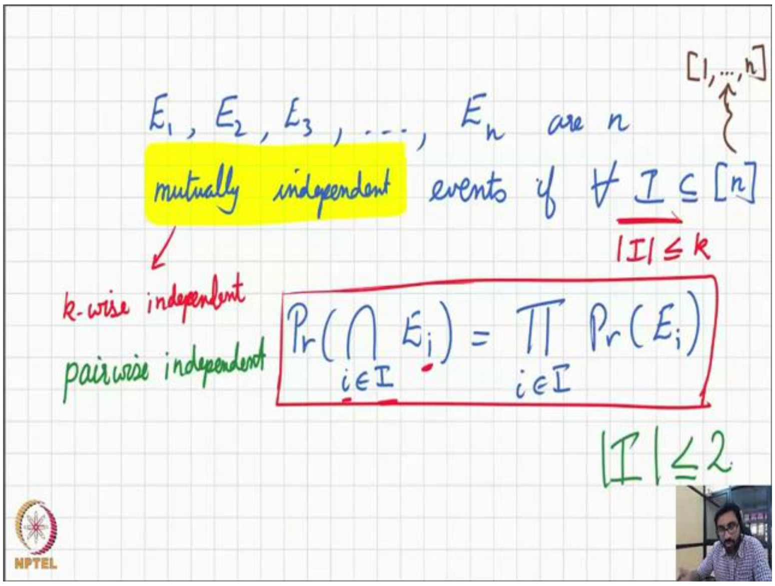
and any values $x_i \in Range(X_i)$

$$Pr\Big(\bigcap_{i\in I} X_i = x_i\Big) = \prod_{i\in I} Pr(X_i = x_i)$$

So, let say we have random variables X 1, X 2 and so on up to X n are mutually independent. Again if for every 'I' that this is subset of the index set from 1 through n, and now for any values X i, where X i fall within the range of the individual random variables say uppercase X i.

We have this independence condition holding true which is that the probability of this joint intersection over all these events that is X 1 taking the values little x 1, X 2 taking the value little x 2 and so on, all these events intersected together equals the product of the individual probabilities. So, if this condition holds then these random and these this condition must hold for every - I subset of n, and for all combination of values x i in the range of the random variable X i for all i. So, what we have seen so far is just a recollection of independence the way we know it in particular we are talking about mutual independence. And we want to relax this notion; and in order to do that, we have to modify the definitions just a little bit.

(Refer Slide Time: 08:42)



$[1, \ldots, n]$

$E_1, E_2, E_3, \ldots, E_n$ are n mutually independent events if $\forall I \subseteq [n]$

$|I| \leq k$

k-wise independent

pairwise independent

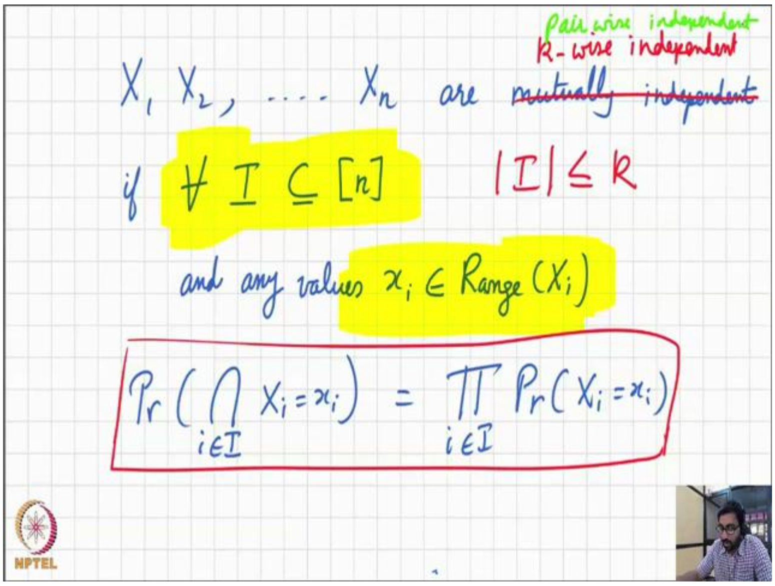
$$\boxed{\Pr\left(\bigcap_{i \in I} E_i\right) = \prod_{i \in I} \Pr(E_i)}$$

$|I| \leq 2$

So, this slide we have defined the mutual independence of events recall this. And now we want to suitably relax this notion of mutual independence; and in its place, we are going to define something called k-wise independence. And the key to defining k-wise independent events is to limit the level of independence. Now, here this independence condition given here must hold true for every - I that the subset of the indices from 1 through n and that is the requirement for mutual independence.

For k-wise independence that is what we are going to relax, we are going to say that this independence condition has to only hold for subsets I, whose cardinality is at most k. When we impose that restriction on I, we get k-wise independence. And our particular importance will be pairwise independence, for which we will require the independence condition to hold for all subsets of size at most 2. And as you can imagine this notion of k-wise independence or pairwise independence quite naturally extends to random variables independence random variables as well.

(Refer Slide Time: 10:12)



pair wise independent

k- wise independent

$X_1, X_2, \ldots X_n$ are ~~mutually independent~~

if $\forall\, I \subseteq [n]$ $\quad |I| \leq k$

and any values $x_i \in \text{Range}(X_i)$

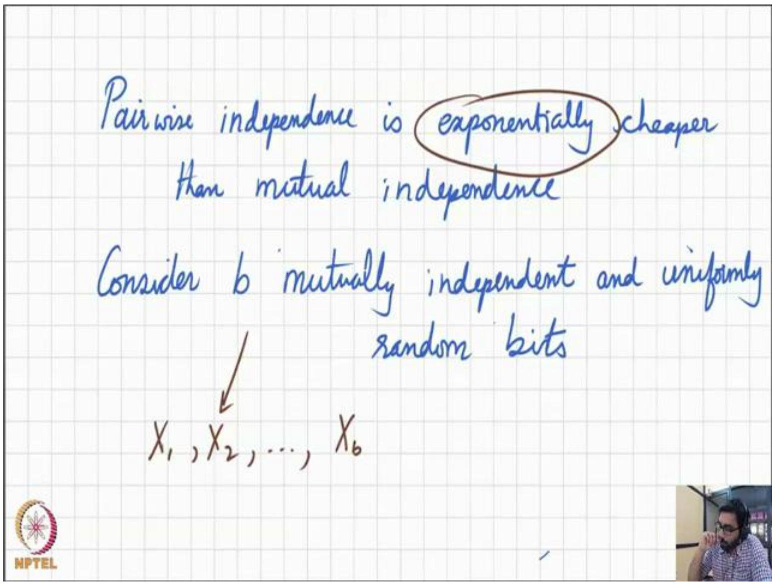
$$\boxed{\Pr\left(\bigcap_{i \in I} X_i = x_i\right) = \prod_{i \in I} \Pr(X_i = x_i)}$$

There again let us go in and make the necessary changes. So, here we are interested in defining a set of random variables to have a limited form of independence; in particular we want as we did with events define these random variables to be k-wise independent. And as before, we want this independence condition to hold for all subsets I of the index set, and with the limitation being that the cardinality has that we care about is at most this k.

When the cardinality of this I goes beyond k, then we have no restriction on whether those events, those individual values for those many random variables obey the independence condition that we have here, everything else states the same. So, every time we restrict our requirement of this independence condition to hold only for subsets of a certain size at most k, then we get k-wise independence. And of course, as before we can define our random variables to be pairwise independent, if the independence condition holds for subset whose cardinality is at most 2.

You may wonder why do we even care about pairwise independence. As it turns out pair wise independent random bits are lots cheaper in terms of the amount of randomness that we need to have in order to generate them.

(Refer Slide Time: 12:09)



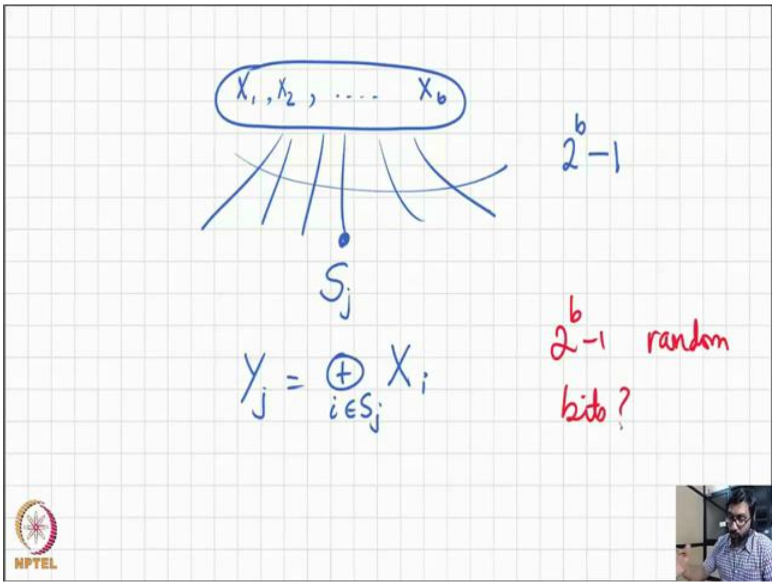
Pairwise independence is exponentially cheaper than mutual independence

Consider b mutually independent and uniformly random bits

$X_1, X_2, \dots, X_b$

So, what do we mean by that. Let us consider b mutually independent and uniformly random bits. And we denote them X 1, X 2 and so on up to X b. From this, we would like to generate pairwise independent bits. And we would like to generate lots of them; in particular, we would like to generate an exponential number of them. Let us see how can we do that.

(Refer Slide Time: 12:38)



$X_1, X_2, \dots \; X_b$

$2^b - 1$

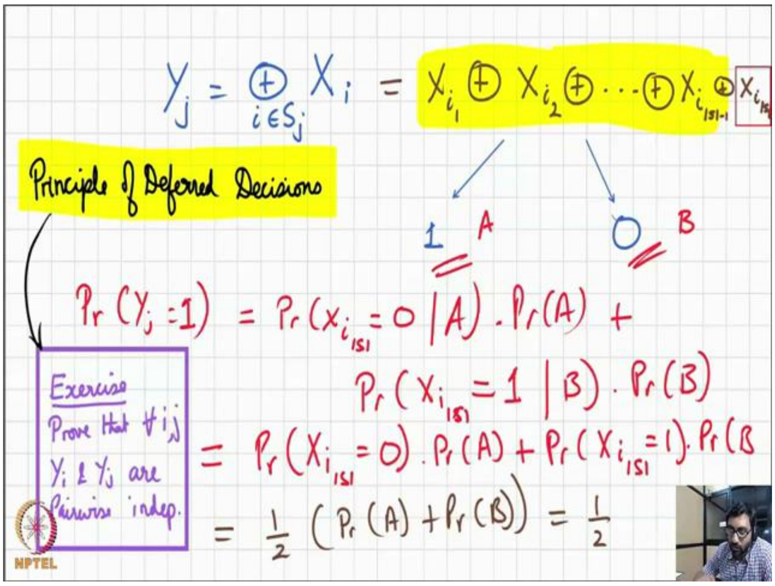
S_j

$Y_j = \bigoplus_{i \in S_j} X_i$

$2^b - 1$ random bits?

Now consider these b bits X 1, X 2 and so on up to X b, we know that the cardinality of the power set is 2 raise to the b. And if we get rid of the empty set from the power set we will have 2 raise to the power b minus 1 subsets of this set. And now take each subset so let us say we take this subset S j; and using this subset S j, let us generate random bit and called that Y j, and Y j is simply the exclusive or over all bits in S j. So, in this manner, we have now generated 2 raise to the b minus 1 random bits, but how random are they. Let us for example, try to prove that they are uniformly at random.

(Refer Slide Time: 13:48)



$$Y_j = \bigoplus_{i \in S_j} X_i = X_{i_1} \oplus X_{i_2} \oplus \cdots \oplus X_{i_{|S|-1}} \oplus \boxed{X_{i_{|S|}}}$$

Principle of Deferred Decisions

$$1 = A \qquad 0 = B$$

$$Pr(Y_j = 1) = Pr(X_{i_{|S|}} = 0 \mid A) \cdot Pr(A) + Pr(X_{i_{|S|}} = 1 \mid B) \cdot Pr(B)$$

$$= Pr(X_{i_{|S|}} = 0) \cdot Pr(A) + Pr(X_{i_{|S|}} = 1) \cdot Pr(B$$

$$= \frac{1}{2}\left(Pr(A) + Pr(B)\right) = \frac{1}{2}$$

Exercise

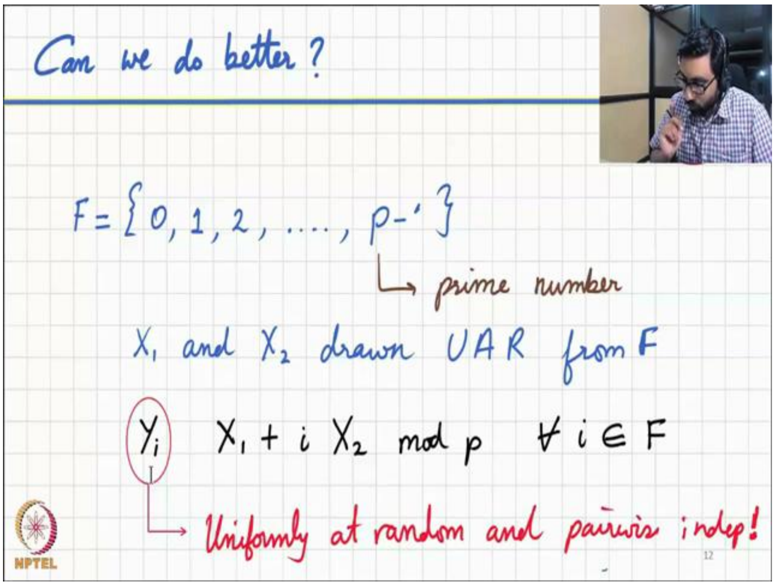
Prove that $\forall i, j$ Y_i & Y_j are Pairwise indep.

Now, if we expand this out we get something like this X i 1 exclusive OR with X i 2 exclusive OR with subsequent elements in the sequence, this is the sequence and so on up to X i cardinality of s. So, here we will be going to do we are going to isolate this one bit. Now if you consider all of these bits all, but the last bit the exclusive OR of those bits is going to be either 1 or 0. Now, what is the probability that Y j equals 1. There are two possibilities either the first set of bits where all one the exclusive OR was 1, or the exclusive OR was 0. And for simplicity, let us call this event A and this event B. So, the probability that Y j equals 1 is equal to probability the last bit is equal to 0 given event a time is the probability of A plus in similar fashion the probability that the last bit is 1 given event B times the probability of the event B.

The fact that the original bits were independent means that this these conditional probabilities we have written here are really not conditional probabilities. So, you can write them as unconditional probabilities, so the expression comes out in this manner which it is written. So, these probabilities we know their values they are just nothing but a half each. So, the expression simple becomes half times the probability of A plus the probability of B of course, the probability A plus the probability of B is nothing but 1 leading to half. Of course, the similar argument will also help you see that the probability of Y j equal to 0 is also going to be a half. So, this idea of focusing on the last bit is called the principle of deferred decisions which is a widely used technique there is well worth knowing.

In fact, here is a little exercise for you what we have shown so far is that these random bits these Y j(s) are uniformly at random, but we also need to show that they are pairwise independent and that is your exercise for all i and j prove that the y i(s) and y j(s) are pairwise independence. And in fact, you can do this again using the principle of deferred decisions that is going to be your exercise.

(Refer Slide Time: 17:28)



Can we do better?

$F = \{0, 1, 2, \ldots, p-1\}$

$\hookrightarrow$ prime number

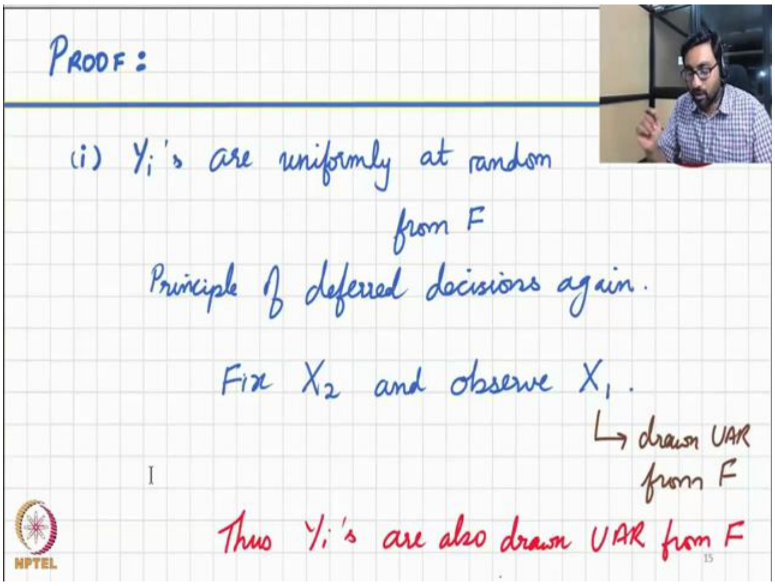
X_1 and X_2 drawn UAR from F

$\boxed{Y_i}$ $X_1 + i X_2 \mod p$ $\forall i \in F$

$\hookrightarrow$ Uniformly at random and pairwise indep!

The question now is can we do better, and as it turns out we indeed can and here is how we can do better. Let us consider the integers the set of integers from 0 to p minus 1, where p is a prime number. And now from this set, let us draw X 1 and X 2 both uniformly at random. So, we only need to draw two numbers from this set. And then with those two numbers, we generate p random numbers, here is how we do that Y i is the i th such random number and we generate that by computing X 1 plus i times X 2 mod p. As it turns out these Y i(s) that we generate are drawn uniformly at random from 0 to p minus 1 and more over they are pairwise independent that is what we need to show.

(Refer Slide Time: 18:50)



PROOF:

(i) Y_i's are uniformly at random from F

Principle of deferred decisions again.

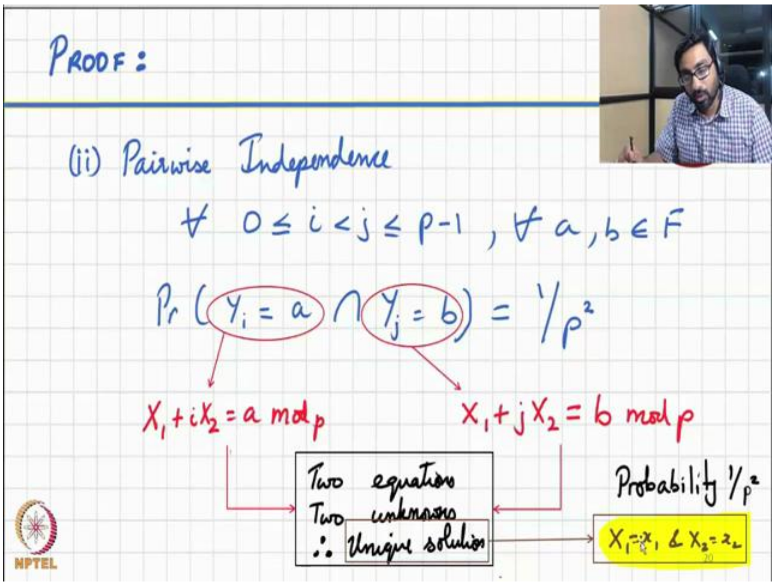
Fix X_2 and observe X_1.

↳ drawn UAR from F

Thus Y_i's are also drawn UAR from F

Well, let us see why that is a case to begin with let us prove that the Y i(s) are drawn uniformly at random from f again without much surprise. We end up using the principle of deferred decisions here, we fix X 2 and then we observe the where random variable X 1. And recall the X 1 is drawn uniformly at random from the set f. And under this context it is easy to see the Y i(s) are going to be also drawn uniformly at random from F, because we take we to whatever i times X 2(s) that we have we add X 1 and then we take mod p. So, the Y i(s) are going to be drawn uniformly at random from F as well.

(Refer Slide Time: 19:54)



PROOF:

(ii) Pairwise Independence

$\forall\ 0 \leq i < j \leq p-1,\ \forall\ a, b \in F$

$\Pr(Y_i = a \cap Y_j = b) = 1/p^2$

$X_1 + iX_2 = a \bmod p$

$X_1 + jX_2 = b \bmod p$

Two equations
Two unknowns
∴ Unique solution

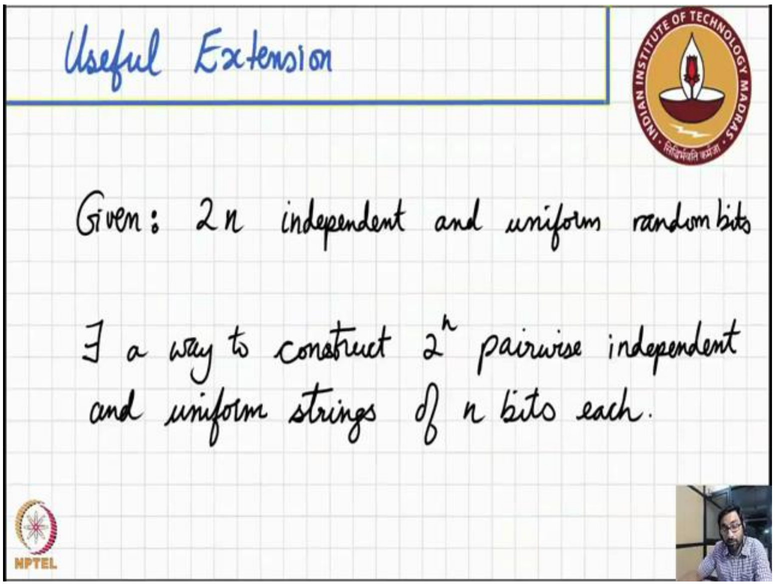
Probability $1/p^2$

$X_1 = x_1$ & $X_2 = x_2$

The more interesting thing will be to prove that these Y i(s) are pair wise independent. So, how do we do that. Well for a pair of indices i and j are ranging from 0 to p minus 1, we need to show that the probability that y i equals some specific value a and y j equals some other specific value b is it is exactly 1 over p square. So, this is what we need to show. And let see how we can prove that. Later observe that we event that Y i equal to a is equivalent to X 1 plus i times X 2 equal to a the whole mod p, and similarly the event Y j equal to b is equivalent to X 1 plus j times X 2 equal to b mode p. So, what do we have here, we have two equations and we have two unknowns. And therefore, there is a unique solution for X 1 and X 2. And recall that p is a prime number. So, let us just fix X 1 equal to lower case x 1, and X 2 equal to lower case x 2 as the as the unique solution for these two random variables X 1 and X 2.

What is the probability that the random variable X 1 would take on the value lower case x 1 and that will that is with probability 1 over p, which we have already seen in the first part of the proof. And similarly what is the probability that X 2 equals lower case x 2, again that is going to be with probability 1 over p. And both X 1 and X 2 are independent of each other, and therefore the probability that both those events will occur is exactly 1 over p square, which is exactly what we want. So, what we have seen so far is that independence itself is expensive; and it is hard to get in a computer. So, we have relaxed the notion of independence to pairwise independence or more generally k-wise independence. And in so doing, what we have realized is that we are able to generate a lot of random numbers, random bits, there are pairwise independent and drawn uniformly at random.

(Refer Slide Time: 23:05)



Useful Extension

Given: $2n$ independent and uniform random bits

$\exists$ a way to construct 2^n pairwise independent and uniform strings of n bits each.

In fact, we can go even further, so there is useful extension where in given to and independent and uniform random bits, we can actually generate 2 power n pairwise independent and uniform strings of n bits each. And in a little while, we will see that this can be leveraged to do sampling in using far fewer rand truly random bits.