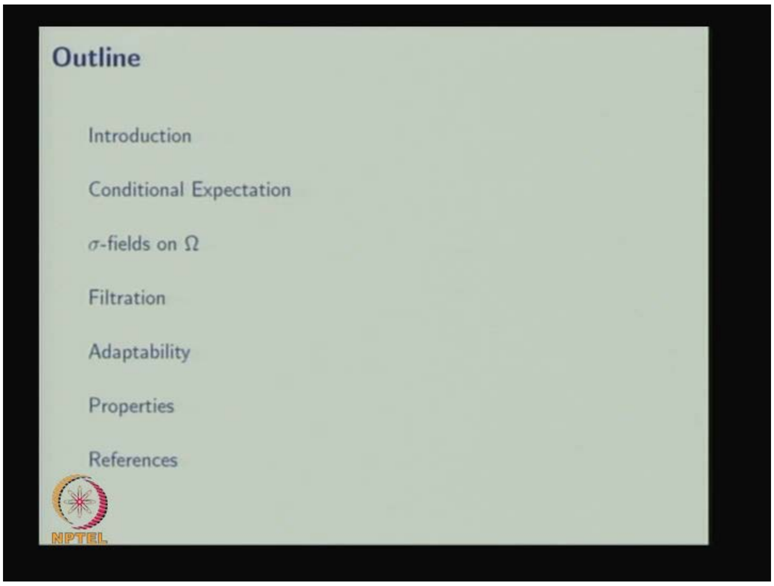
Stochastic Processes
Prof. Dr. S. Dharmaraja
Department of Mathematics
Indian Institute of Technology, Delhi

Module - 6
Martingales
Lecture - 1
Conditional Expectation and Filtration

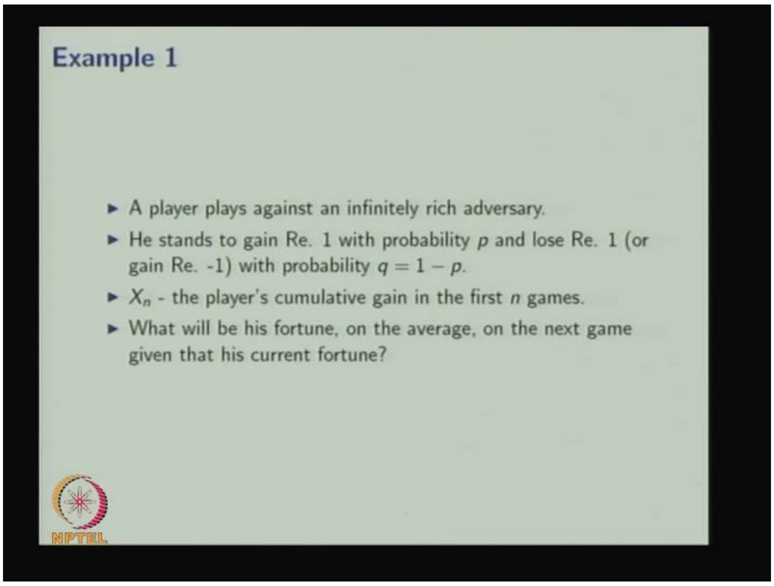
This is stochastic process module six, Martingales, lecture one, Conditional Expectation and Filtration. In the last five modules we have discussed stochastic processes, few properties and then discrete-time Markov chains and continuous-time Markov chains. In this model we will discuss an important property of stochastic processes, Martingale.

(Refer Slide Time: 01:00)



In this lecture, conditional expectation, few important properties, sigma fields on omega filtration, conditional expectation of a random variable, given sigma field, few important properties along with simple examples will be discussed.

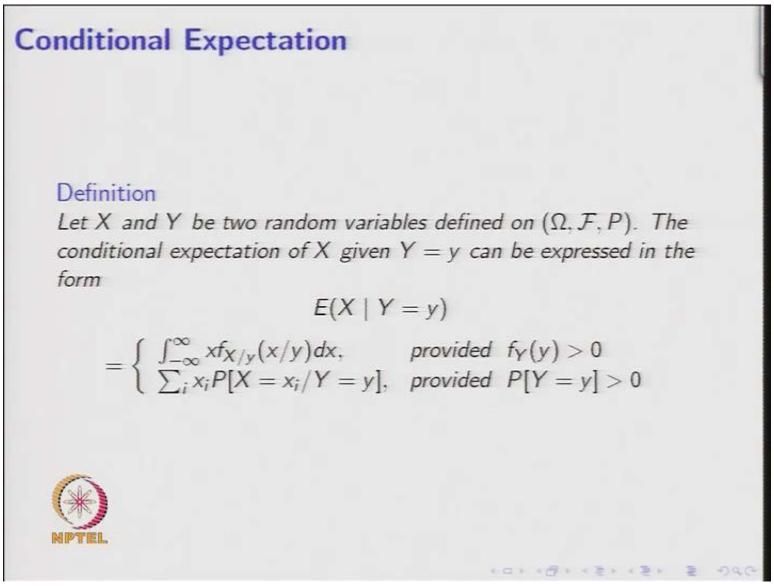
(Refer Slide Time: 01:28)



A stochastic process is often characterized by the dependence relationship between the members of the family, a process with a particular type of dependence through conditional mean, known as martingale property. This property has many applications in probability. An example, a player plays against an infinitely rich adversary, he stands to gain rupees 1 with probability p and lose rupees 1, is equivalent of gain rupees minus 1 with the probability q, that is, 1 minus p.

Let X n be the player's cumulative gain in the first n games. The question is, what will be his fortune on the average on the next game given that his current fortune. We need the knowledge of martingale to solve this problem. Martingale concept involves the conditional expectation of a random variable given sigma field. Hence, we will introduce the conditional expectation.

(Refer Slide Time: 03:01)



Here is the definition of the conditional expectation. Let X and Y be two random variables defined on probability space omega, F, p. The conditional expectation of the random variables X given the random variables Y takes the value small y, can be expressed in the form expectation of X given Y is equal to small y. Both, the random variables are defined on the same probability space, omega is the collection of possible outcomes, F is the sigma field and p is the probability measure. This (( )) is a probability space, so both the random variables defined on the same probability space.

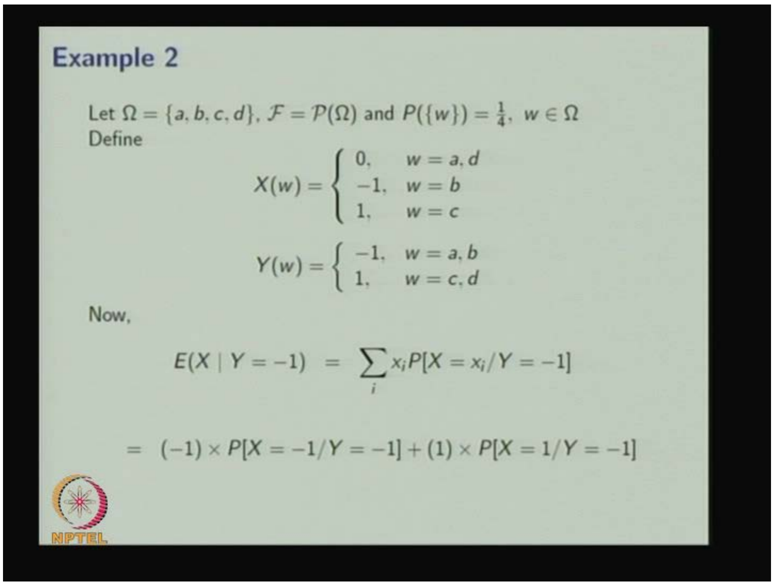
And we are, we are defining the conditional expectation of the random variable given the other random variable takes the value small y, that can be defined in the form. If both the random variables are continuous, then we can make out as integration minus infinity to infinity X times the conditional probability density function of the random variable X given that other random variables takes the value small y integration with respect to X provided the probability density function of the random variable Y at the point Y has to be greater than 0.

We know that the probability density function is always greater than or equal to 0 in the whole range of y, but here we are discussing the conditional expectation of X given the other random variable takes the value small Y. So, at that point of small y, the probability density value density function value F Y small y has to be strictly greater than 0. If that is the case and we assume, that here both the random variables are continuous, therefore we are having probability density functions and integration is with respect to X, X multiplied by the conditional probability density function of X given y.

Suppose both the random variables are continuous, then this conditional expectation of X given Y can be expressed in the form summation X i's, probability of X takes the value small X i's given that Y takes the value small y. So, this is the conditional probability mass function of the random variable X given the other random variable takes the value Y.

Here also the provided, condition is provided, the marginal probability the mass function for the random variable Y at the point small y has to be strictly greater than 0. In that case we can find out the conditional expectation of X given, the other random variable is small y.

(Refer Slide Time: 06:15)



Now, we present an example. So, here the omega consists of four elements, a, b, c, d and the sigma field is the power set, the omega is finite. So, we can create the power set. The number of elements is going to be 2 power n where n is the number of elements in the sigma omega. So, here the F is going to be the sigma field, which is the power set. So, that is the largest sigma field also and we are defining the probability for each sample that is 1 by 4, where w is belonging to omega.

Now, I am defining two random variables, X and Y, satisfying the condition of random variable, that is, X inverse of minus infinity to X, that is, we closed interval inverse, which is belonging to F for all X belonging to omega, then that is going to be a random variable. So, here you can cross-check whether this real valued function X is the random variable or not, which satisfies the random variable condition. So, X is a random variable. Similarly, Y is also a random variable.

Now, I will define, now I am going to calculate the conditional expectation of X given Y takes the value small, Y takes the value minus 1. So, that is nothing but the summation, the summation of X i's P of X equal to X i given Y takes the value minus 1. That means, I have to find out the possible values of X i's, then find out the conditional probability mass function for those X i's, multiply it, then sum it and sum it over i, that is going to be the conditional expectation of X, given Y takes the value minus 1.

So, X can take the value 0, minus 1 or 1. So, I can compute the way minus 1 times probability of X takes the value minus 1, Y takes the value minus 1, the next X can take the value 1 into probability of X takes the value 1 given Y takes the value minus 1, 0 into conditional probability, you do not want to write, therefore we have only two terms.

(Refer Slide Time: 08:38)

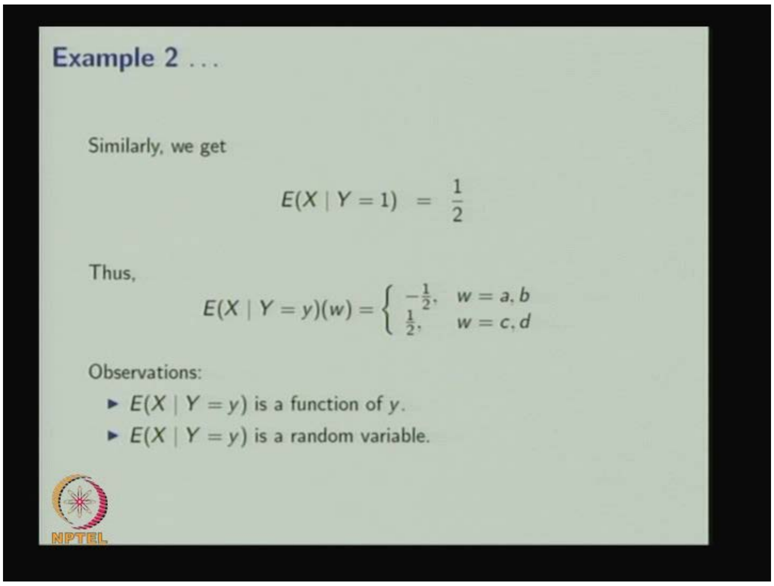
Example 2 . . .

$$
\begin{aligned}
E(X \mid Y=-1) &= -\frac{P[X=-1, Y=-1]}{P[Y=-1]} + \frac{P[X=1, Y=-1]}{P[Y=-1]} \\
&= -\frac{P(\{b\})}{P(\{a,b\})} + \frac{P(\{\})}{P(\{a,b\})} \\
&= -\frac{1/4}{1/2} + \frac{0}{1/2} \\
&= -\frac{1}{2}
\end{aligned}
$$

Therefore the conditional expectation of X given Y takes the value minus 1 that is going to be minus times this conditional probability. You know how to compute, find out the joint probability mass function of X, takes the value minus 1, Y takes the value minus 1, divided by, probability of X takes, probability of Y takes the value minus 1 plus the second term, that is, joint probability mass function of X takes the value 1 and Y takes the value minus 1. You know the probability of X takes the value minus 1 and Y takes the value minus 1, the only possibility is the sample is a.

And similarly, this sample is not possible, that is why the empty set, therefore the probability of sample a, that is, 1 by 4. The probability of Y takes the value minus 1 is possible with a and b, therefore it is half 1 by 4 plus 1 by 4, that is, 1 by 2 empty set probability 0. And the denominator probability is 1 by 4 plus 1 by 4, that is 1 by 2, therefore the simplification gives the expectation is equal to minus 1 by 2.

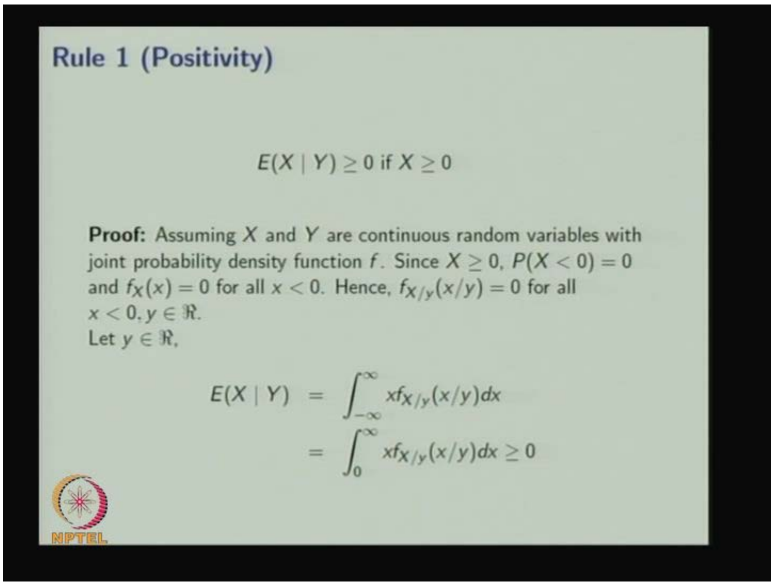
(Refer Slide Time: 10:09)



Similarly, you can find out the conditional expectation of X given Y takes the value 1, the same thing. Now, I can write it in the compact form, expectation of X given Y takes the value small y for all the possibility, that is going to be minus of, if w is equal to a and b, this is going to be half, if w is going to be c or d. That means, based on w, the value changes. w is nothing but the samples; w is belonging to omega.

So, the observations are the conditional expectation is a function of Y. The conditional expectation of X given Y takes the value small y, is a function of y. Not only that, the conditional expectation is a random variable because for all possible values of w you will get the different values. Therefore, this is the random variable whereas expectation is a constant; conditional expectation is a random variable

(Refer Slide Time: 11:28)



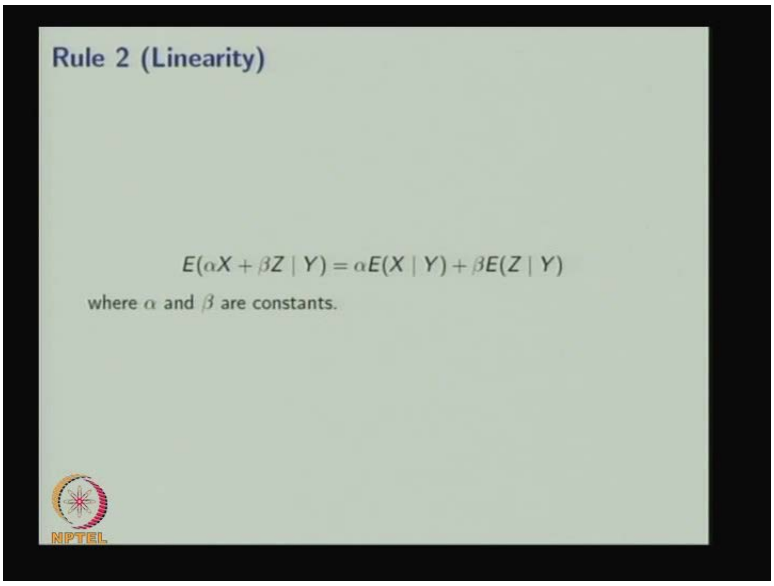
Few important rules on conditional expectation will be used to verify the given stochastic process has a martingale property. So, we list the rules with few assumptions. We have presented the proof, but these rules can also be proved without these assumptions. The rule 1 says, if X is greater than or equal to 0, then the conditional expectation of X given, whatever be the random variable Y, that is always going to be greater than or equal to 0 if X is greater than or equal to 0 with probability 1. Whenever X is non-negative, then the conditional expectation is also going to be a non-negative, that is, positivity property.

So, the proof is given here. Assuming, that X and Y are continuous random variables with joint probability density function F. I am going to give the proof. Similarly one can give the proof, assuming both the random variables are discrete random variables with the joint probability mass function also.

Since X is greater than or equal to 0, the probability of X is less than 0 is equal to 0 and the probability density function of X is going to be 0 for all X is less than 0. Hence, the conditional distribution of X given Y, that is also going to be 0 for all X is less than 0, whatever be the Y belonging to the real.

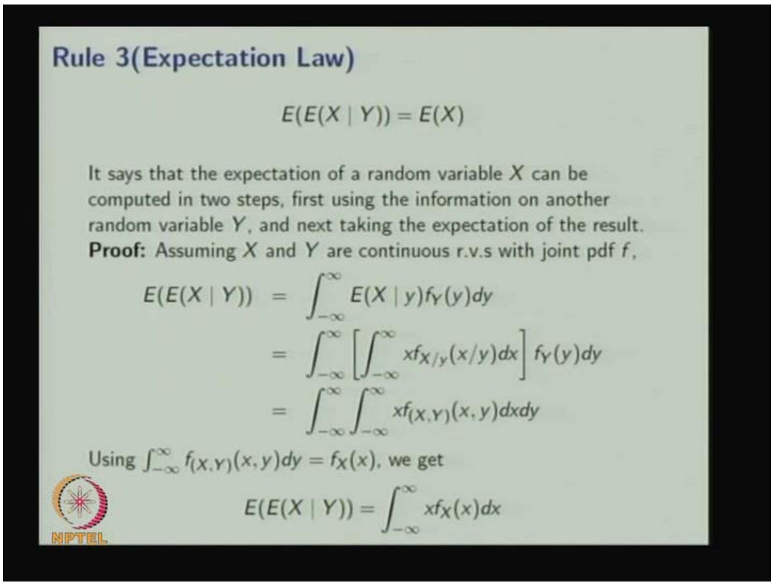
So, when, let Y belonging to the real, the conditional expectation is going to be minus infinity to infinity X times the conditional probability density function and this conditional probability density function is going to be 0 for X is less than 0. Therefore, the integration exists only from 0 to infinity because minus infinity to 0, the density function is 0. Therefore, the conditional expectation is going to be 0 to infinity X times the condition probability density function and you know, that the probability density function is always greater than or equal to 0 in the whole range. And X is here, we are integrating from 0 to infinity. Therefore, this quantity is always going to be greater than or equal to 0. So, this concludes, if X is greater than or equal to 0, the conditional expectation of X given Y, that is also greater than or equal to 0.

(Refer Slide Time: 14:15)



The rule two, it is a linearity property. The linear property says, the expectation of alpha times X plus beta times Z given that Y takes the value some small Y, that is same as alpha times the conditional expectation of X plus beta times the conditional expectation of Z where alpha and beta are constants. It is similar to the linear property of expectation. The same thing holds good for the conditional expectation also, therefore no need to give the proof.

(Refer Slide Time: 14:51)



The third one that is the expectation law, expectation of conditional expectation is same as expectation. It says, the expectation of random variable X can be computed in two steps. First, using the information on another random variable Y and next, taking the expectation of the result. You can visualize in the other way round. The expectation of the X can be computed as a expectation of conditional expectation of X given Y. That means, you take any random variable Y, as long as the conditional expectation possible find out that conditional expectation, then find.

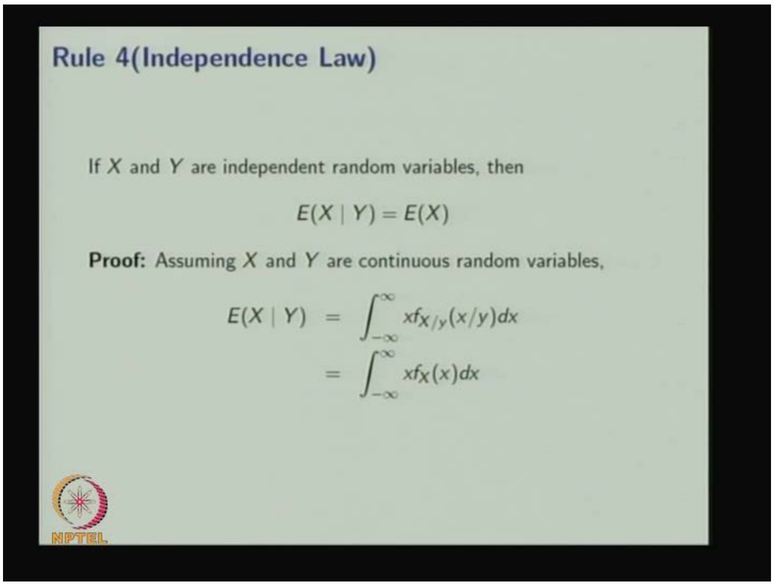
Since the conditional expectation is a random variable, so find out the expectation of that random variable, that is, inverse the expectation of X. The proof is given with the assumption, both the random variables are continuous, the joint probability density function F. So, I am starting with the left hand side. The expectation of expectation X given Y that is same as, you know how to compute the expectation. Here the provided condition is expectation exist.

Similarly, in the definition of conditional expectation also you have to make the assumption, the expectations exists. Then, you are finding the conditional expectation. So, this expectation exists, therefore minus infinity to infinity expect conditional expectation. This is a random variable and you are finding the expectation of that, therefore you want to play the probability density function for the random variable y because this expectation of X given Y is a function of Y. Therefore, you should multiply the probability density function of Y integrate with respect to Y between the limits minus infinity to infinity. And by the definition, the conditional expectation is nothing but X times the conditional probability density function integration with respect to X between the limits minus infinity to infinity. So, you substitute that.

Now, you can come to the conclusion, the integration of minus infinity to infinity the joint probability density function of X and Y is nothing but the marginal distribution. So, here the integration is with respect to Y, therefore you get the marginal distribution for marginal probability density function of x. So, here it is a conditional probability density function multiplied by the probability density function. Therefore, this product will give the joint probability density function of X and Y. Any joint, any two random variables joint probability density function can be written as the product of marginal distribution into the conditional distribution. So, using that I am getting the joint probability density function.

Now, this integration is F of X, therefore the one integration and this much will give marginal, therefore it is minus infinity to infinity X times, that is going to be the marginal probability density function of X, therefore you will get expectation of X. So, the right hand side, right hand side is going to be the expectation of X. So, you can find out the expectation of X by computing the conditional expectation with some other random variable, then find the expectation. So, this rule has lot of importance.

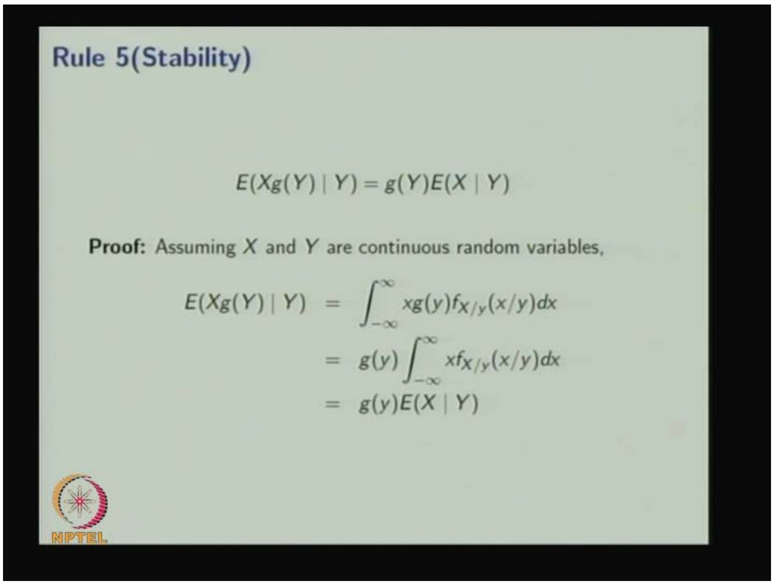
(Refer Slide Time: 19:00)



The next one is independence law. If two random variables are independent, then the conditional expectation and the original expectation are, both are same. The proof is assuming both the random variables are continuous. So, the conditional expectation, this is by definition and you know, that both the random variable are independent, then the conditional distribution is same as the marginal. Therefore, you can replace this way the marginal distribution.

So, X times the F of X, that is nothing but the expectation of X. That means, if the two random variables are independent, then the conditional expectation is not a random variable, it is a constant because the right hand side expectation of X is a constant, therefore this is also a constant.

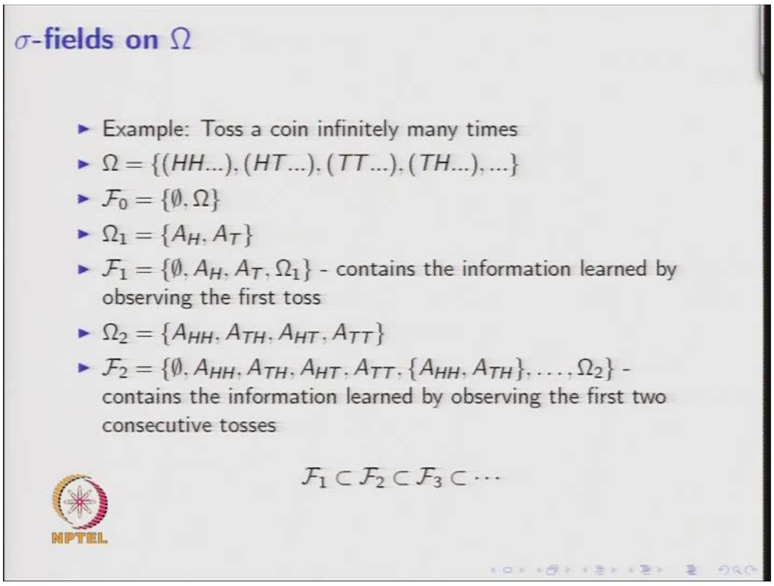
(Refer Slide Time: 19:56)



The next rule, it is stability. Suppose, you have the function g and the g of Y is also going to be a random variable that means, g is a Borel measurable function. So, the expectation of X times G of Y given Y, that is same as the g Y will be out G Y times the conditional expectation of X given Y. That means, later we are going to use the property called known is out, that means, the expectation of X times g of Y given Y takes the value something, some small y that means, this is going to treat as a constant.

So, g of Y has to be treated as a constant. Therefore, the constant will be come out, therefore the g of Y times the expectation of X given Y. The same thing, I have written in the proof with both the random variables are continuous.

(Refer Slide Time: 21:10)



Now, we introduce the sigma fields on omega through example because this is very important concept for the martingale. Example start with the tossing a coin infinitely many times, tossing a unbiased coin infinitely many times.

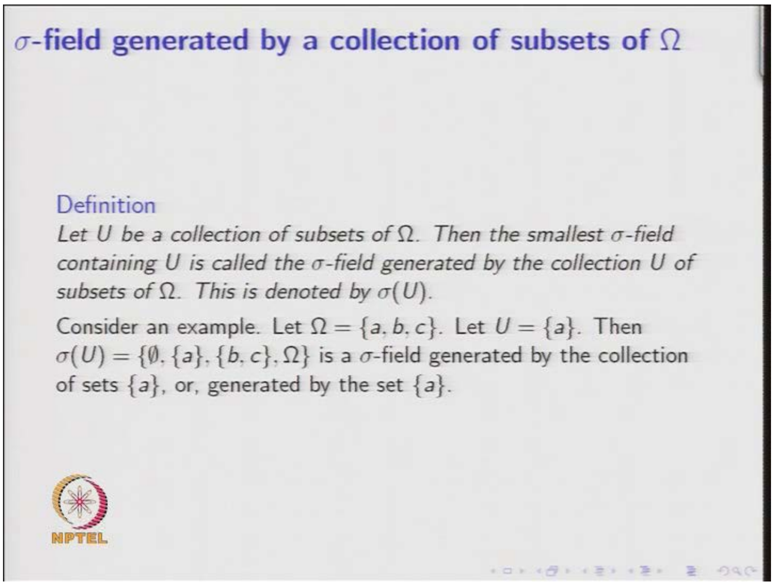
Let omega be the collection of possible outcomes HH and so on, HT and so on TT and so on TH and so on. Let F naught be the trivial sigma field, consists of two elements, empty set and the whole set. F 1 is the smallest sigma field containing in omega 1; F 2 is the smallest sigma field containing in omega 2, containing the information learned by observing the first two consecutive tosses.

If you observed, you will find the omega 1 contained in, sorry, F 1 contained in F 2, F 2 is contained in F 3 and so on. Since we are tossing an unbiased coin infinitely, many times, if you find out the limit of F n, that is going to exist, that is going to be F infinity in a notation and that in the notation we can make out it is F. So, this consists of the information learned by infinitely many tosses observation, that is, the sigma field F. So, this is the way one can create the sigma fields on omega.

So, omega is consisting of all the possible outcomes in infinitely many tosses, observation over the infinitely many tosses. Whereas, the omega 1 consists of only two elements, therefore we are creating a first sigma field on omega 1 and after you create four elements, the possible outcomes, after framing the possible outcomes into the four elements you get the omega 2. So, using omega 2 we are creating a largest sigma field F

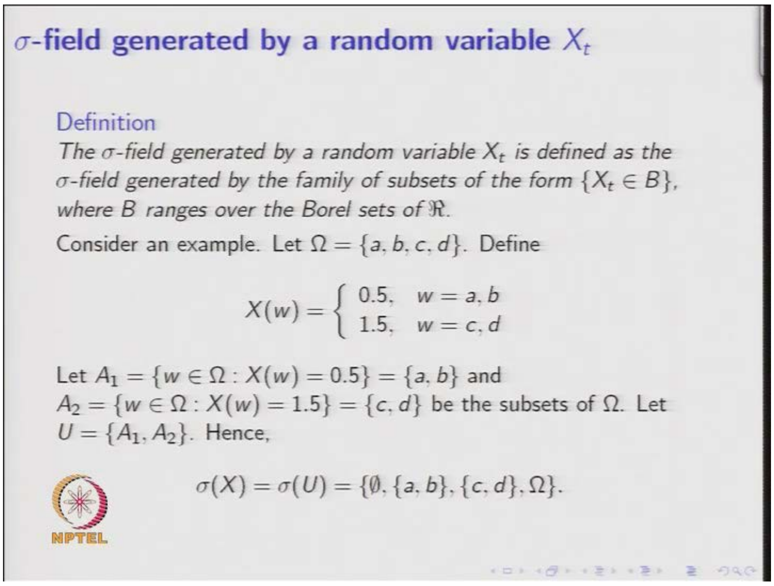
2. So, like that you can create omega 3, F 3, omega 4 F 4 and so on and all those F i’s satisfies this property and the limit exist as n tends to infinity. This sigma field is going to be denoted by the letter F. So, this is the way one can create the sigma fields on omega.

(Refer Slide Time: 24:40)



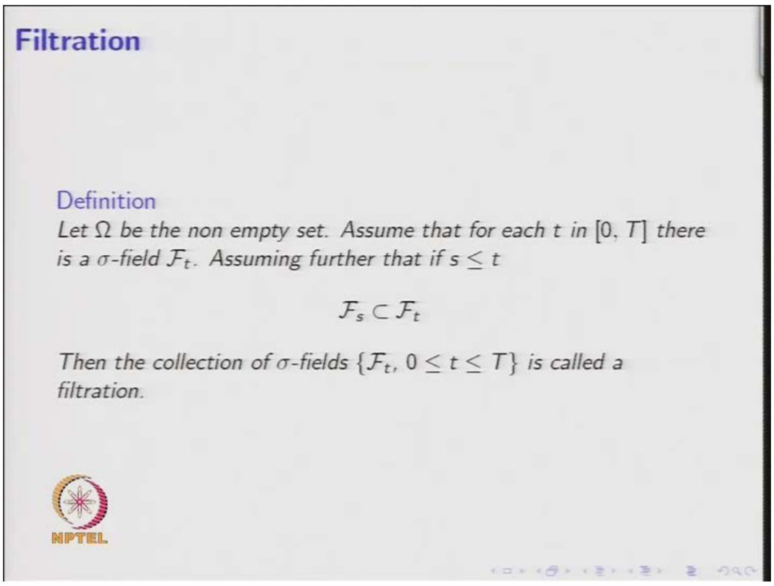
Now, we present the sigma field generated by the, by a collection of subsets of omega. Let U be a collection of subsets of omega, then the smallest sigma field containing U is called the sigma field generated by the collection U of subsets of omega. This is denoted by sigma of U. Consider an example where omega is equal to (a, b, c). Let U is a, then sigma of U is an empty set element a, element (b, c). The whole set is a sigma field generated by the collection of sets a.

(Refer Slide Time: 25:34)



Consider an example where omega is equal to (a, b, c, d). Define the random variable X, which takes the value 0.5 for w is equal to (a, b). It takes the value 1.5 and w takes the value (c, d). Let A 1 be the set, which takes the value, which is the collection of possible outcomes in which the X of w takes the value 0.5. Therefore, it is (a, b). Similarly, let a 2 be the set and the collection of w belonging to omega in which X of w is equal to 1.5. Hence, it is (c, d) be the subsets of omega. Let U is equal to (A 1, A 2), hence the sigma field generated by the random variable X, that is same as sigma of U. That is empty set. (a, b) is the one element, (c, d) and the fourth element is a whole set.

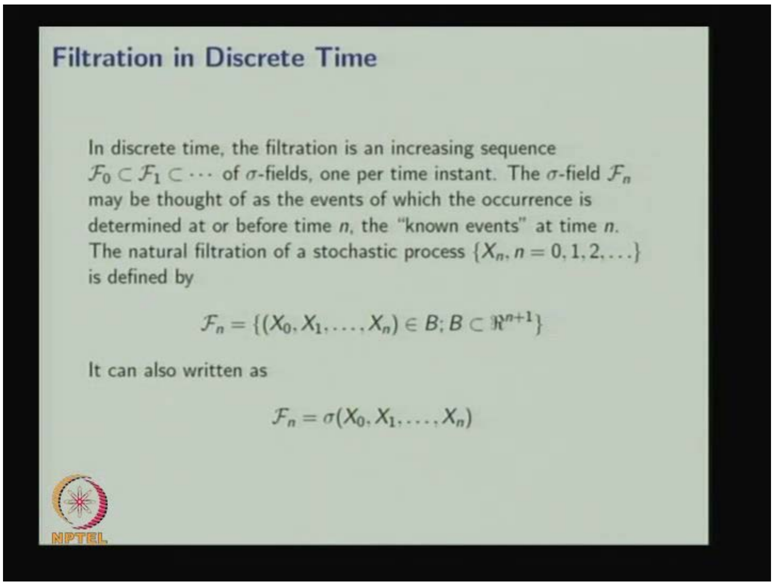
(Refer Slide Time: 26:55)



Assume, that for each t in (0, capital T), where capital T is a positive real number. Then, for each t between the interval 0 to capital T you are creating a sigma field F of t and those sigma fields F of t for possible values of t in between the interval 0 to capital T, it satisfies the condition F of s is contained in F of t.

If this condition is satisfied over the interval 0 to T by s and t, where s is lesser than or equal to t, then thus collection of, this collection of random variables, sorry, this collection of sigma fields is called the filtration.

(Refer Slide Time: 27:49)

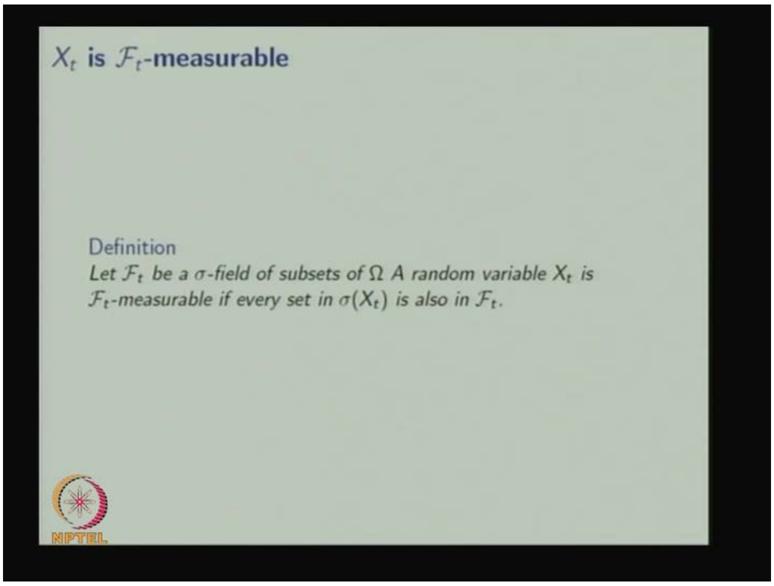


The definition of the filtration in real time is as follows. In discrete time the filtration is an increase in sequence F not contained in F 1, contained in F 2 and so on of sigma fields, one per time instant. The sigma field F n may be thought of as the events of which the occurrence is determined at or before time n, the known events at the time n.

The natural filtration of a stochastic process X n is defined by F n, is the collection of, n dim, n plus 1 dimensional random variables belonging to B where B is contained in R n plus 1. This is also written as F n is a, sigma fi, sigma field generated by n plus 1 random variables or you can think of the random vector with n plus 1 dimensions. So, the F n is a sigma field created by the random vector X naught to X n or the random variables X naught and X 1 and so on till X n. So, this is the filtration in discrete time.

Let us generate a filtration for this example for discrete situation. So, take the same example, tossing an unbiased coin infinitely many times and the omega 1 is having two elements. This is the sigma field on omega 1 and the F 2 is the sigma field on omega 2 and also satisfied sigma. The F 1 is contained in F 2, which is contained in F 3, therefore this collection of, the collection of random, the collection of sigma fields is called the filtration. So, this is the example of creating sigma field in real time.

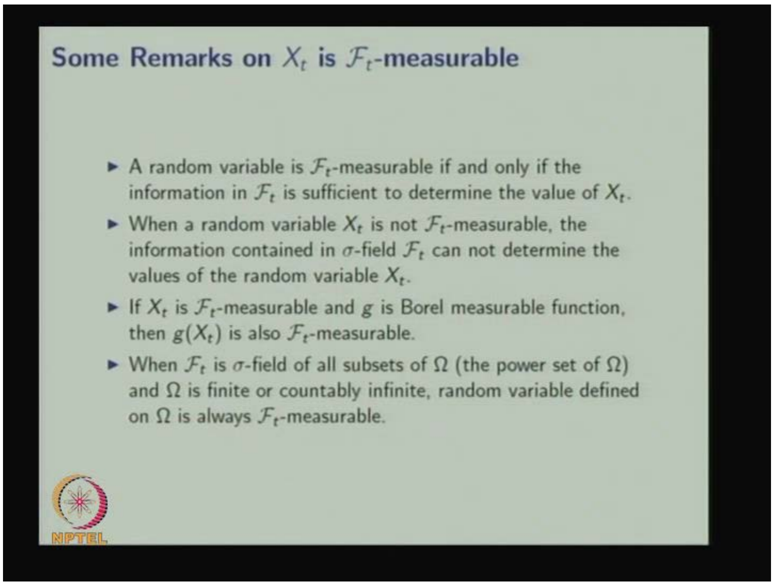
(Refer Slide Time: 30:37)



To verify the given stochastic processes, Markov, sorry, has a martingale property. Each random variable X t must be F t measurable. So, you will see the definition of when the random variables X t is going to be the F t measurable. For that the definition is as follows. Let F t be a sigma field of subsets of omega A. Random variable X t is F t measurable if every set in sigma field generated by the random variable X of t is also in F t.

We have a non-empty set, that is a collection of possible outcomes and we have created the sigma field on omega and the random variable is said to be F t measurable. The random variable X t is said to be F t measurable if the sigma field, if every set in sigma field generated by the random variable is X t is also in F of t. If this condition is satisfied, then this random variable is going to be called it as F t measurable. Obviously, F of t is contained in sigma of X of t. So, here every element every set in sigma field generated by the random variable X t is also in F of t that means this is other way round.

(Refer Slide Time: 32:28)



So, if that property also satisfied, then this random variable is F T measurable. Some remarks on X t is F t measurable. The first remark, a random variable is, a random variable X t is F t measurable if and only if the information in F of t is sufficient to determine the value of X of t. The F of t is nothing but the collection of information up to the time. So, whenever we say the random variable is F t measurable if and only if the information in F t is sufficient to determine the value of X of t.

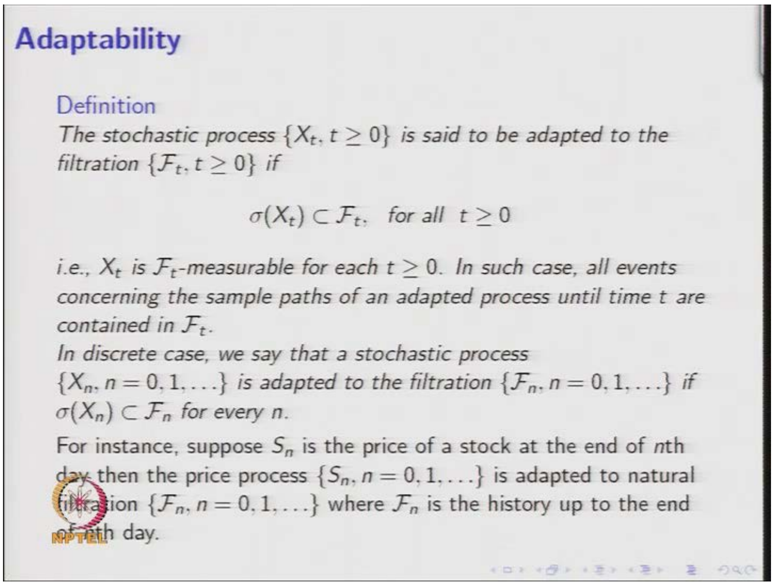
The second remark, when a random variable X t is not a F t measurable, then the information contained in sigma field F of t cannot determine the values of the random variable X t. Whenever the random variables is not F t measurable, the conclusion is the information contained in the sigma field F of t, cannot determine the values of the random variable, whereas X t is a random, is F t measureable if and only if it has the sufficient information to determine the value of X of t.

Third remark, if X t is F t measurable and g is a Borel measurable function, then g of X t is also F t measurable. We know, that X is a random variable then and g is a Borel measurable function, then g of X is a random variable. So, here we are saying, if X is X t is F t measurable and g is a Borel measurable function, then g of X t is also F t measurable. Obviously, g of X t is also a random variable since X t is F t measurable, then g of X t is also F t measurable.

The fourth remark, we are not giving the proof for these remarks. The fourth one, when F t is sigma field of all subsets of omega, that is, the power set of omega, the omega is finite or countably infinite, then the random variable defined on omega is always F t measurable.

Whenever the sigma field F t is the power set for the largest sigma field, in that case the random variables defined on omega is always F t measurable. We know that whenever F t is largest sigma field, then any real valued function is going to be a random variable. Here, whenever F t is a sigma field, which is a largest sigma field or the power set of omega and additional condition and omega is finite or countably infinite, then the random variables is always F t measureable.

(Refer Slide Time: 35:52)



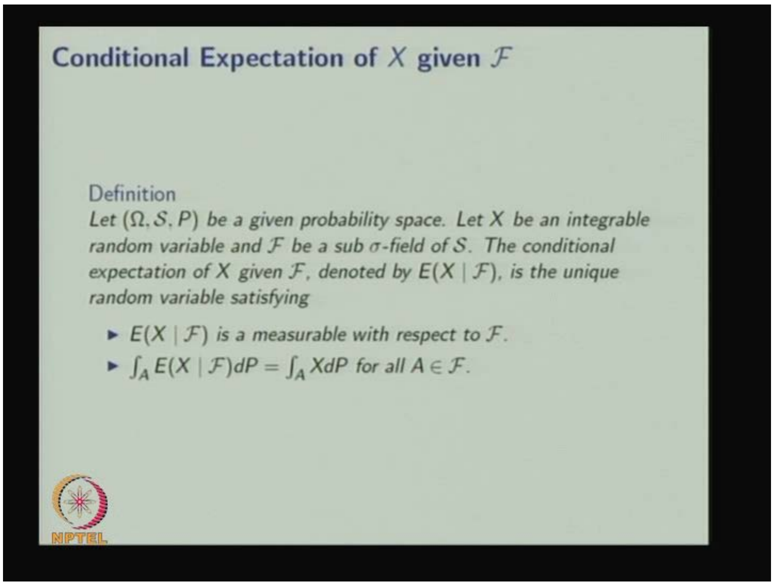
Based on X t is F t measurable, we discuss the adaptability of the stochastic process. The stochastic process X t over the t greater than or equal to 0 is said to be adapted to the filtration F of t if the sigma field generated by the random variable X t, which is contained in F of t for all t greater than or equal to 0.

So, this condition is nothing but X t is F t measurable. Whenever X t is F t measurable, then the collection of random variable X of t is adapted to the filtration F of t. If this condition is not satisfied, in that case X t is not F t measurable. If X t is not F t measureable, then the collection of random variables, the collection of random variables X of t is not going to be adapted to the filtration.

Suppose, it is, X t is F t measurable, in such cases all events concerning the sample paths of adapted process until time t or contained in F of t, it says, F of t has a, has information up to the time t. Whenever X t is F t measurable, then all events concerning the sample paths of the adopted process until time t or contained in F of t. It is same meaning here. The sigma field is generated by the X of t, which is contained in F of t.

Whenever the stochastic process is adapted in a discrete case we say, that stochastic process X suffix n, n can takes the value 0, 1, 2 and so on is adapted to the filtration F suffix n. So, instead of F of t I am using X suffix n for discrete type and similarly, the random variable is also discrete type instead of t. I am using small n, if the sigma field generated by the random variable X n, which is contained in the sigma field F n, for every n, it has to be satisfied for every n, then only this collection of stochastic, sorry, this collection of random variable or this stochastic process is adapted to this filtration F n. For instant, suppose, X n is the price of the stock at the end of nth day, then the price process S n, n is equal to 0, 1 and so on. This adapted to the natural filtration F n, n, 0, 1 and so on where F n is the history up to the end of nth day.

(Refer Slide Time: 39:02)

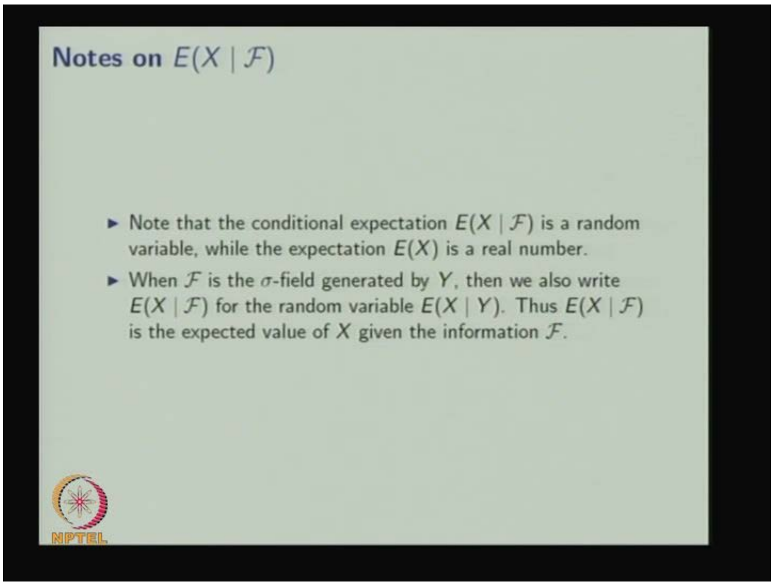


Now, we see the definition of conditional expectation of random variable X given sigma field F. Earlier we have defined the conditional expectation of the random variable X given other random variable Y takes the value small y to the definition as follows. Let omega S P be a given probability space, let X be the integrable random variable and F be a sub-sigma field of S. So, here I am defining a probability space with the omega and this is the sigma field S and P is the probability measure and the random variable, which is integrable and F be the sub-sigma field of S, the conditional expectation of X given F.

So, to distinguish the sigma fields F and S, I am giving S is a sigma field and F is the sub-sigma field and we are defining the conditional expectation for the random variable X given the sub-sigma field of S, that is F, that is denoted by X, expectation of X given F. It is the unique random variable, the way I discussed the conditional expectation is a random variable.

So, here the conditional expectation of X, given the sigma field, that is also a random variable. It is a unique random variable satisfying the conditional expectation of X given sigma field is measurable with respect to the sigma field F, this is measurable. Also, the integration over any set A where A is belonging to F. The expectation of X given F integration with respect to the probability measure P, that is same as the integration with respect to the probability measure P of the integrant is E X integration over A, both are one and the same. Note, that conditional expectation given the sigma field F is a random variable satisfying these two properties and we are defining the conditional expectation given sub-sigma field of S where this is the probability space.

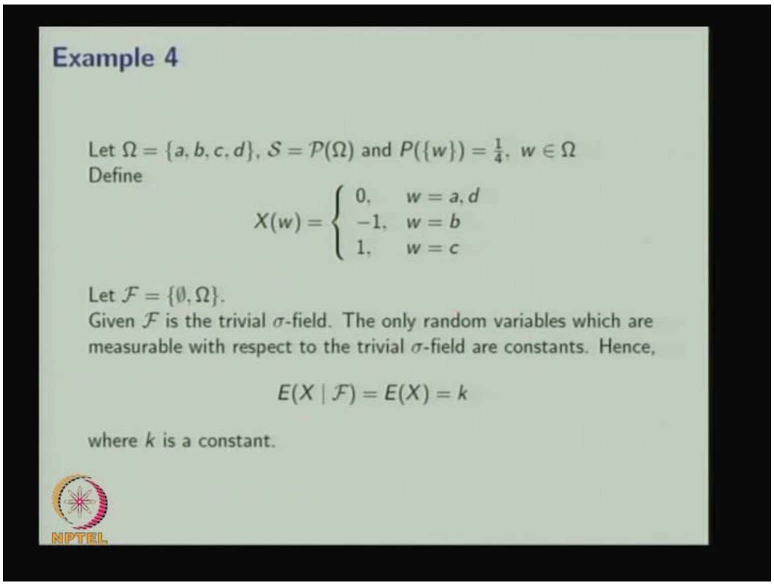
(Refer Slide Time: 42:04)



Notes on conditional expectation of X given the sigma field F. Note, that the conditional expectation, expectation of X given F is a random variable while expectation X is the real number. Similarly, conditional expectation of X given other random variables, is also a random variable, not a constant.

When F is the sigma field generated by Y, then we also write the conditional expectation of X given F for the random variable X, expectation of X given Y. Whenever F is the sigma field generated by Y you can replace F by Y. Thus, expectation of X given F is the expected value X given the information F. That means, you can replace the random variable Y by F whenever the F is the sigma field generated by the random variable Y, that is same as expectation, expected value of X given the information F. Whenever we say the sigma field, that is nothing but the information

(Refer Slide Time: 43:38)

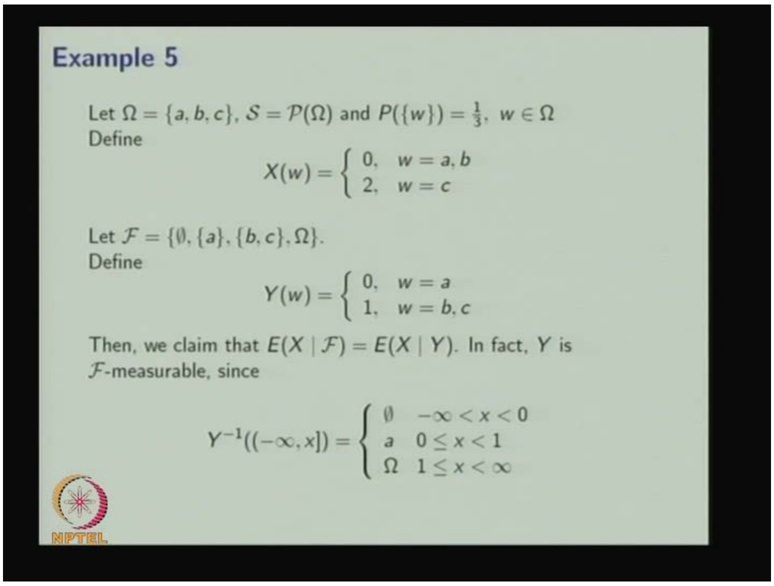


We discussed the conditional expectation of the random variable given sigma field through two examples. The first example is as follows. Omega consists of four elements, S, that is the largest sigma field power set on omega in the P of w’s is equal to 1 by 4 where w belonging to omega. So, therefore, this is the set function probability measure. We are defining the real valued function, you cross-check, this is a random variable.

Let F as the sigma field is a real one, which consists of two elements, empty set and the whole set given F is the trivial sigma field. The only random variables are measurable with respect to the trivial sigma field are constant. The only random variable, which are measurable with respect to the trivial sigma field are constant. Hence, the expectation of X given sigma field F where F is the trivial one that is same as the expectation of X that is equal to constant. That constant you can find out by using the probability and the possible values of 0 minus 1 and 1, you can find out the constant.

So, here the conclusion is, whenever the sigma field is a trivial one, then the conditional expectation over the trivial sigma field, that is a constant and that constant is same as the expectation of X because it is no more random variable.

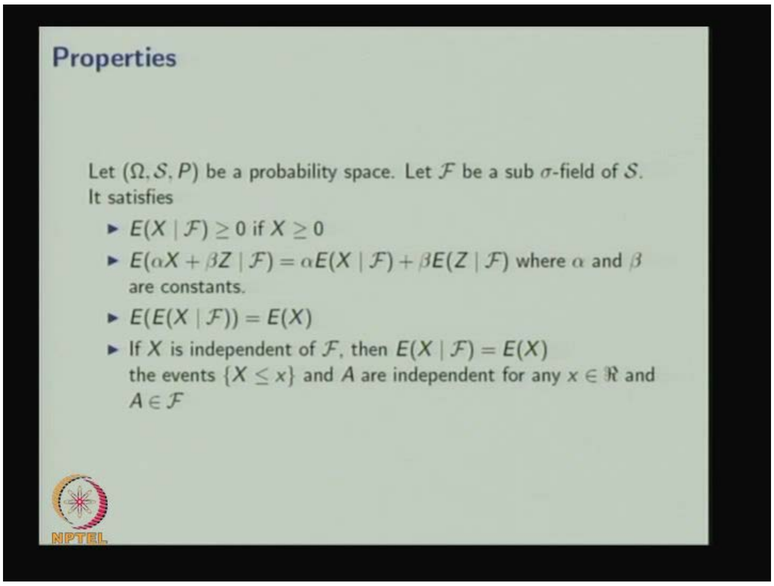
(Refer Slide Time: 45:33)



The second example is as follows. Here the omega consists of three elements, S is the largest sigma field; probability measure is defined on omega in each sample itself if the probability 1 by 3, X is the random variable, F is not the trivial one here. F is the sigma field, which is not a trivial one.

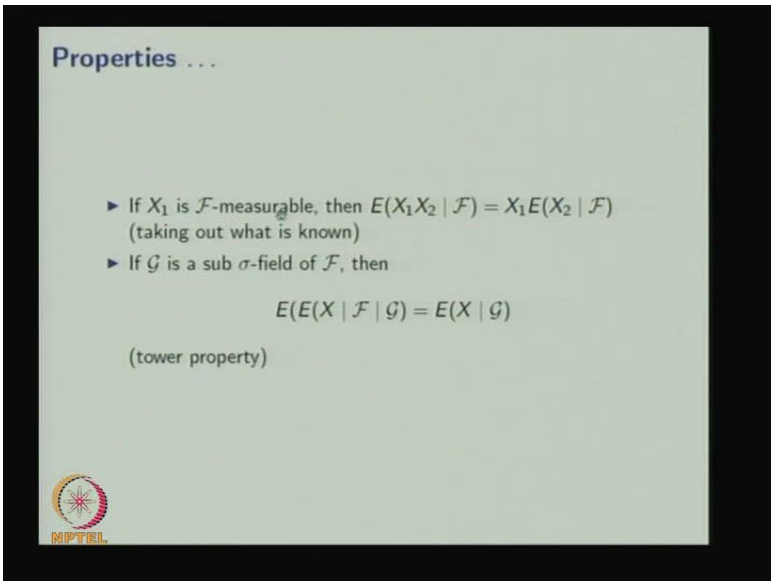
And I am defining another random variable Y and that takes the value 0 or 1 and here, I am claiming, that the expectation of X given F is same as expectation of X given Y because Y is the sigma field generated because F is a sigma field generated by the random variable Y. If you, if you create the sigma field generated by Y, you may land up empty set element a, element b and c, the whole set and that is same as F. Therefore, you can replace expectation of X given F by expectation of X given Y. Here, Y is the F measurable. You can check, Y is the random variable also by finding the inverse images.

(Refer Slide Time: 47:12)



The other properties are, the first one, the conditional expectation is always greater than or equal to 0 if X is greater than or equal to 0, then the linear property similar to the conditional expectation, as I have discussed earlier. Then, this also I have discussed. Instead of sigma field I have discussed with the random variable, if they are independent, then both are same.

(Refer Slide Time: 47:47)



The last two properties, if X 1 is F T measurable, then the multiplication, the X 1 taken out, which is what is known. So, since X 1 is known because X 1 is F measurable, so X 1 will be coming outside, X 1 times the conditional expectation. If g is a sub-sigma field, then this expectation, conditional expectation is same as the conditional expectation and this is called the tower property and I am not going to give the proof of this. Here is the reference.

Thanks.