

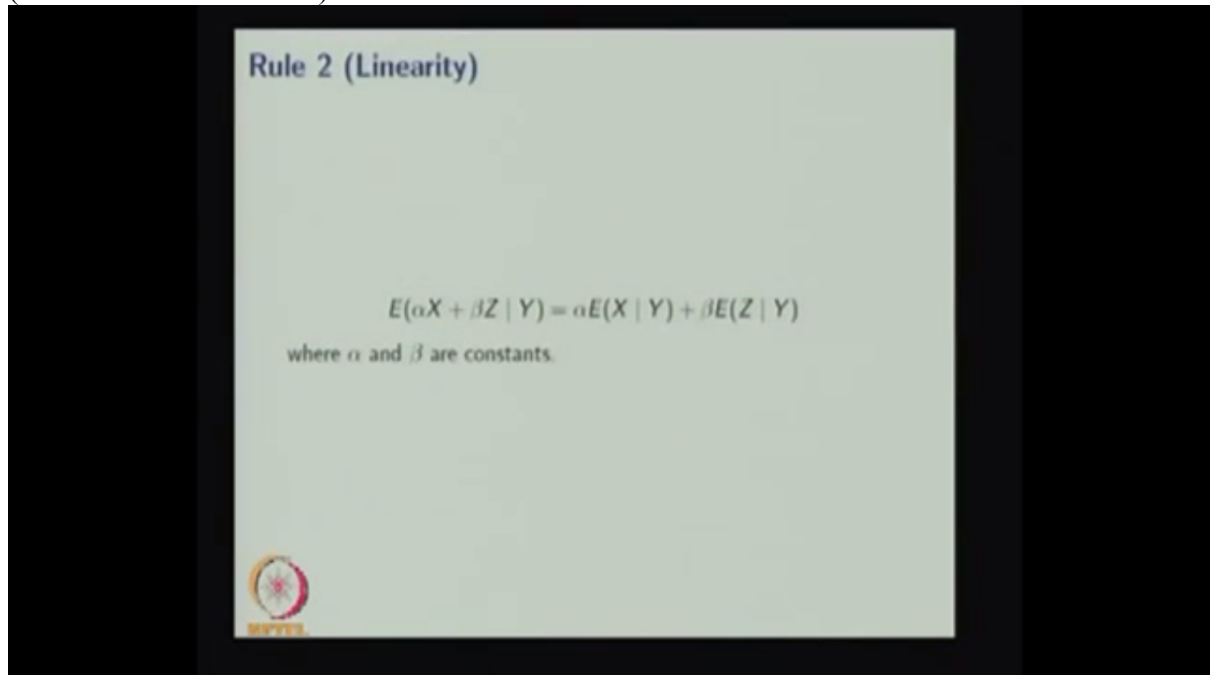
**Video Course on  
Stochastic Processes**

**Filtration in Discrete time**

by

**Dr. S Dharmaraja  
Department of Mathematics, IIT Delhi**

(Refer Slide Time 00:04)



The Rule 2, it is a Linearity property. The linear property says the expectation of alpha times X plus beta times Z given that Y takes a 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's 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 00:41)

### Rule 3(Expectation Law)

$$E(E(X | Y)) = E(X)$$

It says that the expectation of a 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.

**Proof:** Assuming  $X$  and  $Y$  are continuous r.v.s with joint pdf  $f$ ,

$$\begin{aligned} E(E(X | Y)) &= \int_{-\infty}^{\infty} E(X | y) f_Y(y) dy \\ &= \int_{-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} x f_{X|Y}(x/y) dx \right] f_Y(y) dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x f_{(X,Y)}(x,y) dx dy \end{aligned}$$

Using  $\int_{-\infty}^{\infty} f_{(X,Y)}(x,y) dy = f_X(x)$ , we get

$$E(E(X | Y)) = \int_{-\infty}^{\infty} x f_X(x) dx$$

The third one, that is the Expectation Law. Expectation of conditional expectation is same as expectation. It says the expectation of a 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 around. Expectation of  $X$  can be computed as a expectation of conditional expectation of  $X$  given  $Y$ . That means you can 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 same as the expectation of  $X$ .

The proof is given with the assumption both the random variables are continuous with the joint probability density function  $f$ . So I'm starting with the left hand side expectation of expectation  $X$  given  $Y$  that is same as you know how to compute the expectation. Here the provided condition is expectation exists. Similarly, in the definition of conditional expectation also you have to make the assumption the expectation is exist. Then only we are finding the conditional expectation.

So this expectation exists. Therefore, minus infinity to infinity conditional expectation, this is a random variable and you are finding the expectation of that. Therefore, you multiply 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 with the probability density function of  $y$ , integrate with respect to  $y$  between the limits minus infinity to infinity and by 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 or 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  $f(x)$ . Therefore, the one integration and this much will give a marginal. Therefore, it is a 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 a lot of importance.

(Refer Slide Time 04:49)

**Rule 4 (Independence Law)**

If  $X$  and  $Y$  are independent random variables, then

$$E(X | Y) = E(X)$$

**Proof:** Assuming  $X$  and  $Y$  are continuous random variables,

$$\begin{aligned} E(X | Y) &= \int_{-\infty}^{\infty} x f_{X|Y}(x/y) dx \\ &= \int_{-\infty}^{\infty} x f_X(x) dx \end{aligned}$$

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 variables 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 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 5:45)

**Rule 5(Stability)**

$$E(Xg(Y) | Y) = g(Y)E(X | Y)$$

**Proof:** Assuming  $X$  and  $Y$  are continuous random variables,

$$\begin{aligned} E(Xg(Y) | Y) &= \int_{-\infty}^{\infty} xg(y)f_{X|Y}(x/y)dx \\ &= g(y) \int_{-\infty}^{\infty} xf_{X|Y}(x/y)dx \\ &= g(y)E(X | Y) \end{aligned}$$

The next rule is Stability. Suppose you have the function  $g$  and  $g(Y)$  is also going to be a random variable, that means  $g$  is a Borel measurable function, so the expectation of  $X$  times  $g(Y)$  given  $Y$  that is same as the  $g(Y)$  will be out,  $g(Y)$  times the conditional expectation of  $X$  given  $Y$ .

That means the later we are going to use the property called known is out. That means that the expectation of  $X$  times  $g(Y)$  given  $Y$  takes a value something, some small  $y$ , that means this is going to be treated as a constant. So  $g(Y)$  has to be treated as a constant. Therefore, the constant will be come out. Therefore,  $g(Y)$  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 06:59)

## $\sigma$ -fields on $\Omega$

- ▶ Example: Toss a coin infinitely many times
- ▶  $\Omega = \{(HH\dots), (HT\dots), (TT\dots), (TH\dots), \dots\}$
- ▶  $\mathcal{F}_0 = \{\emptyset, \Omega\}$
- ▶  $\Omega_1 = \{A_H, A_T\}$
- ▶  $\mathcal{F}_1 = \{\emptyset, A_H, A_T, \Omega\}$  - contains the information learned by observing the first toss
- ▶  $\Omega_2 = \{A_{HH}, A_{TH}, A_{HT}, A_{TT}\}$
- ▶  $\mathcal{F}_2 = \{\emptyset, A_{HH}, A_{TH}, A_{HT}, A_{TT}, \{A_{HH}, A_{TH}\}, \dots, \Omega\}$  - contains the information learned by observing the first two consecutive tosses



$$\mathcal{F}_1 \subset \mathcal{F}_2 \subset \mathcal{F}_3 \subset \dots$$

Now we introduce sigma-fields on  $\Omega$  through an example because this is very important concept for the Martingale. Example start with 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  $\mathcal{F}_0$  be the trivial sigma-field consists of two elements empty set and the whole set.  $\mathcal{F}_1$  is the smallest sigma-field containing in  $\Omega_1$ .  $\mathcal{F}_2$  is the smallest sigma-field containing in  $\Omega_2$  containing the information learned by observing the first two consecutive tosses.

If you observe, you will find the  $\Omega_1$  contained in, sorry,  $\mathcal{F}_1$  contained in  $\mathcal{F}_2$ ,  $\mathcal{F}_2$  is contained in  $\mathcal{F}_3$  and so on. Since we are tossing a unbiased coin infinitely many times, if you find out the limit of  $\mathcal{F}_n$ , that is going to exist and that is going to be  $\mathcal{F}$  infinity in notation and that in notation we can make out it's  $\mathcal{F}$ .

So this consists of the information learned by infinitely many tosses observation. That is the sigma-field  $\mathcal{F}$ . 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 larger sigma-field  $\mathcal{F}_2$ .

So like that you can create  $\Omega_3\mathcal{F}_3$ ,  $\Omega_4\mathcal{F}_4$  and so on and all those  $\mathcal{F}_i$ 's satisfies this property and the limit exists as  $n$  tends to infinity. This sigma-field is going to be denoted by the letter  $\mathcal{F}$ . So this is the way one can create the sigma-fields on  $\Omega$ .

(Refer Slide Time 10:29)

## $\sigma$ -field generated by a collection of subsets of $\Omega$

### Definition

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  $\sigma(U)$ .

Consider an example. Let  $\Omega = \{a, b, c\}$ . Let  $U = \{a\}$ . Then  $\sigma(U) = \{\emptyset, \{a\}, \{b, c\}, \Omega\}$  is a  $\sigma$ -field generated by the collection of sets  $\{a\}$ , or, generated by the set  $\{a\}$ .



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  $\sigma(U)$ .

Consider an example where Omega is equal to  $\{a, b, c\}$ . Let  $U$  is  $\{a\}$ . Then  $\sigma(U)$  is a empty set, element  $a$ , element  $\{b, c\}$  and the whole set is the sigma-field generated by the collection of sets  $\{a\}$ .

(Refer Slide Time 11:25)

## $\sigma$ -field generated by a random variable $X_t$

### Definition

The  $\sigma$ -field generated by a random variable  $X_t$  is defined as the  $\sigma$ -field generated by the family of subsets of the form  $\{X_t \in B\}$ , where  $B$  ranges over the Borel sets of  $\mathbb{R}$ .

Consider an example. Let  $\Omega = \{a, b, c, d\}$ . Define

$$X(w) = \begin{cases} 0.5, & w = a, b \\ 1.5, & w = c, d \end{cases}$$

Let  $A_1 = \{w \in \Omega : X(w) = 0.5\} = \{a, b\}$  and  $A_2 = \{w \in \Omega : X(w) = 1.5\} = \{c, d\}$  be the subsets of  $\Omega$ . Let  $U = \{A_1, A_2\}$ . Hence,



$$\sigma(X) = \sigma(U) = \{\emptyset, \{a, b\}, \{c, d\}, \Omega\}.$$

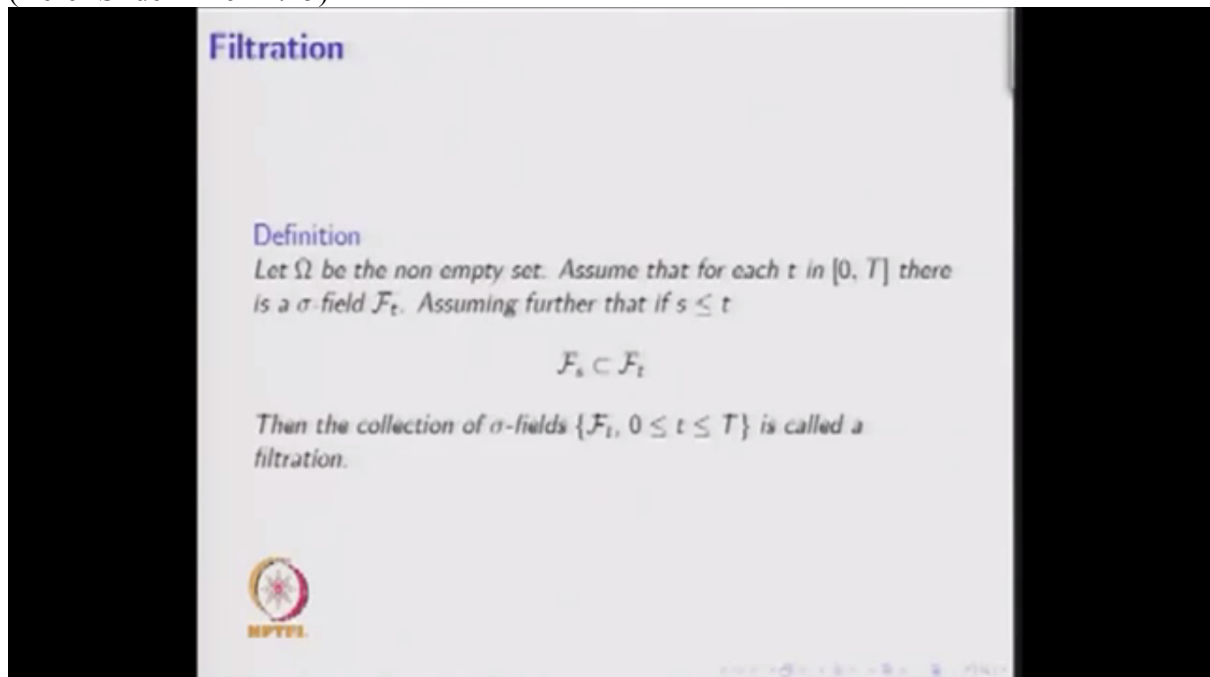
Consider an example where Omega is equal to  $\{a, b, c, d\}$ . Define the random variable  $X$ , which takes a value 0.5 for  $w$  is equal to  $a, b$ . It takes a value 1.5 when  $w$  takes a value  $c, d$ .

Let  $A_1$  be the set which takes the value -- which is the collection of possible outcomes in which the  $X(w)$  takes a value 0.5. Therefore it is  $\{a, b\}$ .

Similarly, let  $A_2$  to be the set for the collection of  $w$  belonging to  $\Omega$  in which  $X(w)$  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  $\sigma(U)$  that is empty set,  $\{a, b\}$  is one element and  $\{c, d\}$  and the fourth element is the whole set.

(Refer Slide Time 12:45)




**Filtration**

**Definition**  
Let  $\Omega$  be the non empty set. Assume that for each  $t$  in  $[0, T]$  there is a  $\sigma$ -field  $\mathcal{F}_t$ . Assuming further that if  $s \leq t$

$$\mathcal{F}_s \subset \mathcal{F}_t$$

Then the collection of  $\sigma$ -fields  $\{\mathcal{F}_t, 0 \leq t \leq T\}$  is called a filtration.

 NPTEL

Assume that for each  $t$  in  $[0, T]$  where  $T$  is a positive real number, then for each  $t$  between the interval 0 to  $T$ , you are creating a sigma-field  $F$  of  $t$  and those sigma-fields  $F$  of  $t$  for possible values of  $t$  between the interval 0 to  $T$ , it satisfies the condition  $F_s$  is contained in  $F_t$ . If this condition is satisfied over the interval 0 to  $T$  by  $s$  and  $t$  where  $s$  is less than or equal to  $t$ , then this collection of -- this collection of random variables, sorry, this collection of sigma-fields is called the filtration.

(Refer Slide Time 13:38)

## Filtration in Discrete Time

In discrete time, the filtration is an increasing sequence  $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \dots$  of  $\sigma$ -fields, one per time instant. The  $\sigma$ -field  $\mathcal{F}_n$  may be thought of as the events of which the occurrence is determined at or before time  $n$ , the "known events" at time  $n$ . The natural filtration of a stochastic process  $\{X_n, n = 0, 1, 2, \dots\}$  is defined by

$$\mathcal{F}_n = \{(X_0, X_1, \dots, X_n) \in B; B \subset \mathbb{R}^{n+1}\}$$

It can also be written as

$$\mathcal{F}_n = \sigma(X_0, X_1, \dots, X_n)$$



The definition of filtration in real time is as follows. In discrete time, the filtration is an increasing sequence  $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \mathcal{F}_2 \subset \dots$  of  $\sigma$ -fields, one per time instant. The  $\sigma$ -field  $\mathcal{F}_n$  may be thought of as events of which the occurrence is determined at or before time  $n$ , the "known events" at time  $n$ .

The natural filtration of a stochastic process  $X_n$  is defined by  $\mathcal{F}_n$  is the collection of  $n+1$  dimensional random variables belonging to  $B$  where  $B$  is contained in  $\mathbb{R}^{n+1}$ . This is also written as  $\mathcal{F}_n$  is a  $\sigma$ -field generated by the  $n+1$  random variables or you can think of a random vector with  $n+1$  dimension. So the  $\mathcal{F}_n$  is a  $\sigma$ -field created by the random vector  $X_0$  to  $X_n$  or the random variables  $X_0$  and  $X_1$  and so on till  $X_n$ . So this is the filtration in discrete time.

(Refer Slide Time 15:23)



## $\sigma$ -fields on $\Omega$

- ▶ Example: Toss a coin infinitely many times
- ▶  $\Omega = \{(HH\dots), (HT\dots), (TT\dots), (TH\dots), \dots\}$
- ▶  $\mathcal{F}_0 = \{\emptyset, \Omega\}$
- ▶  $\Omega_1 = \{A_H, A_T\}$
- ▶  $\mathcal{F}_1 = \{\emptyset, A_H, A_T, \Omega_1\}$  - contains the information learned by observing the first toss
- ▶  $\Omega_2 = \{A_{HH}, A_{TH}, A_{HT}, A_{TT}\}$
- ▶  $\mathcal{F}_2 = \{\emptyset, A_{HH}, A_{TH}, A_{HT}, A_{TT}, \{A_{HH}, A_{TH}\}, \dots, \Omega_2\}$  - contains the information learned by observing the first two consecutive tosses

$$\mathcal{F}_1 \subset \mathcal{F}_2 \subset \mathcal{F}_3 \subset \dots$$



15:24



16:12

15:24/16:12

Let us generate a filtration for this example for discrete situation. So take the same example, tossing a unbiased coin infinitely many times and the  $\Omega_1$  is having two elements. This is the sigma-field on  $\Omega_1$  and the  $\mathcal{F}_2$  is the sigma-field on  $\Omega_2$  and also satisfied sigma --  $\mathcal{F}_1$  is contained in  $\mathcal{F}_2$  which is contained in  $\mathcal{F}_3$ . Therefore, this collection of -- the collection of random -- the collection of sigma-fields is called the filtration. So this is an example of creating sigma-field in real time.