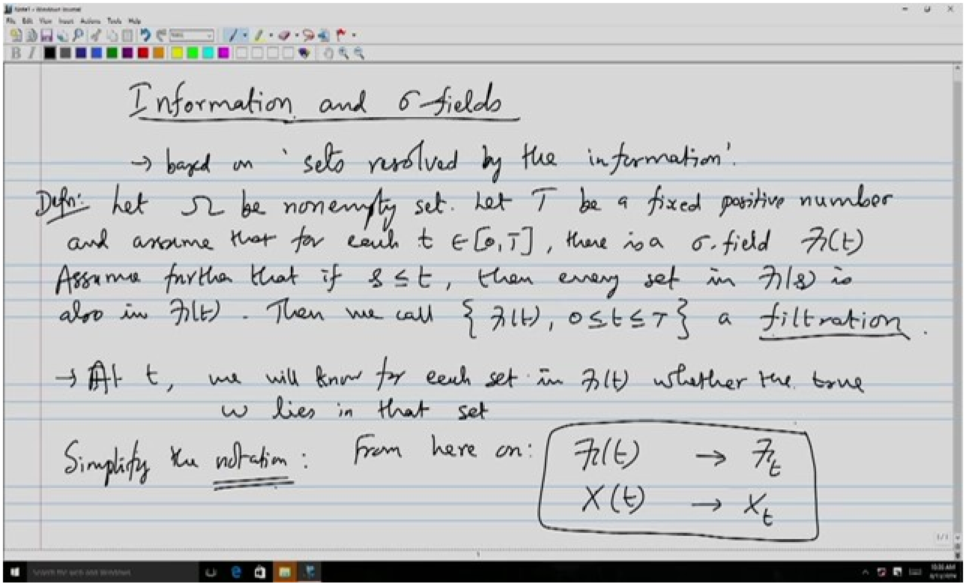
Mathematical Finance

Lecture 26: Filtrations, Independence, Conditional Expectation

Professor N. Selvaraju[1] and Professor Siddhartha Pratim Chakrabarty[1]

[1]*Department of Mathematics, Indian Institute of Technology Guwahati, India*

(Refer Slide Time: 00:33)



Hello everyone, the next concept that we will see is this information and sigma fields or sigma algebras, again you can draw parallel to what we have done in discrete time model, we will continue with the continuous time version of what we have done, for understanding you can always go back and dry parallel with discrete that is it.

As, you see this norbitrage pricing theory, which is used to price the derivative securities is based on contingency plan, right, so to price this derivative security what we do is you know we try to find the initial wealth that is needed to set up hedge for the short position, right, so which means to completely replicate the things, right. Now, the hedge must specify what positions we will take and that requires the information from the present time till the future point of time, till the maturity, so to say, right, and this to make the contingency plan we need a way to mathematically model the information and which our future decisions will be based, right. And in binomial this we made an analogy to the coin toss and then the each time period we said that it is as if the coin, one more coin toss is getting added and how things are information is gathered and that is the way we model the discrete time.

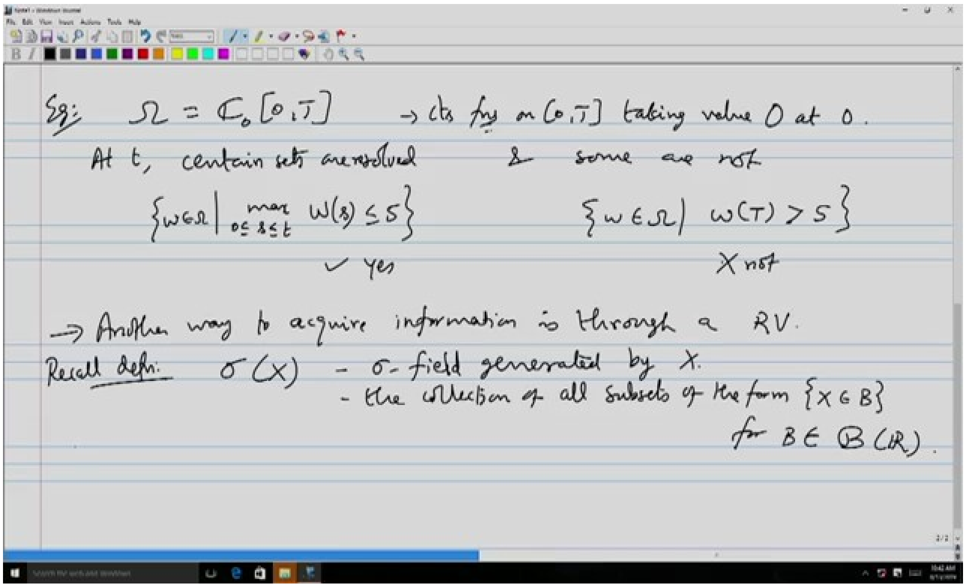
In continuous time also it is along similar line but we need a little bit more sophisticated model, way to model the information and that is based on basically in continuous times sets resolved by the information, right, so this is same in discrete but it is little bit complex here that is all. So, now what is the filtration, so this filtration would consist of sets resolved by the information available up to time t. So, the definition is thus follows, you have a omega be non-empty set and let T be a fixed positive number and assume that for each t in this interval [0,T] there is a sigma field or sigma algebra which we call $\mathcal{F}_t$, assume further that if you pick two time point s and t such that s is less than or equal to t, then every set in fs is also in $\mathcal{F}_t$, then we call the collection, which is $\mathcal{F}_t$ a filtration.

Just like in discrete time what we have here an increasing collection of rather than now sequence, it is a collection of sigma field it is family, because t is belongs to 0 to T in a continuum manner, so what we have is collection of increasing sigma fields is what then we call it as a filtration which is much like the similar t. So, this collection this filtration tells us the information we will have at all future times.

More preciously if, more preciously suppose at t, we will know we will know for each set in f of t, if you pick we will know whether the true omega lies in that set or not that is what would be so this is why the information or the sets resolved by the information at time t means, what we will know for each set you pick and you pick a true omega by it would be able to say further for each set weather this omega belongs to this or it does not belongs to this with certainty.

So, that is what the sets will be of the nature of the sets in this sigma filed. Now, before we go further let us simplify the notation, so from here on what we will do instead of saying this as a function of t, so what we will do is we just make this as a subscript, so that you know for readability it is just easier way of denoting but the meaning is a function of t. Similarly, suppose, if we have a process something X random variable later we are going to see, so this we will be denoting it in this manner. So X_t right this $\mathcal{F}_t$ X_t rating as a function of t there is only one then obviously we will put it in this, suppose if it is more than 1 then ofcourse we make spend it with 2. So, this is for the notation right, so this is you know you just keep than in mind from now onwards when we write f sub t the subscript t which means f of t is a function of t, I mean we are looking at as a function of t, so just for simplicity we will write it, which is common in stochastic process area, so it is nothing unusual that you know we are doing.

(Refer Slide Time: 07:58)



Now, let us look at an example. Let us take this collection of what we call the set of all continuous functions on 0 to T taking value 0 at 0, so this is what it is so suppose if there is a function f which is a continuous function and 0 1, defend on 0 1, which had 0 to T and f of 0 is 0 that is what, so any such f is what then this collection is. Now, at t if you look at, set and sets subsets are resolved by time T which means a typical omega here we means is a continuous function in this omega, and whatever the sets are resolved so we put that into Ft and we make a sigma field. If you simply put whatever sets that are all resolved by a time T then that will a sigma field, so let us not get into the details but that is you know you can take it.

Now, which means that all the sets that can be described by the path up to time T of those continuous functions they all would be resolved by this thing, so set and sets are resolved and some are not, say, for example, if you take set of all omega such that so if I look at something like this say max or mean or sum or at any particular point of time but the time is omega hour of s is say less than or equal to 5, suppose if I look at here for sum omega which is a continuous functions here.

So, we denoted in a usual way omega, so this is result whereas if I look at a set of this form, set of all omega such that my omega at time T is say greater than 5, so, this is not, this is yes, this is resolved, so this is the kind of sets that we are putting inside this Ft to form this, so this is basically you can see there is a path of this functions, the function values observe between 0 to T if anything it can be described by the set, right.

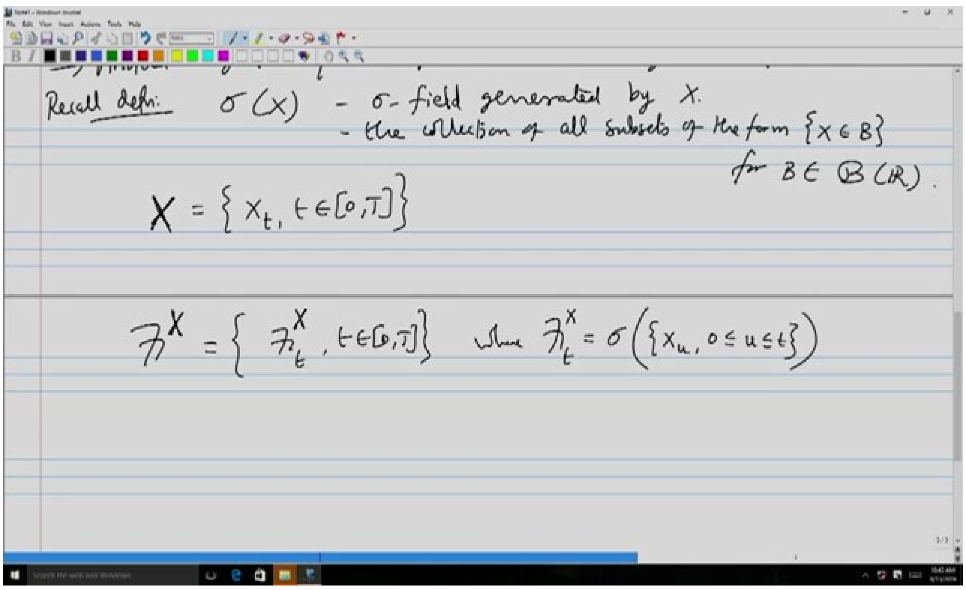
So, then this will all be in this form like this are all the sets that are resolved by the thing, so this example provides the simplest setting in which we may construct later on the Brownian motion. If we could assign probability is to each of this sets in Ft as we said, you know, when we talk about probability being assign in a continuous time model, it is actually sets in the sigma field rather than the atoms, and this if you assign only the probabilities then, of course, all this things will become the parts of the Brownian motion that we are going to define it will later.

Recall, atoms concept here we are not saying because rather we are directly dealing with the sets in the sigma filed typically as happens in a continuous time or infinite probability spaces and we work with all the set in Ft directly rather than the atoms and atoms if you typically take and specially the sets which have passed in probability, because atoms are generally have a probability of 0 in such situations.

This is one way of describing the information through, another way to acquire information is through random variable, so, recall the definition that we have written for a random variable X the quantity sigma of X, what is this, this is the sigma filed generated by the random variable X or which is nothing but the collection of all subsets of the form X belong to B for B a borel set.

So, you take a borel set B and you collect all those omegas, which gives the value of X in the borel set B and you collect all such sets that will be the sigma generated by X, this is what we have seen. This definition which we already know we have seen discrete time this is exactly same as in the definition. Now, extend this two again in discrete time also we have extended but you know there it was it is more convenient because you had a sequence of random variables, now you extend this to a family of random variables.

(Refer Slide Time: 14:31)



So, suppose if, we have say X, which is nothing but we denote the collection so this is sometimes the boldly we can write X_t, t belongs to some T right so this is the set T or I can make this as for our convenience in a usual way this collection X_t, right. Now we can define the filtration given by this or filtration generated by this family of random variables rather then a sequence of random variables as follows.

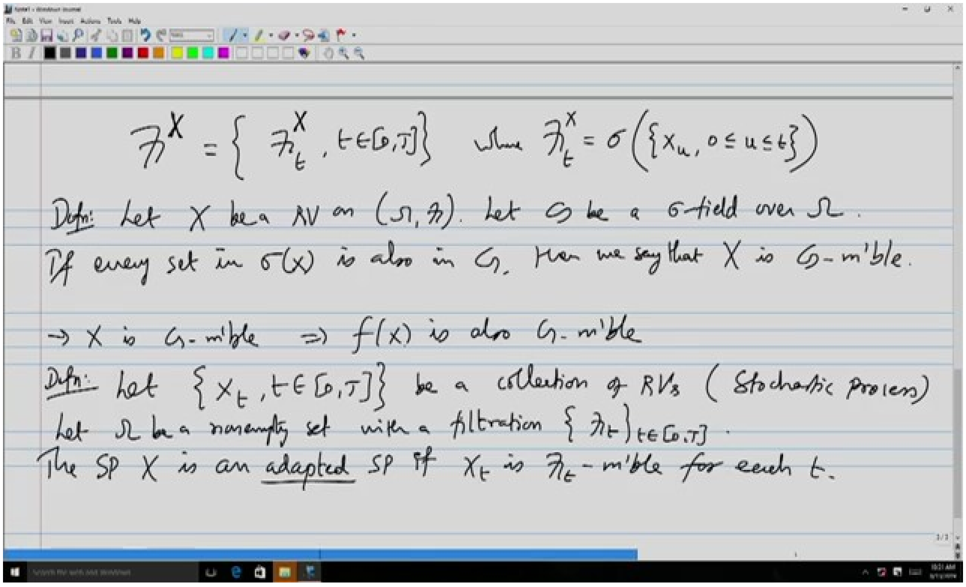
So, what we denote, we denote this by this $\mathcal{F}^X$ whenever we write this we mean that this is a sigma field or filtration generated by X which is in our case is nothing but $\mathcal{F}_t^X$, so $t \in [0,T]$.

I can write it something like this (where what it meant is), so in discrete you recall that will be said that this one is sigma of $X_0, X_1, X_2, \ldots$. We have extended. Now, again similar thing but now since we do not have this X0 X1 you have into the serer infinite number of points in between so we have writing it in this family.

So, this is the family and the sigma generated by this particular family is what my $\mathcal{F}_t(X)$ what are that mean this is the smallest sigma field with respect to which each of this random variable Xu for you less than or equal to t, is measurable, right that is what is the other way when you want to say sigma generated by a random variable or random variables right so this is what we would mean.

So, later on, we would be considering mainly the filtration generated by some family of random variables, so, that is nothing but stochastic process, so, this is some collection of random variables or some random process or stochastic process which is what this.. This is the most common filtration that we will be using it later on.

(Refer Slide Time: 17:42)



So, this is the filtration basically what we will do, right, now connected with this we also recall the definition of measurable T, now when do we suppose, you have a X which is random variable on $(\Omega, \mathcal{F})$, recall that you do not need the probability to define X, let G be a sigma field over omega, which means sigma field of subsets of omega.

If every set, this is also in G then we say that X is G measurable what is mean the measurability that means that if and only if information is G is sufficient to determine the pre size value of the random variable X, that is what measurability would mean in practice right, and σ_X, the sigma field generated by the random variable X, is nothing but the information content of the random variable, it has exactly same information as you have observing the random variable X.

So, it means that if each of the sets there in G means that G also has atleast that much information then we say X is G measurable so if we know G then you know X precisely the means what is the value of the random variable, there is no more randomness essentially in the case. Now, if X is G measurable then you can also see that f of some function of X is also G measurable where the function, we again recall that this is the borel measurable function right, so all smooth and a function that we are going to consider this.

Now, this gives also rise to another definition, next which is, now you pick a collection as we see a collection be a collection or a family of random variables, that means it what, which is a stochastic process, so what do you mean stochastic process I mean we have seen discrete time but again you may recall even if not there is nothing big about it is a, we have a family of collection indexed by a parameter T here right.

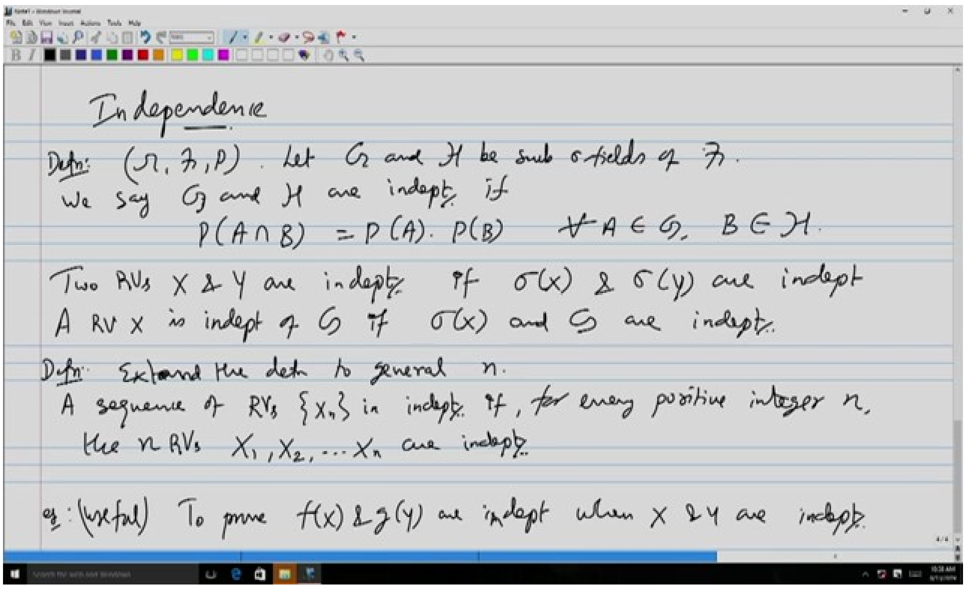
So, this is just a family. So, this family is what generally referred to a stochastic process and depending up on the nature of the relationship between the random variables in this collection, so we give a name to stochastic process, we already have seen martingale, markov process, (so which are all depending up on what is the nature of the relationship between the random variables that is what you have.)

Now, we pick that omega be non-empty the usual assumptions non-empty set which is essentially the sample space here with a filtration say $\mathcal{F}_t$ where t belongs to even if we do not write we always assume that this is, then the stochastic process X which means this collection is adapted stochastic process or random process if X_t is $\mathcal{F}_t$ measurable for each t.

So, what do you mean because from now on we may say that some process is an adapted process, adapted process means that in that process if I take tth element of the process, then that tth element is correspond to the filtration f measurable with respect to the Tth element in the filtration, which is $\mathcal{F}_t$ which is what we call the process of that nature is adapted process right.

So, in discrete time also we have seen the wealth process, the portfolio process, the asset price process, all of them were adapted, means that once you reach up to that point of time then you know the value exactly of that prior to that you will not know, right, so that, is what it is. So, this what in this case so again in continuous time model also you have the asset price process, portfolio process, wealth process all of them would be adapted process, so this is the definition of the adapted process when we use this here.

(Refer Slide Time: 23:39)



So, the next concept that we will look into that is independence, so you have a random variable, you have a sigma field now on the one extreme if this random variable is measurable with respect to the sigma field that is given, now the other extreme is that the random variable is independent of the sigma field given, now, what we saw in the case of measurability?

We saw that, the information in G will determine completely the random variable X but in the case of independence the information in G is of no value to determine the value of X, the intermediate case is what will take it up, will later is what the conditional expectation case. So, let us first look at the other extreme, which is independent and since you already know from the earlier theory that independence is connected with probability measure.

So, by changing the probability measure the independence nature of the either between the random variables are between two sets are anything might be broken, so, you need to keep that thing in mind. Now, let us define what we may have, so as usual we have a probability space and let us take two sub sigma fields of is of F, say let G and H be sub sigma fields of F.

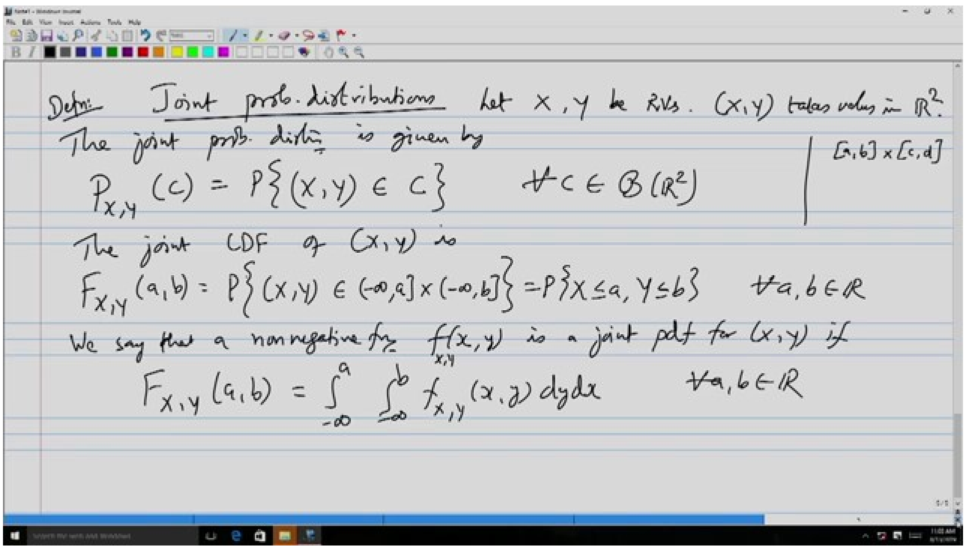
Then, we say this two sigma fields, G and H are independent if P, this is similar to independence of two even, but now when we talk about independence of two sigma will be mean the set from one sigma field and an arbitrary set from the other sigma field right, these two should be independent for all such arbitrary sets, that is what the sigma field is mean.

So, extending this you can also talk about (you know) two random variables say X and Y are independent if σ_X and σ_Y are independent. So, this is what we mean, when we say two random variables are independent. similarly a random variable of X is independent of sigma field G, if σ_X and G are independent, so there is nothing to this case. now you can also extend the definition to general n, so you can not talk about n random variables being independent of each other and n sigma field and so on extending same thing that you have extend.

But more importantly we say a sequence of random variables say X_n is independent if for every positive integer n, the n random variables say from this collection which is $X_1, X_2, ..., X_n$ they are independent, so if a collection of random variables, if this is true for each finite set when you pick, so this claim above that we use here that we are talking about independence of random variable in terms of the underlying sigma field generated by the random variables is much more strong, mean not necessarily strong in the sense of anything extra that it gives, but it is easier to view it through this way, then it will be useful to prove many things.

Say, for example, this way of thing is useful, if you want to prove say, for example, to prove $f(X)$ and $g(Y)$ are independent whenever X and Y are independent, so if we talk in terms of the sigma of fields, the prove of this is simple. Straight away it follows from there (but otherwise) but it is difficult to check in practice.

(Refer Slide Time: 30:02)



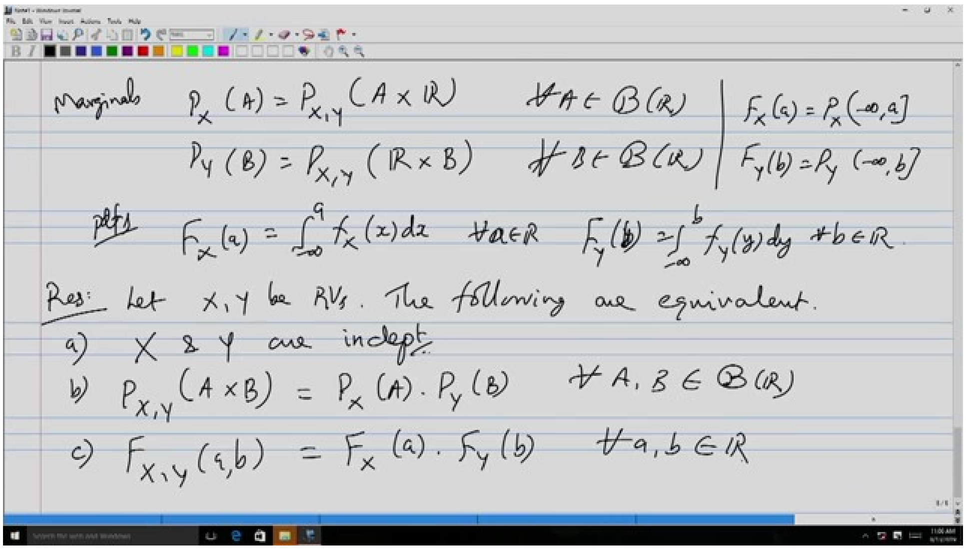
So, that is where your useful definition of independent, which you might be aware would come into play, which requires the knowledge of what we call joint probability distributions. So, this is you might have idea about this probability but let us take the simple case two, but that is exactly similar in the case of more than two. So, let X, Y be random variables, the pair, so the pair so which means if I look at this so this takes values in R^2, so if I consider this as a pair so this are random vector is what generally called or a two dimensional random variable which means that each of those one dimensional random variable that we have.

Now, the joint probability distribution is given by which we call $P(X,Y)$ of some set C as defined by this pair taking values in the set C for all C in this collection, what is this collection, this is the collection of Borel sets over R^2, but we have not seen this yet so far, but it is much similar to the Borel sigma field over R. In R what we did we do we took close intervals of the form ab and we generated a sigma field that is what Borel sigma field over R.

So, in this case you know R^2 you can take rectangles in R^2 and we generated sigma field over R^2 with generating class as being the rectangles in R^2, so that is what you will end up with the Borel sigma field over R^2. So, this is much like you know say, for example, you can take sets of the form this, which is nothing but the rectangle and this is basically the Cartesian product that we are talking about so this will represent some, you know, rectangle in R^2, so now you take this as a generating class and it generate the sigma fields sigma field over such rectangle so then what you get is the Borel sigma field.

So, this is just that thing that you have extending the idea of here. Now, you also know the joint CDF or the cumulative distribution function of this two dimensional random variable some a, b as probability of this pair belonging to minus infinity, this set products Cartesian products with this collection, so this is nothing but probability of X less than a, while then b for all a, b in or is what is the equivalent way in which you know can also represent as the joint CDF. Now, we also say that non negative function say $f(x,y)$, if you want subscript you can also put x, y is joint pdf for x, y, if you can write this as a, b so this is the joint case.

(Refer Slide Time: 35:05)



And similarly, you could also have the marginal, so this are all ideas that we are just recalling, so $P_X(A)$ is nothing but from A cross R for all A in Borel set of R similarly $p_Y(B) = P_{X,Y}(R \times B)$ for all $B \in B(R)$. So equivalently you can also define $F_X(a) = P_X(-\infty, a]$ and $F_Y(b) = P_Y(-\infty, b]$.

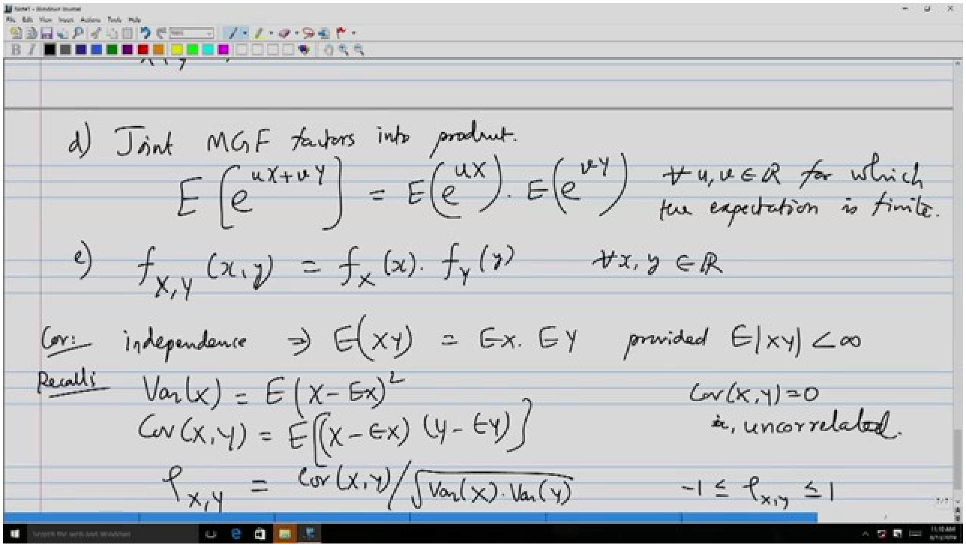
Now, given this definitions, now we have the result which tells us you know how to check for independence in practice. So, let us pick two random variables, let X and Y be random variables, then the following whatever we are going to write or equivalent, what are they, we say that

a) X and Y are independent,

b) the joint probability distribution, if it factors as here for all sets A, B which are Borel sets,

c) the joint CDF factors joint CDF factors.

(Refer Slide Time: 38:10)



d) which is basically what we call the joint moment generating function factors into product. So, what it that mean so this MGF you would be aware or even if not in the let us write in the original form itself, so, this is

$$E[e^{uX+uY}] = E(e^{uX})E(e^{vY})$$

As you know MGF may not exist for a particular random variable, so we need to worry about the existence, so whenever this exist so this is true that the MGF factors into its product of MGF of the marginal.

e) the joint pdf again the density may not exist but whenever it exist, so the joint pdf factors into product for all x, y in R. As a corollary to this will also see the

$$E(XY) = E(X)E(Y)$$

Now, you recall the definition of what we call variance of X which is nothing but expectation of X minus expectation of X the whole square.

And you have covariance of X and Y which is

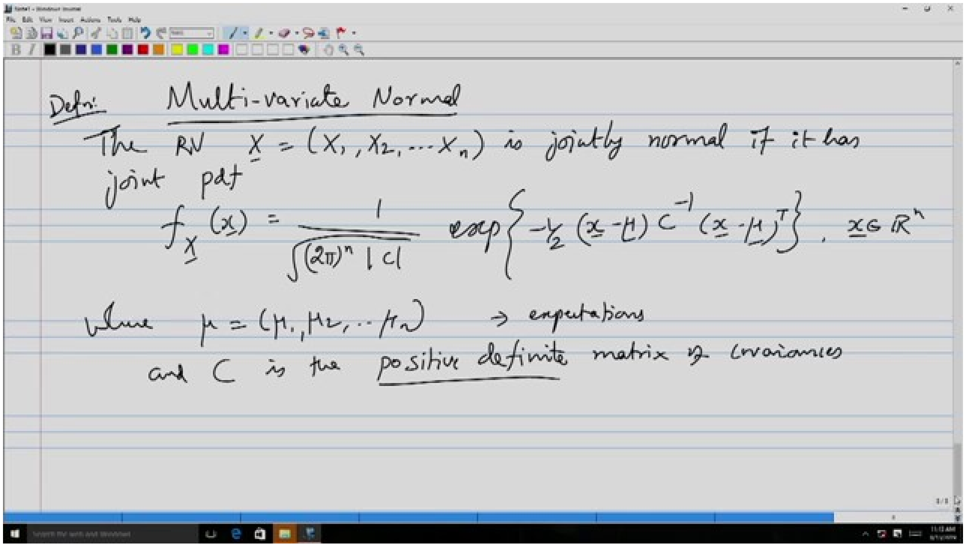
$$Cov(X,Y) = E[(X - E(X))(Y - E(Y))]$$

Also the correlation coefficient of X and Y which is

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)var(Y)}}.$$

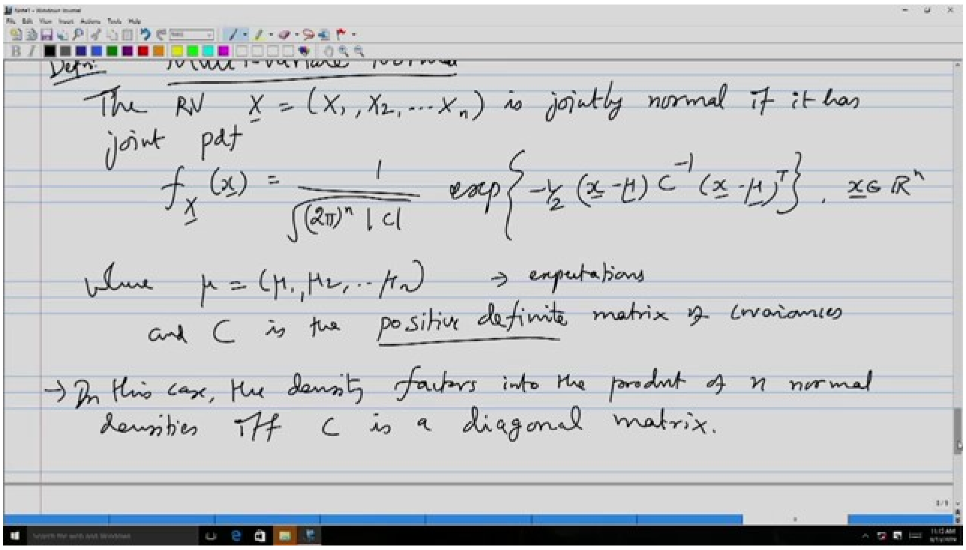
Now, this is whenever covariance of X and Y equal to 0 that is what we call it as uncorrelated, so in general, so what this the corollary which we just wrote before means is that the independence implies un correlation but un correlation need not imply independent, of course you can find examples easily, so this is one way implication but as far as the independent checking is concerned you can talk either in terms of the pdf in terms of the MGF in terms of distribution function CDF or in terms of the distribution measure itself any one them you can use to check for independence of two random variables X and Y, for checking this, this is what we would be using it.

(Refer Slide Time: 42:40)



Now, so we will also need the idea of you know multi variate normal we know the univariate number distribution, so we will also need this idea as we go forward so let us take the random variable $X = (X_1, X_2, \ldots, X_n)$ is what we call jointly normal or this as a joint normal distribution, if it has joint pdf which we write as above.

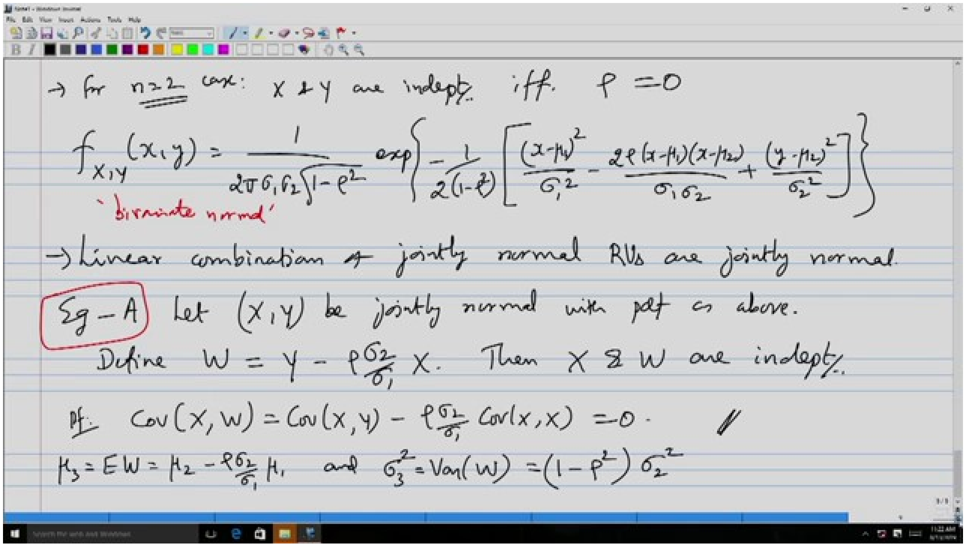
(Refer Slide Time: 45:18)



Now, if you look at this multivariate normal and if you look at when this multivariate normal will be or this n random variables are independent you can easily observe that the density if I look at here so in this case the density factors into product of n normal densities, if and only if C is a diagonal matrix, we assume positive definite just to ensure that its inverse exist and the random variable are jointly normal.

Now, in this particular case, if this is a diagonal matrix then you would see that this the quantity that we have written with this exponent, so will be the determinant and the quantity that we will we have written into the exponent they all will be factoring you can write is as simply as some of the individual element without anything so we will be able to see that then the this factors into the product of marginal densities of the normal pdf that you have.

(Refer Slide Time: 46:39)



Now, special case, so for n is equal to 2 as you look at it so this factorization or this into the products happens so this is X and Y in for n is equal to 2 case X and Y are independent, if and only if my $\rho(X,Y) = 0$. So this is a special case with respect to the normal that in normal case that this un correlation and independence are equal, simple so this is a additional advantage in general this is not true but in case of joint n number distribution this is true.

Now, you can easily see if I put rho is equal to 0 this quantity become 0 and this is know you can see easily it factors into the 2 quantities here, so this becomes 0, this becomes 0 so this becomes you know product of 2 standard normal densities. That is what we means and we say this one, so, in case of normal, we also know if I take linear combinations linear combinations of jointly normal random variables are jointly normal and since independent normal is jointly normal.

So, in order to get a generic or a much more general multi variate normal one way is you know you take independent normal and take linear combinations of those independent normal to arrive at the required multi variate normal on the other hand if you have a multi variate normal distribution then it can be written as a linear combination of independent normal random variable.

So, that is what will be helpful as we go along because whenever you have dependent normal, but they are jointly normal but they are not independent then you can write it down in terms of as a linear combination of linear normal random variables and while dealing with another advantage with a normal distribution is that if the normal distribution is characterized by its mean and variants and the covariance is in general in case of multi variate normal means an covariance then it is enough if you work with the means and covariance terms it need not it is equivalent to working with the case of the densities.

So, dealing with means is sufficient, so now let us consider an example, this example we will keep repeating so let us call this as an example A. Now what we have, so let X Y be jointly normal with pdf as above, since, this is a two dimensional case so what we mean is this a by variate normal distribution we call it and this the density of the bivariate normal, so this is called bivariate normal general bivariate normal density now bivariate normal distribution.

Now define random variable W, as this $W = Y - \rho(\sigma_1/\sigma_2)X$ then the claim is X and W are independent. we can see by looking at the covariance of X and W so which we can easily see is covariance of X and Y minus 4 times this of covariance of X, X which is equal to 0, so covariance of X and W is 0 and hence X and W are independent, and (what is its mean let us) that is done, so this is prove is only this much is done.

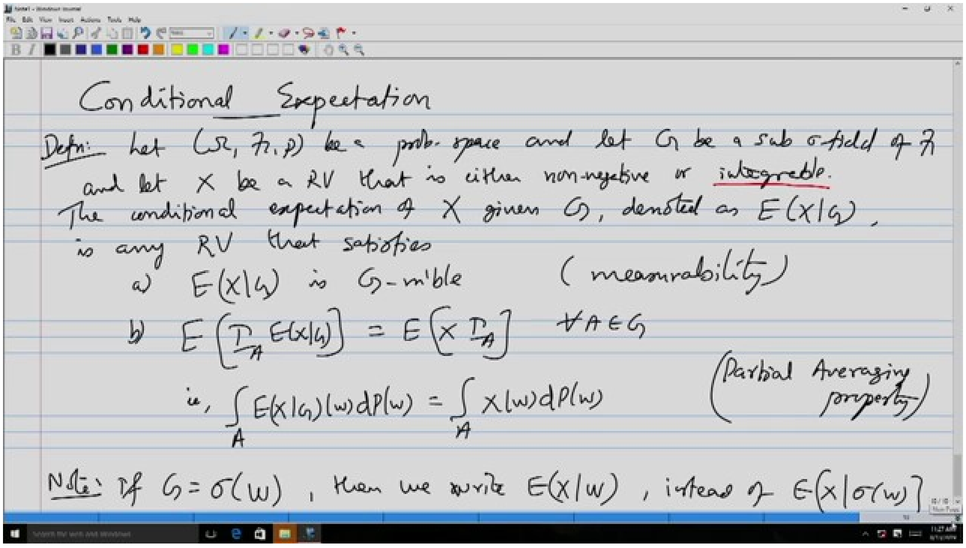
Now, further we see that

$$E(W) = \mu_2 - \rho\rho(\sigma_1/\sigma_2)\mu_1$$

and

$$var(W) = (1 - rho^2)\sigma_2$$

(Refer Slide Time: 53:50)



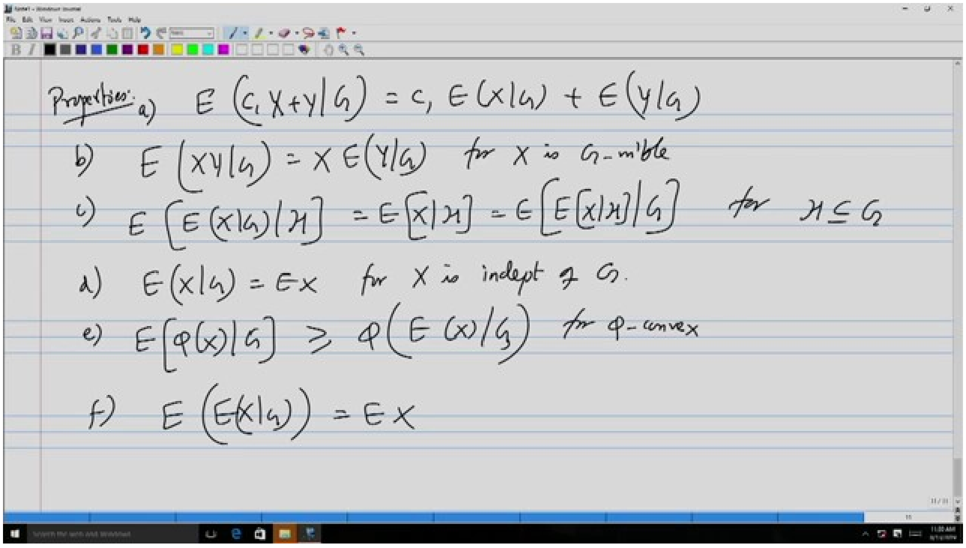
Now, this is the other extreme independence and how do we check and in case of multi variate normal like how this is related in this case, now we will go to our general case of a relationship between random variable and sigma field where X is neither measurable with respect to G nor independent of G then what we have we have to really now estimate we can not determine the precise value of X in that case what we have to do is that we can we have some information which is useful to know about X.

So, which means that you will use this information to estimate X, which is what is the more common case, so we can not precisely calculate X but we can estimate X and that estimate is what we call conditional expectation in discrete time also we said in what sense we said this is the best estimate which is the mean square either sense, so that is true here also and we also arrived at the definition in discrete we define it slightly differently at we arrive at a characterization, which is equivalent and we said that the characterization is going to be the definition, which we will going to take in the characterization and that is precisely we will state it now.

So, what we have, so we have a probability space and let G be a sub sigma field of f and let X be a random variable that is either non negative or integreble though we state this thing for both the cases we will mostly assume that this is integreble case but this is true for non-negative. Then the conditional expectation of X given G, denoted by X given G, which you know will keep as it is, it may call it as some Y.

If it is the disconvenient is any random variable that satisfies two condition,
a) $E(X|G)$ is G measurable this is called the measurability property,
b) $E[I_A E(X|G)] = E(XI_A)$ This is called partial averaging property. So this is the partial, you recall, in discrete time also we said this two properties only defines the conditional distribution. Now, note if our G is simply a sigma field generated by some random variable say W, then we write this as $E(X|W)$ instead of $E(X|\sigma(W))$, but what we mean is this one only but we simply write this in this form. Now, as far as the existence of such a quantity is concerned yes. As long as you are expectations of X is we assume integrable so expectation of X is finite, such a conditional expectation exist and unique almost surely, so that prove ofcourse we will not needed.

(Refer Slide Time: 58:53)



Quickly , recall the properties that we have seen in discrete time, which is also would good to in continuous time,

a)the first is the linearity property,

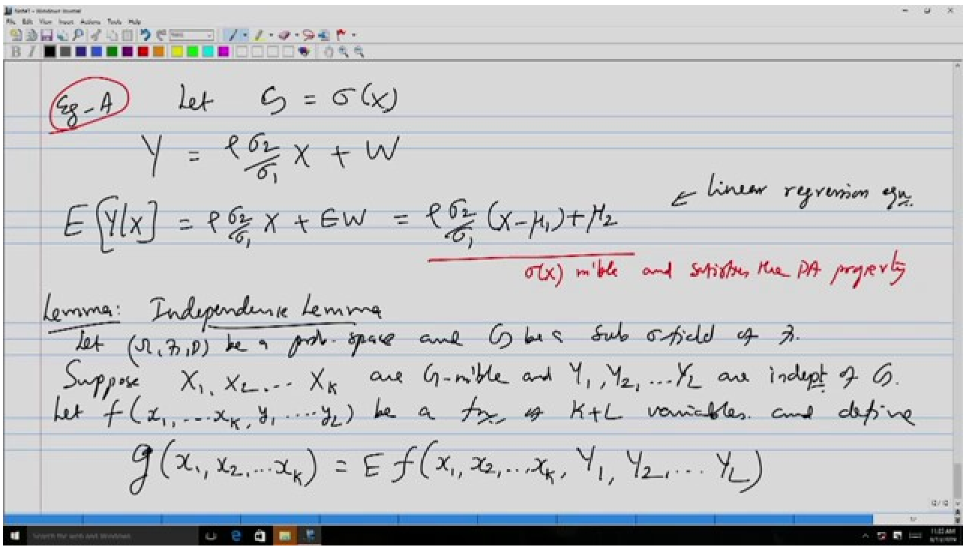
b)then taking out what is known property which means X is G measurable then you can take it out from the estimation,

c) the third is basically the tower property or the iterated conditioning property which essentially means expectation of expectations of X given G is expectation of X given H, which is also would be same as expectation of X given H for 2 sub sigma fields H is a sub sigma field of G.

d) the expectation of X given G, is expectation of X for X is for all random variable X, which is independent of G and E expectation of $\phi(X)$ given G is greater than or equal to ϕ of expectations of X given G for ϕ convex, and this is what is known as the Jensen's inequality as we have seen.

e)Now, we also see the expectation expectations of X given G as simply expectation of X which means this estimate if is given G is an on by estimate for quantity that we have there is expectation of X.

(Refer Slide Time: 61:33)



Now, let us continue or look at the example A that we have considered earlier and let my G in this particular case would be given by this, we know that we have $Y = \rho(\sigma_1/\sigma_2)X + W$

So, what we have is precisely expectation of Y given X this notation which we have already explained it is Y given the sigma fields generated by X is what we looking at it, so this is: $E(Y|x) =$

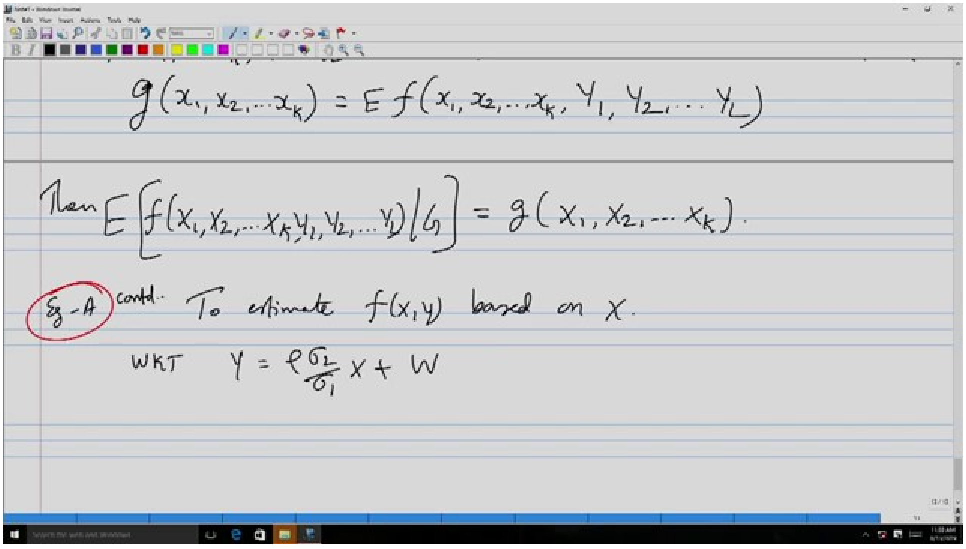
$\rho(\sigma_1/\sigma_2)X + E(W) = \rho(\sigma_1/\sigma_2)(X - \mu_1) + \mu_2$ First you have used linearity property, then X is measurable with respect to X and W is independent with respect to X, all this properties we will use, then you will end up with this.

So, this essentially means this one and this is what we call the linear regression equation and you have noticed that this quantity is expectation of Y given X is simply this quantity, which is, you know, sigma X measurable and satisfies the partial averaging property, you have to take any particular set and then look at weather this satisfies this and see.

So, this is what we have, now we also known the lemma in discrete time which is also we will re state here for convenience which is called the independence lemma which helps us to compute or evaluate the condition expectation of X given G in an much easier way, so what we have we just recall that you have omega f P be a probability space and G be a sub sigma field of $\mathcal{F}$, now we pick some collection of random variables say $X_1, X_2, ... X_K$ are G measurable and $Y_1, Y_2, ..., Y_L$ are independent of G.

Now let $f(x_1, ..., x_K, y_1, ..., y_L)$ be a function of $K+L$ variables and define this functions $g(x_1, x_2, ..., x_k)$ to be expectation of $E(f(x_1, x_2, ..., x_K, Y_1, Y_2, ... Y_L))$.

(Refer Slide Time: 66:44)



Then, what is the claim, expectation of

$$E(f(X_1, X_2, ..., X_K, Y_1, ..., Y_L)|G) = g(X_1, X_2, ..., X_K).$$

So, now, example A will continue (previously also this is example A continued so there is also we are now continuing) continuing this, so what we want to know, so we want to estimate what is our objective, so our objective is some function of X and Y we want to estimate which is based on the random variable X.

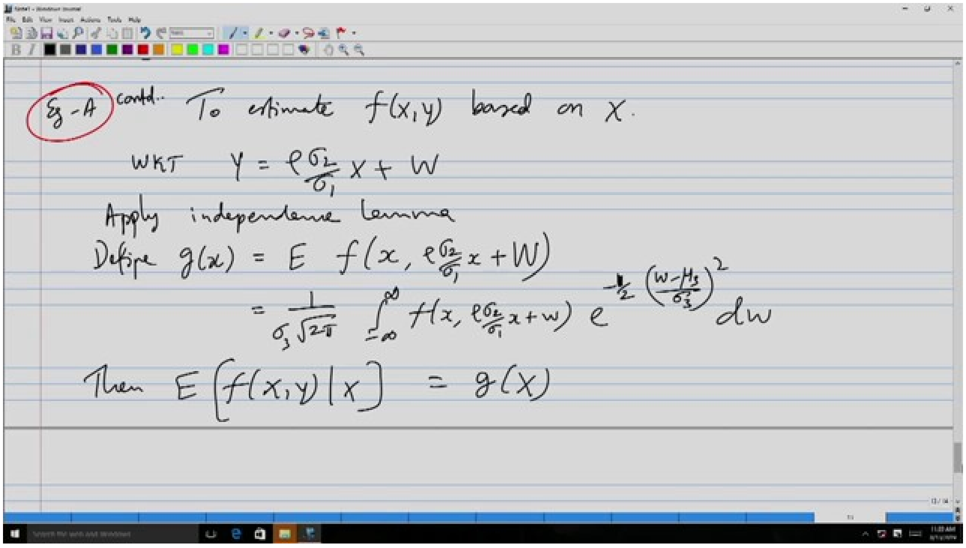
Now we cannot use independence lemma directly in the situation because based on X ofcourse X is measurable with respect to X or rather σ_X but Y is neither independent of X nor measurable with respect to X but then how do we estimate, but we know a way because this Y can be expressed as a linear combination of W and X where W and X have the required property.

So, we can use that idea to estimates this. So since we know that

$$Y = \rho(\sigma_2/\sigma_1)X + W.$$

where, now if I instruct this $f(X,Y)$ if I replaced this Y by this quantity then the random variable involved in the function or X and W where one of them is measurable with respect to g the other is independent of g than I can use my independence lemma to estimate or to compute this quantity.

(Refer Slide Time: 69:11)



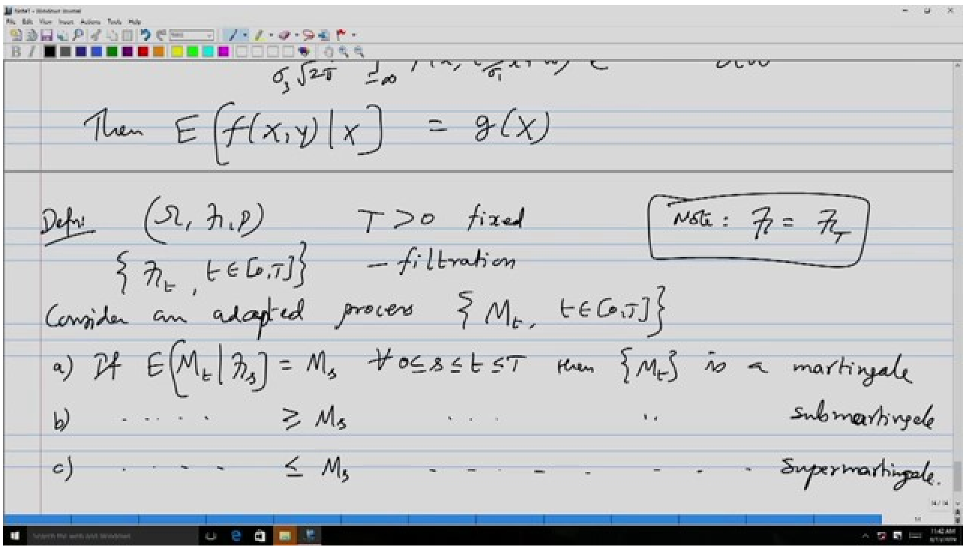
So, now we can apply the independence lemma, so define

$$g(x) = E[f(x, \rho(\sigma_2/\sigma_1)x + W)].$$

Now what is this expectation. It can be easily seen as above. So this is the function g so if I evaluate this quantity then I will get the function g and I can see that this one

$$E(f(X,Y)|X) = g(X).$$

(Refer Slide Time: 70:41)



Now, this so we have now all the ingredient to define the process that we might use, so what we have now, you have a probability space and T greater than 0 is fixed positive number and you also have a filtration connected with f and essentially note, so basically here you can note, this f is generally this f at capital T, so this is what we have. Now consider an adapted process we have already taken adapted process, say Mt again as we said, it is a function of t that is what we have.

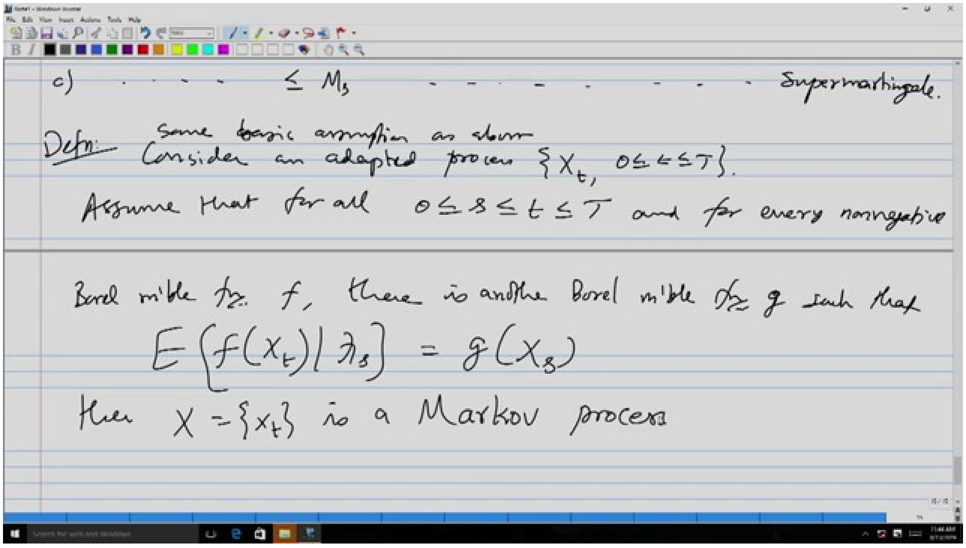
Now,

a) if expectation of Mt given fs is M_s for all $0 \leq s \leq t \leq T$, then this process M_t is a Martingale,

b) if you know this one is greater than or equal to M_s for all then this is a sub martingale and

c) this is less than or equal to M_s then this is a super Martingale, even from the discrete time notions you know what is a meaning of Martingales, which neither tend to increase nor fall the process the expected values, the average value, whereas in the cases of martingale it tend to increase and if super martingale case it tends to fall, so this the same definitions that we have seen it in case of discrete but only thing is not time is continued, so this is we have.

(Refer Slide Time: 73:35)

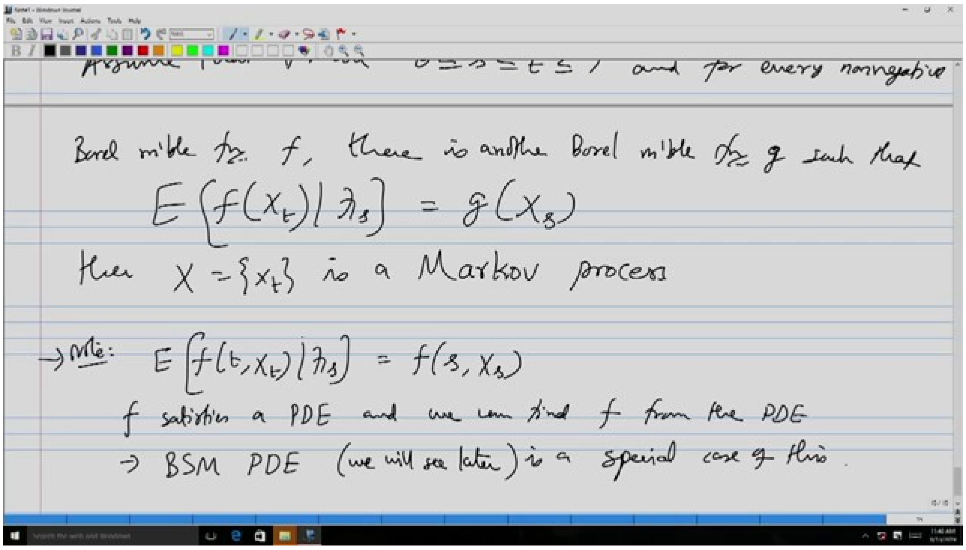


Now, the one is again, so you have this with the same basic assumptions about probability space time and filtration. Now, consider an adapted process, so now let us call this as the stochastic process with mean, so, $0 \le t \le T$. Now, what we will assume that for for all $0 \le s \le t \le T$ and for every non negative Borel measurable function f there is another Borel measurable function $g$such that expectation of

$$E(f(X_t)|\mathcal{F}_s) = g(X_s),$$

then this process X which is nothing but this collection is a Markov process.

(Refer Slide Time: 75:43)



So, note that, just like in the discrete time case this f may depend on t as well, so if we write in that case the above this not quantity then we do not need a g separately because the t dependent serial bring it inside that quantity itself, so you can write this the above as $f(t, X_t)$ given $\mathcal{F}_s$ because now the t dependency comes in, so you can write this in this form. Then f satisfies a partial differential equation, if this is true then we can write the above is this which is the Markov property, then this f will satisfy a

PDE and we can find f from the PDE by solving it the Black Scholes model PDE that we will encounter we will see later is a special case of this.

So, remember here, what we are dealing with conditional expectation when we are looking at this Martingales, Markov process and so on and their condition, so we need to find a g if it is know the Markov process then a function of Markov process given a filtration is given by another function of the Markov process at that point of time. Now, that satisfies this kind of relationship if we have then this satisfies PDE.

By solving the PDE you get the function or Markov, then hence, you have obtained the required solution to this expectation. So, we have tried to estimate this, which is again given by in terms of function f. Now if we find f we have done, right. So that is what we will use, so this much is what, then we have it now for the case of information, conditional expectation and other things. Next we will see in the next lecture, about the Brownian motion concept. Bye.

(Refer Slide Time: 78:02)