Mathematical Finance
Risk-Neutral Pricing in Discrete-Time
Lecture 22: Examples of Conditional Expectations, Martingales

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

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

Hello everyone, in the last lecture we have seen about the properties of conditional expectations. There are few minor points that we will want to look at that and then we go on to see examples, especially in the context of binomial model. We have seen that in some sense the conditional expectation is the best, in what sense?

(Refer Slide Time: 1:06)

Recap: Conditional Expectation

X on $(\Omega, \mathcal{F}, P)$ $\qquad \mathcal{G} \subseteq \mathcal{F}$

$E(X|\mathcal{G})$ is the best predictor for X given $\mathcal{G}$.

$\rightarrow$ $E(X|\mathcal{G}) = \underset{Z \,\in\, \mathcal{G}\text{-m'ble r.vs.}}{\operatorname{argmin}} E(X-Z)^2$

$\rightarrow$ If Y is a discrete RV, i.e. takes values $y_1, y_2, \ldots$

we let $B_i = (Y = y_i)$ $\qquad$ B_i's are disjoint & $\bigcup_i B_i = \Omega$

$\sigma(Y)$ $\qquad E[X|\sigma(Y)] = E(X|Y)$

We have seen given an X on some probable which is defined on some probability space and you have a sub-sigma field of $\mathcal{F}$, then $E(X|\mathcal{G})$ is the best predictor for the X given the information $\mathcal{G}$.. So, that is what we said. So this also let us to look at this conditional expectation in some other way as if you can also have a different look or different way to view this conditional expectation which, you could also define and we said that this conditional expectation we define one way, and we characterized this and then we said that we characterization is what we are going to take it as a definition later on, right.

That could also be taken as a definition and given that this is the best predictor in the sense that we have shown, which is the mean square error sense. We could also define $E(X|\mathcal{G})$ to be, you know. Suppose if I take this quantity X I have picked, $\mathcal{G}$ I have picked, now I pick some Z, which, is essentially the Z belonging to the class of $\mathcal{G}$ measurable random variables.

So, we could also define the conditional expectation as the minimizer of this quantity among all $\mathcal{G}$ measurable random variables. It is an alternative view which one can think about as this conditional expectation, but as you see like this requires the existence of second moment as far as X is concerned to talk about this part whereas we defined it even if you look at the characterization based upon the assumption that the first moment of X exists, okay.

This is alternative view in some books or in some material like you would find that this is defined as the minimizer of this particular quantity among all $\mathcal{G}$ measurable random variables. It is the $E(X|\mathcal{G})$. There is one more interpretation which of course, if you are comfortable with the linear space idea or linear vector space ideas then one can also look at the set of all random variables defined on $(\Omega, \mathcal{F}, P)$ is a vector space or a linear space and $\mathcal{G}$ is a subspace and the orthogonal projector of X on to the subspace $\mathcal{G}$ is what is the conditional expectation.

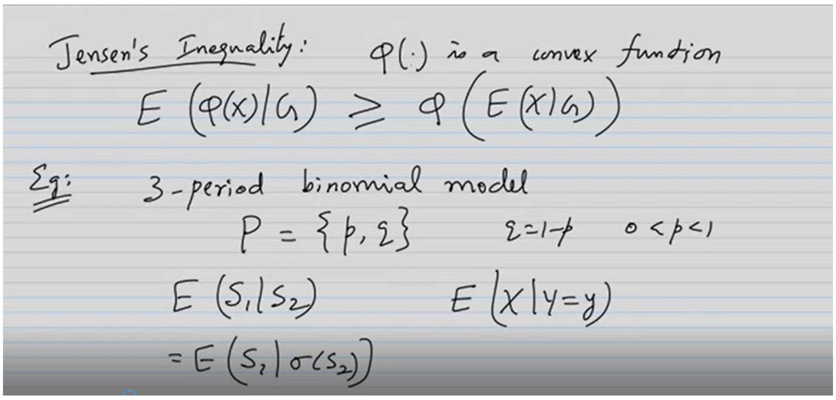
So, if you are comfortable with that view you can also explore that a bit, right. So, this is what is the conditional expectation is, right. Now we also can see that if you have a discrete random variable, say if Y is a discrete random variable that is it takes values say Y_1, Y_2, and so on then, we let this B_i to be the set Y_i. Now,if I take this B_i to be the set of all omegas that gives the value of random variables equals y_i. Now, I see that this B_j satisfies the properties that earlier we have described that this has positive probability, they are disjoint and their union is Ω.

Then B_i's are disjoint and they are union over i of B_i is Ω. Then if, then we know that we have already defined that, right. Then, if $\sigma(Y)$, if we pick which is nothing but the sigma of field generated by the random variable Y are equivalently the sigma of field generated by this classes of set B_i's. Then this is a sigma of field which we have already seen right and it is this smallest sigma field with respect to which this random variable Y is measurable right.

Then if you pick this as you know in our case, suppose if I pick my $\mathcal{G}$ to be this then we write the conditional expectation of X given this particular sigma field $\sigma(Y)$ which is essentially $E(X|Y)$. So whenever, we write in this form $E(X|Y)$, the conditional expectation given sigma field that is information content and Y the information content which is given not by this random variable Y and the sigma field, $\sigma(Y)$, here they are one and the same.

And whenever the sigma field with respect to the which we take the conditional expectation is generated by a random variable, we can simply denote by this $E(X|Y)$. But you should understand that when we write $E(X|Y)$, we essentially mean the sigma is generated by Y in this place.

(Refer Slide Time: 8:20)



And going forward, we will also state one important result, which is just the extension from the classical or mean usually probability theory ideas to the conditional expectation ideas, that is basically from the property of $E(X)$ and the similar property for conditional expectation which is also true.

If a function ϕ is convex function, then we will have the following inequality holds

$$E(\phi(X)|\mathcal{G}) \geq \phi(E(X|\mathcal{G}))$$

So this inequality will be true, this is what called as Jensen's inequality, which is also true with respect to ordinary expectation and this is also true with respect to conditional expectation. Of course, this we take without proof and the proof is similar to how you prove in the case of ordinary expectation. So we will not worry about that too much. So this is what one more property that you know we may recall, okay and we will be using this quite frequently.

Now let us take example, we will consider the 3-period binomial model for this example. Now since now, we are going to talk about expectations and so we need probabilities to be associated. So let the probability distribution that we consider with respect to the binomial model be denoted by P, which is essentially the distribution is given by these two elements p and q because it is binary.

We only need to assign probabilities for an up moment which is p and for a down moment, which is $q = 1 - p$. We also assume for the sake of clarity and expectations that this is strictly between 0 and 1.

Of course, if P is either 0 or 1 then you see that you know it collapse to a single (tree) and to keep that as a tree we assume this way and then so this is what is probability being assigned. So that, essentially mean, if we are corresponding to the coin dice experiment so each coin has probability of p of occurring head and q of occurring tails and the coin dices are independent so you get this as assignment for each and every node of this particular binomial tree.

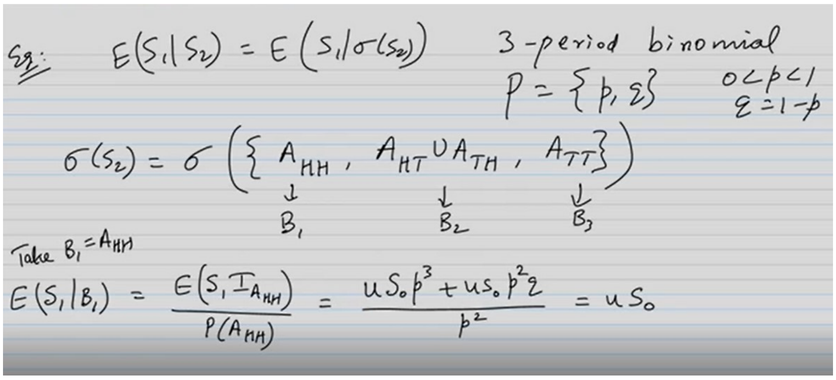
So we have this 3-period binomial model. Now let us look at the quantity that you know, which you can easily associate from your ordinary probability theory, you can easily translate to this particular way of looking at the conditional expectation. So let us take the simple example of $E(S_1|S_2)$. Go back to your basic probability theory where you would have seen quantities something like expectation of $E(X|Y=y)$ and you know how to compute this. Just refresh those ideas, right.

So, you have joint probability distributions of X and Y, mean you can assume or you can be comfortably within your discrete random variable case. There is a joint distribution of X given Y from which you get the model distribution of Y, and you compute the conditional distribution of X given $Y=y$ and take the expected value of that conditional distribution which will give you this particular value, that is what you have it, okay. So you can easily associate.

Now, when we view this as not for a particular value of this random variable Y but in a generic case then we will call this as $E(X|Y)$ and that is a random variable that also you would have seen right, that is precisely this. So, that is the way that you compute the conditional expectation in that case. Of course, that computation can still work.

You can see with this example and here we will compute this using the way how we have defined this particular case right. Recall, this in our sense is essentially $E(S_1|\sigma(S_2))$, right. Now what is this sigma field generated by.

(Refer Slide Time: 13:54)



Okay, so we have this 3-period binomial model with the probability measure P and we are trying to compute $E(S_1|S_2)$, which is essentially $E(S_1|\sigma(S_2))$. Now, as you know $\sigma(S_2)$ is the sigma field generated by the random variable S_2 and this has 3 sets, as atoms which is what these P_i's are. What are they?

So essentially speaking, this is Sigma field generated by the sets you know $A_{HH}, A_{HT} \cup A_{TH}$ and A_{TT}. So this is the sigma field generated by these 3 atoms and where, we know A_{HH} is the set of all omegas for which the first two elements are head-head which means only since this is a 3-period binomial model, only third element is different which is either H or T. So it has since this omega has 8 elements. So 2 elements are there in A_{HH} and 2 elements are there in A_{TT} and the remaining 4 elements are there in this.

So this suppose, if I call this as my B_1, this as my B_2, this as my B_3. Now you can take first on the set B_1. You know you recall the definition, either you can write it as $E(S_1|S_2)$, as $E(S_1|B_i)$. So, we can take one by one B_i and then we can try to compute that. So now first take say B_1which is A_{HH} and you look at $E(S_1|B_1)$ which is essentially $E(S_1|A_{HH})$.

This would be then $\frac{E(S_1 I_{A_{HH}})}{\mathsf{P}(A_{HH})}$. You can see on the 2 elements in this set A_{HH} are HHH, and HHT and in both my S_1 value is simply uS_0 and the probability of each of those elements is what would give me the numerator which in this particular case is $uS_0p^3 + uS_0p^2q$. The first one is corresponding to the three H and this is HH and T for that element and probability of A_{HH} is simply p^2 which is what would result in uS_0. So this is one way right, now if I have to use my partial averaging property alternatively, suppose if I have to use the partial averaging property.

(Refer Slide Time: 17:48)

PA property:

$$E\left[E(S_1|S_2)I_{A_{HH}}\right] = E(S_1 I_{A_{HH}})$$

$$= P(A_{HH})\, E(S_1|S_2)(\omega), \quad \omega \in A_{HH} \qquad = p^2 u S_0$$

$$= p^2\, E(S_1|S_2)(\omega), \quad \omega \in A_{HH}$$

$$\Rightarrow E(S_1|S_2)(\omega) = uS_0 \quad \forall \omega \in A_{HH}$$

$\rightarrow$ If $S_2 = u^2 S_0$ then $S_1 = uS_0$

lly on $B_3 = A_{TT}$, $E(S_1|S_2)(\omega) = dS_0 \quad \forall \omega \in A_{TT}$

Then partial averaging property because we are looking at, the same thing we are trying to derive through this partial averaging property. So what does that says? It simply says

$$E(E(S_1|S_2)I_{A_{HH}}) = E(S_1 I_{A_{HH}})$$

Now, if I consider the left hand side right, so this you know A_{HH} is an atom for the sigma field and on B_i's this conditional expectation is a constant. right, so you can write this as probability of, you can take this out essentially because this will be a constant on A_{HH} and this that $E(I_{A_{HH}})$ which is $\mathsf{P}(A_{HH}$ and $E(S_1|S_2)(\omega)$ for $\omega \in A_{HH}$ which is essentially $p^2E(S_1|S_2)(\omega)$ for $\omega \in A_{HH}$, this is the left hand side.

And on the right hand side, this quantity is simply equal to on this what is the expectation of S_1 on the set A_{HH}. So this is essentially you would easily see that this is $p62uS_0$. We have just shown it in the previous step. So this is. Now by equating these two from this implies my $E(S_1|S_2)(\omega) = uS_0$ for all $\omega \in A_{HH}$. So this is what you have obtained. But intuitively is if you see also, now when you see on the set A_{HH} the S_2, right, so the S_2 is simply u^2S_0, which means two heads are appearing and two up moments are there.

And S_1 given that there are when the value of S_2 is given to be, so equivalently if I have to say that my $\omega \in A_{HH}$, the equivalent way of saying is my $S_2 = u^2$. So equivalent way is that if my $S_2 = u^2S_0$, only on this set, this is the set A_{HH} all my ω which is in A_{HH} gives me $S_2 = u^2S_0$. So given this then my S_1 has to be, other than this nothing else can come and this is what normally you would have seen with respect to the simple probability.

Now, you can try to compute this right, given an S_2 of this, what is the S_1 of this? Even from the ordinary probability theory then you will get this value to be uS_0. So this is on the set A_{HH}. Similarly, on B_3 which is A_{TT} right, so you can also see that $E(S_1|S_2)(\omega) = dS_0\ \forall\ \omega \in A_{TT}$. So this is exactly same, similar way that carry out what you have done for the set A_{HH} or B_1 and B_3 for this also, this will be true.

(Refer Slide Time: 22:30)

Take $B_2 = A_{HT} \cup A_{TH}$

$$E(S_1|B_2) = \frac{E(S_1 I_{B_2})}{P(B_2)} = \frac{E(S_1 I_{A_{HT}}) + E(S_1 I_{A_{TH}})}{P(A_{HT} \cup A_{TH})}$$

$$= \frac{uS_0 p^2 q + uS_0 pq^2 + dS_0 p^2 q + dS_0 pq^2}{p^2 q + pq^2 + p^2 q + pq^2}$$

$$= \frac{uS_0 pq + dS_0 pq}{2pq} = \frac{1}{2}(u+d)S_0$$

$$E(S_1|S_2)(\omega) = \frac{1}{2}(u+d)S_0 \quad \forall \omega \in A_{HT} \cup A_{TH}$$

Now, take the other set which is $B_2 = A_{HT} \cup A_{TH}$. Now, in this case you can compute $E(S_1|B_2) = \frac{E(S_1 I_{B_2})}{\mathsf{P}(B_2)} = \frac{E(S_1 I_{A_{HT}}) + E(S_1 I_{A_{TH}})}{\mathsf{P}(A_{HT} \cup A_{TH})}$. Now, this if I do the calculation with respect to the first expectation on the set A_{HT}, the $E(S_1 I_{A_{HT}}) = uS_0 p^2 q$ corresponding to HTH, and $uS_0 pq^2$ corresponding to HTT. Similarly for the other expectation on A_{TH}, we get $dS_0 p^2 q$ for THH and $dS_0 pq^2$ for THT.

Now, $\mathsf{P}(A_{HT} \cup A_{TH}) = p^2 q + pq^2 + p^2 q + pq^2$. You can easily see that the final result is $1/2(u+d)S_0$. Thus, you can use the partial averaging to arrive at the same. So thus, you have $E(S_1|S_2)(\omega) = 1/2(u+d)S_0 \ \forall\ \omega \in A_{HT} \cup A_{TH}$. So this is what you compute, right.

So now can, you have pretty much computed the conditional expectation of S_1 given S_2 right and you would see that, so this is, this conditional expectation of S_1 given S_2 is random only through the deponents of S_2 and that is why S_1 given S_2 will be, right.

(Refer Slide Time: 25:51)

$$E(S_1 | S_2 = x) = g(x)$$

where $g(x) = \begin{cases} uS_0 & \text{if } x = u^2 S_0 \\ \frac{1}{2}(u+d)S_0 & \text{if } x = udS_0 \\ dS_0 & \text{if } x = d^2 S_0 \end{cases}$

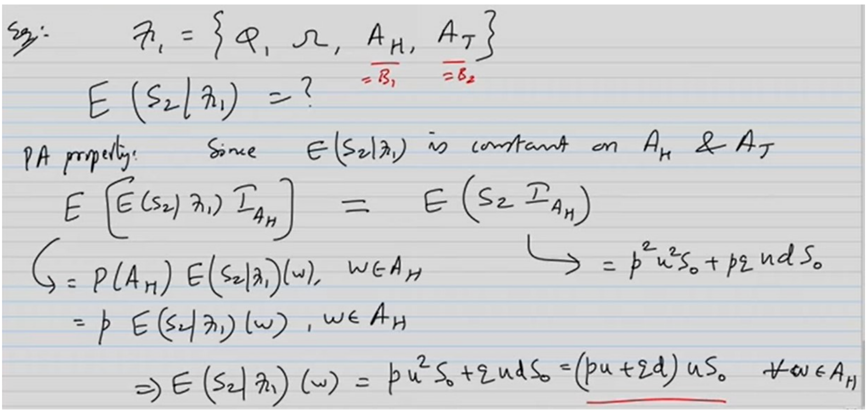
$E(S_1|S_2) = g(S_2)$ where g is as above.

We can also write this just to connect to what you might have been already comfortable. $g(x) = uS_0$ when $x = u^2 S_0$, $= 1/2(u+d)S_0$ when $x = udS_0$, and $= dS_0$ when $x = d^2 S_0$. And now, this form you would have seen in the ordinary probability theory, right. So this quantity is what? So $E(S_1|S_2)$ is nothing but $g(S_2)$ where g is as above.

So the function g, so this is what is you are computing. So essentially in ordinary probability theory for the fixed value of S_2 that is the first step that you normally do is you compute what is this. Now you vary this x for all possible values, so you get a function g, now, make that function as a function of the random quantity because this x is essentially $u^2 S_0$ or udS_0, or $d^2 S_0$ is what the values of S_2. So make that as the random. So $g(S_2)$ is what this function. So this function is what then you are computed as a function of S_2 is what this $E(S_1|S_2)$ that is what we say, this is a random only through the dependence on S_2.

So, this example then you can easily correlate or related to what you might have studied in your ordinary probability theory where $E(X)$ given under the random variable particular value. You can easily relate and then you can come to see what we have doing it here.

(Refer Slide Time: 28:05)



Next, we will see another example, again in the binomial model and that is going to we take a bit forward a little bit, okay. So let us take, so instead of taking expectation of in the binomial model itself one random variable given under the random variable. Now, where is our definition? If it is exactly the same then why do we need this different setup. No, because we need the set up because we are talking about random variable, expectation of a random variable given a general, a generic sigma field which might not be or may not have been generated by random variable.

So, we need we are just generalizing in that sense, okay. So, recall in our models so as again you can assume a 3-period model or for this particular example, even you can go beyond does not matter because there were notation takes care of that. Suppose, if I take my $\mathcal{F}_1$ this is the sigma of field that you have it here, okay. Now what we are interested? We are interested on $E(S_2|\mathcal{F}_1)$, what is this quantity? Is what we are interested, okay.

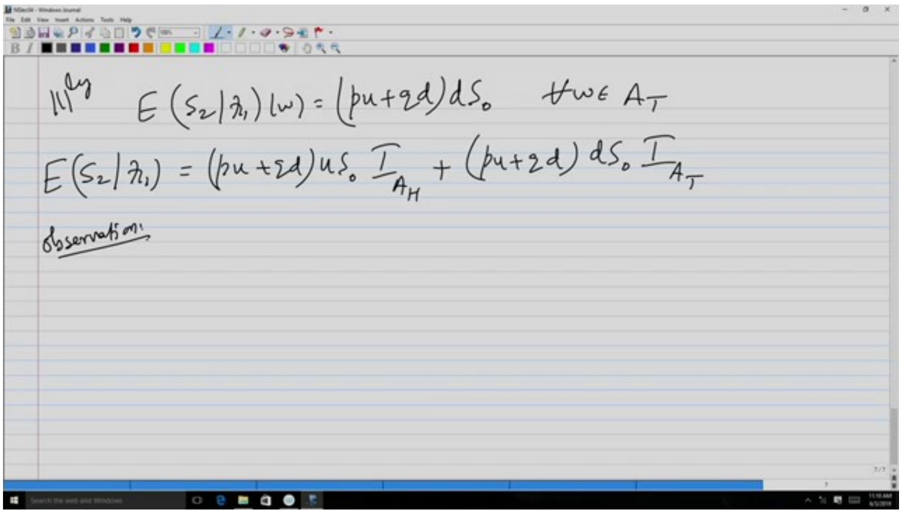
And F_1 is this quantity where these are my B_1 and B_2, right. So as far as this particular sigma field is concerned there are two B_1, B_2, and which is what forms the atoms for the sigma field and we are interested in $E(S_2|\mathcal{F}_1)$. Again in the binomial model you can for understanding purpose one-two step binomial model, for this particular step is sufficient but even if you generalize, generally take any n step binomial model, this whole notation and everything goes through.

Okay, now what we know? We know that this particular quantity we have to compute and this particular quantity is constant on each of this B_1 and B_2. That is what we know. And we are going to use that property. So by the partial averaging property what we are saying property since, this $E(S_2|\mathcal{F}_1)$ is constant on A_H and A_T. So we will take first $E[E(S_2|\mathcal{F}_1)I_{A_H}]$, so we see that this must be equal $E(S_2 I_{A_H})$. This is the partial averaging property on the set, A_H. And on this set, this $E(S_2|\mathcal{F}_1)$ is constant.

So again the left-hand side would imply that this is $\mathsf{P}(A_H)E(S_2|\mathcal{F}_1)(\omega)$, for $\omega \in A_H$. This is $pE(S_2|\mathcal{F}_1)(\omega)$, for $\omega \in A_H$. Similarly in the right-hand side what you will have. This is S_2 on A_H, right. So, S_2 on A_H means, we talked A_H is only the first coin toss is head, second coin toss could be anything, right, and given that the first coin toss is head, second coin toss is head or tail, the two possibilities.

So if I have to use that idea then this will be u_0^S with probability p^2 and udS_0 with probability pq. If I equate these two then $E(S_1|\mathcal{F}_1)(\omega) = (pu+qd)uS_0$ So, this is for all $\omega \in A_H$.

(Refer Slide Time: 34:06)



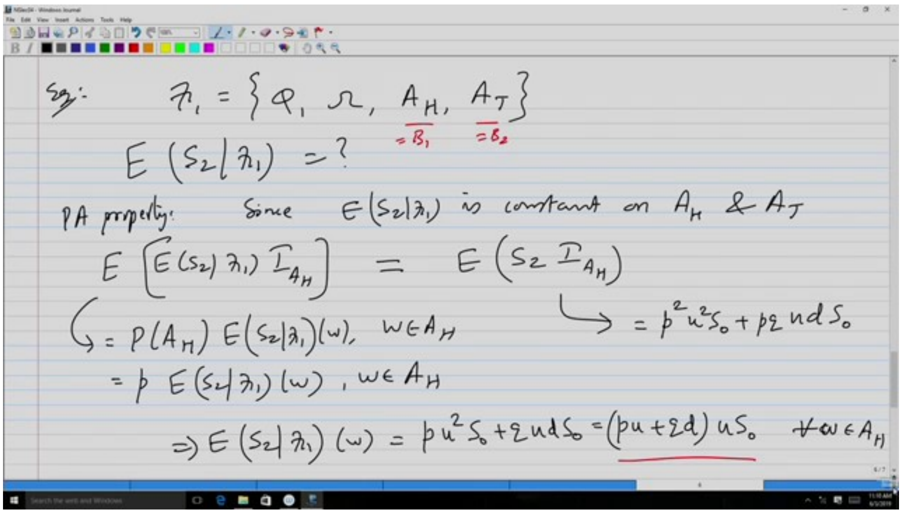
Similarly, you need to show this:

$$E(S_2|\mathcal{F}_1)(\omega) = (pu+qd)S_0 \;\forall\; \omega \in A_T$$

So this is, what is the conditional expectation, so I mean if I have to write in our original form, so I can write this $E(S_2|\mathcal{F}_1) = (pu+qd)uS_0 I_{A_H} + (pu+qd)dS_0 I_{A_T}$. So, this is the expression for this conditional expectations that you have here.

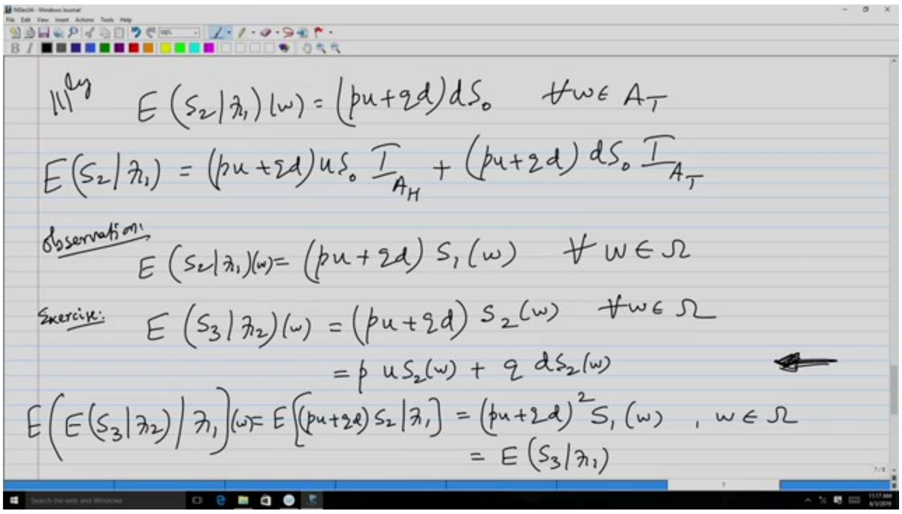
Now, certain observations are in order because that is what we are going to use here. You can see that in this $E(S_2|\mathcal{F}_1)(\omega)$. For example for $\omega \in A_T$, we wrote $(pq+qd)$ as dS_0. This dS_0 is nothing but my S_1 at the down node right. Similarly if we look at the previous quantity.

(Refer Slide Time: 35:50)



This uS_0. it is the value of S_1 at the up node, right.

(Refer Slide Time: 35:56)



So combinedly, you can write both these as simply $E(S_2|\mathcal{F}_1) = (pu+qd)S_1(\omega)$ $\forall \omega \in \Omega$. So, this is the simple expression, right. So this tells something more that the $\mathcal{F}_1$, which is what we are trying to do. We are trying to estimate, remember conditional expectation as the best estimator. So among all $\mathcal{F}_1$ measurable random variables, we are looking for the best estimator for S_2 which is what this $E(S_2|\mathcal{F}_1)$ means: given all the information $\mathcal{F}_1$, when we are trying to estimate S_2, right, is in a way, the all information that is required are summarized in the value of S_1 that is what this pretty much says.

So, the complete information is $\mathcal{F}_1$ and here we say, this estimation requires only the value of S_1, not the complete information. What is relevant from the complete information is the value of S_1. So, this is what this says. Now similarly, you can extend the same, you can do exact same steps that we have done so far to compute $E(S_2|\mathcal{F}_1)$.

So this is an exercise, you can do a similar thing and you can arrive at $E(S_3|\mathcal{F}_2)$. Again, this is a random variable, so I can put an omega here. So $E(S_3|\mathcal{F}_2)$ you can also show that this is nothing but $S_2(\omega)$ for all ω. See extending the same idea of $E(S_2|\mathcal{F}_1)$ to $E(S_3|\mathcal{F}_2)$, you will see that exactly the same expression comes but S_1 be replaced by S_2, okay.

You can also see that this equation can slightly be written in a slightly different way also, that this could be written as sum p times okay, I mean, this is first thing is what you observe is, in a similar way, you can uptime this. Now but either this or this can also be written as suppose this is $\mathcal{F}_2$, so now if I write explicitly suppose if I assume that it is some 3-period model or so then I can write this as $puS_2(\omega) + qdS_2(\omega)$, right. What is this u S_2?

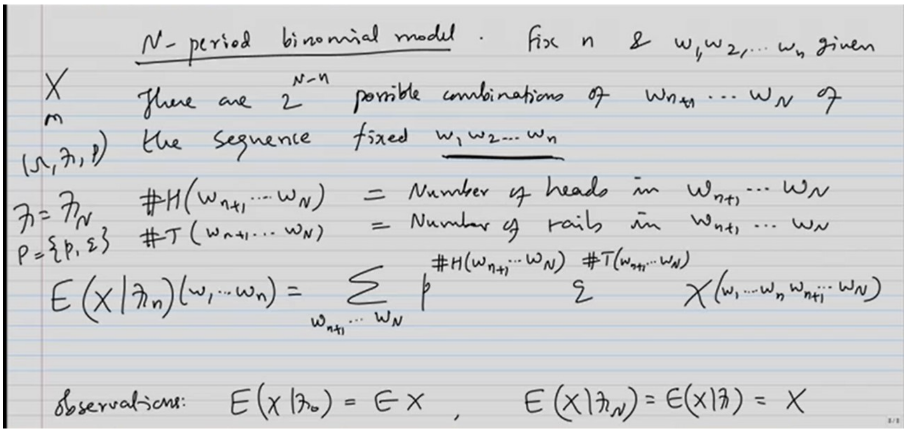
It is basically S_3 if I look at it and with omega up ended with because this omega when we write suppose if I assume if I have 2 period or 3 period model simpler case so this will be up to 2 then the third element of this omega would be head and here the third element of this omega is tail right in the second component that we are write. So you can see that you can also write it in this form.

Now further, you can also notice, suppose if, I take this quantity right $E(S_3|\mathcal{F}_2)$. I further take $E[E(S_3|\mathcal{F}_2)|\mathcal{F}_1]$. Now, I will get $E[(pu+qd)S2|\mathcal{F}_1]$, now $pu+qd$ is constant so I can take out and then I see $(pu+qd)E(S_2|\mathcal{F}_1) =$, so in net result is my $(pu+qd)^2 S_1(\omega)$ for $\omega \in \Omega$ and you can see that by the tower property, this is nothing but $E(S_3|\mathcal{F}_1)$.

You can see that this particular quantity, the left-hand side of this equation by the tower property if I apply it is nothing but the $E(S_3|\mathcal{F}_2)$. So what we have written here, is simply S_3 given $\mathcal{F}_1$ which has this property, right. So this you can observe and this part is what you can observe with respect to the binomial model that you what you have as conditional expectation and this definition can be generalized and written as in the textbook that we have we are using which is series volume 1 if you look at the conditional expectation.

Definition in the context of N period binomial model is nothing but is the generalization of this where S_3 and $\mathcal{F}_2$. Now you take any N period general model and a random variable x which is defined on the full probability space and given any fn intermediate then what is the step that you have is what is precisely that quantity that you have it here, right. So that is a general definition and if for the sake of completeness you want we can also write down that definition which is the general.

(Refer Slide Time: 42:52)



Suppose, if I take in general N-period binomial model and if I fix, fix an n and if I pick ω_1, ω_2, and so on ω_n as given and this both these for the moment given and fixed. Suppose if I assume then there are 2^{N-n} possible combinations of ω_{n+1} and so on of ω_N that you have here. So, this is as a sequence are is you know if that if you look at it. So, this is what you will have it here. From the sequence, this is basically a fixed value of ω_1, ω_n. So any fixer value up to ω_n that you fix this ω_1 to ω_n either head and tail take a particular this particular things you keep it fixed then the remaining ones you will have these many possibilities.

Now, suppose if I denote the number of head of $\omega_{n+1}, \ldots, \omega_N$ be the is equal to the number of heads or the number of up moments in this sequence and similarly number of tail in ω_{n+1}. This is a notation that if we use number of tails in $\omega_{n+1}, \ldots, \omega_N$ and if you look at, if look at go back and look at my $\mathcal{F}_1$ and $\mathcal{F}_2$ and so on that you would what have seen. So for a fixer value of omega, for each fixed one the variations will give the atoms. So the atoms will be as per the variation of this 2^n elements that you know $\omega_1, \ldots, \omega_n$ that you may have here right. So these are variation that you have it here.

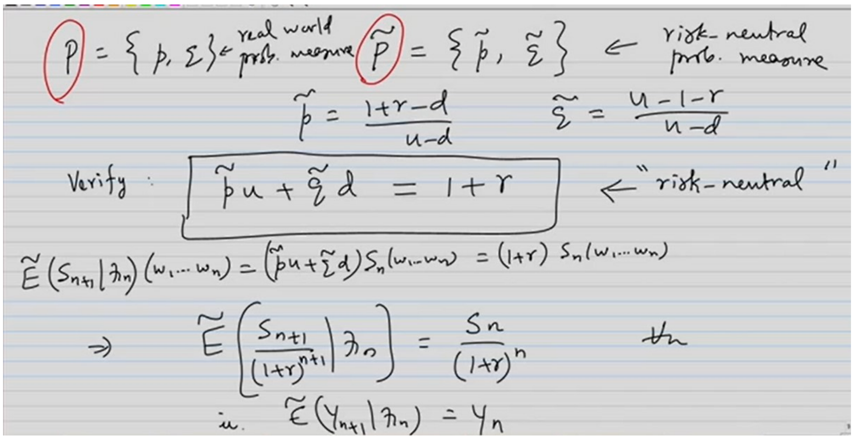
Then, if I pick you know a random variable X on my full space omega where my $\mathcal{F}$ is same as my $\mathcal{F}_N$, right. This is the full sigma field that you have, complete information case. Then this particular quantity and if I also assume my probability to be p and q then the $E(X|\mathcal{F}_n)$ right. so since this is $\mathcal{F}_n$ is what is our sigma field. So appropriately you know we denote only $\omega_1, \ldots, \omega_n$ because this is what it would depend on. This could then be written if I use the previous ideas as a sum over probability of number of heads in ω_{n+1} to ω_N and q number of tails in ω_{n+1} to ω_N times $X(\omega_1, ..., \omega_n, ..., \omega_N)$ where the sum is over all ω_{n+1} to ω_N.

So, for all variations, so you can just whatever we have computed earlier $E(X|\mathcal{F}_1)$ or $E(X|\mathcal{F}_2)$ that we have seen. So, you can easily see that it falls into this. So, this is what is the definition. If you want to specifically write the conditional expectation in the context of the N-period binomial model for use then one can directly use this definition. This is what is what will turn out to be the quantity if you apply the usual process, right.

But what we have defined so far is any general sigma field $\mathcal{G}$ that is applicable here, right. Now one couple of simple observations you can make which is trivial $E(X|\mathcal{F}_0)$ as you know is simply $E(X)$, and $E(X|\mathcal{F}_n)$ and if it is same as $E(X|\mathcal{F})$ because we have taken $\mathcal{F}_n$ to be $\mathcal{F}$ which is usually the case and since X is defined on $(\Omega, \mathcal{F}, \mathsf{P})$, so this is X is measurable with respect to $\mathcal{F}$.

So you can see that this too follows easily and this is small property that you can look into this case, okay, so this what is all the examples of this conditional probabilities that you know that you might think and this definition is basically with respect to N-period binomial model, the definition of conditional expectation you can write in this form, okay. This is just the additional but even without the even to the basic definition will work, but this is a simplified definition which you can directly use it in the case of binomial model that you have it here okay. Next, we look at the case of certain observation that they have made of course we will come back to that.

(Refer Slide Time: 49:41)



So, you recall in the binomial model case you had defined not just P but you also used another probability distribution $\tilde{P} = \{\tilde{p}, \tilde{q}\}$, where $\tilde{p} = \frac{1+r-d}{u-d}$ and $\tilde{q} = \frac{u-1-r}{u-d}$.

This P is what we call the real-world probability measure and we call this $\tilde{P}$ as risk-neutral probability measure, right. Now, it is easy to observe that

$$\tilde{p}u + \tilde{q}d = 1 + r$$

So this is, so what this says you know just look at this and what are the u and d? These are the up factors and down factors in the binomial spacing model and the expected value of this factors, up factor and down factor suppose if I call this is u and d. I can always represent it as S_{n+1}/S_n. So which is what this up factor and down factor would mean , right.

So the left-hand side is nothing but the expected increase in the underlying risky asset like underlying stock suppose if I call this S as a stock price process then this is nothing but the expected increase in the stock prices is what is given by the $\tilde{p}, \tilde{q}$. $\tilde{p}u + \tilde{q}d$ under the risk-neutral measure, right and on the right-hand side is basically you can think of it as the increase in the risk-free asset. So, if you have 1 rupee invested in risk-free will grow to $(1+r)$. If we invest 1 rupee in the stock then it will grow on an average to the quantity which is given by the left side. Now for this particular probability measure $\tilde{P}$, this quantity is equal to this and hence the name that from where this name the word risk-neutral came. You could now understand that you know it neutralizes the additional risk in a way by making the expected increase equal to the risk-free. So under this measure, both these assets are on an average is going to give the same growth, right. So that is what is the risk-neutral word that it comes here.

Now, where does this helps? Recall we have written under this measure suppose if I take under $\tilde{p}$ earlier we have written the $E(S_2|\mathcal{F}_1)$, $E(S_3\mathcal{F}_2)$ and so on. In general, you can write $E(S_{n+1}|\mathcal{F}_n) = (pu+qd)S_n$. Now here you could write under $\tilde{P}$, $\tilde{E}(S_{n+1}|\mathcal{F}_n)$. Now this will be a function of ω_1 to ω_n. It will be equal to $(\tilde{p}u + \tilde{q}d)S_n(\omega_1, \omega_2, ..., \omega_n)$.

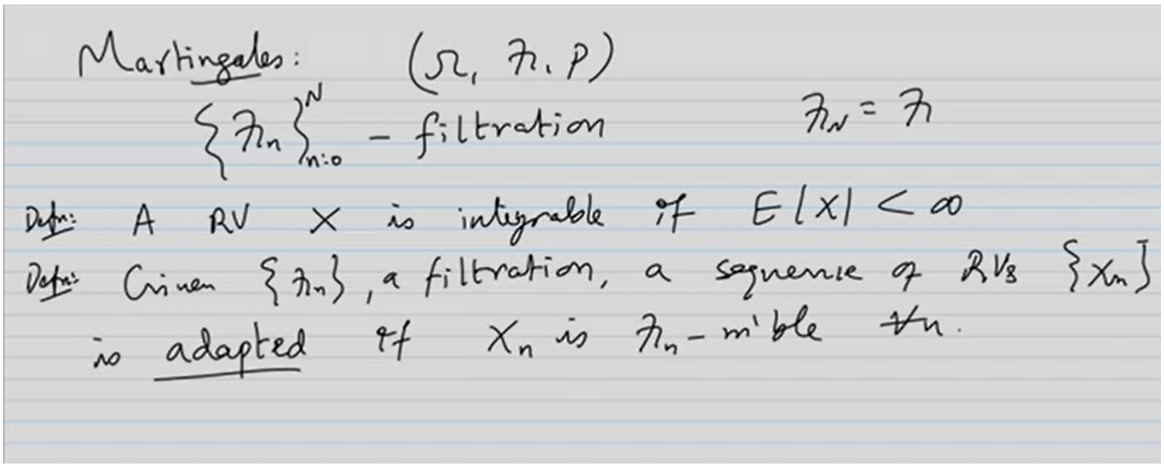
Now under this measure remember I have used an $\tilde{E}$ here the tilde means that it is a expectation under probability measure $\tilde{P}$. So there are two probability measures here, one is P the other is $\tilde{P}$, right. So this is under $\tilde{P}$, the risk-neutral measure. So this will be $(1+r)S_n(\omega_1, ..., \omega_n)$. Otherwise, this gives rise to this property: $\tilde{E}\left(\frac{S_n}{(1+r)^n}|\mathcal{F}_n\right) = \frac{S_n}{(1+r)^n}$.

So now, if I look at this so what this says is that, so what we are looking at? We are looking at the best predictor of this particular random variable given some information, that is the meaning of this conditional expectation. Now what this says is that, the conditional expectation of, the conditional expectation under the $\tilde{P}$ measure of $\frac{S_n}{(1+r)^n}$ given the information up to the time n is nothing but the present value of this. Suppose if I call this process as some Y_{n+1} if I call. so $E(Y_{n+1}|\mathcal{F}_1) = Y_n$.

So under this measure or under any measure for that matter but in this particular case it so happened for the stock price process this property. This, if it happens, then this property is what we call as the

Martingale property. So when we define next, so we have to going to define is this martingale's and so you see here that it has this, so it has a name that property of this type, right, that we have here is essentially what we call the Martingale property. So you see here certain property and you are now observing that this is satisfying this. Now what is this called? This is what is called the martingale property.

(Refer Slide Time: 57:03)

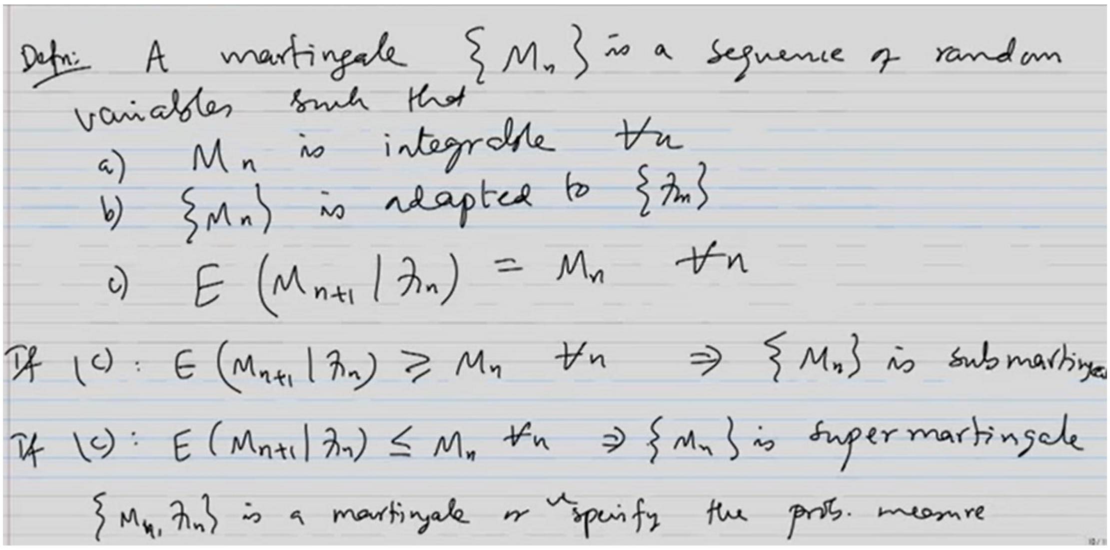


So, now let us define what is Martingale? Now to define Martingales, we need to first assume as usual there is always an underlying probability space $(\Omega, \mathcal{F}, \mathsf{P})$ okay and there is filtration. Now let us assume that this n runs from 0 to N fro some discrete filtration. If I assume finitely so it is some $\mathcal{F}_N$ would be this $\mathcal{F}$. So there is a filtration and we have random variable defined on this probability space, okay.

So we take a sequence of random variables, right. This filtration is generally we are saying this is need not be the complete filtration that we talked about in the case of binomial model. So it could be any filtration, right which leads to this and this the Fsub capital N is what is my F so that is the full probability space and the random variables are defined on that probability space, okay.

Now we need a few definitions here, one which is we say that a particular random variable X we say it is integrable if my expectation of this is finite. Now another definition is given a filtration F_n, a filtration, we say that a sequence of random variables ,say X_n, is adapted, (this word is adapted) when we say if X_n is $\mathcal{F}_n$ measurable for all n. So there is a particular terminology that we will use, so we say an adapted process. So this sequence of random variables is what is generally called as more generally for more generic case is called a random process or a stochastic process and this process is adapted to this particular filtration $\mathcal{F}_n$ if X_n is $\mathcal{F}_n$ measurable for each n.

(Refer Slide Time: 59:45)



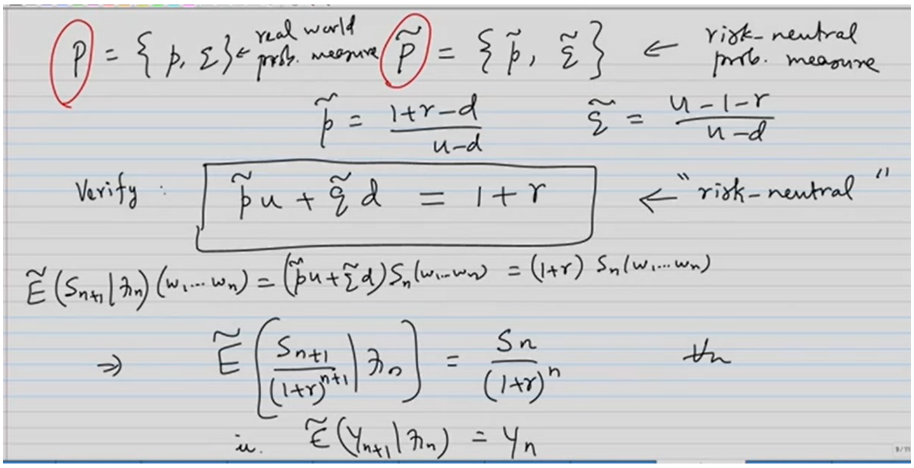
Now once I have these two definitions, now we can define Martingale. So a martingale, what is Martingale? Suppose if I call this as some $\{M_n\}$ is (we are talking about now discrete time martingale) sequence of random variables such that
(a) M_n is integrable for all n,
(b) this process M_n adopted to this sequence $\mathcal{F}_n$ and
(c) $E(M_{n+1}|\mathcal{F}_n) = M_n$, for all n.
The first is the integrability property which ensures that this conditional expectation exists and then M_n is adopted property which shows that this M_n will be measurable with respect to F_n and the third is Martingale property.

So the third is what is the crucial property which is called is martingale property (okay). Now this martingale, this is what is called the martingale and if the property (c), if you have in this form rather than equality here, if this is greater than or equal to M_n for all n then we say that M_n is sub martingale and if instead this property (c), we have $E(M_{n+1}|\mathcal{F}_n) \leq M_n$ for all n, then this $\{M_n\}$ is called super martingale, okay.

So this is what the martingale, sub martingale, and super martingale. So what it says if you look at it in a martingale setup, the best predictor of next random variable in the sequence given the information up to the current which is Fn is the current value itself whereas in the case of sub martingale and super martingale, you are saying on the one side which means that it is at least the current value of and this is utmost the current value of the process, okay.

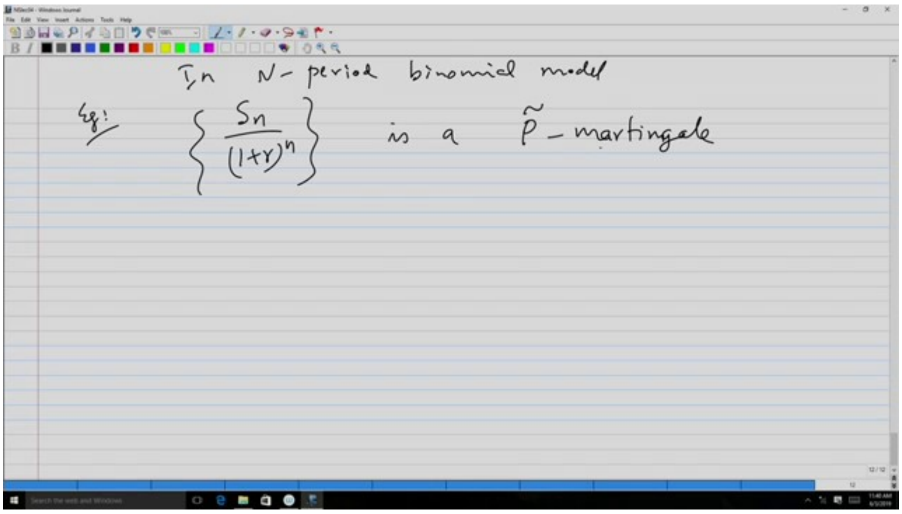
So this is what is the and here if you see depending upon the convenience we may simply say $\{M_n\}$ is a martingale whenever the underlying $\mathcal{F}_n$ and the probability measure is understood or sometimes we say $\{M_n, \mathcal{F}_n\}$ is a martingale or you know if or sometimes we specify the probability measure also because as we saw just in the previous example right.

(Refer Slide Time: 63:31)



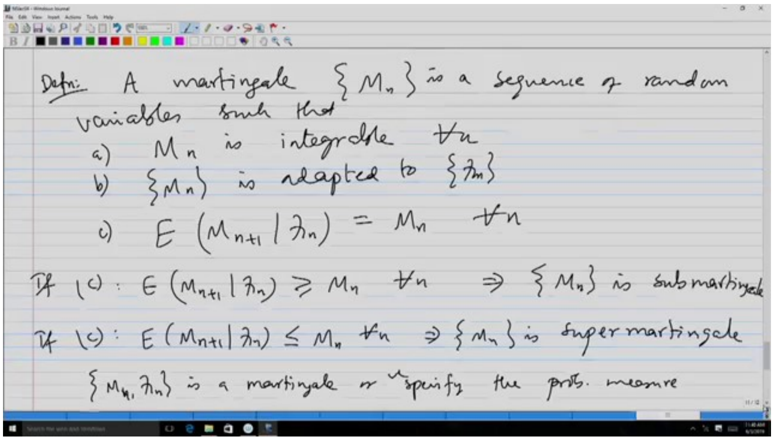
If you go back to this particular example, you see under $P = p,q$ if you compute, under P this is not going to be true. Here, $(pu+qd)$ is not in general is going to be equal to $(1+r)$, and hence under the measure P this equality that we have written here need not be true and hence under P this need not be martingale but under $\tilde{P}$, since this is going to be equal to $(1+r)$ and hence this particular quantity is going to be $\frac{S_n}{(1+r)^n}$.

(Refer Slide Time: 64:07)



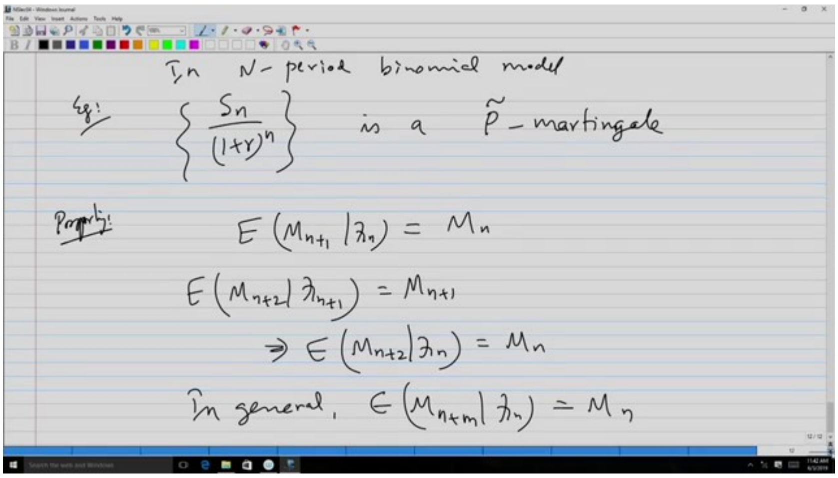
So, we will say this particular example if we have to say so this $\{\frac{S_n}{(1+r)^n}\}$, now say in N period binomial model what you are being observing is that this quantity is $\tilde{P}$ martingale or a martingale under that probability measure right. So that is what we will say. So, this is one example so which we have seen in this and we have define now what is a martingale. It is nothing but a stochastic process with certain properties, right. In the stochastic we classify it or we give names to various stochastic processes depending upon the properties that it has.

(Refer Slide Time: 65:01)



And if in particular if it has this property then we say this is the martingale or if it has either of these two properties then we say accordingly either sub or super martingales that we have, right.

(Refer Slide Time: 65:12)

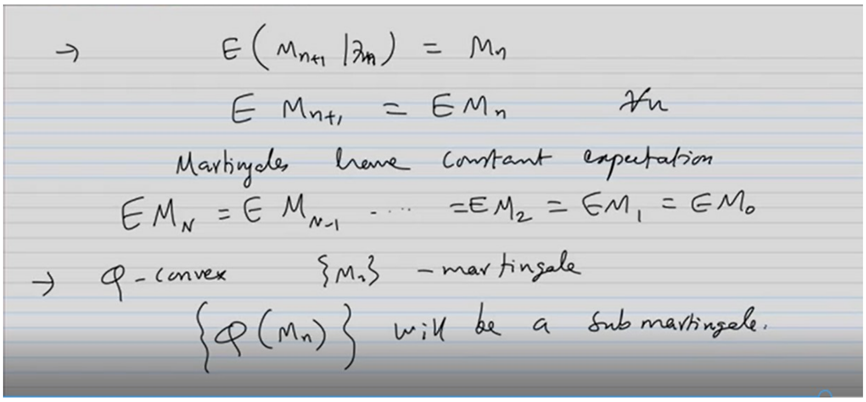


So this is what it is and one example in the case of N-period binomial model is this is a $\tilde{P}$ martingale which is not in general, it may not be P martingale or martingale under the real-world probability measure. It will be a martingale under $\tilde{P}$, the risk-free probability measure and this is how this probability measure you know is constructed so that you know this becomes a martingale that you have it here, right. Now, this there is one step property that we have seen a couple of properties that we see quickly before we close which is the one step property we know. Quickly we will close this with a couple of properties that we can observe so that we no need to repeat it again.

So this is basically what we have observed is $E(M_{n+1}|\mathcal{F}_n) = M_n$. This is true for any n and hence $E(M_{n+2}|\mathcal{F}_{n+1}) = M_{n+1}$. Now, take the conditional expectation of this quantity with respect to $\mathcal{F}_n$ further, so you would see that $E(M_{n+2}|\mathcal{F}_n) = M_n$. So what we are saying is $n+2$, n to $n+1$, this is one step property $n+2$. So in general $E(M_{n+m}|\mathcal{F}_n) = M_n$ which is what we call it as the multistep property that you have it here, okay.

So this is true across any step given $\mathcal{F}_n$, if you want to predict the value of the process M, M step ahead then the best predictor is still the current one, is what this property one would observe.

(Refer Slide Time: 67:23)

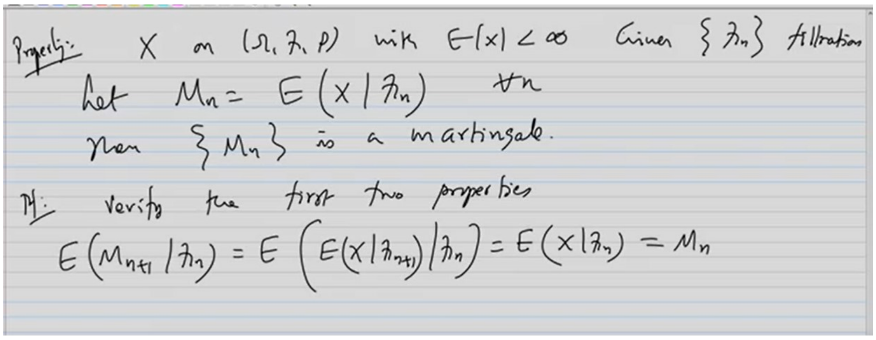


The other thing is that if I take again, the other property if I look at it $E(M_{n+1}|\mathcal{F}_n) = M_n = M_n$. Now I take expected value on both sides what I will get is $E(M_{n+1}) = E(M_n)$, this is for all n. So what this means is that martingales have constant expectations overtime period, right.

So which means $E(M_n) = E(M_{n-1})$ Thus, $E(M_2) = E(M_1)$ which is equal to $E(M_0)$, and if M_0 is not random then this will be equal to simple some constant value. So this will be true this is what you know you will have.

The third property that you can observe from Jensen's inequality is that, if I take a convex and if I take martingale, if I take a martingale then you can easily observe that this process $\{\phi(M_n)\}$ be what. if I apply for a couple of convex function and for a martingale if I insert and I can look at this. This will be a sub martingale, right.

(Refer Slide Time: 69:12)



So this is one other property which we can easily see is which is what will be very helpful for us when we go ahead is that suppose you fix X is a random variable on $(\Omega, \mathcal{F}, \mathrm{P})$ with $E(X) < \infty$, let you define $E(X|\mathcal{F}_n)$, this is what you define for all given any and so and also you are given a filtration $\mathcal{F}_n$ then this you define. Then what is the claim: $\{M_n\}$ is a martingale.

So you can verify the first two properties again you will use the Jensen's inequality to verify the integrability property and since this definition itself X given $\mathcal{F}_n$. So measurability is immediate. For the integrability property, you will use Jensen's inequality. We can look at the third property which is $E(M_{n+1}|\mathcal{F}_n) = M_n$. You can look at this is $E(M_{n+1}|\mathcal{F}_n) = E(E(X|\mathcal{F}_{n+1})|\mathcal{F}_n) = E(X|\mathcal{F}_n) = M_n$.

So this is the martingale property and hence the result follows and this is an important example. In the set, you can also take an example and you can see how one can construct the martingale and most of the martingale that you basically encounter are of this nature. So they are actually given as conditional expectation of some random variable given a filtration, a sigma field in the filtration would define a martingale. So most of the martingale that you encounter can be represented in for some random variable X and for some sigma field in this form. So that is why this thing. So whenever you encounter any such random variable basically what you are looking at it as $\mathcal{F}_n$ increases, $\mathcal{F}_n$ is a filtration so it the information accumulates, so there is an increase in sub sigma fields of $\mathcal{F}$ and which means if you will think about this information so as you get more and more information, you are trying to estimate the

X and those estimates for X given more and more, more and more information that you will add will form a martingale right. So that is what this pretty much says okay. So we will see this with respect to binomial model and other things later. We will see in the next lecture. Bye.