Lecture - 06
Conditional Expectation-I

Welcome to the first lecture of the second week of this course. Today we are going to discuss something called conditional expectation. Conditional expectation is the first most important thing, that I need to discuss with you. The remaining part, that I was discussing with you is something I was, expecting that you had almost some idea about. But here we start with something very new usually not done in standard courses of probability. Standard courses mean courses which engineers usually take or economists take, etc.

(Refer Slide Time: 01:13)
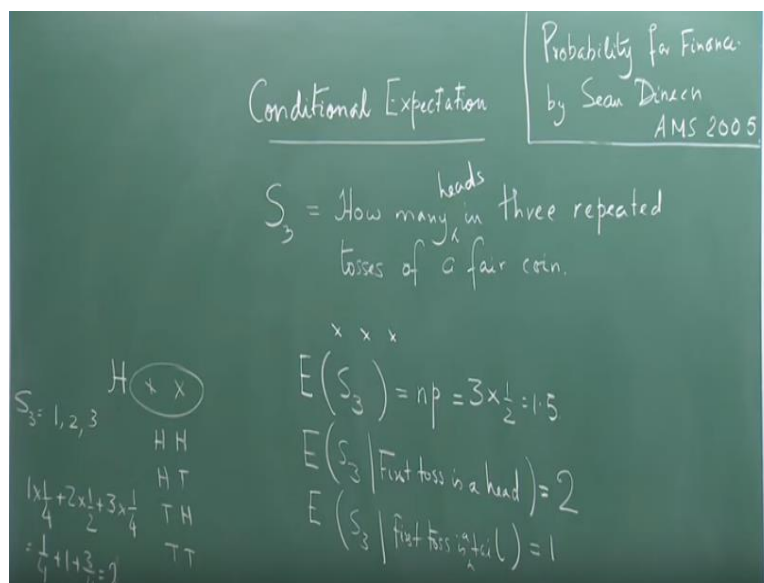
So, the exposition that I found most interesting was from a book called Probability for Finance by Sean Dineen published by American Math Society in 2005 under the series Graduate Texts in Mathematics. So, what does conditional expectation mean? We will see, a very simple example. Suppose, I take this S is a random variable or S is a random variable, which tells me, how many times head appear, when I toss a coin three times. So, how many heads in three tosses of a fair

coin, repeated of course each, first toss, second toss, third toss, three repeated tosses of a fair coin.

Of course, you know this is a binomial random variable. We have already know, we know that, if N be the number of success and Np is the expectation, given that p is the probability of success, sorry, how many heads, so just make a little correction, how many heads in, three repeated tosses of a fair coin. So, it tells you that how many successes you have. So, by success, I am meaning the appearance of a head. So, I toss any of the three can come; head, head, head, head, tail, head and all these things.

So, you know that expectation of this because, this is a binomial random variable, is nothing but n into p where probability of success is half and n is 3 so 3 into half which is 1.5. What happens if I now ask you this question, find the expectation of S3? find the expected number of heads that you expect if you make three repeated tosses of a coin, but with the knowledge now, that the first toss is a head?

So, now we are going to see how many tosses, you have three repeated tosses, so how many heads would appear if you already know, how many heads you are expecting rather, if you already know that the first toss is a head. What is the answer. This is interesting.

You will see now I know that the first toss is a head. I really have to look into the second two positions, this is given. So, here I can have head, head; head, tail; tail, head; and tail, tail. So, means, now I can have only 1 head appearing remaining two are tail. I can have only two heads appearing there is no tail at all so this one comes all here head comes and tail comes then tail comes and head comes.

So, if S3 is 1, if S3 can take value 1 it can take value 2 it can take value 3. It can never take value 0 because, I know the first one is 0 right. So, when I am talking about 1 means these two tail has come. So here among this four, now I do not bother about it, it is already a known fact so I just have to look into this diagram. So, when tail comes both are tail so among these four choices I have only one choice, so there my probability is one fourth. So, 1 into one fourth, right.

Now if, I have the situation, where I have 2 heads which has come in three so among these two, one has been head; so, either head, tail; tail, head. So, any one of the two should appear with probability of the 4, this will appear two times, so the probability would be 1/2. You see the whole working changes. Now 3 means I have got this. So, among all the 4 I have one chance. So, it is one fourth plus one plus three fourth. So, it gives me one, so does this answer sorry 2, 1 plus 1 2, 1 plus 1 2, sorry this answer is 2.

So, I leave you with the to show that, leave you with this question, that if the first toss is tail then what is my expected occurrence of heads? So, that would decrease. So, here my expectation increases. When the first one is head, I am expecting that among the 2, 3 at least 2 would be head. Here because the first one is tail I am expecting at least among the 1 there will be 1 which will be head. So, you see once more information is available my expectation changes. This is called conditional expectation essentially.

Now, how can I put this idea into a more rigorous form. See, this idea is being developed on a scenario, where your omega is finite, your sample space has finite elements. So, your Sigma-algebra or the set of all events, is the power set of that finite set. So, let us look into this scenario.

(Refer Slide Time: 07:55)

(Ω, F, P) be a probability space.

$card(\Omega) < +\infty$ $F = 2^{\Omega} = $ Power set of Ω.

$P(\{\omega\}) > 0, \forall \omega \in \Omega.$ Let $A \in F, 0 < P(A) < 1$

$$E[X \mid A] = \sum_{\omega \in \Omega} X(\omega) P(\{\omega\} \mid A).$$

$$P(\{\omega\} \cap A) = \begin{cases} \dfrac{P(\{\omega\})}{P(A)} & \text{if } \omega \in A \\ 0 & \text{if } \omega \notin A. \end{cases}$$

So, we have, instead of u I am writing F because, that is what is usually written in finance books. So, let this be a probability space. So, what is more interesting to know here, is that I take that this is finite. The cardinality, when total number of elements of omega is finite, that is so my F is nothing but the power set. So, F is the power set of omega.

Now, assume that strictly greater than 0 for all omega in the sample space. So, every sample point, every outcome of random experiment the probability is strictly greater than 0, just like a head or tail scenario. Let A be an event, such that probability of Aa is strictly bigger than and we can take it less than 1 or strictly less than or equal to 1 does not matter much. So, this is what I have assumed and under this assumption I have to define what is called a conditional expectation.

So, I am looking for the conditional expectation of a random variable X, given that the event A has occurred. This is in some sense conditioning using a Sigma-algebra because A belongs to a Sigma-algebra and soon we will come to that, how we actually condition over a Sigma-algebra. This is nothing but the very definition, it will be X of omega and probability of omega given the conditional probability of omega given that the event A has occurred. This is exactly what will happen. This is exactly the definition.

So, what is probability omega intersection A. That is probability of omega by probability of A, if omega is in A and is 0 if, omega is not in A. This is quite simple to understand because, you see what happens. When omega is in A, omega intersection A is omega, so you have just P of omega by P of A, and when omega is not in A, omega intersection A is the null event, empty, impossible event, so that time this will become 0 because, this will be the null event.

(Refer Slide Time: 12:17)

Conditional Expectation

$$E[X|A] = \sum_{\omega \in A} X(\omega) \frac{P(\{\omega\})}{P(A)}$$

$$= \frac{1}{P(A)} \int_A X \, dP$$

$$\left[\int_A X \, dP = \sum X(\omega) P(\{\omega\}) \right]$$

$$E[X|A^c] = \frac{1}{P(A^c)} \int_{A^c} X \, dP$$

So now coming to this board again. I am rubbing off what is in here. That is why sliding boards are so important and this place has space for sliding boards actually. So, now I can write down the definition by putting in this in this place. So, what I finally get is expectation X given that the event A has occurred is summation omega element of A, X omega, P of omega divided by P of A, and this I can write as, 1 by P of A and symbolically, you know that this sum is actually an integral if you were not considering, the discrete distribution or discrete setup, then this would be an integral.

So, essentially a sum, summation X omega, P omega omega element of A but this we will always symbolically even if we have sum we would symbolically write this as. So, this is just a symbolical writing for the expression sorry is just a symbolical writing for the expression X Omega, P of omega, symbolical writing.

Of course, you can move to non-finite spaces, which we will soon do, but this is just a useful trick to understand it. So, if you want to write about, any other set, what about the compliment, if I know the compliment of A has occurred. Then also, this is nothing but, the same you can actually prove that, they are you can just replace A with.

Now let us look at the Sigma-algebra which the event A can generate, that is the smallest Sigma-algebra which contains the set A, it should contain A, it should contain the empty set, it should contain the whole set, it should contain its compliment. So, Sigma-algebra generated by A the empty set, the whole set A and A compliment.

(Refer Slide Time: 15:05)

So, if B, so we are just writing it B is element of A so some sigma element B from this Sigma-algebra with element P B greater than 0, so we can write. So now observe that, whenever B occurs, what does this means B occurs, some omega belonged to B has occurred. When you have random experiment, what is the meaning of B has occurred. When I say, I throw a dice, odd number has occurred. If 1 occurs odd comes over the face, then the event that odd number has occurred has occurred.

So, I can actually view this as a function. So, for any omega that I take in B, I can define a function like this, rather a random variable like this. So, for every omega in B, that is over the whole set B, whole event B, this is the value. So, it remains constant over the whole event B, so if I look at it like that, that I am looking at a function which is constant over various events, then the conditional expectation itself can be viewed as a random variable.

So, you see because it does not matter whatever, whenever any omega occurs, B has occurred actually. So, for any omega in B this should be the story.

So, whatever once you know that, B has occurred so for whatever omega you are taking this should be the story, this should be the answer. So, this idea, that you can actually view, the conditional expectation itself as a random variable is a very-very fundamental idea and is a very helpful idea. So, this is itself a random variable r.v, where r.v. is just a short shorthand for random variables. So, this will allow us to make shift from the standard probability space where we do not bother about the finiteness of omegas and all those sort of things.

(Refer Slide Time: 17:40)
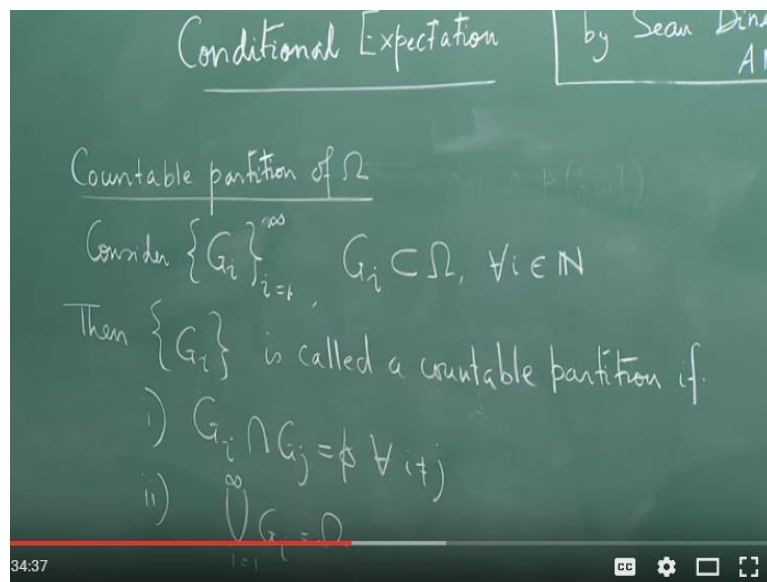
So, first we will talk about a countable partition. Now, what is the meaning of a countable partition of omega. So, consider a sequence of sets, consider Gi, i is equal to 1 to infinity right, 1 2 3 by G1 G2 Gn so consider Gi where G i is a subset of omeg for all i in set of natural numbers. Then this Gi is called a countable partition, if number 1, Gi intersection Gj is equal to phi for all i not equal to j, and number 2, union of Gi, i is equal to 1 to infinity, should give me back omega the whole sample space.

So, this is called a countable partition and what we would now require, is a Sigma-algebra generated by that countable partition that you take elements of that countable partition, when I basically, take, do not take intersections, because intersections would ultimately generate either the, if you take intersection of the same set it will generate the same set itself or it will generate the empty set.

So basically, if you want to take any non-empty set of a Sigma-algebra generated by the set, you just have to construct unions of this set using some subset of n could be finite could be infinite whatever right.

So, what we will be interested is, in a Sigma-algebra generated by a countable partition, which will be a sub Sigma-algebra of F, naturally because, we are taking a subset, when taking a subset of A and generating a Sigma-algebra.

(Refer Slide Time: 20:32)

So, consider a Sigma-algebra G, subset of f and G, G being generated by a countable partition. Then, we define it like this, and the conditional expectation of the random variable X, conditioned on the Sigma-algebra G, is defined as like this, at any omega right, say omega n or

whatever, any omega is PGn integral Gn X dP because this is a countable because this is a generated by a countable partition say G i.

So, given any omega it must lie in one of the Gi s because omega is the element of omega and this is a countable partition of omega so it must lie in one of the Gi s. So, suppose take any omega then this will happen. This is the way I define how you compute this random variable at the point omega if omega is element of G n, good.

I will now, been a random variable, it is important to know that because I am claiming it to be a random variable, you should be able to prove that, I leave this as an exercise to you, in your assignments, so of course and this is true. You are going to prove this fact that this is measurable. An interesting thing is that if x is integrable random variable, that is it has a finite expectation then this random variable also has a finite expectation that is the, conditional expectation, conditioned on a sub Sigma-algebra of F which is generated by a countable partition, has also got to be integrable. That can be proved, but I would not rather prove this fact now.

So, this would be a part of the exercise. If X is integrable, I am expecting that, you know what is integrability though we have given some basic definitions. So, if X is integrable then, integrable means it has a finite expectation because expectation is expressed in terms of an integral, that is integral X dP this is finite, then E is integrable. That is this random variable also has a finite expectation. This is important.

Now, once we know this, we are now going to state a very important property, of this random variable. This property essentially characterizes and this idea actually allows us to move beyond Sigma-algebras, which are Sigma-algebras of this type, that is you will now, once you know this you can, once we know this result what we are going to state, using this idea, we can give a very general definition of a random, conditional expectation as a random variable that is we can define the conditional expectation, conditioned not just by a Sigma-algebra generated by a countable partition, but by any Sigma-algebra which is a sub Sigma-algebra.

So, you are writing this important property, give a very small proof of this and then we will wind up today's discussion. So, this section is the first part of conditional expectation and tomorrow we are going to discuss the general case and the properties of conditional expectation that is it.

(Refer Slide Time: 25:43)
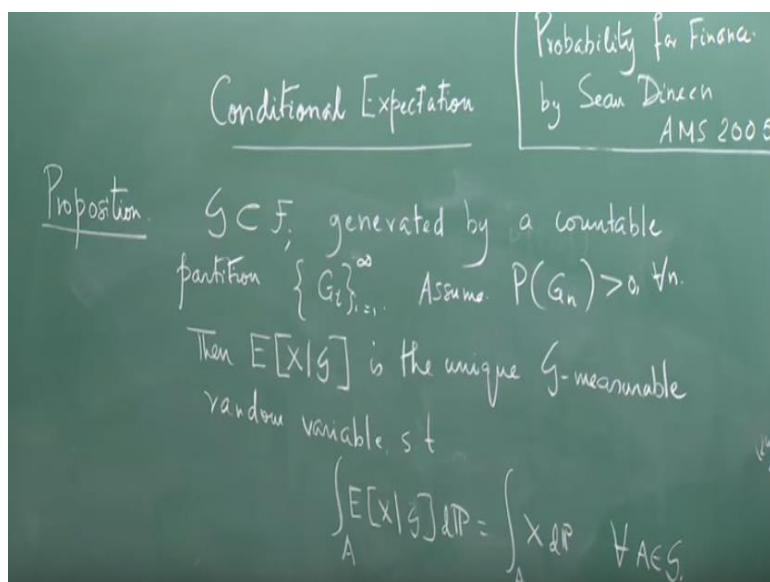
So, what does this proposition say? So again, G is a sub Sigma-algebra of F, generated by a countable partition, Gi 1 to infinity. Assume, of course this definition when you are writing, you are assuming that P Gn is strictly bigger than 0 for all N because. without that you cannot write this.

Assume that P Gn that is probability of each of these pieces of the partition is greater than equal to 0 for all N. Then EX by G this conditional expectation is the unique G measurable, of course when I am talking about measurability here, I am essentially talking about G measurability, is the unique G measurable random variable such that, random variable on the probability space of course, such that for all A in G.

So, on G, integrating over a set of G, subset of G, integrating over any element in G, any set which belongs to G, over X is same as, integrating over the conditional expectation. So, it is essentially it says that, over G conditional expectation and X are almost the same thing. The over

such as Sigma-algebra, X and conditional expectations, they behave in a almost similar fashion. The integrals are same and this idea is actually used to extend, to a higher basically, when you go to a case when you do not take this countable partition you take this as a definition of condition expectation.

You see this is how mathematics develops because, you do it for some simple case and then you know that, it is not so easy to talk about the general case, so you pull it up by taking the, what you know for the simple case as a general definition and that definition will always work if the case is simple. So, we will come to that later on. It has lot of links with certain things called Radon-Nikodym theorem. So, let us give a proof of why this is happening.

I am not going to prove, the case for uniqueness because, that is very simple. So, to prove the uniqueness, we really have to use take any other Y, which for this the same thing happens, which will tell me that integral of this variable minus Y dP would be 0. So, which means that this function is 0, what does that mean, function is 0 almost, that is exactly. So, that is the how you receive your uniqueness. So, we will just do the existence of this that okay if I define the random variable in this way the way which we have defined in this particular case then this must be equal to this. So, proof and with that we will end today's discussion.

(Refer Slide Time: 30:25)

So, I am just taking with this first with the Gn, let us see why we are doing so. First, you have to understand why I am just taking first with Gn because, if A is any non-empty set in G, then A is always written as, union M subset of N Gn where M is a subset of N, could be finite could be infinite whatever.

So, take some Gn from there and you are writing this G M, M is capital M with a set. It contains the indexes which is subset could be just 1 2, 3, 4, 5 so any subset of M. So, every A can be expressed like this right because the intersection you could take intersection of the empty right that is why you cannot. So, every non-empty set has to be expressed like this okay.

Let us see what happens on this Gn. Now you observe, that on this Gn for every part, if you look at the definition this is a fixed quantity. You take any omega on element of Gn, whatever omega you take it will give you the same value. It is a constant on that Gn part. So, I can take to be some omega integral Gn dP, that omega is element of G n.

Now what is this. This is nothing but probability of Gn but I can again use this, I will write down the definition here, probability of Gn integral Gn X dP probability of Gn because this is nothing but probability of Gn integral d P Gn is probability of Gn. So, what I will get here is integral Gn X dP.

Now if I am talking about integral A, so I can write this as integral union of Gn M subset of N because, these are all disjoint, I can actually sum them up sum each of the integrals. So, it is summation M subset of N and summation of all the indexes of N integral Gn and this summation, what is this, for each individual piece this is nothing but this. So, this is nothing but summation M subset of N integral Gn X dP, which is nothing but, integral union of Gn, union of Gn is nothing but A which is integral A X dP and that is the answer that is the proof.

So, with this proof we end our discussion here today.

Thank you very much.