STOCHASTIC APPROXIMATION: THEORY AND APPLICATIONS

Dr. Gugan Thope

Department of Computer Science and Automation

Indian Institute of Science

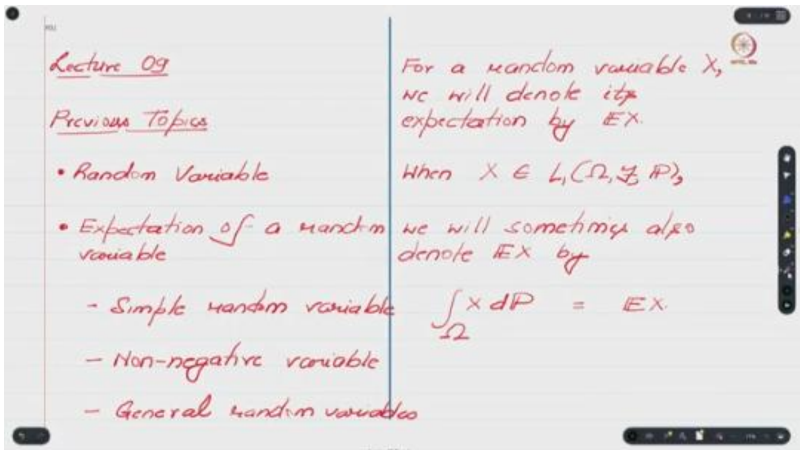
Lecture 9

Conditional Expectation: A Formal Introduction

Hello and welcome, everyone. Namaste. So, today we will be talking about Lecture 9, and let us do a quick recap of what we have done so far before we go ahead. So, this second week has been on preliminaries, and I have been rushing a bit so that we can cover some quick background, and you know, we can get yourselves well-versed with the topics that will be needed. From next week onwards, we will be going a bit more slowly.

Ah, with regards to the kind of topics that will be playing a key role in our design and analysis of stochastic approximation algorithms, okay? In the first three lectures of this week, we have looked at the concept of probability spaces, then we have looked at the concept of random variables, and in the last class, we looked at the concept of the expectation of a random variable. So, in measure theory, this expectation plays a very, very key role, right? And the way we defined expectation was we first defined expectation for, ah, what we referred to as a simple random variable, then we extended this idea to the expectation of non-negative random variables. In particular, we invoked what is called the measurability theorem, which showed that any non-negative random variable can be expressed as a limit of a sequence of simple random variables. And finally, we extended this expectation idea to general random variables.

And in this class and the next, we will move on to this concept called conditional expectation, which will play a very key role in understanding what are martingale differences, and recall that martingale differences are the way we will view the noise, ah, that you know is present in your stochastic approximation algorithms. And in particular, by understanding conditional expectation, we will understand the definition of martingale differences; in particular, we will also understand the conditions under which these martingale difference sequences converge, all right. So, ah, with that, ah, background and introduction, let us begin, ah, today's class. So, you know, we will first recall the notation that we had introduced: for any random variable X, we said that we will denote its expected value as the expected value of X, and today we will, you know, introduce one more notation to denote this expected value of X. In particular, when X belongs to L1(Ω, F, P).



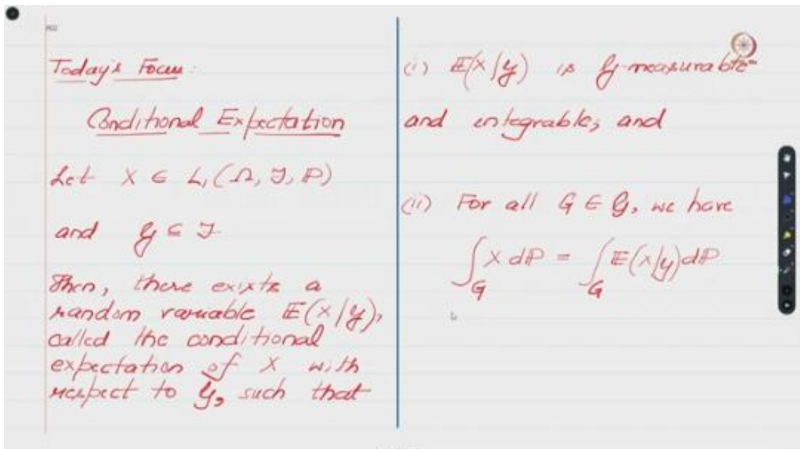
$$X \in L, (\Omega, \mathcal{F}, \mathbb{P})$$

So, recall what this notation means. These three things here together constitute what is called a probability space, and L1 here means the collection of random variables that are integrable with respect to this probability measure, right? So, let us say we have one such random variable; then we will denote its expected value of X, sometimes also using this notation.

$$\int_{\Omega} X d\mathbb{P} = \mathbb{E}X$$

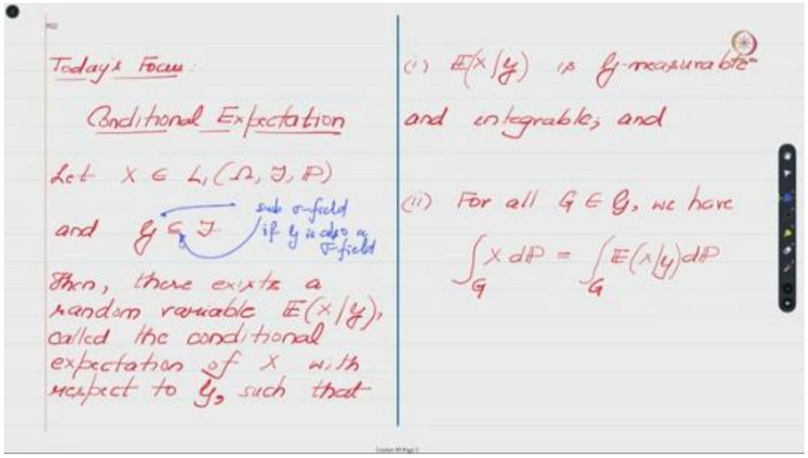
So, this is just a notation. You do not have to necessarily think of it via this integration that you must have studied in your undergraduate classes and so on and so forth. At this point, this is just a notation.

So, instead of writing the expected value of X, I will denote it in this fashion, and this quantity over here is the sample space. This is the random variable that we have, and this P over here is this P over here. And later on, we will see how this notation plays a very important role in some of our discussions. So, as I said at the beginning, our goal today is to understand this concept called conditional expectation. So, first, let us define what conditional expectation is in its full detail. So, suppose we have been given some random variable X, which is, you know, integrable with respect to this probability space. In particular, let us presume that it is integrable with respect to this probability measure that we have.



Furthermore, let us presume that we have a sub-sigma field. So, let us understand what this means. So, notice that we have some parent sigma field over here, which is denoted by calligraphic F, and what we are saying is that we have a subset of calligraphic F. Right? And this subset—so what is F made up of?

Recall it is made up of the subsets of omega. Now G is a subset of F, meaning it is a different collection of subsets of omega with the property that G—this collection G—is a subset of this collection F. And when I say G is a sub-sigma field, I imply that—so let me write it down here—we will call G a sub-sigma field. So, I think I made a mistake; I should say sub-sigma field if G is also a sigma field, and the 'sub' here refers to the fact that it is a subset of some parent sigma field. So, what is the notion of a conditional expectation?



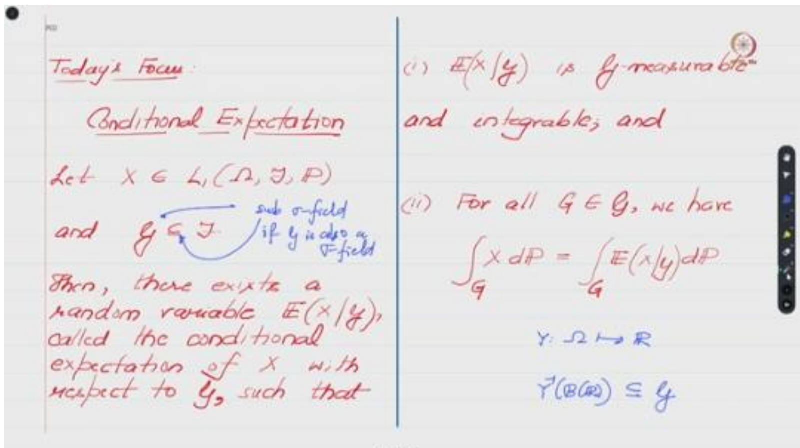
Well, you know, we will talk about the existence soon, but the conditional expectation is a random variable. So, notice that conditional expectation is a random variable, and we will denote this random variable using this notation. So, what is conditional expectation? It is a random variable denoted by this notation over here. And formally, it is referred to as the conditional expectation of X with respect to this sigma field G, such that the following two properties hold.

So, what are these two properties? The first property is that this random variable should be G-measurable and integrable, and the second property is that for every set G in this collection calligraphic G, your conditional expectation should satisfy this property over here.

$$\int_G X \, d\mathbb{P} = \int_G \mathbb{E}(x|y) \, d\mathbb{P}$$

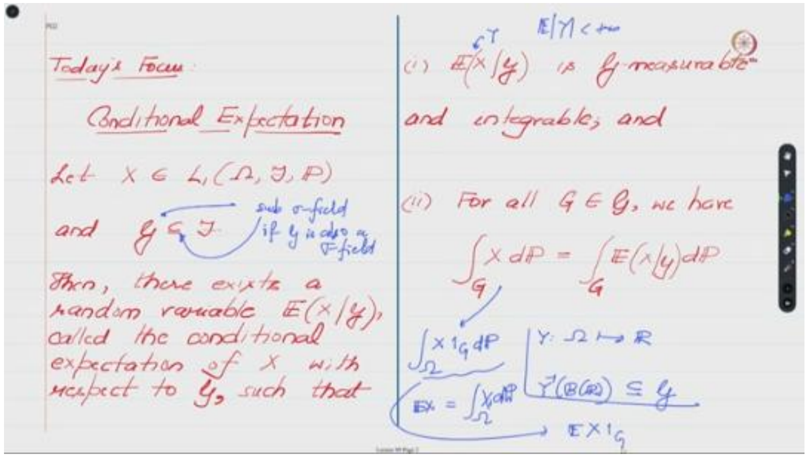
So, let us go over it one by one: what does the first property mean, and what does the second property mean? So, recall that if you have a random variable Y, which is defined on—let us say it is defined from omega to R—and, you know, we will say this random variable Y is G-measurable if Y inverse of B of r—so, recall B of r is the Borel sigma algebra defined on R, right. So, B of r is a collection of subsets of R; Y inverse of B of r will be a collection of subsets of omega, and we will say Y is G-measurable if Y inverse of B of r is a subset of G.

Is this okay? So, in that setting or in that sense, I said that this is a random variable. So, you can think of this random variable for the time being as some y, and we require that this random variable be g-measurable and integrable. Integrable just means that if you denote this by y, right, then we require that the expected value of Y be less than infinity, right. So, we require that.



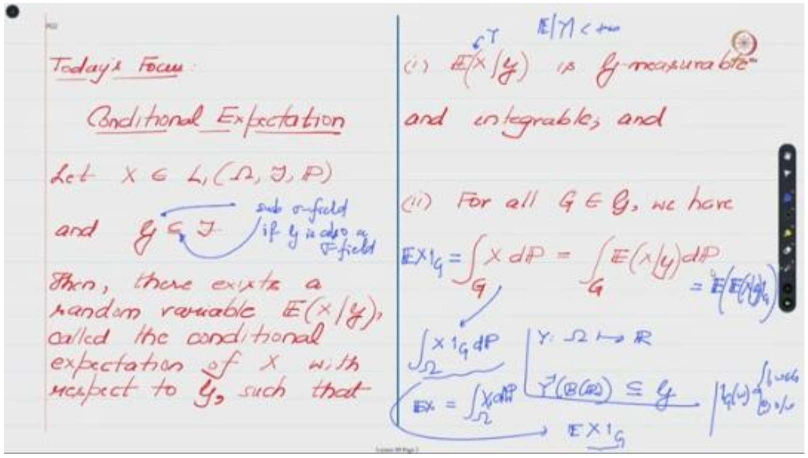
So, I hope with this the first property is clear. Now, let us try to understand the second property. So, the second property in words means that your conditional expectation is an approximation to X with respect to the collection G. In particular, this approximation, you know, satisfies this relation. So, let us go over this relation.

So, on the left-hand side, we have integral x dp g, and the way to interpret this is integral x indicator g and here you have omega. So, recall that I had told you that we can denote the expected value of X or we will alternatively denote the expected value of X via this notation. So, in that spirit, this expression that we have basically means the expectation of X indicator g. So, x is some random variable, indicator g is also some random variable. Recall that what indicator g means: indicator g of little omega, right, indicator g of little omega is 1 if omega belongs to g and 0 otherwise, right. So, the indicator function is defined in this fashion, and x times 1g means that it is a product of these two random variables, and now we are taking its expectation, right.

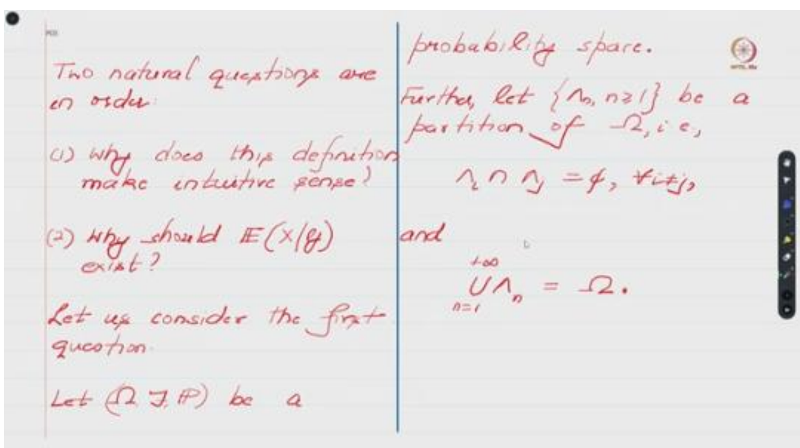


So, this notation that we have over here basically means the expectation of the product of X and the indicator G. So, now let us try to interpret this relation that we have over here. So, the left-hand side is the expected value of X indicator G. This is what is there on the left-hand side, and on the right-hand side, what we have is the expected value of X, this random variable—the expected value of X given G indicator G. So, the left-hand side is some expectation, and the right-hand side is also some expectation, and we require that these two expectations be the same.

So, note that if X is a real-valued random variable, which is indeed the case over here, then this quantity will be a real number, and this quantity also will be a real number, and we require that these two real numbers be the same. And what this condition insists is that these two things should be the same for every G in calligraphic G. So, if such a relationship holds, we will say that this equals this. So, let us go over the definition of conditional expectation again. So, conditional expectation—for that, we start with some parent probability space, and then we have a sub-sigma field, and we will say the conditional expectation of X with respect to this sub-sigma field G is denoted by this, and it is a G-measurable random variable which satisfies this condition over here.



So, two natural questions should arise right: what is this definition, and what is, you know, the intuition behind this definition? And the second is, why should this conditional expectation even exist, right? So, we are saying that if you have given a random variable X and a sigma field G, then, you know, we—I mean, the conditional expectation exists; that is, there is a random variable which is G-measurable and satisfies some condition. So, you should ask, why should such a thing exist? So, this question we will answer in the next class.

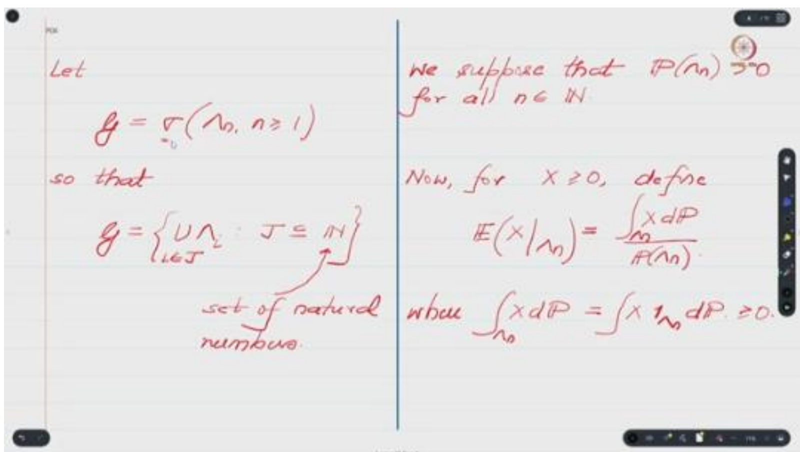


In today's class or in today's lecture, we will focus on answering this question—that is, we will try to give some intuition of what this definition is and why it makes sense, right. So, towards that, let us consider a probability space omega fp right, and to give you an intuitive notion of what this conditional expectation is trying to do. Let us consider the case where we have a partition of omega. So, what does this mean? You have a collection of sets or subsets of omega, that is, you have lambda 1, lambda 2, lambda 3, and so on, and you have a lambda n for every n in the set of natural numbers, right?

And we will say it partitions omega if, on the one hand, their union is the whole space omega, right? And on the other hand, they are pairwise disjoint.

$$\bigcup_{n=1}^{+\infty} \Lambda_n = \Omega$$

And let us define a sigma field which consists of these elements in the partition that we just defined. So, let G be the smallest sigma field consisting of all these elements in this partition or all the sets that constitute this partition. So, since this is a sigma field.

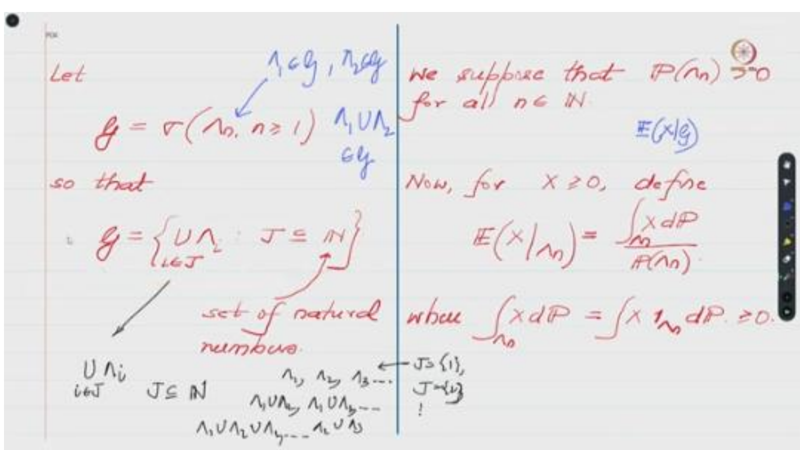


So, this is a sigma field, then you know we saw that a sigma field should satisfy few properties, and you know since lambda 1 belongs to this collection, lambda 2 also belongs to this collection, right? And since G is a sigma field, we also require that lambda 1 union lambda 2 should belong to G. Right. And based on this analysis and the fact that this collection forms a partition, one can show that G actually has this specific description—that is, every element in G will be of this form. So, let us try to understand this. So, what this means is that every element in G can be expressed as the union of lambda i's.

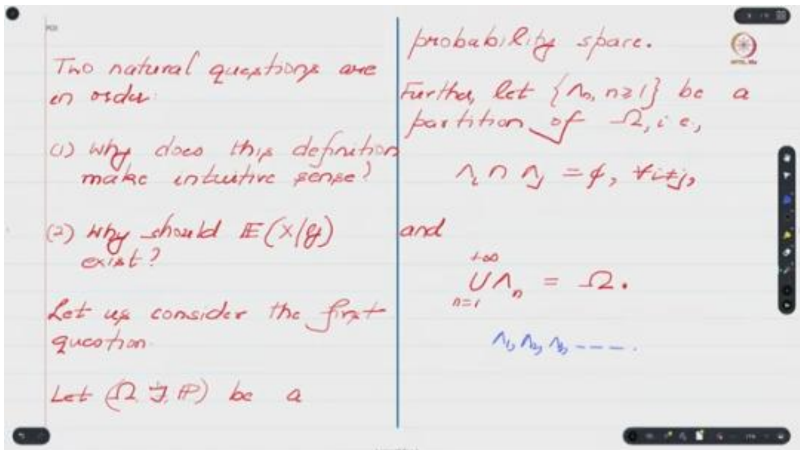
That is, we were earlier given a partition of omega. Then, every element in G, which now is a sigma field, has this form, right? That is, it is a union of lambda i's where I lies in this index set G J, and this J over here is a subset of this set of natural numbers. That means in this collection, you will have you know lambda 1, you will have lambda 2, you will have lambda 3, and so on. So, all these sets you get by choosing j equals 1, j equals 2, and so on. But you also have lambda 1 union lambda 2, you also have lambda 1 union lambda 3, you also have lambda 2 union lambda 3, and you can similarly construct index set J's for these unions. Furthermore, you have lambda 1 union lambda 2 union lambda 3, and so on.

So, imagine any union of these lambda n's. All those unions will be present over here, right? So, on one hand, you know the description of G is very easy, right? But on the other hand, one may ask, 'Oh, what happened to the complements of lambda n's, right? What happened to finite intersections of lambda n's? Where do they appear?' So, I will leave that as an exercise to figure out, you know, why does G only have this, and still we continue to claim that G is a sigma field. The hint is that your collection lambda n forms a partition and, in particular, it satisfies certain properties because of which you know the complements and finite intersections can also be expressed in this fashion over here.

And, to illustrate our example, we will presume that the probability of lambda n is strictly bigger than 0. So, now let us say we have been given some random variable X, and again, for simplicity, let us presume that this random variable X is non-negative. The question that we are interested in asking is, what can we say about the conditional expectation of X with respect to G, right, where G is this sigma field over here. So, that is the question. So, let us recall what the setup is. We have some probability space over here.



So, we have this calligraphic F, which is some parent sigma field that we have been given. So, now what we do is we consider a partition of omega, and I should highlight that we presume that each lambda n itself belongs to F. And on top of that, this collection forms a partition of omega, which means that they are pairwise disjoint and their union is omega. And then we saw that, you know, the sigma field consisting of all these lambda n has this form, and given a random variable X, our question is what is the conditional expectation of X given G. So, now, you know, let us define this notation over here, right?



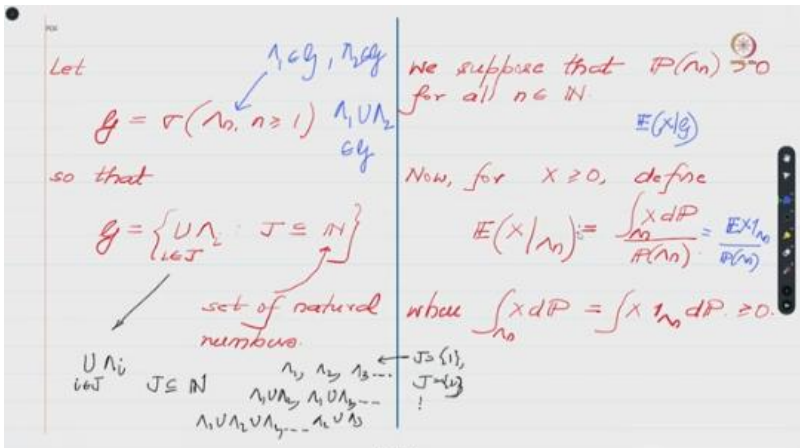
So, let us say the expected value of X given lambda n. So, this is a notation that we are introducing. Let us say this equals the integral of X dP over lambda n divided by P of lambda n. Again, this notation over here, as I emphasized before, equals the expected value of X indicator lambda n. Right, and this is divided by the probability of lambda n, right? And we have presumed that these probabilities are positive, hence this division makes sense, right? And this quantity over here will be non-negative, recall the monotonicity property of expectation. And that monotonicity property applies here because this is a non-negative random variable, and this is a non-negative random variable as we have presumed.

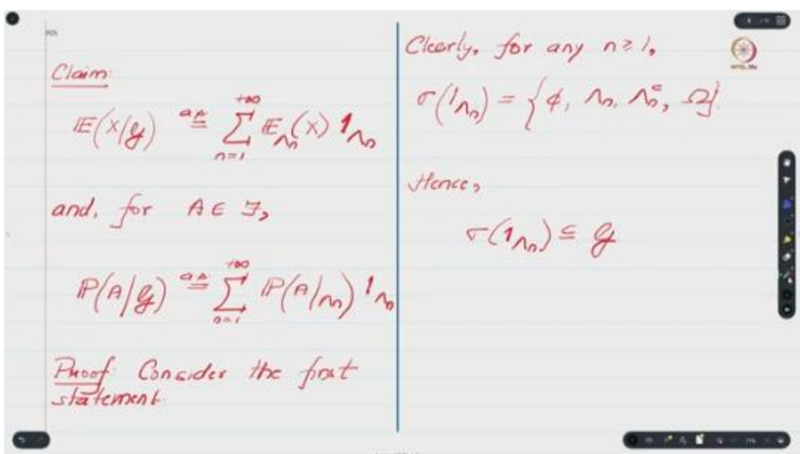
So, this is non-negative, this is non-negative, and this ratio will also be something that is non-negative.

$$\mathbb{E}(X|_{\Lambda_n}) = \frac{\int_{\Lambda_n} X \, d\mathbb{P}}{\mathbb{P}(\Lambda_n)}$$

$$= \frac{\mathbb{E}X1_{\Lambda_n}}{\mathbb{P}(\Lambda_n)}$$

So, let us define it in this fashion, and since I am defining it, let me put these dots to emphasize that this is a definition, right? And our claim is that this conditional expectation that we spoke about when we introduced the definition of conditional expectation, this notation over here actually equals this quantity over here, right? And here above this equality, I have written a dot s, which is supposed to be interpreted as almost sure. And I will talk about this later on.





So, at this point, let us skip it. So, the way to understand this claim is that this conditional expectation, which we defined a few slides back, has this form that is given over here. So, this quantity over here is a real number that we introduced here. So, this is the real number that we introduced here, and this indicator lambda n is basically your indicator function or your indicator random variable. We are saying that this conditional expectation actually equals this countable sum that we have over here. Is this okay?

And in addition, okay, for any A in F, recall that this calligraphic F was the parent sigma field that was given to us. So, let us consider any A in F; then one can say that the conditional probability of A given G, right, also has some form like this. Now, the probability of A given lambda n—this probability quantity—is something that you must already be familiar with, based on your undergraduate probability courses and things like that. So, this is something that you must already be familiar with, and what we are saying is that the conditional probability of A with respect to a sigma field is actually this countable sum when your G is generated by these lambda n's. And again, recall that I had said the conditional expectation is actually a random variable.

So, these quantities over here and here are real numbers, and these quantities are random variables. So, recall that a random variable is a function that is measurable with respect to calligraphic F, and one can indeed check that these are measurable with respect to F because lambda n belongs to F. So, this is a real number, this is a function, and basically this is a countable sum of elements of this form, right? And similarly, this quantity is a countable sum of elements of this form, and because of these indicators, we have a random variable over here. So, the summary of this claim is that in this special case, where G is a countable—I mean, G is the sigma field consisting of this partition of omega—the conditional expectation has this special form. In particular, it involves terms that are familiar to us.

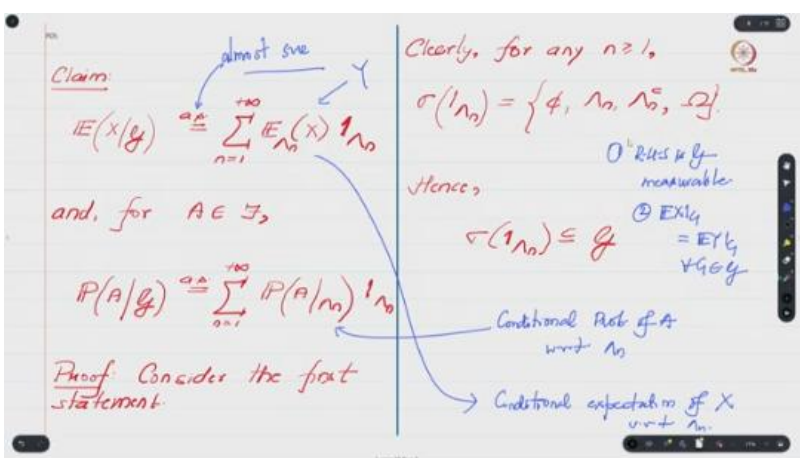
And let me just highlight what the familiarity or relation is. So, this is the usual conditional probability that you must have seen. So, this is the conditional probability of A with respect to the set lambda n or the event lambda n, and this quantity that we have over here is basically the conditional expectation. of X with respect to lambda n, right. So, these are things that you must have studied already, right? And now we are saying that look, this conditional expectation that we have defined actually has such a relation to known quantities, right.

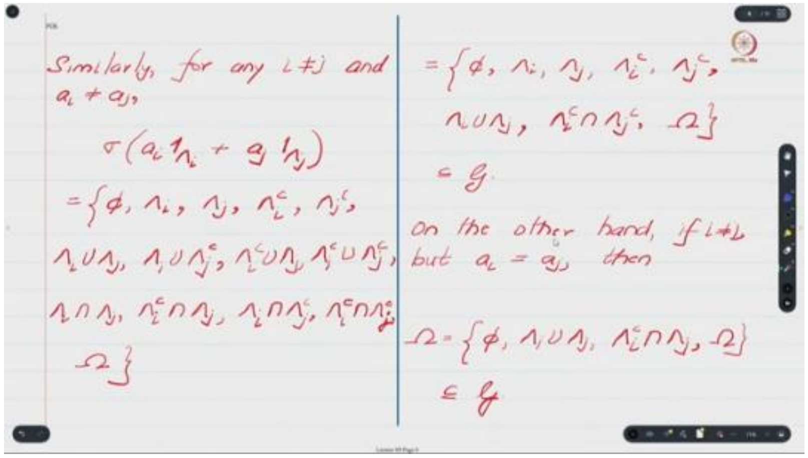
Okay. So now what we are going to do is, in the rest of today's class, we are going to actually prove this relation that we have over here. Right. And my claim is that, you know, once we prove this, we get this for free. So we will focus on this.

Right. So let's consider the first statement. Right. And let us first try to understand what the sigma field is that is generated by this random variable indicator lambda n. This is something that we have seen before, and you must also be familiar with it. It is very easy to see that the sigma field generated by indicator lambda n basically consists of the empty set, the whole sample space, lambda n, and lambda n complement, right.

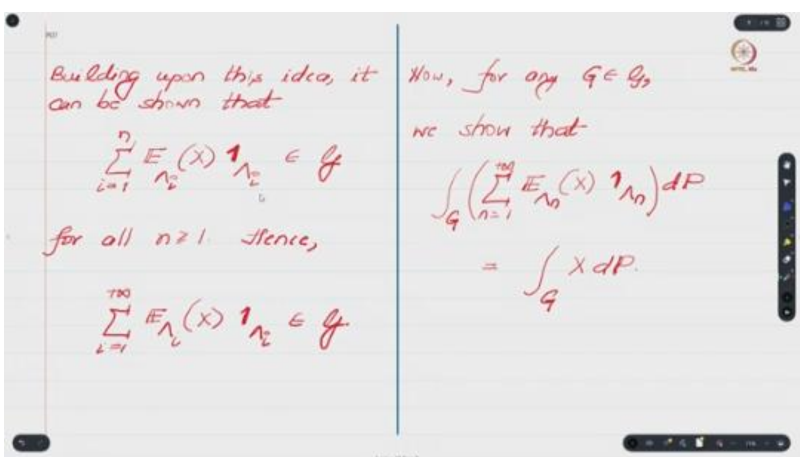
And because G consists of all these subsets of omega, it is trivial to see that sigma of indicator lambda n, right, is a subset of G. Is this okay? Alright. So, with this background, let us try to summarize our strategy on how to show this, right? So, this term over here equals this if two things are true. First is that this right-hand side, okay, is G-measurable, right?

Furthermore, if I let us say denote this as Y, we also require that the expected value of X indicator G equals the expected value of Y indicator G for all G in calligraphic G. So, in order to prove this claim, we will try to show these two things. And we are now focusing on showing this first part, and towards that, it is trivial to see that the sigma field generated by indicator lambda n is actually a subset of G. In other words, your indicator lambda n is G-measurable. Now, with that same understanding, one can easily see that if you take any linear combination of these indicators—that is, you have indicator lambda i, indicator lambda j—and you consider any linear combination. Right. And it is very easy to check that the sigma field generated by this random variable will actually be what is specified over here, right?





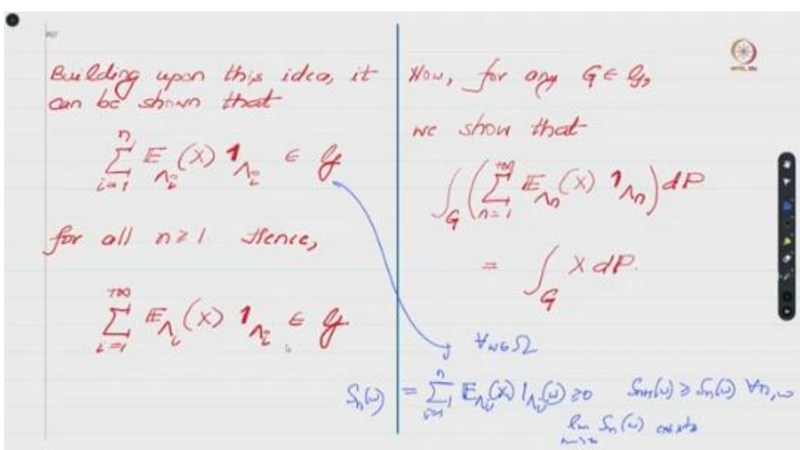
So, you can actually look at this, and because these lambdas are part of a partition, one can actually see that this sigma field has this simplified form. You can easily verify it. And because of this simplified form, the sigma field generated by this linear combination is a subset of G. So, I have actually considered two cases over here. The first case is where A i is not equal to A j, that is this. And the second case is where A i equals A j. So, in either of these two cases, it is very easy to see that the sigma field generated by this random variable is a subset of G, which means that this linear combination is again measurable with respect to G. And building upon this idea and some properties of sigma fields, one can show that if you take any finite sum that is measurable with respect to G, which is what this notation over here means.

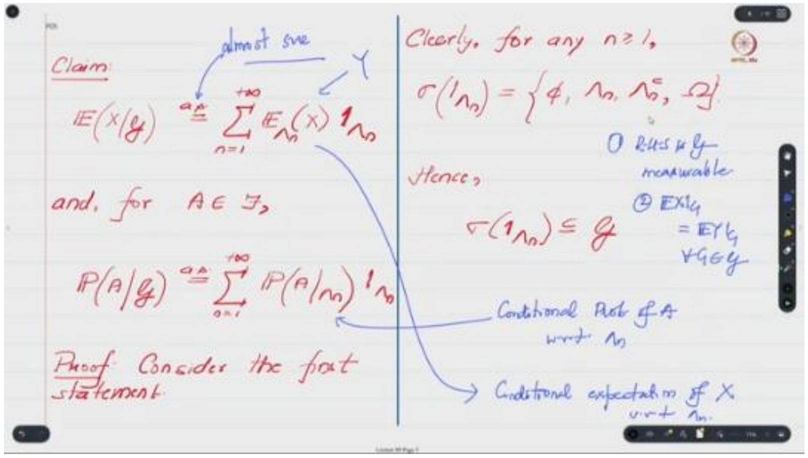


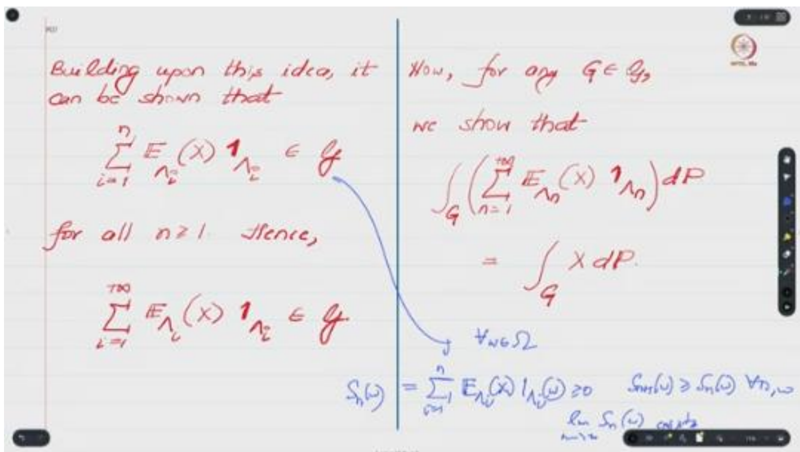
This notation means that the random variable on the left is measurable with respect to G for every n greater than or equal to 1. And then, via some properties, it can be shown that the limit of these random variables is also measurable with respect to G. So, first, let us understand why such a limit should exist. Well, this quantity over here is non-negative. This quantity is a non-negative function.

Hence, this sum that we have, as a function of n, is monotonically increasing. For every omega—so let me just write it—for all omega in capital Omega, your i equals 1 to n. Expected value of lambda_i times the indicator of lambda_i of omega. So, let us call this some S_n of omega. This is, first of all, non-negative, right? And also, your S_{n+1} of omega is greater than S_n of omega for all n, right? And also for all omega. And hence, one can use this monotonicity to conclude that the limit of S_n of omega exists.

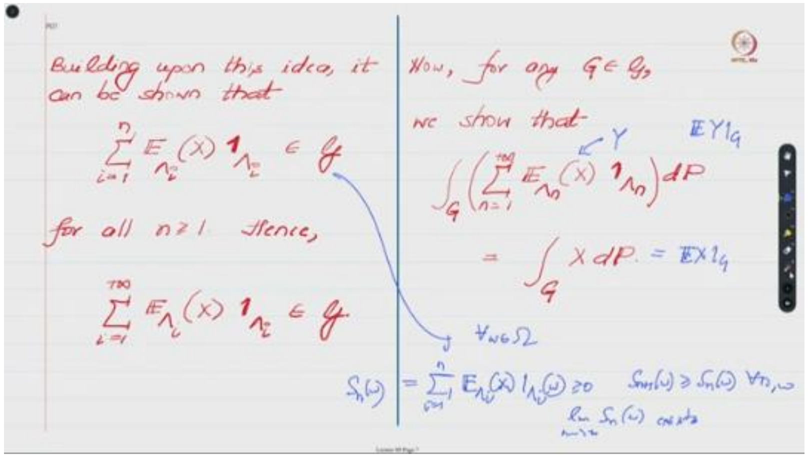
So, this is true, and you know from the properties of sigma fields that one can then conclude that because this limit exists, this limit must be measurable with respect to this sigma field G. So, this completes the first point that I raised. The first point was that, in order to show these two things are the same, we have to show that the right-hand side is measurable with respect to G, right? So, we have finished that verification, and now what we will do is verify that this these two things are one and the same, where the Y quantity over here is basically the right-hand side. So, toward that verification, we have to formally show that if you give me any little g in calligraphic G—so G is a sigma field.

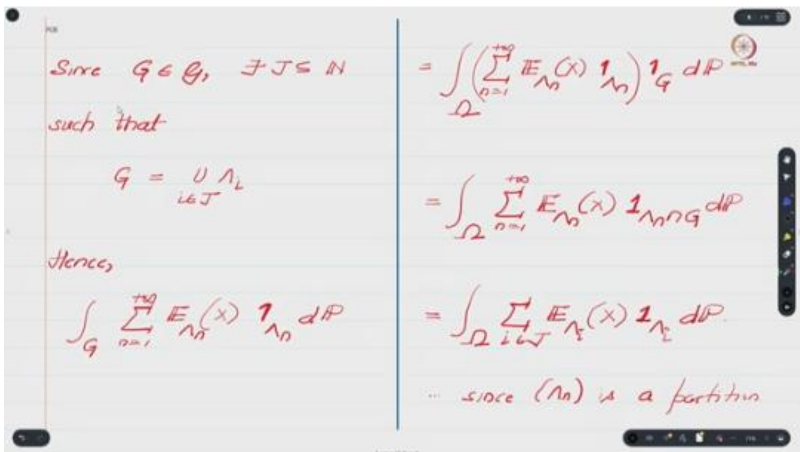






So, give me any element g in calligraphic g. So, I need to show that this relationship holds, right. So, this is the—so, if you recall that we had denoted this as y. So, the left-hand side is basically the expected value of y indicator g, and the right-hand side is the expected value of x indicator g. So, this is what we need to show, and let us proceed towards showing that. Since we want to show that this holds for any G, let us consider one such G and try to verify it. So, because G belongs to calligraphic G, and this calligraphic G was a sigma-field made up of a partition, we saw that every element in G has a special structure—that is, it is some union of lambda i's. And one can see that because G is in capital G, it must also have this form for some index set J, where J is a subset of this set of natural numbers.



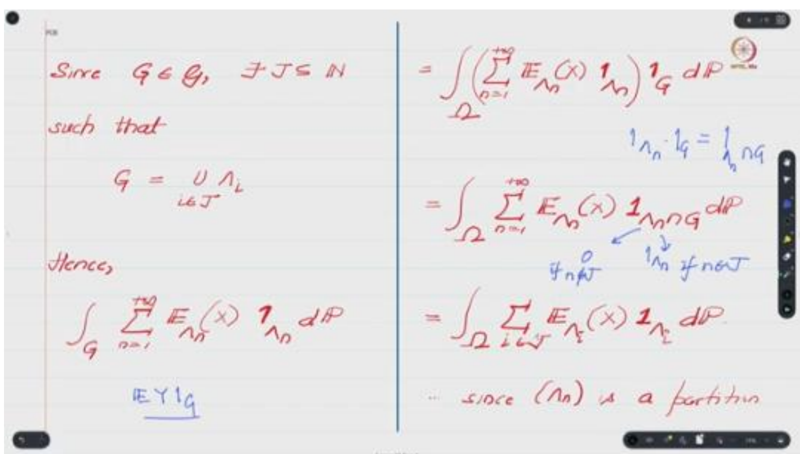


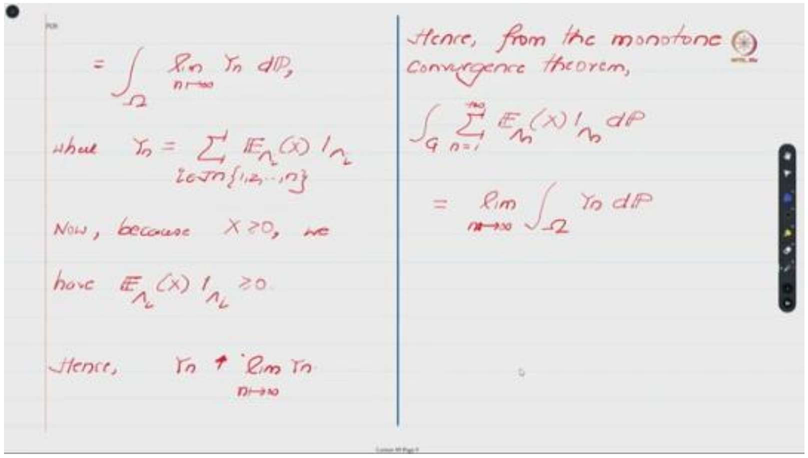
Is this okay? So, because of that fact, this quantity over here—okay, which I, as I told you, is the expected value of Y, you know, indicator—sorry, maybe I should—yeah, Y indicator G. So, because of this relation, notice that I have written this expectation over here—this expectation as the whole expectation—and I have written an indicator over here. The whole expectation, and I have written an indicator over here. This is precisely what I have written over here.

So, we have this indicator over here, and because everything is non-negative, I can actually pull this thing inside. It is easy to check that if you have two indicators and look at their product, this actually equals the indicator of lambda n intersection g. So, this product of indicator lambda n and indicator g equals the indicator of—maybe I will write it a bit neatly—this is the indicator of lambda n intersection g. So, you see that I have written exactly that over here. Right. And because this g has this form, and you know these elements in these lambda i's are part of a partition, one can check that this indicator over here—right—will be 0 if you know this n that we have over here does not belong to j. And if n belongs to J, then this indicator will basically be the indicator lambda n because lambda n intersection G then only will be G. So, one can show that this is 0 if n does not belong to j and equals this if n belongs to j, right.

So, with this interpretation, one can see that this infinite sum that we have over here actually is this sum. Is this okay? So, we have this sum. Now, keep in mind that this j that we have itself could be an infinite subset of the set of natural numbers. For example, j could be made up of the set of all even numbers or the set of all odd numbers, and so on and so forth. So, this could still be an infinite sum.

So, we started with all n's, but now we are left with, you know, this sum where i belongs to j. So, this jump over here happens because your lambda n is actually a partition. So, one can, you know, think of this expression that we have over here, right, this expression that we have over here, as some limit of these y n's, where y n is defined in the following way, right. So, you take j, you intersect it with the first n elements of your natural numbers. And look at this sum.

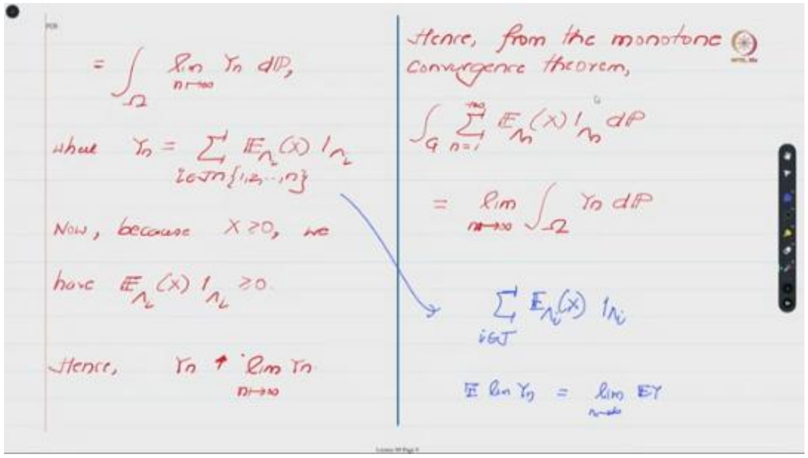


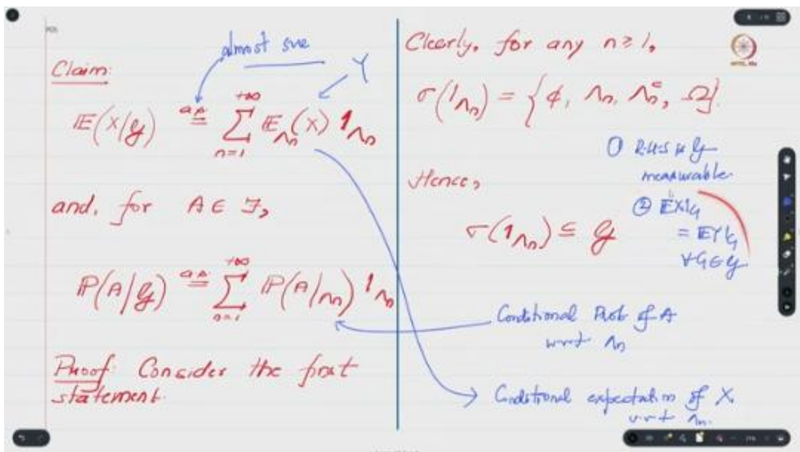


So, in the previous slide, we had, you know, sum i in j expected value of X given lambda i indicator lambda i. So, what we are doing is we are trying to express this sum as a limit of partial sums. So, what we do is we take this J right, and for example, if your capital J includes all the set of even numbers. So, you look at a random variable Yn right, where you know you only consider those indices in J which are between 1 and n and define this corresponding sum related to this sum that we had. And because of everything being non-negative, your Yn actually is monotonically increasing, and hence its limit exists, and this limit is precisely this quantity that we have over here. Is this okay?

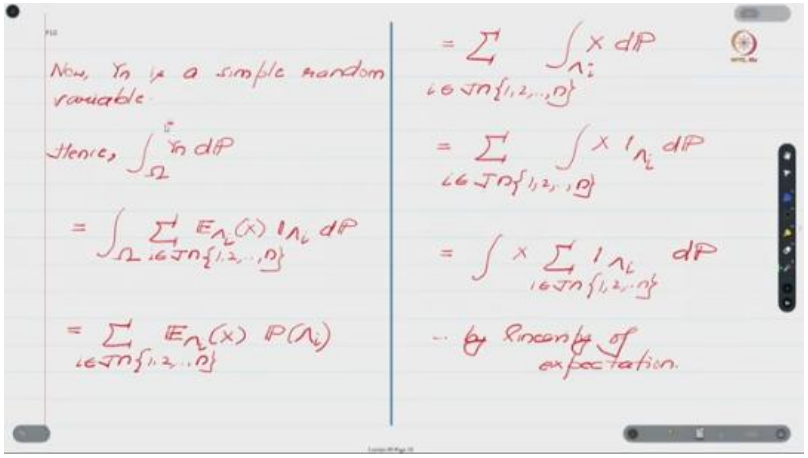
So, one can show that these Y_n's are non-negative and they are monotonically increasing, and this is basically an expectation of the limit, right. So, this is like the expectation of the limit of Y_n, where Y_n's are monotonically increasing, and then by appealing to the monotone convergence theorem, one can conclude that the expectation of the limit is basically the limit of the expectation. So, we can conclude that the expected value of the limit of Y_n is basically the limit of the expected value of Y_n. So, we can easily show this. So, this quantity equals this due to the monotone convergence theorem.

Is that okay? So, now we have to ask how we work with this. So, let me quickly recall the goal for you. Our goal is basically to show this part. This part is what we are trying to show, where this Y is defined over here.





So, let us proceed. So now, let us ask: our goal now is to, in some sense, simplify our understanding of this expression over here. So towards that, let us first try to understand what this indicator ω Y_n dP is. So, recall what Y_n is: it is basically this finite sum that we have over here. So, this is a finite sum because the indices are a subset of the set of natural numbers from 1 to n, which is why this is a finite sum.

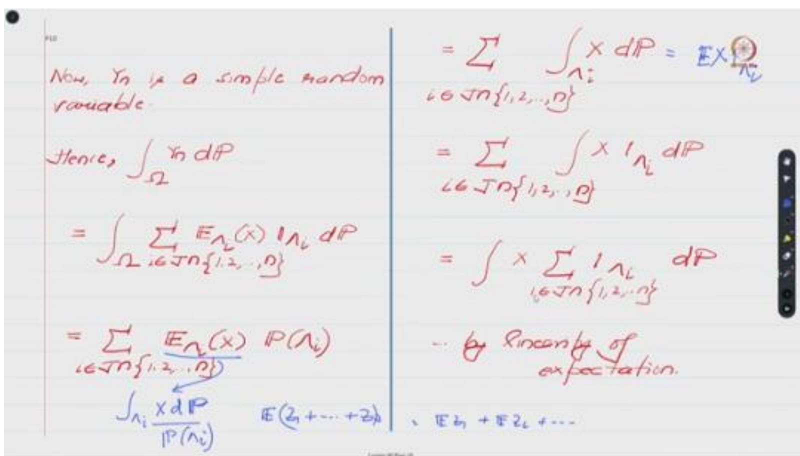


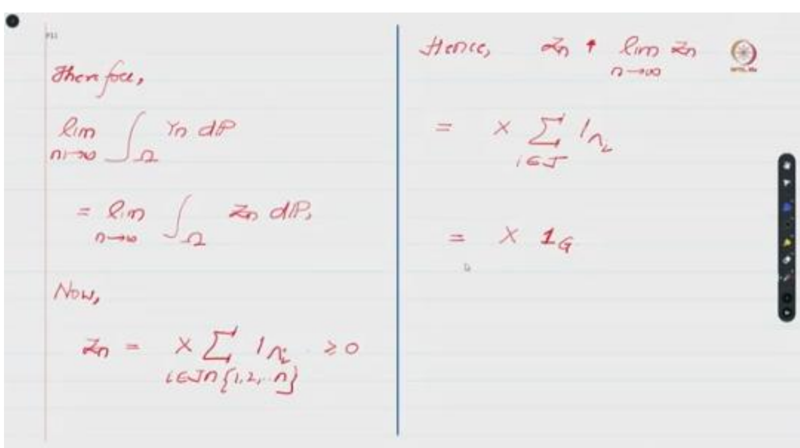
And so, since you have a finite sum over here, we can apply the linearity of expectation. So, recall what the linearity of expectation says: if you have, you know, some Z1 plus Z2 all the way up till Zn, then this equals the expected value of Z1 plus the expected value of Z2 and so on. Is this okay? So, one can, you know, take the expectation and then one can see that, you know, each of these is actually a simple random variable because this quantity over here is a real number and this is an indicator. So, this quantity that we have is actually a simple random variable; hence, its expectation must be, you know, whatever is the scalar times the probability of the event with respect to which the indicator is defined.

So, you have the probability of lambda i over here. Is this. So, this quantity, by linearity of expectation and the fact that this is a simple function, one can conclude that this expectation is basically this real number times the probability of lambda i. Now, recall what is, you know, this quantity. So, this quantity is basically, you know, the integral of lambda i x dp, in other words, divided by the probability of lambda i. So, this is what we have, right? And since we are multiplying these two quantities, right, and this quantity has this probability of lambda i in the denominator. So, this expression and this expression will cancel off, and you will be left with something like this, right? And this quantity, one can, you know, quickly interpret as this is like the expected value of X indicator lambda i. So, this is there, and whatever sum you had, that sum continues to exist over here.

So, you know, I have just moved this lambda i over here and I have expressed this expectation as the expectation of X indicator lambda i. I have forgotten to write omega here, but whenever I do not write anything here, it is to be interpreted that we have omega over here. And again, you know, by linearity of expectation, this is like the expectation of X indicator lambda i. So, you have a finite sum over here. Hence, by invoking linearity of expectation, you know, I can interchange the sum and the expectation. So, this is exactly what I have done over here. Is this okay?

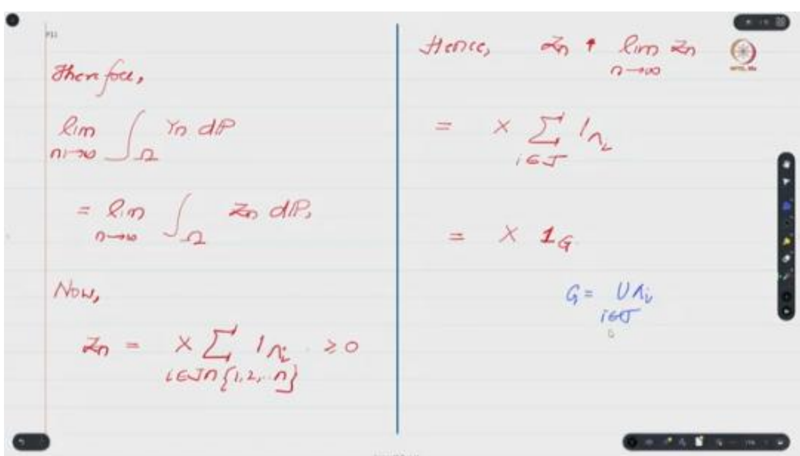
The sum you can see comes inside and, you know, because of this X being common in all these random variables, I can pull out this X. and you know I can look at the sum of indicator a i is this ok. So, one can finally conclude that this you know expected value of y n is basically you know the expectation of this random variable that we have specified over here alright. So, now recall what we were interested in doing. We were interested in looking at this limit, and we have now come up with a simplified expression for this, which is given over here. So, let us denote this quantity that we have as Zn.

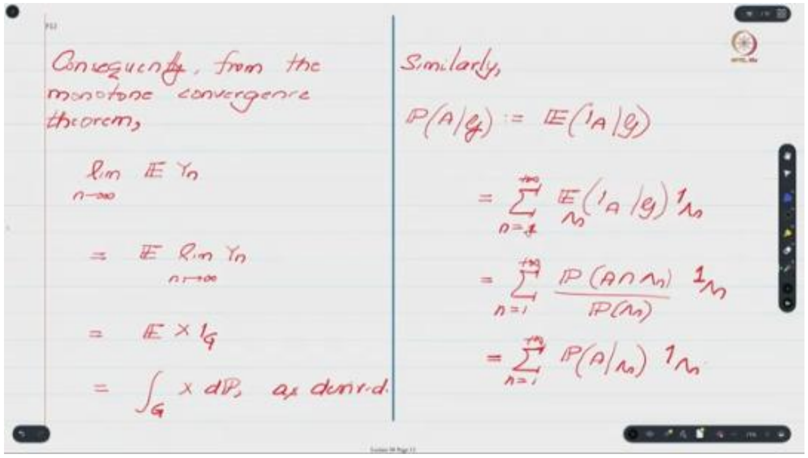




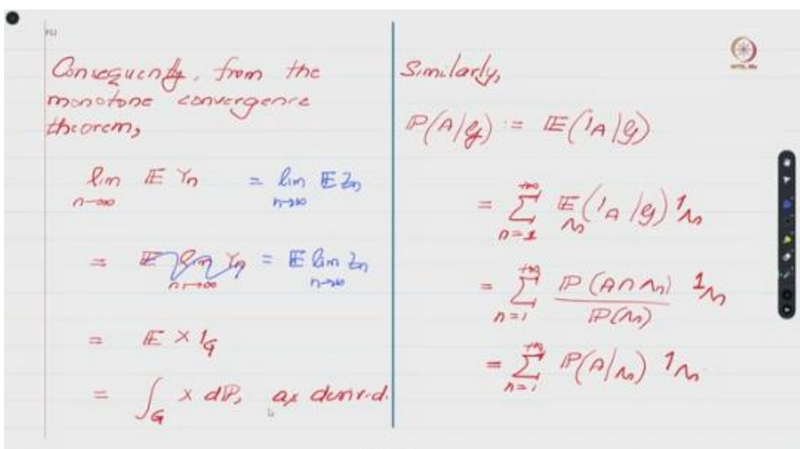
So, if you write this as Zn, then you know this expectation is basically the expectation of Zn, where Zn is defined over here, right? And again, since X is non-negative, these things are non-negative. Your Zn is actually a non-negative random variable, and hence you know Zn converges to some limiting value, and that limiting value—because you have n over here—will basically be this quantity over here. And by recalling how we had defined G and making use of the fact that these lambda i's form a partition, one can conclude that the limit of Zn is actually X indicator G. So, recall what your G was. G was your union lambda i, i in J. And since you are taking n to infinity, we are basically including every possible lambda i or every i that was there in J. Is this okay?

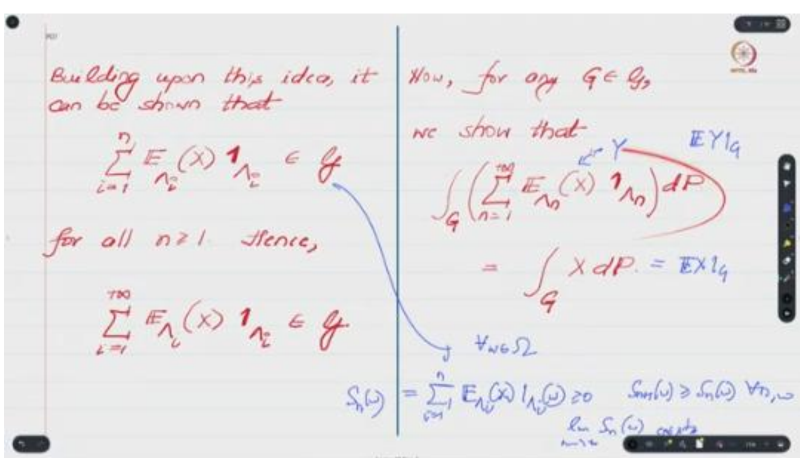
And hence, one can conclude that this is X times indicator J, right? And again, now by invoking the monotone convergence theorem again, one can show that the expected value of Yn's limit is actually the expected value of the limit Yn, right? And okay, let me, you know, do this one more time, okay? So, this was equal to limit n tending to infinity expected value of Zn, right? And from the monotone convergence theorem, one can conclude that this is the expected value of the limit n tending to infinity Zn, right? And this interchange is possible because of the monotone convergence theorem, and we just now showed that the limit of Zn is actually X indicator G, and hence this expectation equals this expectation, and this can be written in this fashion.





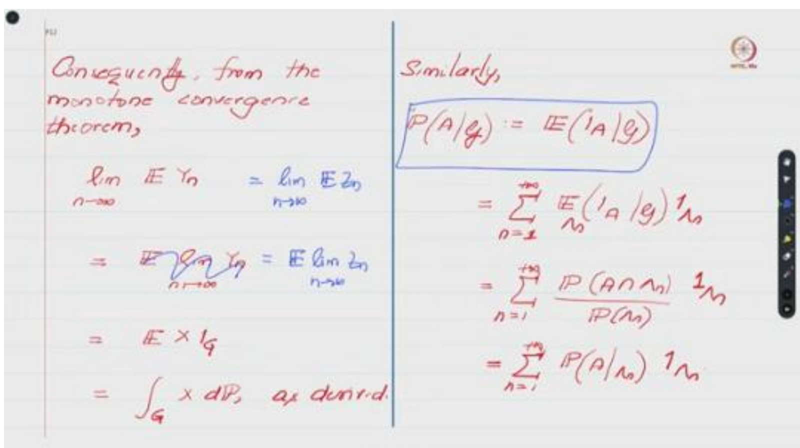
Now, this completes our proof. Why does this complete our proof? This is what we wanted to show on the right-hand side. And if you go back to what we started out with, regarding what we began with, this is the expression that we started with. We did a bunch of algebra and eventually showed that this expression equals this expression.





Is this okay? So, now it remains to show the second part of the claim, which is basically that whatever we have shown for conditional expectations, a similar understanding also holds for conditional probability. And while I have not defined conditional probability so far, I am claiming that the conditional probability of A given G is basically defined as the expected value of the indicator A given G. So, this is the definition of conditional probability. Now, notice that even for conditional probability, we can have an expectation interpretation.

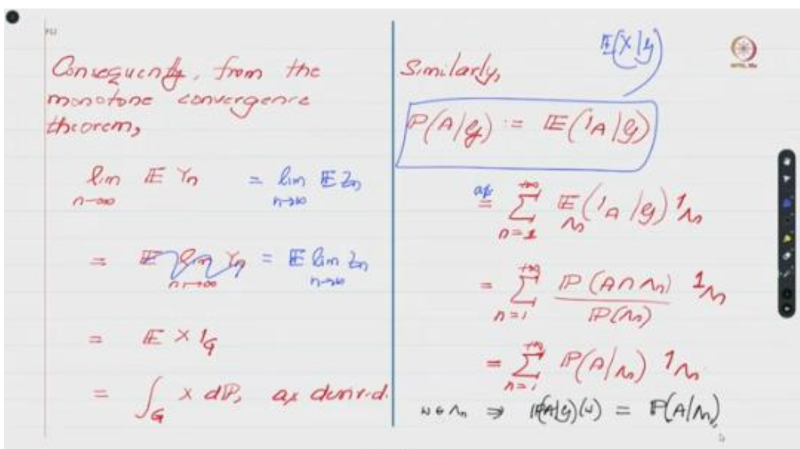
And in the first part of this claim, we have shown some result for the expected value of X given G. Is this okay? And using that relation, one can conclude that wherever there was X, we can substitute the indicator A over there and get this relation that is Probability of A given G is basically the expected value of the indicator A given G, and because G has this sigma-field generated by the partition, that expectation equals this sum. And again, I should emphasize A's over here, and I will talk about what this means.



And from the definition of this quantity, one can immediately check that this expectation is basically this ratio of probabilities, and this ratio of probabilities is the familiar conditional probability that we have seen. So, this is the probability of A given the event lambda n. And one can see that this conditional expectation actually equals this countable sum. So, let us quickly understand what is, you know, the significance of this. So, whenever omega belongs to lambda n, this implies that the probability of A given G.

So, the probability of A given G is actually a random variable. So, that means it is a function. So, if you evaluate the probability of A given G at omega, this will equal the probability of A given lambda n. So, this equality holds because your omega belongs to lambda n, and this will be the case for every omega, right? And recall that your lambda n together form a partition; in particular, the union of lambda n equals omega. So, you give me any omega, there will be a lambda n to which omega belongs, and such a relationship holds. Is this OK?

So, this brings us to the end of this lecture. So, let me quickly summarize what we have done. We have looked at the definition of conditional expectation. In particular, we have tried to interpret the conditional expectation with respect to a sigma field. In particular, we looked at an example where the sigma field is generated by a partition of the sample space, and we showed that in that very special case, this conditional expectation has this countable sum in terms of these conditional expectations that we are familiar with.



So, when I say we are familiar with—you know, in our undergraduate courses on probability—we must have seen the expectation of a random variable with respect to an event. Similarly, we must have seen the probability of an event with respect to another event. So, these things we must have already seen, and now in today's class, we saw this conditional expectation with respect to a sigma field, and we showed that, you know, that is actually related in terms of these known objects. via this countable sum in this special case. So, that is in general; this relationship need not hold, but for the special case where the sigma field is generated by a partition, it has this interpretation.

So, more generally, what is conditional expectation? Well, it is an approximation. So, given a random variable X and a sigma field G, the conditional expectation of random variable

X with respect to the sigma field is basically a G-measurable random variable that best approximates this random variable X. And this approximation is in terms of expectation defined with respect to the sets in this sigma field calligraphic G. So, I hope you will be able to, you know, utilize this lecture, in particular, to understand the topics that we will be needing for further studies.

In the next class, we will be talking about some properties of conditional expectation, which will finish this week's course, and in the subsequent week, we will start going a bit more slowly and in more detail and go over martingale differences. With that, let me stop. Thank you and Namaste.