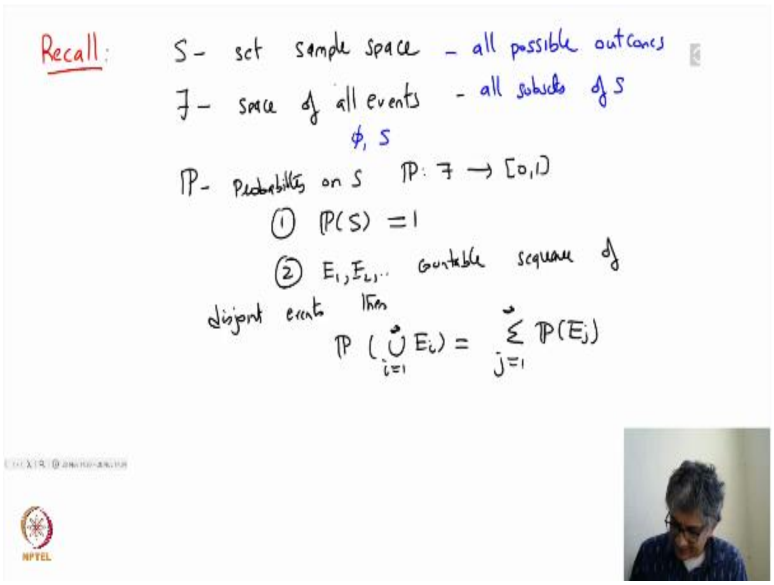
Introduction to Probability – With Examples Using R
Professor Siva Athreya
Theoretical Statistics & Mathematics Division
Indian Statistical of Technology, Bangalore
Lecture 2
Properties of Probability

(Refer Slide Time: 0:21)

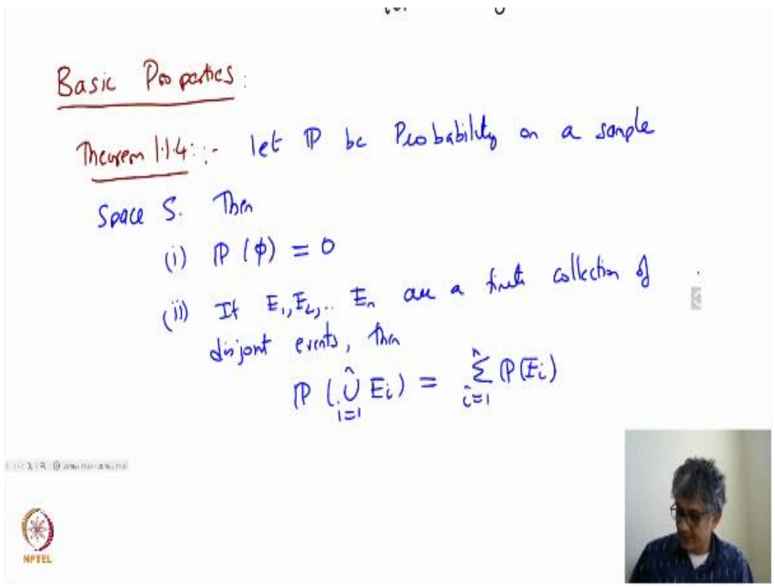


So today, let me just begin by recalling a few facts last time, so what all have we done, we had start off with the set S which is a sample space, this is supposed to be all possible outcomes and we define the space of all events, this is a, so an event E was essentially all subsets of S, so this is we have taking this temporary definition of subset of S, so and also we have to point out that it includes the empty set and the whole space because nothing happens and S is just all possible outcomes, anything happens and then we define P which is the probability on S, probability on S, on this as a function P is a function from the event space to 0,1.

So, it takes any event to 0, 1 and it had 2 axioms, one was that the probability the whole sample space was equal to 1 and the second was that if you take Ei, E1, E2 what is input you take a countable sequence of disjoint events, of disjoint events, then the probability of the union i equal to 1 to infinity of the Ei is the sum from j equal to 1 to infinity of the probability Ej.

And this included several sort of nuances that 1 was that the series on the right hand side here converges and it is equal to the left-hand side and left-hand side just says that the if you have disjoint events then the probabilities of disjoint events add up to give you probability. Excuse me, so now, we will now try to sort of build on this and do some basic properties of probability which I would like to say, before I do examples. So, today I will begin with some basic properties of probability.

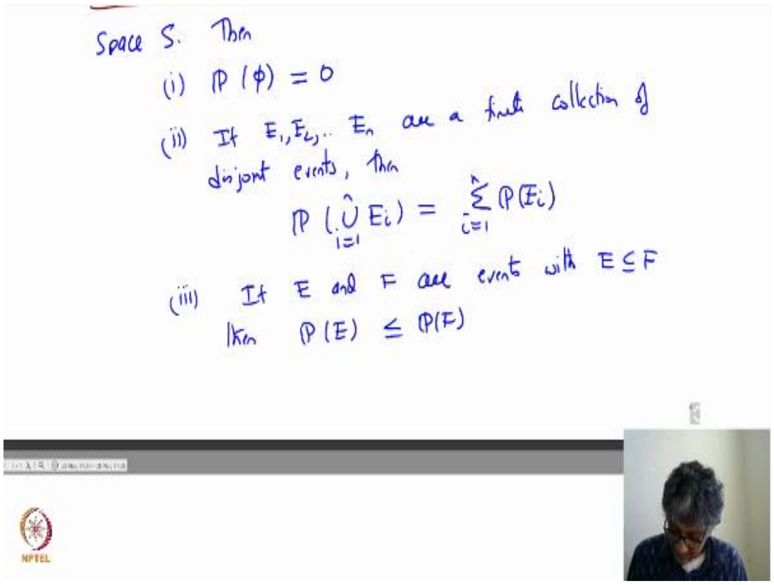
(Refer Slide Time: 2:56)



So, here are some basic properties, basic properties, so this is the first theorem that I would like to show before I do examples, the first theorem that I want to show you is a longest theorem. Let, P be a probability or a sample space S, sorry, I made a mistake in it, let P be a probability on a sample space S, then the first thing is, the first property that probability of the empty set that is nothing happens is 0, the second thing is like I said probability is finitely additive, that is if E1, E2 up to En are a finite collection of events, collection of disjoint events, then probability of union i equal to 1 to n Ei is the sum from i equal to 1 to n probability.

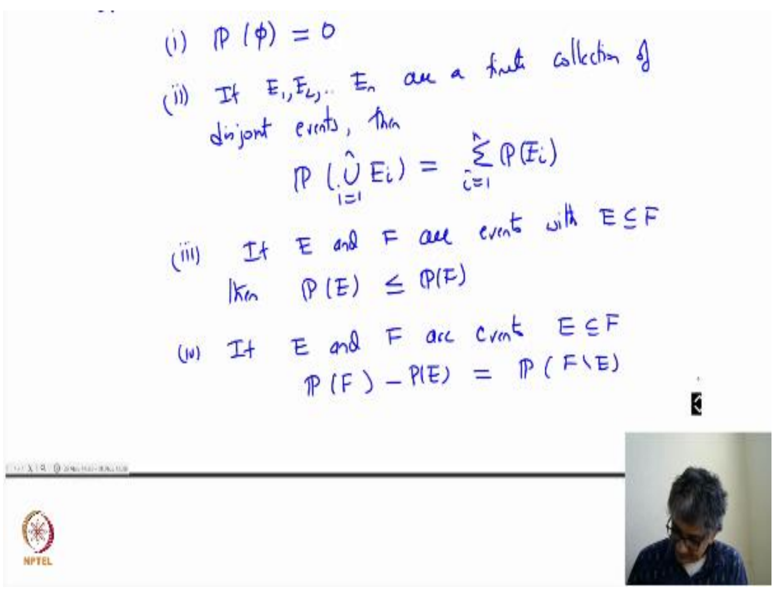
So, this is a little, little different definition of in the axiom but here I require only for countable unions but it is sort of a simple exercise which you will do to show that it also holds for any finite.

(Refer Slide Time: 5:05)



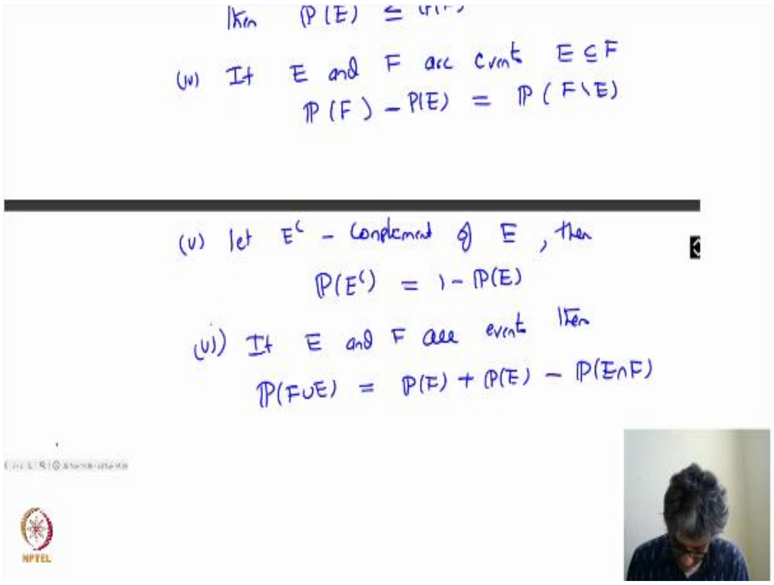
Number 3, was that if E and F are events with E contained in F, then probability of E is less than equal to F, so if an event is larger and contains a smaller event its probability is larger also.

(Refer Slide Time: 5:45)



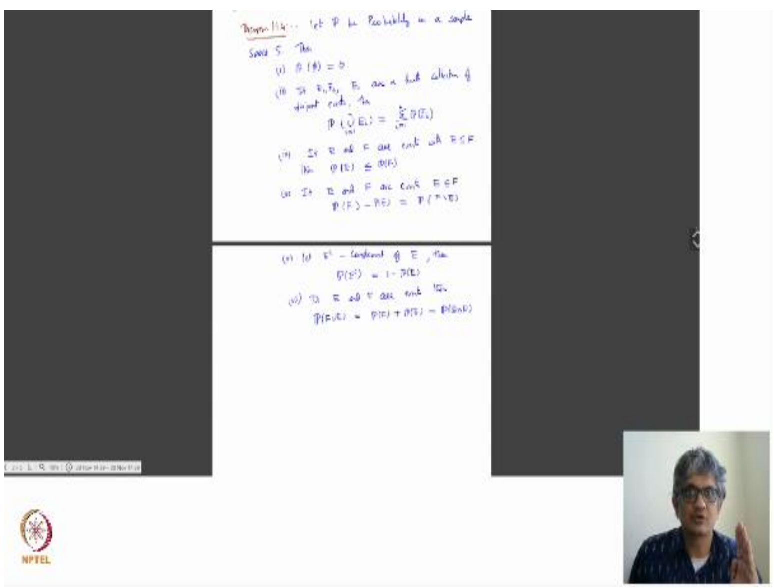
The fourth is, if E and F are events and E is contained in F, then you can actually write the probability of F minus the probability of E is actually the set probability of E, sorry, F removed from E, so F minus E, (())(6:18).

(Refer Slide Time: 6:18)



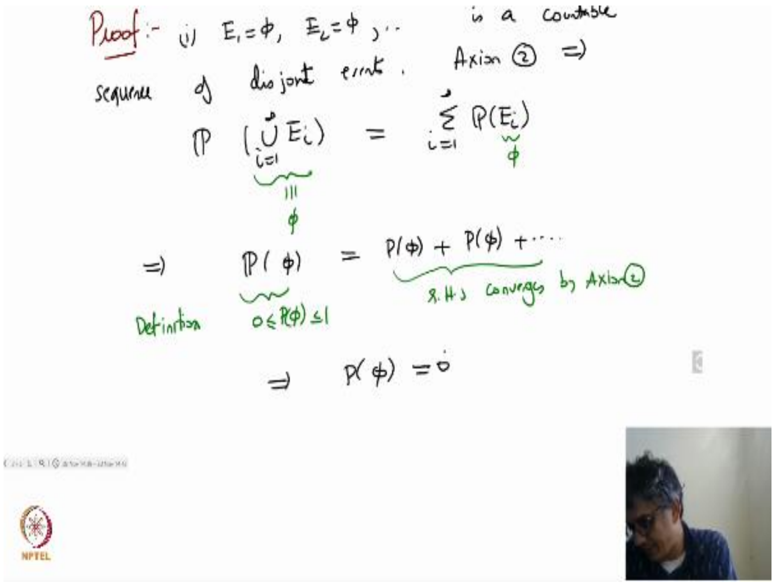
Number 5 is that if E complement, if E sub c is the complement of E, then probability of the complement is 1 minus the probability of E. And the last is if E and F are events, then if you look at the union of F and E that is going to be a probability of F plus probability of E minus the probability of E and F, so this is one theorem that I like to show.

(Refer Slide Time: 7:16)



Let us show in one shot, the whole theorem you can see for one shot, so this is one thing I would like to show in this class that is sorry, about that. So, I would like to show one through 6 and try and show you how the proof goes and this assuming just P is a probability sample space, that means it satisfies those 2 axioms that we start off with.

(Refer Slide Time: 7:51)

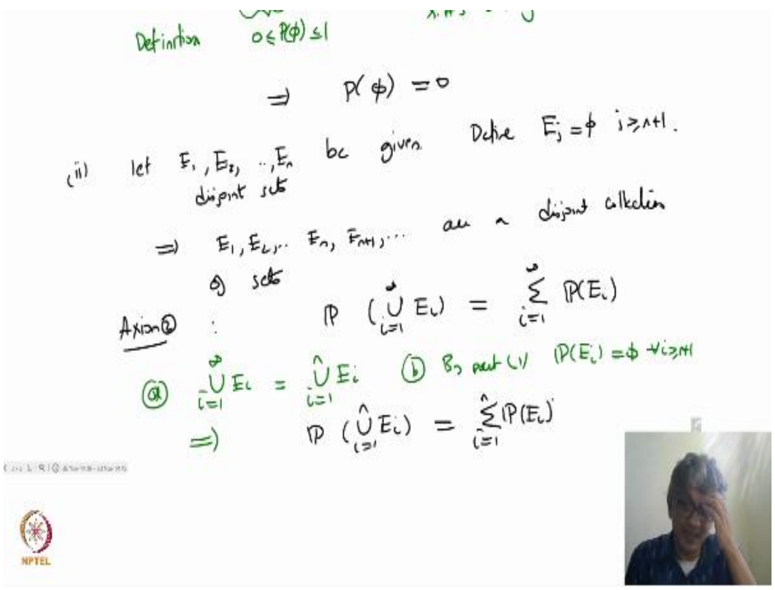


So, let me now prove it, here is the proof, so the first one is that the empty set probabilities, so now this is something, it is a little trivial to understand, if I take E1 equal to the empty set, E2 equal to the empty set and on and so forth, this is a countable sequence of disjoint events by definition, is a countable sequence, countable sequence of disjoint events, once you have this, once you have this, then you can apply axiom 2, so axiom 2 of definition implies probability of the countable union i equal to 1 to infinity Ei is equal to the sum from j equal to 1 to infinity the probability of Ej, I have used i (())(8:54) same i.

Now, if you notice the right hand side, this particular portion right here, if I take union of empty sets, that is the same as the empty set again and these are all empty sets, so in some sense this just implies that the probability of the empty set is equal to the probability of empty set plus probability empty set and so on, but we know that the right hand side converges, converges, we know that by assumption this right hand side should converge, converges by axiom 2, we also know that by the axiom, by definition of probability is between 0 or 1.

So now, it is a simple exercise if you have a, if you have right hand side you have some finite number and right hand side you are adding the same number many, many times, the only way this can be true is that this implies that the empty set has to be 0, because empty set, positive number the right hand side would diverge and we are adding the same thing again and again and it is positive so the left answer has to be equal to 0 and left hand 0 means that probability 0 means that means each of the quantities is 0.

(Refer Slide Time: 10:31)

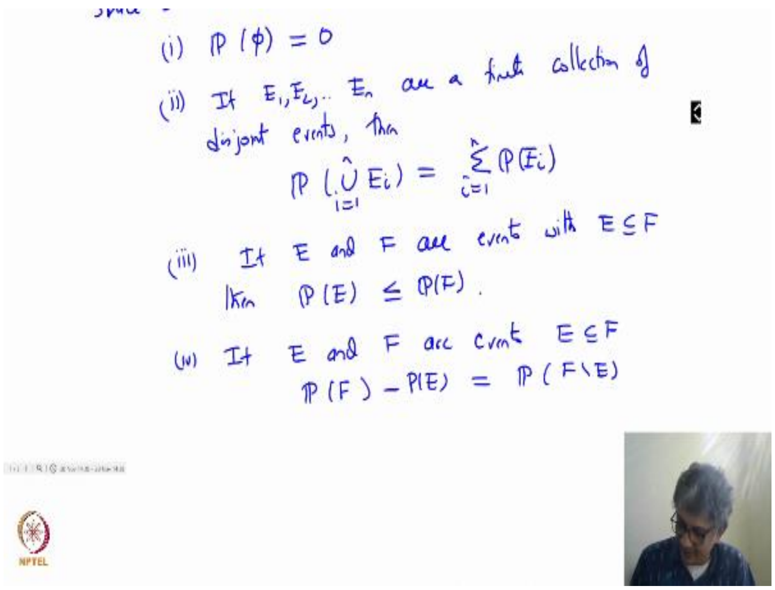


And the second thing is the following, what are the second things, the second thing says that you should get if you take finitely many guys then the probability is sums up, let us do that, so here what I do is let E1 up to E2 up to En be given, they are given to you, so then you define Ej to be the empty set for j began (())(11:06), so you are given E1, E2, En disjoint is given. Now, if you take everybody empty set after Ej n plus 1, then this would clearly imply that E1, E2, En and En plus 1 and so on so forth are a disjoint collection of sets.

So now again by axiom 2, I know that the probability of union i equal to 1 to infinity Ei is the same as the sum from i equal to 1 to infinity of probability Ei, but now you are in luck because you have to make 2 observations, one observation is that, observation a is the following, is the following that union i equal to 1 to infinity Ei is the same as the first n because after that your all, everybody is empty, so this is even i equal to 1 to n of the Ei and you also know this (())(12:23) by part one, probability of Ei is equal to 0 for all i greater than equal to n plus 1 right because after n plus 1 everybody is empty and the probability is 0.

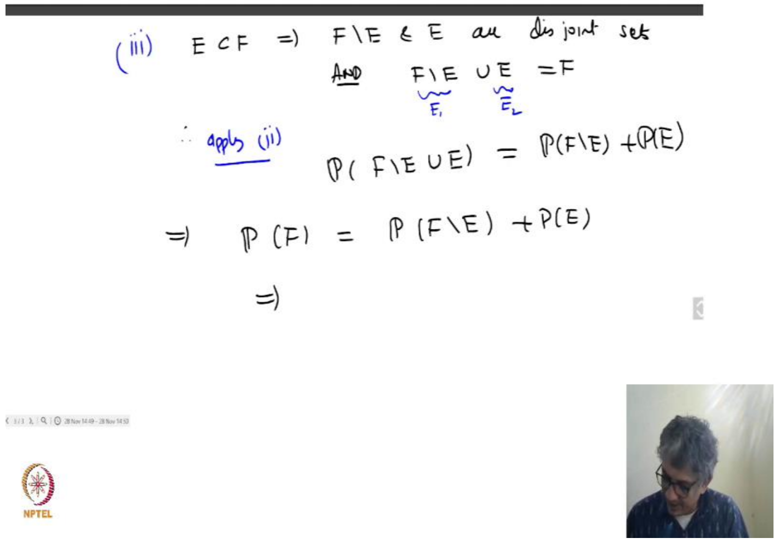
So, these 2 would immediately imply, these observations would imply that the probability of union i equal to 1 to n of the Ei, so the left-hand side is just this, so that is clear so let me erase it properly, left hand side is just this union i equal to 1 to n of the Ei is the same as the sum from i equal to 1 to n of the probability, very nice because after some time everybody is empty, so this guy is 0, so it will affect, this only be the finite sum on the left hand side and similarly here right hand side.

(Refer Slide Time: 13:25)



Now, let me show the where was I, I was trying to show this theorem, so where I was number 3, number 3 is if E and F are events with E contained in F then probability E less than equal to F, probability of F. So, we will do that in a second, so I will try and show 3 now, that is if E is larger than F, F is larger than E then probability E is less than (())(13:49).

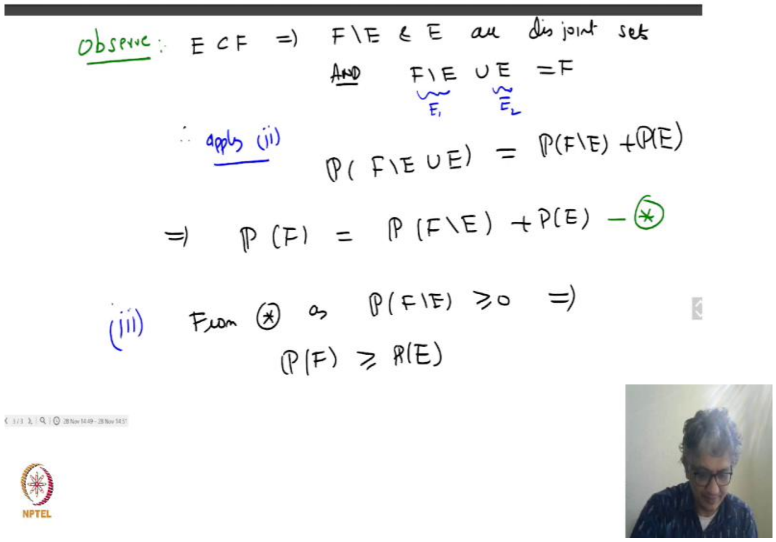
(Refer Slide Time: 13:59)

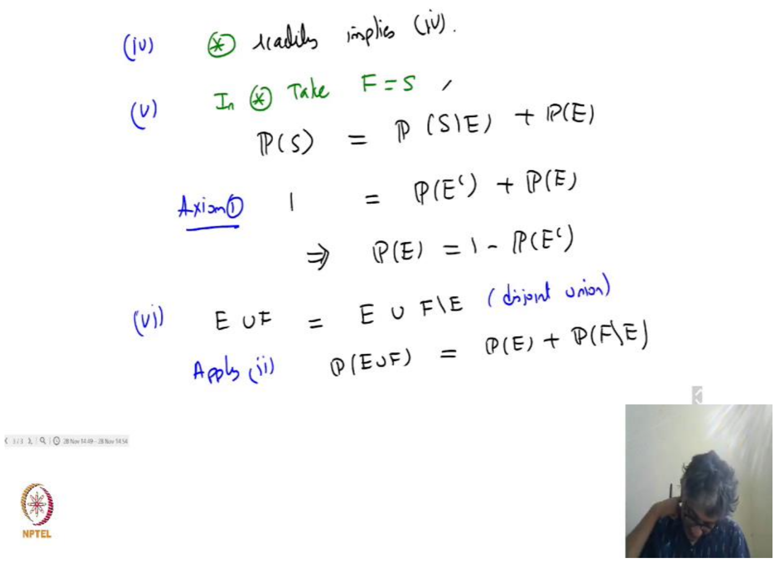


Now, 3 4 and 5 can relate it, so we will try and do it together but I will try and do separately also. so here is 3, so in 3 what I do is, I do the following, so E is contained in F this implies F minus E, that is you remove E from F and E are disjoint by definition and further more you know F minus E union E is the same as F.

So therefore, if I take this to be my E1 and this to be my E2 you apply, apply 2 so what do you get? You get probability of F minus E, sorry this side F minus E union F from the right is the same as probability F minus E plus probability of F but this implies what on the left hand side F minus E union F is just F, so we get F, on the right hand side I get probability F minus E plus probability of E, sorry it is E here.

(Refer Slide Time: 15:55)



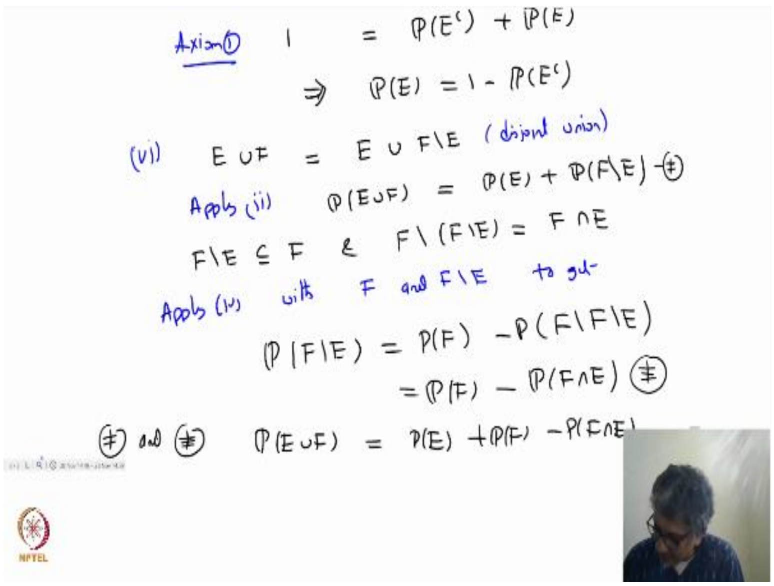


So therefore, this whole thing would imply that, would imply everything in one shot, this calculation, so let me, let me erase this now, let me just not call it 3, so let me just call it, let me just call it observe, so this calculation we will call it, we will observe this calculation that this happens, you apply 2 so you get stuck, okay.

Now, we will finish everything in one shot, so if you want to do 3, for 3, so what, so this is 3, so what I do now, if I took from star as probability of F minus E is always non-negative, so it becomes 0 and 1 this will imply immediately the probability of F is bigger than equal to E, which is what I want, that is F immediate and star readily implies 4. So, four is ready made, so star readily implies, that is what 4 exactly said.

How you do 5? Five again is from star again, in star take, how do we take, take F equal to S, take F equal to S, then 5 is immediate because if you take F equal to S you get probability of S is equal to probability of S minus E, that is probability of S minus E which is E complement plus probability of E, but probability of S is 1 from axiom 1, so I get probability of 1 is equal to this probability of E complement plus probability F. So, this would imply that probability of E is 1 minus the probability F.

(Refer Slide Time: 17:41)



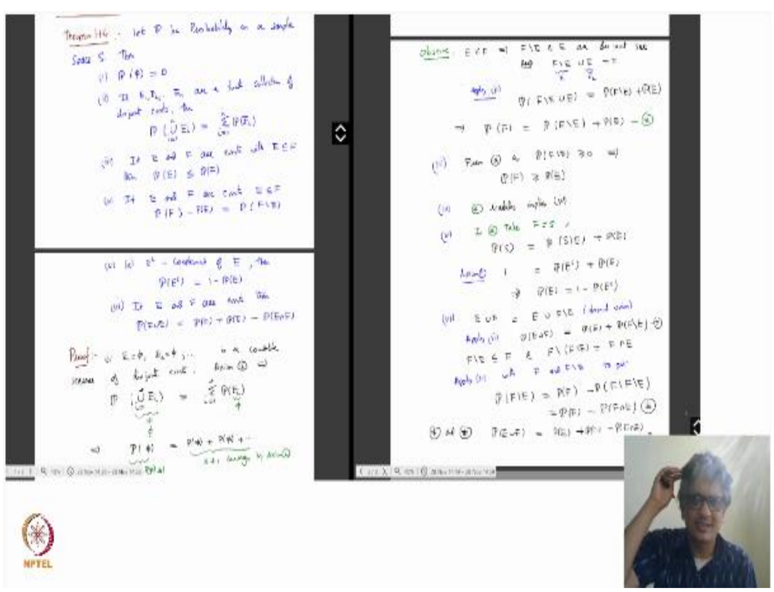
Now, 6 is easy, 6 is easy, 6 is again you just have to just work a little bit with star again and just do it properly, so what you do is you take E union F, so that is everybody in E and everybody or F, so that you can take it as everybody in E or everybody in F outside E, that is the same thing, there is no problem, so now we apply 2, what does 2 say, the 2 says probability of E union F is equal to probability of E plus probability of F minus E, and notice that they are disjoint sets, disjoint which I have used for, because F minus E and E are disjoint by definition.

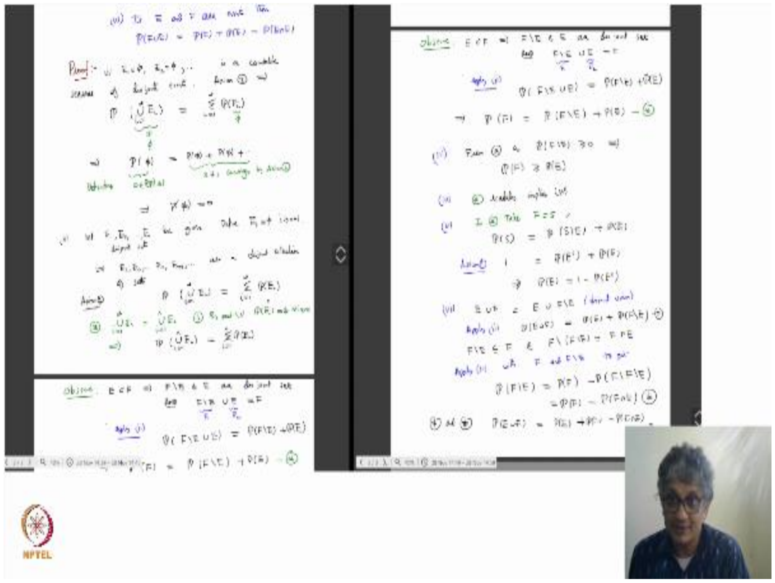
So, now once you have this you are in shape now, right, because you have this statement now, you also know that F minus E, sorry let me write that in black again, you also know that F minus E is contained in F and F minus E, F minus F minus E is the same as those elements in F which are not in, which are not in F minus E, that means they are in E also, that means this is just F and, so therefore, you are now in good shape, now again you have F minus E here, you have this guy here, right, so you can use apply 4, what does 4 say?

4 says that, exactly this probability F minus E is exactly this, so you apply 4, right, with F and F minus E to get probability of F minus E, F minus of F minus E is equal to probability of F minus probability of F minus E. So, to get you apply with F and F minus E to get probability of F minus E is the same as probability of F minus the probability of F minus F minus E, but that is the same as probability of F minus, this is just F and E.

So, putting let us call this as dagger and this as double dagger, so putting dagger and double dagger together what do we get? We get probability of E union F is probability of E plus this person, but this person right here was equal to this person, so same as plus probability of F minus probability F and that conclude it.

(Refer Slide Time: 21:15)





So, let us quickly recap, then make it double so show it to you, so let me go back here, so here is the, let me go back to the previous page, so we start from here, so it is what then make a little smaller so you can see the whole thing in one page, when you come here, here, let me make this little smaller and go to page 3 very nice.

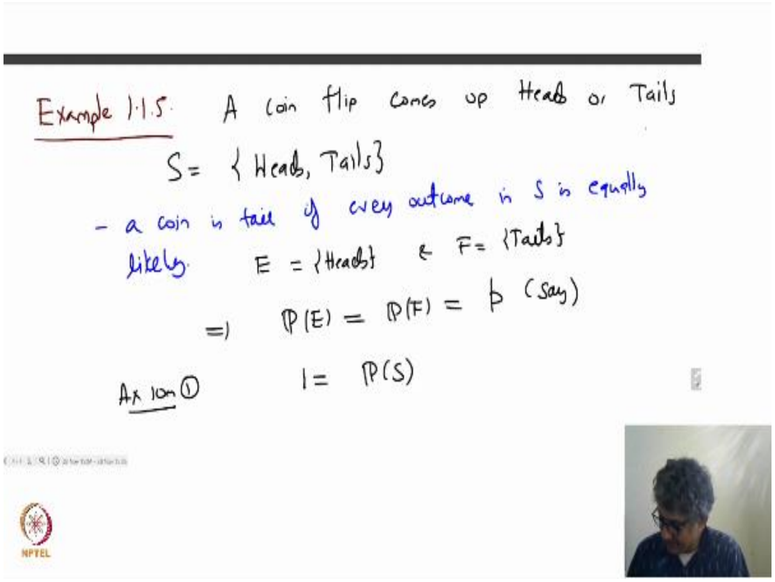
So left hand side clearly tells you that, what you want to show, you want to show probability of empty set of 0 nothing happens, that proof is quite straightforward all you do is take E1, E2 all empty sets accountable sequence and once you do that the axiom 2 implies the probability of countable union and sum of the probabilities and that immediately implies that probability of empty set is equal to the probability empty set, we (())(22:07) of both but the right hand side converges absolutely and we know that the probability of 0, 1 the only way the equality can be true is probability of empty set is 0, that finishes of proof of one.

For the second part, you have to just use part one, so what you do is you take E1, E2 to En be the given sequence of disjoint sets, then you define Ej to be empty for j beginning n plus 1, once you do that you apply the axiom 2 a probability, then you know that probability of the union i equal to 1 to infinity is sum of the probability of Ei's and once you do that you are in business because the countable union is just the finite union because after some time everybody is empty, after n plus 1 for surely and you also know the probability empty set is 0, so this should be probably empty set 0, I am sorry I made a mistake here, which I said empty set is 0, sorry so is equal to 0 for everybody, that means the right hand side sum is only a finite sum.

So, probability union is equal to the finite sum, then once you have this to show 3, 4 and 5 is fairly straightforward, what you do is you make the following observation that if E is contained in F, then F minus E and E are disjoint sets and probability, if you take E1 and E2 and apply 2 you get probability of F minus E union E, there is E1, E2 is probability E1 plus probability E2, once you do this you take your probability F on one side is probability F minus E plus probability of F probability E, using star 3 is immediate because this guy is non-negative, so probability of F is because of probability (())(23:57) that is what star said, 3 said because star is in 3.

4 is exactly what star is because you just get, you just bring it on this side probability F minus E is same as probability F, in 5 you just take probability E to be the whole space basis, for 6 you have to be a little bit careful but not that much you first observe that E union F is E union F minus E, once you do that, you apply 2 with these 2 sets the disjoint unions, so probability E union F is probability of E plus probability F minus but now what you do is you know F minus E is contained in F and F minus F minus E is same as F and E using these 2 facts you apply 4 with the sets F and F minus E, so you look at 4 you get F and F minus E, you get F, F minus E which is exactly what I get, so I plug that in back and star and dagger and double dagger and I get the answer.

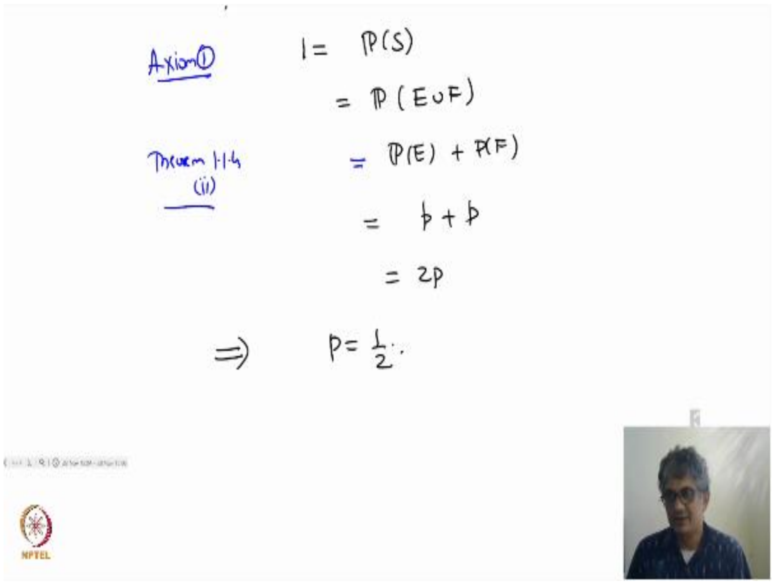
(Refer Slide Time: 25:13)



Now let me do a quick example, I will just close this lecture, 2 examples, so one example is the following, let us go back to our original example, let us call this, I will call this example 1.1.5. Here a coin flip comes up heads or tails, comes up heads or tails that means S is either heads, say heads or tails, so let us try and see how to model, let us say a coin is fair, if each of these outcomes equally likely, let us write that down, so let us say a coin is fair, if every outcome in S, outcome in S is equally likely. So, now we will try and see if we can assign our probabilities to head and tail.

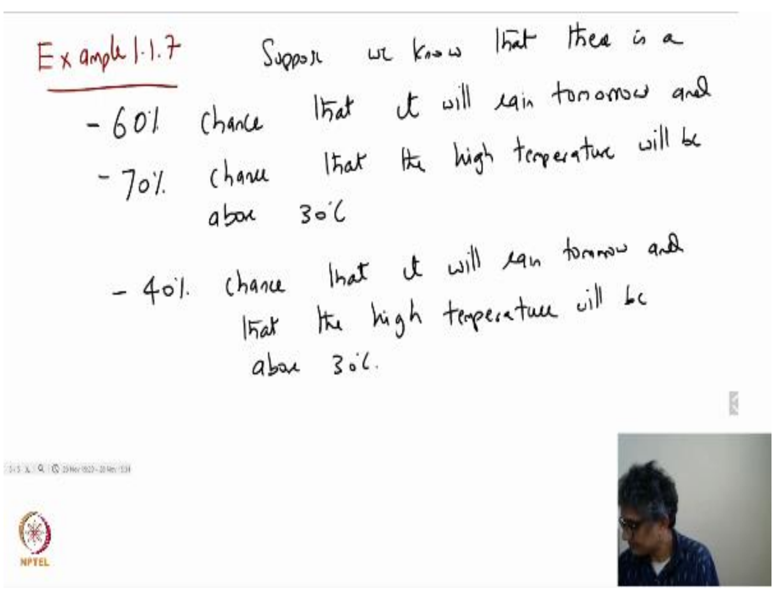
So, now let us say E is heads and F is tails, so they are both single outcomes, so we know that if they are equally likely this implies that let us say probability of E should be the same as probability of F should be same as let us say, let us call the value ps but you also know from axiom 1, that 1 is equal to probability of S that is maximum 1, but you also know that S is the same as E union F by definition. Let me write this in blue so it is clear, this is axiom 1, (())(27:21) but if you use this, if you use the this theorem what is it called, let me call it, what is the theorem called, the theorem is let us say,

(Refer Slide Time: 27:43)



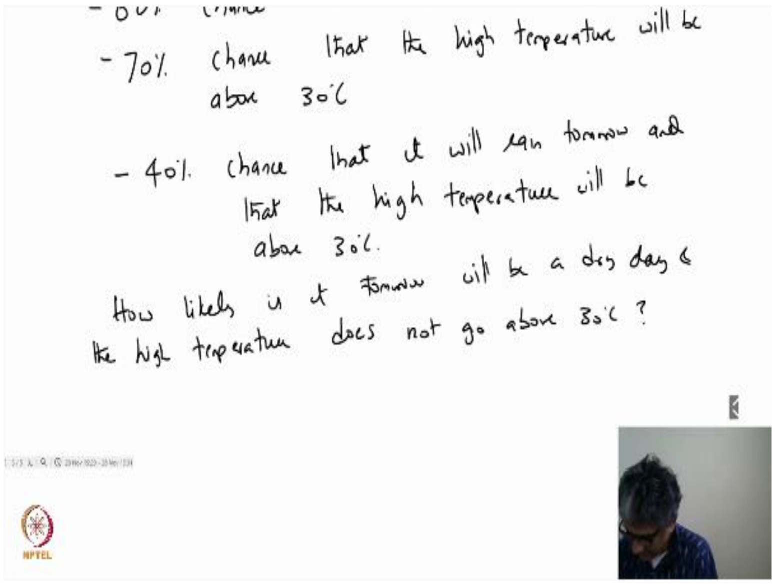
Let me go back up get the number down theorem 1.1.4 2 we come back down here, theorem 1.1.4 2 what do you get? This is equal to probability of E, so let me write down probably that is equal to probability E plus probability of F but each of these was p, this is equal to p plus p and that's equal to 2p, so if you run through the calculation you start off with one you end up at 2p this would be true only if p was equal to half, so that means if you have a fair coin the chance of getting heads is 50 percent, one simple operation.

(Refer Slide Time: 28:22)



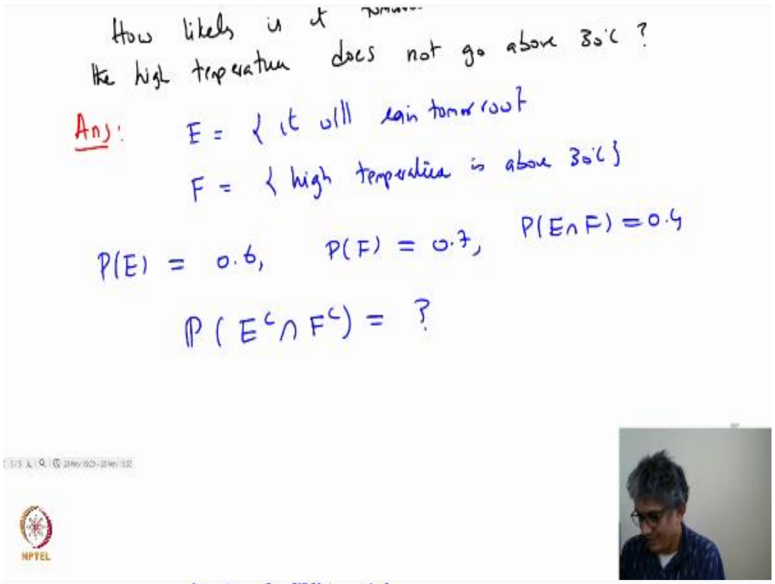
So, let me begin with an example, so you collect the, probably left off from last time so this example 1.1.7, suppose, we know that there is a 60 percent chance that will rain and 70 percent chance that the high temperature will be about 30 degrees and suppose we also know that that there is a 40 percent chance that it will rain tomorrow and that the high temperature will be above 30 degree, so I have these 3 facts.

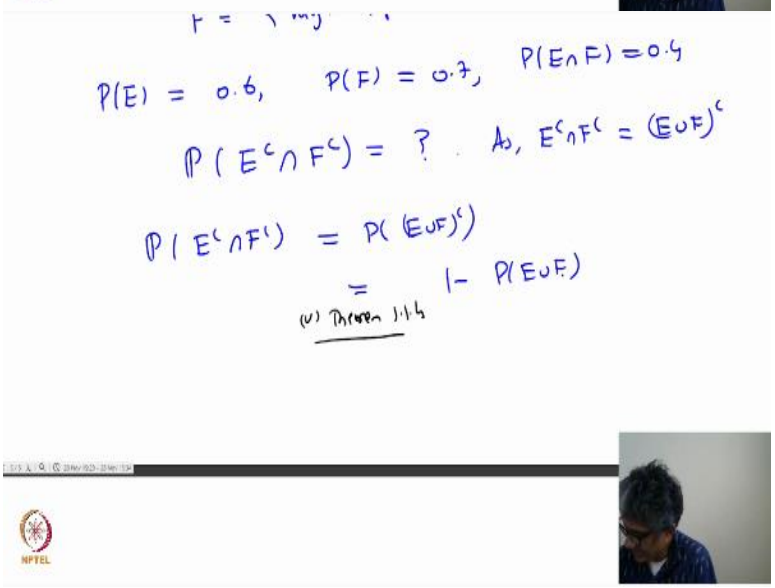
(Refer Slide Time: 30:18)



And the question I ask you is how likely is it tomorrow will be a dry day, there is no rain and how likely is it tomorrow dry day and the temperature, high temperature does not go above 30 degrees.

(Refer Slide Time: 31:13)





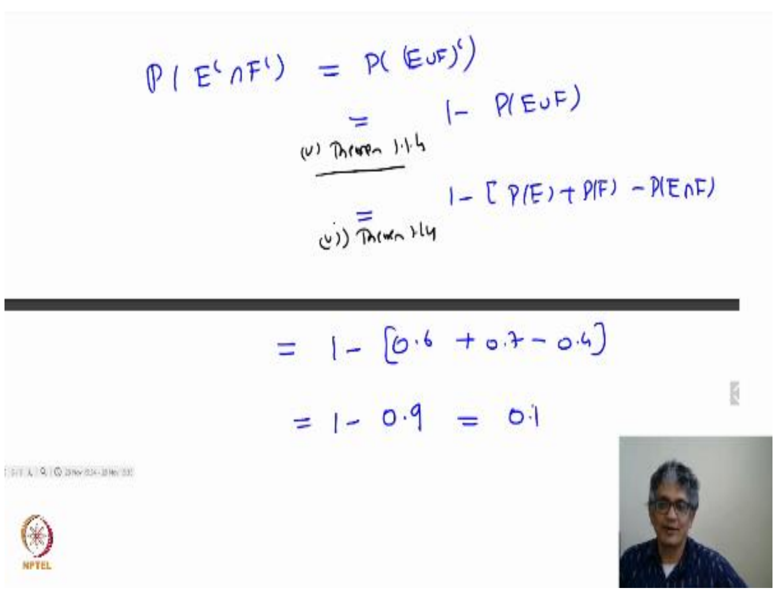
Suppose, I have to solve this question, so let me try and solve this question using the accidental probability and what we did last time, so what you do is you set up properly, you set at E be the event it will rain today, F be the event that the high temperature is above 30 degree, listing that, that are the 2 events E and F, so if you look at the 3 things that I have said you probability of E is 60 percent as 0.6, probability of F is 70 percent is 0.7 and probability of E and F is 40 percent.

So, what we have given is probability of E is 0.6, probability of F is 0.7 and probability of E and F is 0.4 and what are we asking that how likely it will be a dry day tomorrow and the high temperature go above 30, so what is the question asking you is what is the chance that it will be a dry day, that is E complement does not rain and the chance that F does not come up with F (())(32:30) this is what we are asking.

So, how does one do this problem? You look at this, so you get probability of E complement and F complement, that is the same as, so you observe that E complement and F complement is the same as E union F the whole compliment, that is the first observation, right, so as this happens then you have this is equal to this, F complement but this you know from the previous exercise I think it was 5 or let me write down in black so it is clear.

So, it is 5 of theorem 1.1.4, so this will be equal to from 1 minus the chance of probability E union F, now you can use 6 of the same theorem, to get the same as 1 minus inside this probability of E plus probability of F minus probability E and F, so what is that going to be equal to now?

(Refer Slide Time: 33:46)



That is going to be equal to 1 minus this is 0.6, this is 0.7 and this is 0.4, so the whole thing is 0.9, so in fact it is quite high that it will not rain, so this is 1 minus 0.9, so 1 minus 0.9, so the answer is 0.9, so it is quite small that it will not be a dry, it will be a dry day and it will not go up above prediction. Yes.