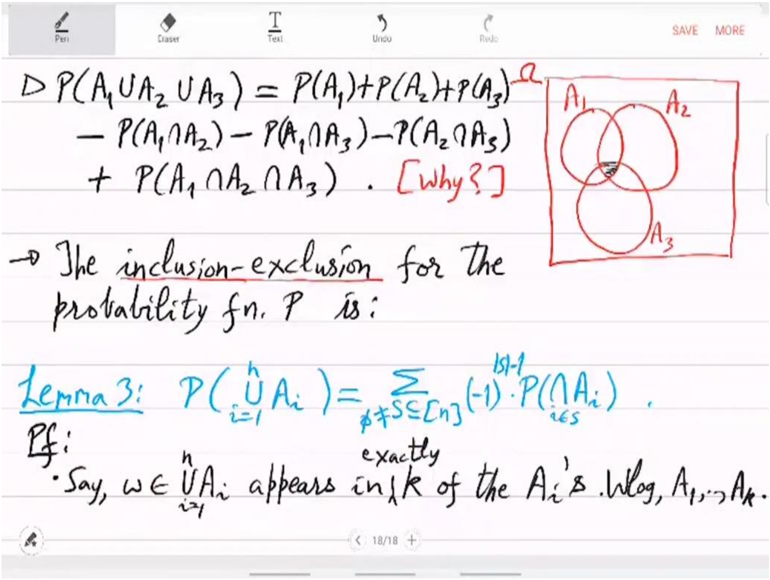
Probability for Computer Science

Prof. Nitin Saxena
Department of Computer Science and Engineering
Indian Institute of Technology - Kanpur

Module - 1
Lecture - 4
Inclusion-Exclusion Principle

(Refer Slide Time: 00:23)



The question is, what about union of 3 events? So, probability of A union B union C. So, for that, the picture is, you have A 1, A 2, A 3. So, that is omega, A 1, A 2, A 3. Now, the problem is that there are now many more intersections possible. It is not just double counting of A 1 intersection A 2, but it is 3 kinds, 2 at a time and 1 kind, 3 at a time. So, there are 4 more possibilities of double counting. So, we have to take care of that.

But you must have seen this calculation before. It is called inclusion-exclusion. So, probability of A 1 union A 2 union A 3 is by inclusion-exclusion. So, first you do the normal thing. So, probability of A 1 plus A 2 plus A 3. And this will be true if A 1, A 2, A 3 were completely disjoint, mutually disjoint. But since they are overlapping in general, so you have to now subtract the double counting. So, what are the double counts?

It is for A 1 and A 2; it is for A 1, A 3; it is for A 2, A 3. Is this correct? So, actually, this would have been correct if this part in the centre was not there. The, an element belonging to all 3, if that element was not there, then we can stop at this formula. Otherwise, what we have done is an element omega that is in all 3, that has been removed from the calculation, because it is counted 3 times, then it is subtracted 3 times. So, we have to add it back.

And this is now correct. So, show this as an exercise. Why is this correct? So, this is actually correct, because of the inclusion-exclusion. So, the inclusion-exclusion for the probability function in general is: Let us write down the lemma first. So, probability of n events, their union; n is at least 1. So, you already know the result for n = 1, n = 2 and also n = 3. What is the general form?

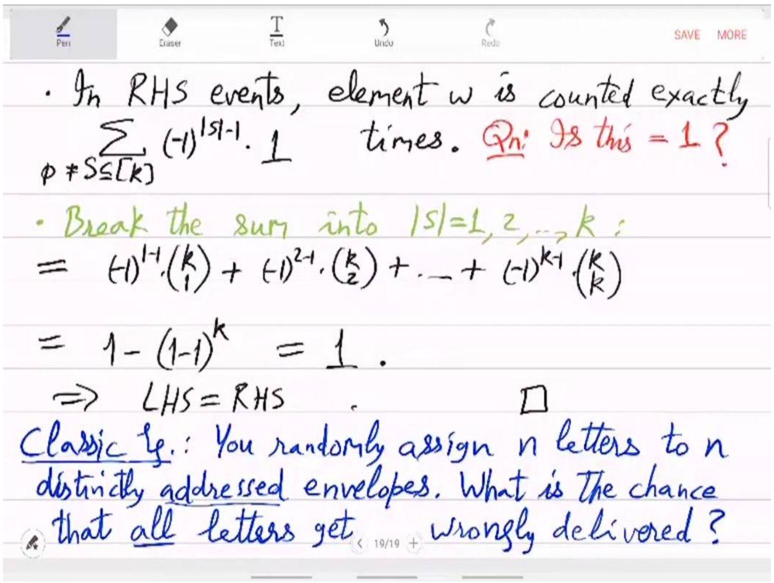
So, it is easy to see that in general form, you will have to look at all possible intersections. You have to subtract them. Then, you over subtract, so, you have to add them. Then you over add, so, you have to subtract them and so on. So, you get a formula like this. Go over all the subsets of 1 to n. Go over all the non-empty subsets. And the contribution will be intersection on these Ai's; intersection Ai for all these i in S.

But of course, you are not just summing it up, you have to sum it up. Sometimes you subtract, sometimes you add. So, there is a sign. And that sign, you can guess is -1 to the S minus 1, size of S minus 1. You can do a sanity check for n = 3. This is correct. You get this F or size of S = 1; you get the first part. Then, for size is equal to 2, you get the second part. Size is equal to 3, you get the third part. And clearly, for n = 2 also, hence it is correct.

So, what is the general proof of this? So, a general proof is simply by looking at an element and seeing that it is correctly counted in RHS. So, say element omega appears in k of the Ai's. Omega will appear in at least one, and say it appears in exactly k of the Ai's. And without loss of generality, you can assume the first ones, A 1 to A k. Why is this without loss of generality? This is simply because of the symmetry of union.

If it is something else, then you can just relabel it, call it A 1 to A k. Both left-hand side and right-hand side in this lemma statement are symmetric, with respect to this operation. And next is:

(Refer Slide Time: 07:20)



In the right-hand side events, omega is counted exactly how many times? So, you have a sum over the subsets, subsets of 1 to n, right? But then, now, beyond k it will not matter, so, subsets 1 to k; non-empty subsets, -1 to the S minus 1. And intersection of these Ai's, in which, each of these Ai's, omega is present. So, that count is only 1. It is counted these many times. That is the expression. Now, this seems to be a complicated expression.

We have, this may even be negative, it may be 0 or it may be positive. What we want to show is that, this is exactly 1. That is the question, is this 1? Because, if it is not 1, if it is anything else, then the count on the RHS of omega is wrong. So, then the formula will be wrong. So, in other words, what you have to, or it suffices to show that this count is 1, for the lemma. Break the sum into size 1, 2, … k.

So, when you do that, what you will get is: For size S = 1, you have k choose 1 many contributions. So, you get -1 to the 1-1, k choose 1. For size S = 2, how many contributions are there? how many summons are there? That is k choose 2 with sign 2-1. And this goes on till size of S = k, in which case you have sign -1 to the k-1 and k choose k contributions. So, what is this expression?

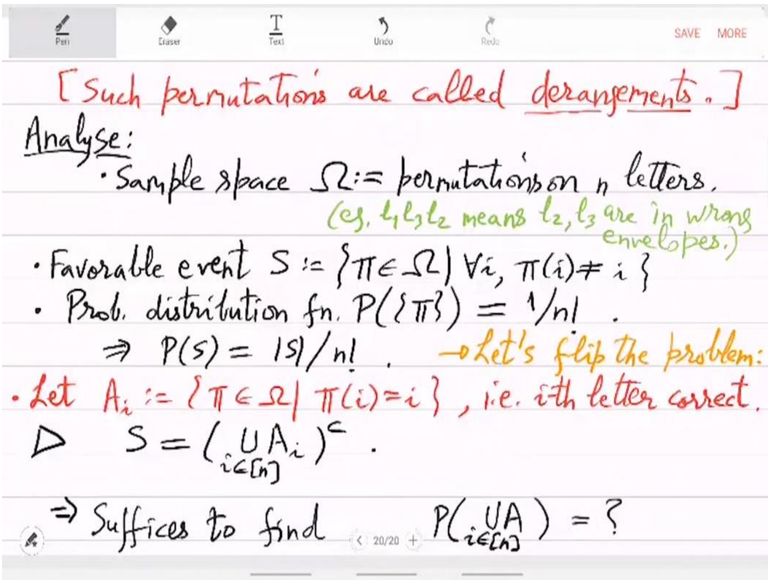
This expression is exactly 1 minus 1-1 to the k, by binomial; if you look at the binomial expansion of 1-1 to the k, you will get a similar expression, which is 1. So, omega has been counted exactly once in the RHS. Hence the formula is correct. So, this means that LHS = RHS. So, that is the principle of inclusion-exclusion. It works both for set union and for probability of that.

Now, this is a very useful formula and you can see its use immediately by a very classic example. So, let us do that next. Or, suppose you randomly assign n letters to n distinctly addressed envelopes. So, there are n letters; each of these is addressed to a different person or location; and accordingly you have the envelopes. Now, obviously, you want the letters to be sent to the correct person, correct address.

But we will be asking the opposite question, what is the chance that every letter is misplaced, wrongly delivered? So, what is the chance that all letters get wrongly delivered. So, this is a very interesting, very old example. It is a classical example. If you try to do this directly, it seems pretty hard, because the way we modelled it before, we modelled by permutations. So, what this question is asking is, how many permutations are there where every entry is in a different position.

The i'th letter is not in the i'th place, it is somewhere else. So, what is the chance of that? So, that seems mind boggling. It is the space is also exponentially large. So, how do we do this? It is a counting question and we will solve it in the probability format.

(Refer Slide Time: 14:17)



So, this, by the way, such permutations are called derangements. So, this calculation will also be count on derangements on n elements. So, let us do the analysis. So, sample space is a permutation on the n letters. I should say permutation. So, every possible permutation on n letter; so, it is n factorial size. That is the sample space. For example, if you look at letter 1, letter 3, letter 2 permutation, then here, letter 1 was put in the right envelope, the first envelope, but letter 3 has been put in second, letter 2 in third.

So, second and third are wrong. They will be wrongly delivered; means, l 2, l 3 are in wrong envelopes. That is the meaning. That is how you model this practical problem as an, using the abstraction of permutations. Next is events. So, events are just subsets of omega. And you are interested in the favourable event S, which are permutations pi, where everything is wrong. So, for all i, pi i is not equal to i; pi is a derangement.

These are your favourable events for this calculation. You are interested in those assignments where every letter was wrong. Now, let us move to the probability distribution. So, probability distribution function P. So, on a set, singleton; so, for a permutation, what is the value? Since every permutation is equally possible and you have this axiom that sum of the probability should be 1, so, you immediately get 1 over n factorial.

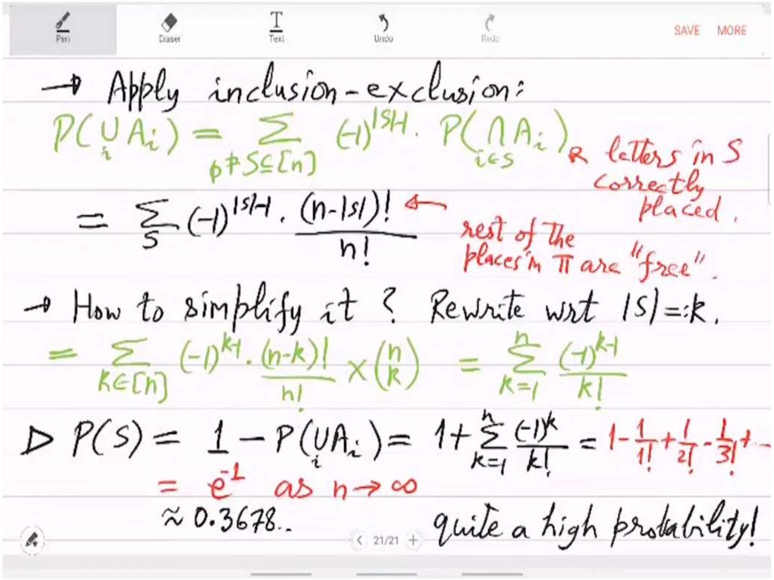
And which means that probability of a favourable event that we are interested in is size of S divided by n factorial. This again follows from the definition, the axiom of probability. So, that is the modelling, but how does it help to actually compute this, either P of S or size of S. Both look equally hard; there is no simplification yet. So, simplification will come if we look at inclusion-exclusion principle. So, let us define event Ai.

So, those permutations that work for i, that are correct for i, pi i = i. That is, i'th letter is correctly delivered. Obviously, we do not want this to happen, but since the original problem looked hard, we have flipped the problem. So, this is what we are doing. So, let us flip the problem, which means that, let us look at those permutations and hence that event where I was correct. We are not saying anything about others.

We are only making a statement or putting a condition on the i'th letter. So, that is correct. These are the events Ai. And do this for all i 1 to n, for every letter. So, now, what is S in terms of Ai? So, look at the union of Ai, which means that at least some letter is correct, and take the compliment, which means that all the letters are incorrect. So, now we start getting ideas. So, now, probability of S gets related to probability of union.

And then on probability of union, we will do the inclusion-exclusion. This means that suffices to find the probability of the Union. Because, once you have that, 1 minus that value will be your answer. So, let us apply inclusion-exclusion.

(Refer Slide Time: 20:37)



So, probability of union is, go over the subsets, non-empty subsets, put a sign and the probability of this intersection. So, what is probability of this intersection? So, here you are saying that i in S, these letters, letters in S, they are correctly placed. So, that is what it means. This part of the permutation is already fixed then. What remains? n minus S. S complement is what remains. And there, anything is allowed.

So, which gives you the idea that the probability has to be n minus size of S factorial divided by n factorial. That is what you get. So, again, the reason for this is; so, rest of the places in pi are free. So, how many permutations are there were these n minus S places are free? Obviously n minus S factorial. So, again, here we are using the axiom of P of S or at the axiom of P on the events. And we are also using that each permutation is equally likely.

So, you get 1 over n factorial times this. But this still looks like a complicated expression. How do we simplify? So, we will do something similar to what we did before, which is rewrite this sum according to size of S. So, rewrite with respect to size of S = k. So, what you get is sum k 1 to n, because it is a, S is a non-empty subset of 1 to n. So, k goes from 1 to n. Sign will be -1 to the k-1 expression.

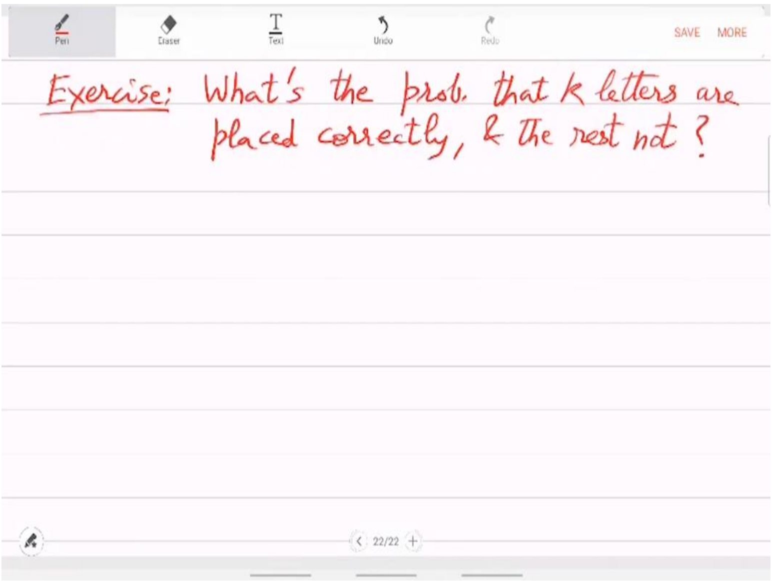
The summoned is easy to write. It is n-k factorial divided by n factorial. But how many such subsets are there of size k. That is n choose k. That is your expression. That is the simplification. Which is equal to all k -1 to the k-1. And n choose k is n factorial over k factorial times n minus k factorial. So, actually, things very nicely cancel out and you get this expression.

So, you get this alternating sign 1 over k factorial expression, which looks beautiful. And this should have; you would get the feeling that this should have a nice value, and indeed it has. So, it is related to 1 over e. So, that is what you get. So, you get that probability of S is equal to 1 minus this thing, which comes out to be 1 plus sigma k equal to 1 to n minus 1 to the k over k factorial, which is equal to 1 minus 1 plus 1 over 2 factorial minus 1 over 3 factorial, so on.

So, what is this expression? This says e to the -1 as n tends to infinity. So, obviously, right now, you wanted an expression for general n, so, that expression you have. But to understand the value that this expression takes, when n is large, say n is more than 10, then it is very close to 1 over e. And what is 1 over e? So, 1 over e is 0.3678. So, this shows that the probability of every letter being misplaced. It is pretty high; it is close to 40%.

Close to 40% is the chance that every letter will be misplaced, even if you have millions of letters, which is very surprising. It is surprisingly high value. So, quite a high probability. But this, nobody would have guessed this in advance. In advance, you would think that, a priori you would think that, well, how can it be that every letter gets misplaced? But you do this calculation, this elaborate calculation, and it is related to 1 over e. So, this is a very nice example. It shows you how inclusion-exclusion helps simplify life. And I will end with this question exercise for you.

(Refer Slide Time: 27:26)



So, repeat this calculation. So, what is the probability that k letters are placed correctly and the rest not? So, this is a simple extension. What is the probability that; so, k is given to you and n is given to you. Now, as a function of k and n, solve this.