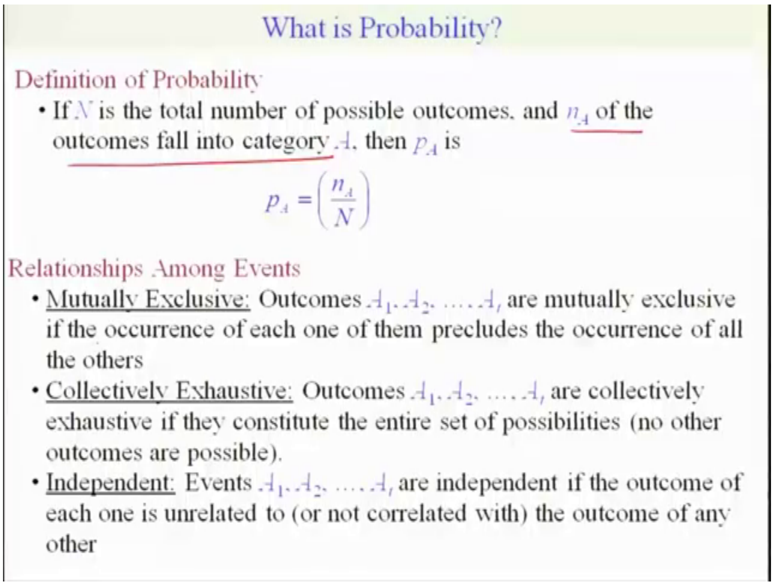
Thermodynamics of Fluid Phase Equilibria
Dr. Jayant K Singh
Department of Chemical Engineering
Indian Institute of Technology, Kanpur

Lecture - 32
Probability abd Multiplicity

Welcome back. We will start a new topic a in this course ah; that is on statistical mechanics. I will not be able to cover the whole range of statistical mechanics in details, but what we are intend to do, is to learn the basics of that. So, that we can appreciate some derivations in later part of this course ah, which are based on statistical mechanics

Now, in order to appreciate statistical mechanical route for connecting the thermo dynamical properties from molecular pictures or molecular information, we must first get a quick recap of mathematical operations; such as the probabilities and using the probabilities to connect to the thermodynamic relations. So, we will start with first probability. So, what is probability?

(Refer Slide Time: 01:10)



By definition if N is the total number of possible outcome and nA of the outcome falls in category A, then a probability of outcome falling in category A is given simply by small nA divided by capital N. This is something which we are all aware of. And I do expect that since as an engineering student, you will have gone through all this exercise in the earlier part of your course and hence they should not be difficult for you to understand it.

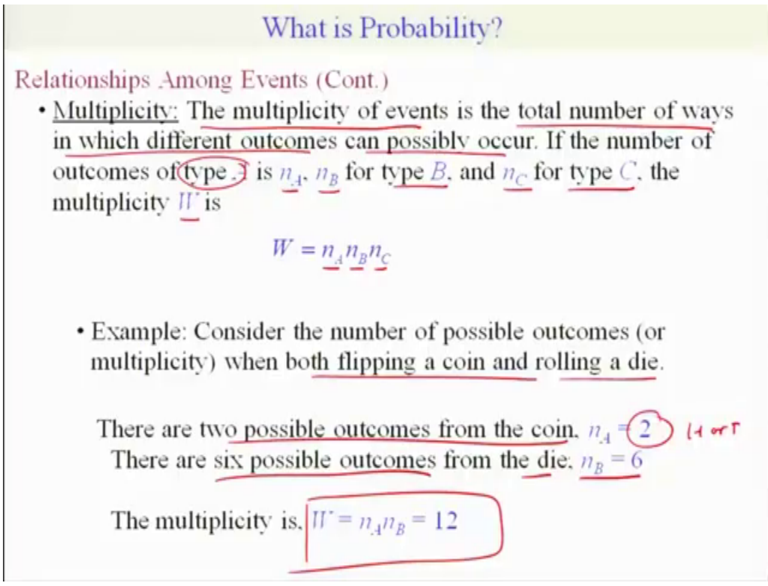
Nevertheless my intention is basically to give a quick recap of the terminology which we are going to use it, and later we are going to make use of this to explain the maximization of multiplicity, in order to connect to the thermodynamic properties. So, in order to get there, we must first cover the basic concepts, as far as this property of A is concerned.

We know this, but then their event can have various different relations and we are aware of this relation; such as mutually exclusive ah. When we say the outcomes A 1 to A t are mutually exclusive, it means simply that if the occurrence of each one of them preludes the occurrence of all others, what does that mean? It means that if A 1 occurs, then none of two till t outcomes can occur.

So, once A 1 is there ; that means, this is the only possible one; that means, this cannot occur at all ok; that is what the words are precludes means. Now in addition we can have collective only exhaustive, which essentially means that the outcomes A 1 to A t, if they are collectively exhaustive; that means, they constitute the entire set of possibilities ; that means, no other outcomes possible other than A 1 to 80.

In addition we can have independent events. These are the events which if they are called independent, if the outcome of each one is unrelated to or not correlated with the outcome of any other. So, these are the possible relationships events, mutually exclusive, collectively exclusive independent. And of course, there could be dependent events also.

(Refer Slide Time: 03:51)



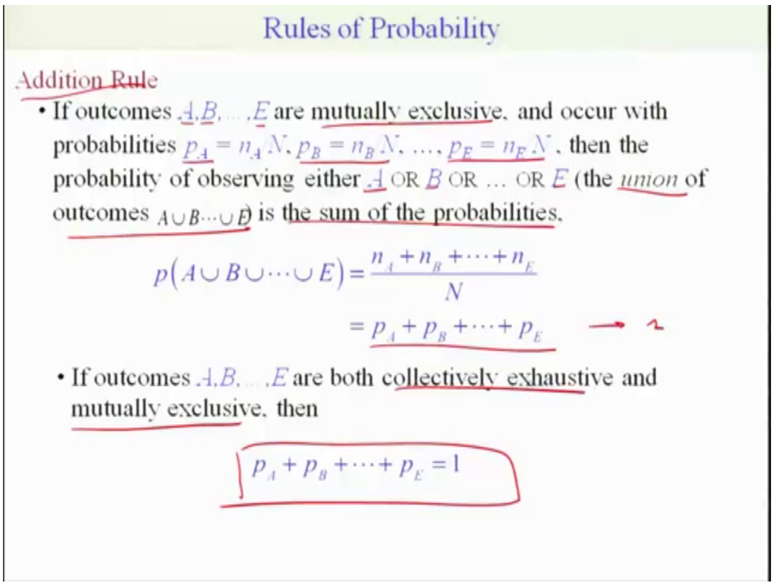
Now, making use of this relation, we can also define the multiplicity. Now the multiplicity of events is the total number of ways in which the different outcome can possibly occur. So, for example, if the number of Outcome of type A event is nA, for type B is small nB, for type C is nC then the multiplicity W having events of type BC A is simply the multiplication of this number outcomes nA nB nC ok.

So, this would be the total number of ways. You can have these events in this form. Now consider the number of possible outcome, when both flipping a coin and rolling a die ok, these are independent events right. Flipping a coin is a completely unrelated to rolling a dice.

Now if you consider the possible outcome of the coin, there are two possible outcome, which is 2; either you can get head or tail and if the die has 6 phase, and let us say assume that is to be fair, then or if the die has 6 phase; that means, you have 6 possible outcome, which is going to be 6.

In that case the multiplicity of such a events, whether when you flip a coin or roll a dice is simply the multiplication of nA and nB ok. So, this is the total number of possible ways you can arrange all these outcomes of coin and the die ok.

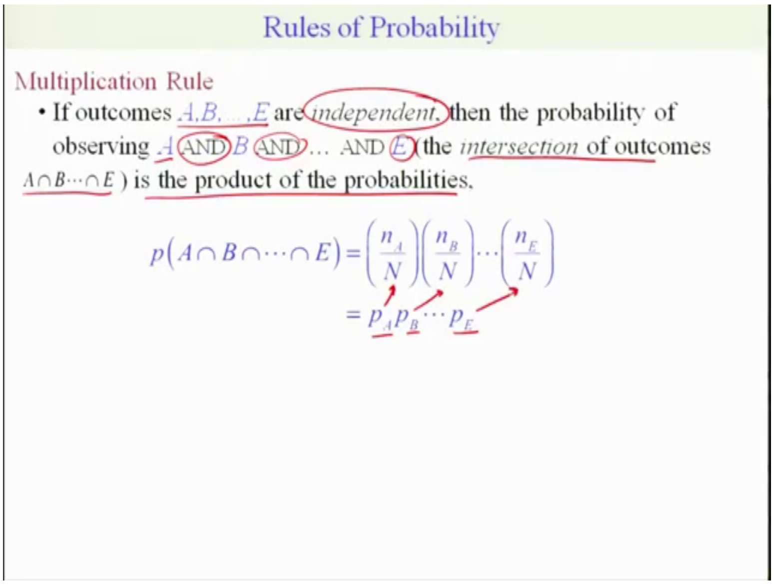
(Refer Slide Time: 05:51)



So, now I am going to just quickly go through the addition and the multiplication rules. If the outcome A B to E are mutually exclusive ok; that means, if one occurs this cannot occur, and they occur with the probabilities p A which is nA by capital N pB which is nB by capital N pE, which is nE by capital N, then the probability of observing either A or B or E; that is the union of outcomes is simply the sum of the probabilities.

. So, simply the sum of the property that is what the addition rule would be. Now if it is collectively exhaustive and as well as mutually exclusive, it simply means that the sum should goes to 1, because that is the only possible outcomes are there. So, this would be the case when these outcomes are collectively exhaustive and mutually exclusive.

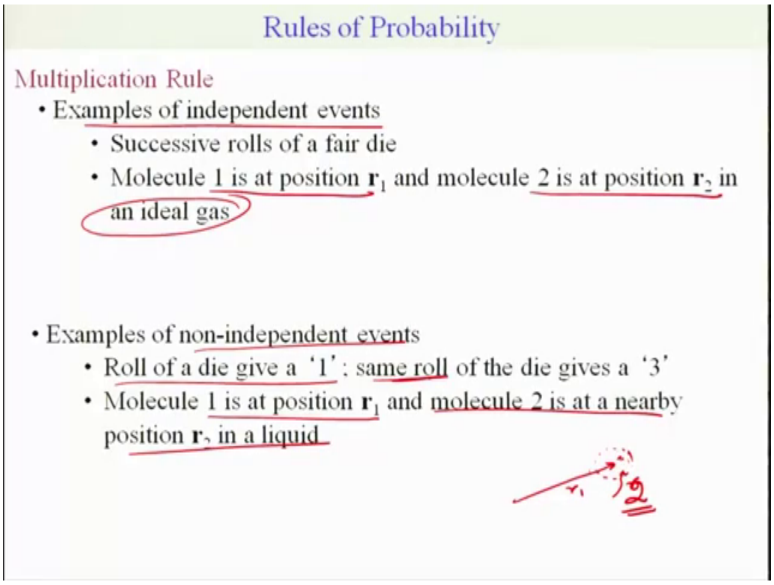
(Refer Slide Time: 06:57)



In addition we have this multiplication rule. If the outcome A to E are independent, then the probability of observing A. and now look at this here way putting and A and B and C and so forth till E, which is nothing, but the intersection of outcome of A given in this way.

In the in the mathematical term is the product of the probabilities, which is simply pa multiplied by pB multiplied by pE, and pE is given by this.

(Refer Slide Time: 07:37)

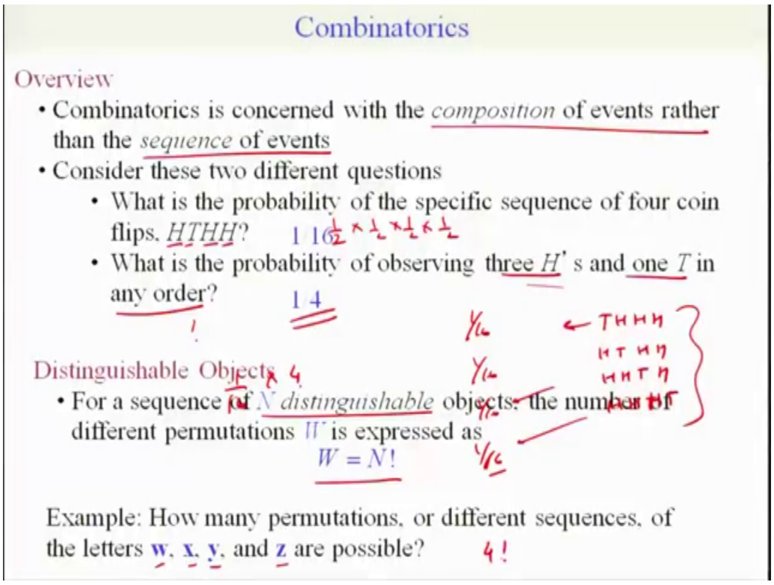


So, this is the multiplication rule here ok. So, what are the example of independent events. So, for example, successive rolls of a fair die are independent ok. There is nothing to do with the previous roll or molecule 1 in position r 1, and molecule 2 in position r 2 in an ideal gas.

When they do not interact at all is completely independent, which means basically the molecule 1 in position r 1 is not related correlated with molecule 2 in position r 2 and, but the non independent events could be an example, would be the roll of a die gives A 1 and the same role of a die gives 3. So, the same here refers to this, the role which has been done here ok.

So, this is like dependent event here or molecule 1 is position r 1 ok, could be let say this is a reference origin and this is molecule 1, and you are asking that molecule 2 is at a nearby position r 2 in a liquid. So, this molecule 2 is is within the vicinity of molecule 1, where this molecule 2 is. So, in this case, this becomes a non independent event, it depends on the position of r 1 and the vicinity of it. In addition the combinatorics are used in probability theories.

(Refer Slide Time: 09:04)



So, let me just go through the overview of that. Combinatorics is concerned with the composition of events rather than the sequence of the events. Now consider two different questions for example, what is the probability of the specific sequence of 4 coin flips. So, specific sequence is H T H H. Now we know that the probability of obtaining the head or tail is 1 by 2.

So, in this case there is a sequence of 4 flip and each one has a probability 1 by 2 is. So, you are going to multiply 1 by 2 and 2 1 by 2, and these are independent event ok, and this should give you 1 by 16. Now the second question is, what is the probability of observing 3 H and 1 T in any order ok. Now this is another question.

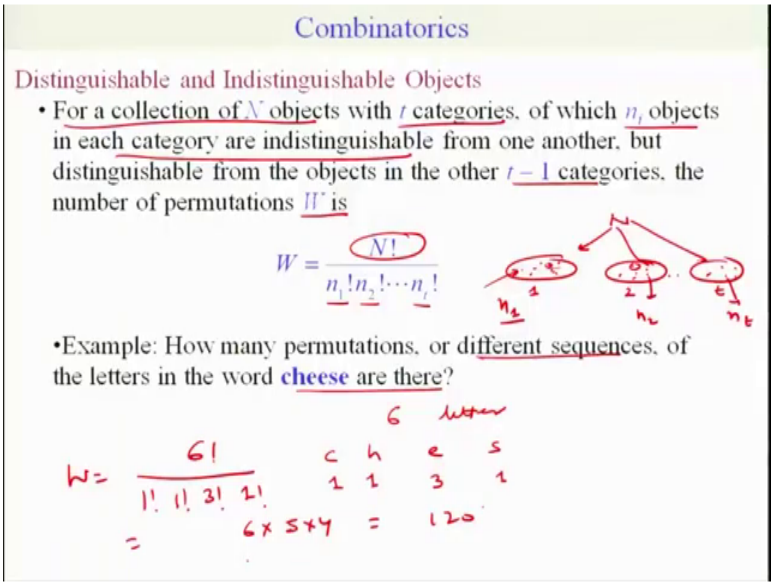
Now in this case, we are not talking about a specific sequence here. So, which means basically if you re-arrange this, you can do that in. So, this is the probability and then you can rearrange this in many ways, and essentially you can rearrange in four ways. So, you can have T H H H H T H H H H T H and at H H T H H T. So, there are four different ways.

So, each one is 1 by 16 ok, and you can add this. This will give us 1 by 4. So, 1 by 16 multiplied by 4 ok, this was going to give you the answer 1 by 4. So, we will we will discuss this, how you the can solve this particular one using permutation or or binomial distribution.

Now there is some the objects or the events or many times are distinguishable and non distinguishable. So, let us consider object which are distinguishable for a sequence of N distinguishable objects. A number of different permutation is simply expressed as N factorial.

So, for example, if you have a distinguishable that is such as y w x w y w x y and z, then how many permutation of different sequences of the letter can be possible, its considering as a distinguishable. You can simply say its 4 factorial, but if it is not distinguishable.

(Refer Slide Time: 12:05)



Then of course, the other issues which we have to take into account. So, in order to express that we will combine this distinguishable, and in distinguishable object together, let us consider collection of N objects with t categories of which ni objects, in each category are indistinguishable from one another ok, but distinguishable from the objects in the other t minus 1 category.

What does it mean, that you have N objects which can be segregated in different categories, and in each these are not distinguishable from each other, but this one is distinguishable from with the other one. So, this is let say category 1, this is 2 and this is t and in each one you have n 1, the object in these cases n 2 object, and is this and t object.

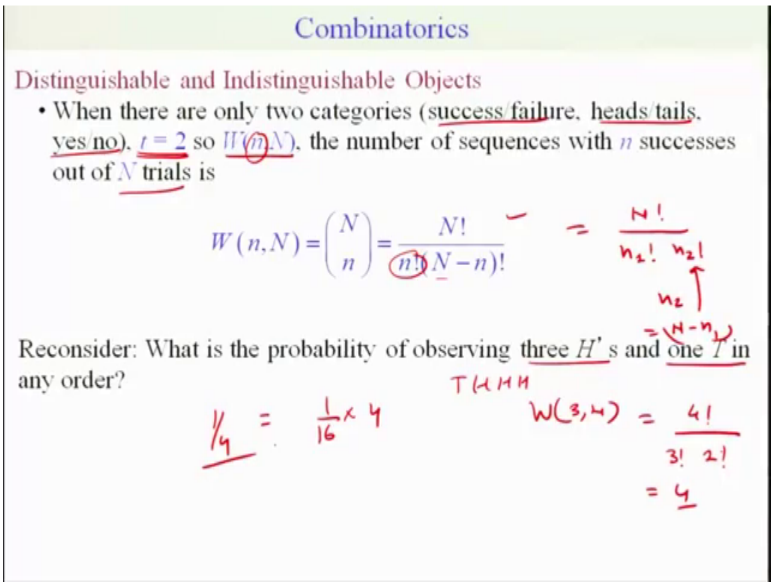
So, n 1 is of type 1 and 2 is of type 2 and so forth nt is of type t. Among nt all are of same type, in this cannot be distinguished from each other, can, they cannot be differentiated from each other. On the other hand an object from t category can be differentiated from object from two category for example. So, in such case what is the permutation ; that means, how many ways you can arrange such an objects.

So, if it is distinguishable, this would have been simply N factorial, and then since these are the groups which are, can be segregated, and among these, these are indistinguishable. So, we are going to divide it so, that we can correct this number of possible ways. So, in this case, it will be n 1 factorial divided by n small 1 n 1 factorial small n 2 factorial and so forth, which is to avoid or to (Refer Time: 13:08) the over counting.

So, this is a permutation of the W, in such a case. Now we can express this understanding in an example. So, for example, how many permutations or different sequences of the letters in the word are there. So, if you look at it, here you have 1 2 3 4 5 6. So, there are 6 letters, and then you have category c h e s this is 1, this is 1 e.

There are three of them and s, there are one. So, in this case you have W as 6 factorial divided by 1 factorial 1 factorial 3 factorial 1 factorial. So, this is the number of possible way which is going to be 6 into 5 into 4 120. So, these many permutations or different sequences or the letter you can arrange it.

(Refer Slide Time: 15:18)

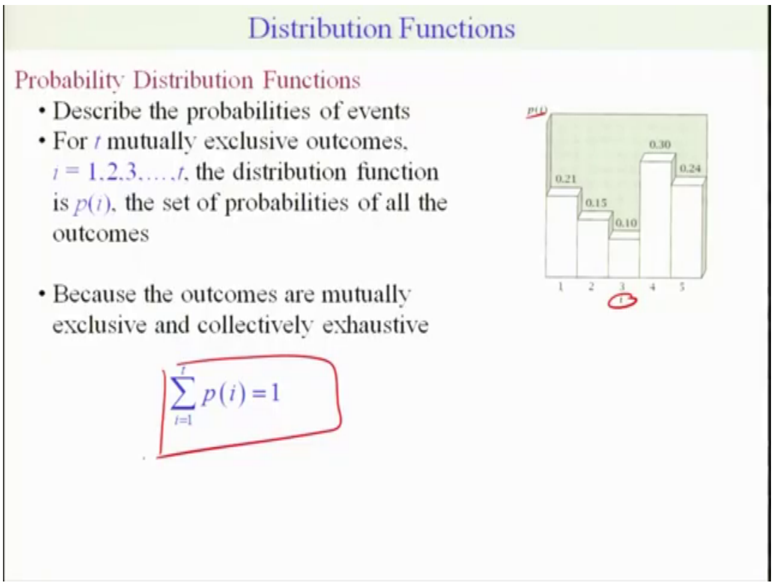


So, let us consider the case where you have only two categories; such as success failure heads, tail yes or no. So, in that case if t is only 2; that means, you can write W small n, which corresponds to a success; that means, the number of sequence with n success out of N trials can be written as simply N factorial divided by N, N factorial multiplied by N minus n factorial.

So, it could have also written something like this N factorial n 1 and n 2, but n 2 is nothing, but N minus n 1, and as you can tuck in here to get this expression. So, this is relevant when you have only two categories ah, outcome of the events. Now let us go back to the earlier question where we said we wanted to find out the probability of observing 3 H and 1 T in an any order.

So, we can actually do the same exercise. So, we have this, 1 by 16 was the probability of getting T H H H, let us assume that and then you can multiply by by having the number of such possibilities, which should be simply W 3 4, which is nothing, but 4 divided by 3 factorial and this is 1 factorial, which is nothing, but 4. So, this is, this would have given simply 1 by 4; the same answer which we got earlier ok.

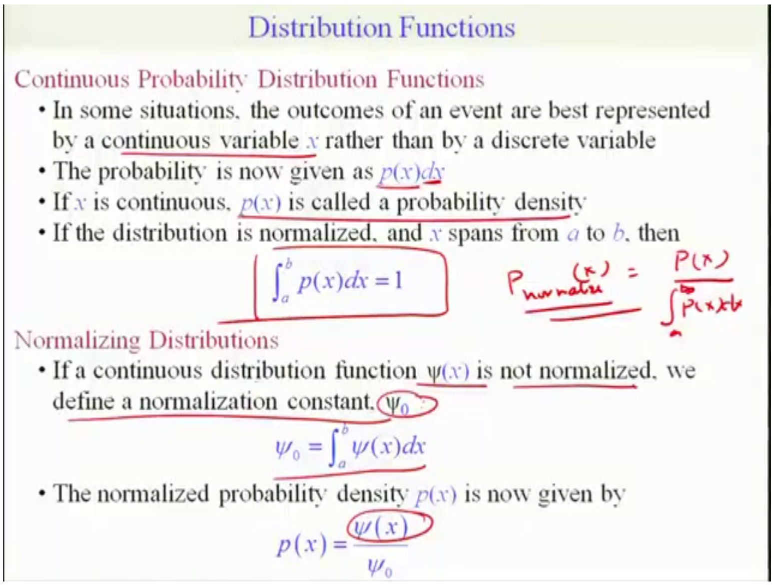
(Refer Slide Time: 17:06)



So, what we have discussed this combinatory, we are going to extend this exercise and express the probability and this form, where this is pi as a function of various different event outcomes. So, I could range from; let us say 1 to 5.

So, let us assume this to be a mutually exclusive outcome ok, and these are the numbers or the values of such an outcome. This is an arbitrary value for a given set of events, which we know ok, just for the sake of examples. Now because the these are mutually exclusive and collectively exhaustive; that is some of this should yield one which would be the case if you add it up here.

(Refer Slide Time: 17:56)

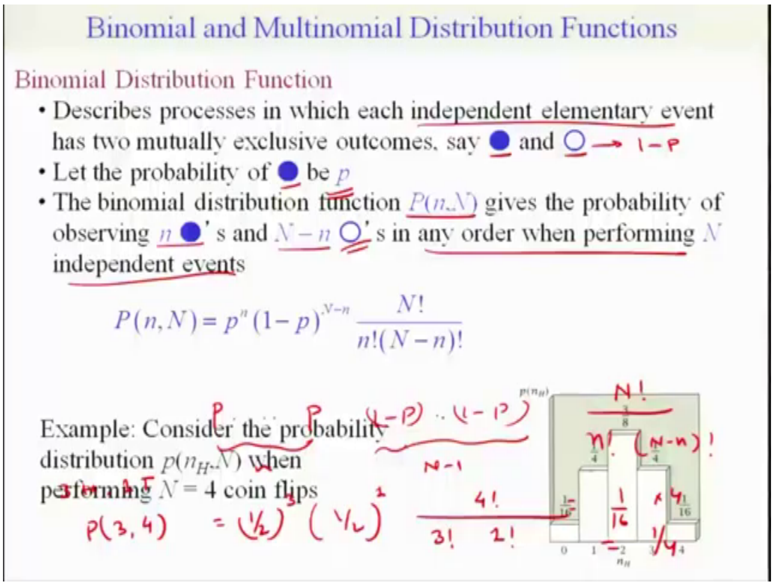


Now, this was a very discrete event for a continuous probability distribution function which would be in many case, this becomes a the case, this is usually where the system is not ah, cannot be categorized in some discrete events. In that case we usually give a probability in a continuous form using a variable x here.

So, probability in that case is the px multiplied by, there is the differential element of x. If x is continuous then we call it px as probability density. Now the probability density, in general the distribution, the probability distribution should be normalized in a, for a given range, and hence you need to make sure that integral of px from a to b, it should be 1 and thus you, if you want to normalize P, normalize value should be. P of x divided by the integral from a to b.

So, this would be a normalized version of it. Now that is what we are saying here normalizing distribution. If a continuous distribution function, let say is not normalized we define the normalized constant which is simply the integral version of this or in integral of psi x from a to b, then we normalize of this function by by dividing psi of x divided by this constant psi 0. So, its the very straightforward exercise. Now we will extend this now this distribution, this continuous distribution function form binomial and multinomial distribution.

(Refer Slide Time: 19:45)



So, let us first discuss binomial distribution. Here we are interested in only two events which are this, mutually exclusive outcomes. So, let us consider here in which independent elementary events are filled circle and empty circle. These are mutually exclusive outcomes. So, let us say for this filled circle event the probability is p.

Now since it is a mutually exclusive, whatever the probability of this can be now written as 1 minus p, because summation of this 2 probabilities should be 1, considering is 2 is to be mutually exclusive and as well as exhaustive ok, collectively exhaustive. So, now the binomial distribution is given as P small n and N is total number of outcomes, and its the outcome of specific type.

Now which we are interested in this case, is the outcome of the events having fields circle ok. So, the binomial distribution gives the probability observing N filled circle and n minus 1 empty circle in any order, when performing N independent events. So, the way we write, is the following that since this is N, small n, the probability of this is going to be P multiplied by P multiplied by P of n times ok.

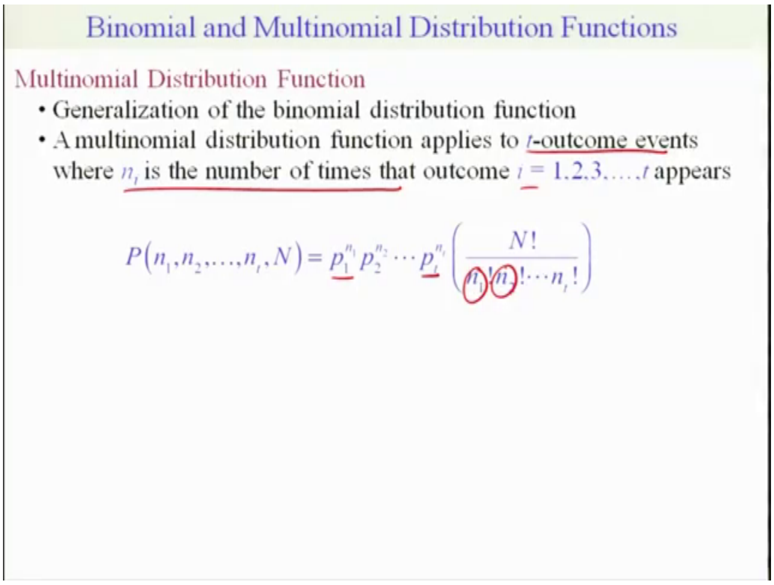
And similarly for the case of empty circle to be 1 minus p n minus 1 ok, considering these are independent events, but now you can arrange this in many ways. So, using their earlier expression of the permutation we are going to consider, these are only two events. So, its going to be N factorial divided by the small n, the number of events or the objects of type circle would be small n factorial multiplied by n minus 1.

The remaining one are of type empty circle. So, this becomes a simple binomial distribution function. Now if you go back to the original question that how many ways or what is the probability of observing 3 heads and 1 tail. So, this will be. So, if you go back to the original question of 3 heads and 1 tail, the probability of head or tail is 1 by 2. So, in this case let us say this is p 1 by 2 and we are talking about 3 head.

So, is P 3 4; that means, 3 heads and 4 is a total number of events here. This would be 1 by 2; again 1 minus p is 1 by 2 n minus 1 is 4 minus 3, which is 1. This is 4 factorial small n is 3 factorial and this is 1 factorial. So, this is going to be 1 by 16 multiplied by 4 is 1 by 4.

So, you get the same expression, as we have discussed it earlier using this binomial distribution function. Now similarly we can find out by changing these 3 to 2 3 to 2, 3 to 1, 3 to 4 and this kind of distribution we can get ok. So, I will not go in to derive this, but its obvious that once you have done this exercise for the case of 3 heads where we got 1 by 4 in the answer, you can use the same exercise to obtain the probability for different cases as shown in this plot.

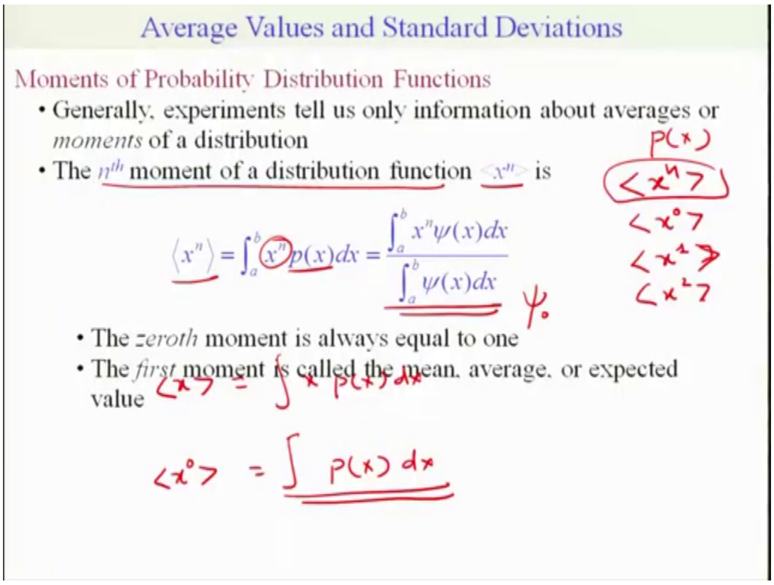
(Refer Slide Time: 23:51)



Now, I can generalize this binomial distribution function to multinomial distribution function. So, a multinomial distribution function applies to t outcome events, where ni is a number of times that outcome I from 1 to t appears. So, instead of two events now, we have n event. So, you are going to multiply p 1 and 1 p 2 and 2 n p t to the power n t these are the t outcome events and then we are multiplying the, the number of ways, we can arrange these events ok, which is going to be n 1 factorial divided by the number of object of type, there are different types.

So, n 1 n 2 and n t. So, this is a multinomial distribution function. Now, given a distribution function, we can obtain the average value and standard deviation, which is a commonly ah, done in statistical mechanism and thus we must know how to calculate it.

(Refer Slide Time: 24:45)

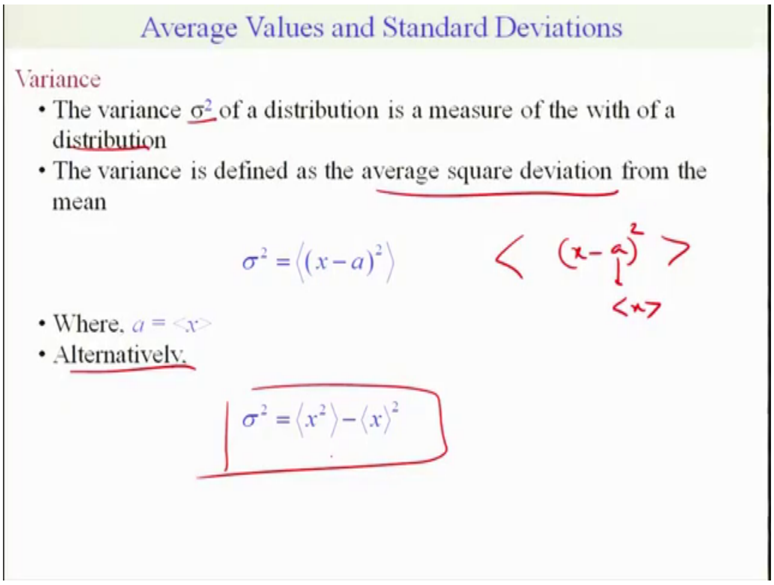


So, we going to use the moments of the probability distribution function to calculate the average or the mean values or in general the standard deviation also, let me just defined that given a function p of x, the nth moment of the distribution is defined as this. So, this is the average x to the power n indicates the nth moment. So, zeroth moment would be this and one moment would be this, second moment would be this ok.

So, let me just describe this little later that how does we explain this, but let me first come up with the definition of xn. So, the moment, nth moment is nothing. The x to the power n multiplied by the probability, use like a weighted average, but then you have considering x to the power n. So, this is a normalized version and that is why you are writing x to power n psi of x, which is a non normalized probability distribution, dx divided by the normalization constant.

So, we have written it, this is x psi 0. So, which means basically, average of x, the mean value of the x using the distribution is going to be x ex dx and; that means, this p of x, dx is going to be simply x 0. So, the zeroth movement is always equal to 1, if it is normalized of course it is going to be 1 and the first moment is called the mean or the average or the expected value.

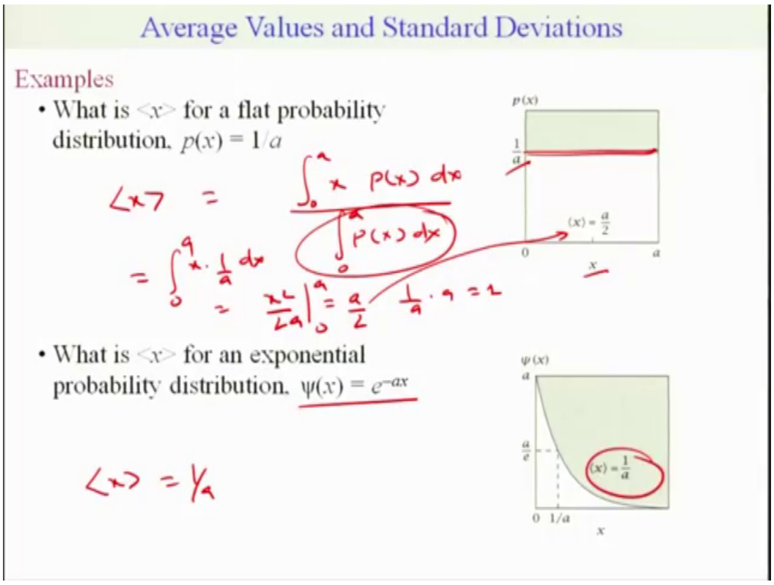
(Refer Slide Time: 26:31)



Now, let me also define the variance, the variance sigma square of a distribution is a measure of the width of the distribution, the variance is defined as the average squared deviation. So, this is the average. So, this means average squared deviation means x minus a, which is, mean value of this square of that.

So, this becomes the variance of that. Alternatively, you can also show, using the properties of the, this angle bracket, which is a mean value that this sigma square is nothing, but the mean value of x square minus square of the mean value x. So, that is a variance.

(Refer Slide Time: 27:19)



Let me end this lecture by just doing this, final exercise of evaluating the, the average value of x for a given flat probability distribution and as well as for non flat probability distribution. So, this is the probability distribution given to us, which is 1 by or a as, a function of x, which is a flat; that means, it is a constant and we need to find out x here.

So, this is going to be x p of x dx divided by p of x dx. So, I do not know whether it is a normalized. So, we have to first make sure that is normalized. So, this is nothing, but 1 by a times, a is 1. So, certainly, this is normalized. Now, this can be shown, this is as 0 to a x, p x is 1 by a dx, which is nothing, but x square by 2 a and x is nothing, but 0 to a.

So, this is nothing, but a by 2. So, that is what we can get here. Similarly, you can show this ah, that for such a system or such a probability distribution, the average x is nothing, but 1 by a here, in the similar, exercise.

So, with this I believe that ah, we have got the basis of the foundation for us to use this understanding, to take it to the next stage, where we are going to, show that maximizing multiplicity that is W, can be related to the entropy and from there we can connect to the macroscopic ah, macroscopic properties.

So, I will see you in the next lecture, and continue with this exercise.