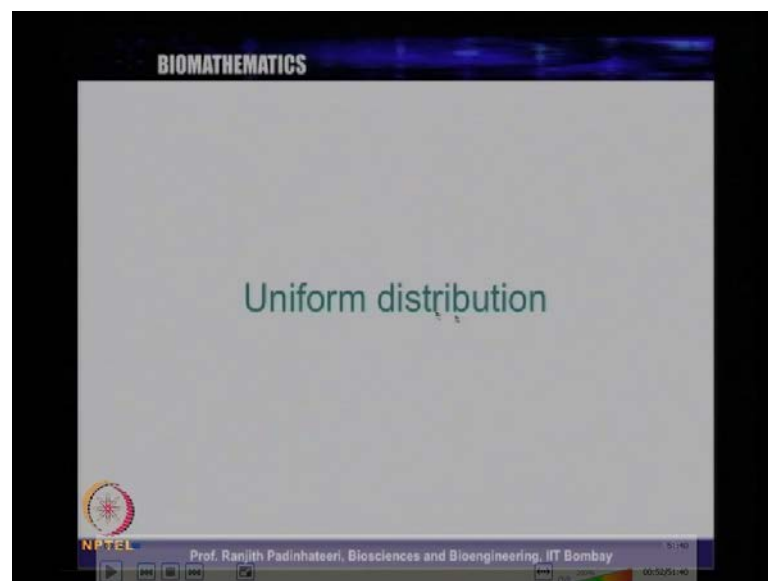


**Biomathematics**  
**Prof. Dr. Ranjith Padinhateeri**  
**Department of Mathematics**  
**Indian Institute of Technology, Bombay**

**Lecture No. # 26**  
**Uniform Distribution**

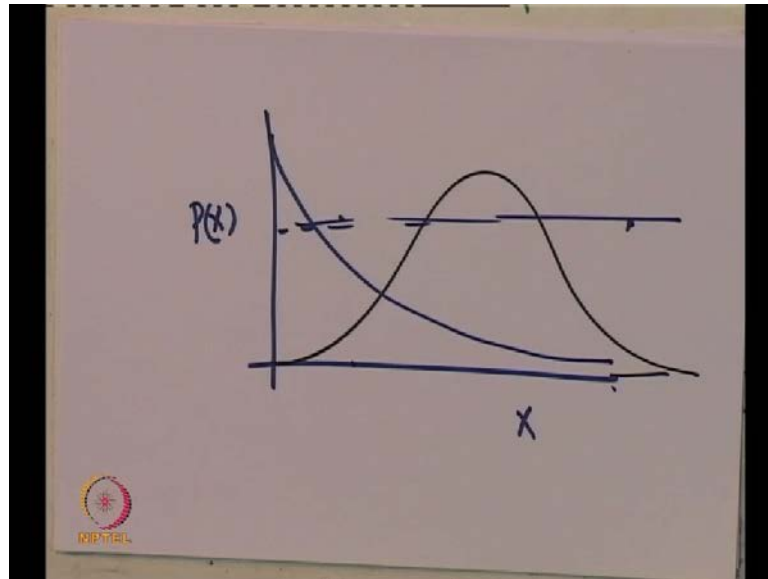
Hello, welcome to this lecture on biomathematics. We have been discussing statistics as a team for few lectures and we will continue discussing well more about statistics in coming lectures. So, today we are going to discuss two distributions something about uniform distributions and poisson distribution; so the general topic is statistics.

(Refer Slide Time: 00:50)



First we are going to discuss something called uniform distribution; what is uniform distribution? We have heard of normal distribution; we heard of explanation distribution. So, what is uniform distribution? Uniform distribution as the name itself suggests that everything is same – uniform; everything is equal. The probability is equal; so this is essentially the probability distribution.

(Refer Slide Time: 01:28)



Previously, we had exponential if we have look at hear  $x$  and  $p$  of  $X$  we had exponential distribution this look something like this we had normal distribution which looked like this. Now, what is the uniform distribution uniform distribution is itself mean that it cannot be it has to be uniform everywhere. **If I**, what is the mean for any value of  $x$  the  $p$  of  $x$  should be same like for all values of  $x$ . So, essentially uniform this name with itself would suggest something like this; uniform everywhere, same. So, this value the probability is same for this value; the probability is same. Such distributions are called uniform distribution. We will come to this picture later but, let us look at some examples. To begin with let us look at some examples; examples and biology, general example that you see around that will have a uniform distribution.

(Refer Slide Time: 02:37)

BIOMATHEMATICS

## Uniform distribution

- Events that are equally likely
  - Results from tossing a coin
  - Results from throwing a die
  - N bp random sequence

NPTEL Prof. Ranjith Padinhateeri, Biosciences and Bioengineering, IIT Bombay

Let us, as we said, if there has been equal probabilities any number any set of events that are equally likely equally probable will have a uniform. So, that is the key here; events that are equally likely that means, if the probabilities are equal then you will see uniform distribution. Simplest thing you can imagine is results from tossing a coin. If you toss a coin you will get either head or a tail; if it is head as you know we have only two events - two events here; so let us discuss the example of tossing a coin.

(Refer Slide Time: 03:26)

Tossing a Coin

---

H      T

Total 2 outcomes

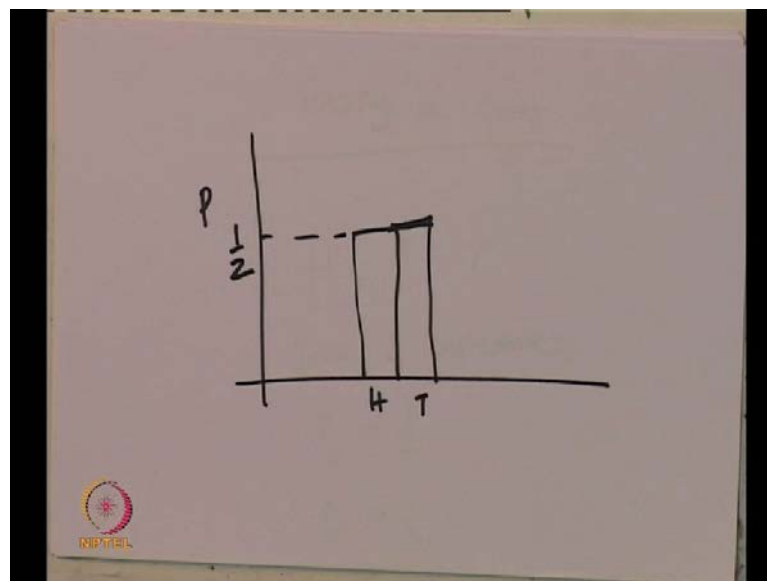
$P_H = \frac{1}{2}$

$P_T = \frac{1}{2}$

NPTEL

Tossing a coin there is only two outcomes: either head or tail. If you toss a coin you will get the **our** sign which we called 1 side of sign, we can call it as head and other side can be called as tail so you will get either head or tail. That means you will get either 1 side or the other side. There are total two outcomes, either head this is one; this is another one; 2 outcomes head or tail. Now, both of them are equally probable; that means, half of the time will be head, half of the time. Probability of having head or the first outcome is half and probability of finding the second outcome is half; so these are equally likely. So this is equal probability - head and tail have equal probabilities.

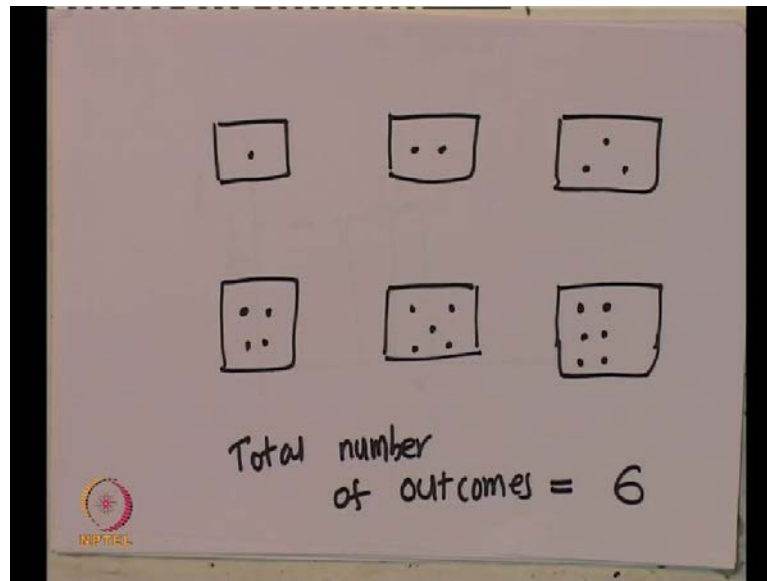
(Refer Slide Time: 04:37)



If you have a distribution, if you have a histogram of head and tail, so probability this is the head and this is tail, this probability of getting head will be half and probability of getting tail also will be half. So, this is half; this is we can call it uniform this is an example of uniform distribution.

The next example that I discuss here is results from throwing a die. So, you know die **you what die is die** will have like 6 sides we play. Playing people use dies so it will have 6 sides with 1, 2, 3, 4, 5, 6, either return or not so there will be either 6 dots or 1 dot or 2 dots or 3 dots.

(Refer Slide Time: 05:36)

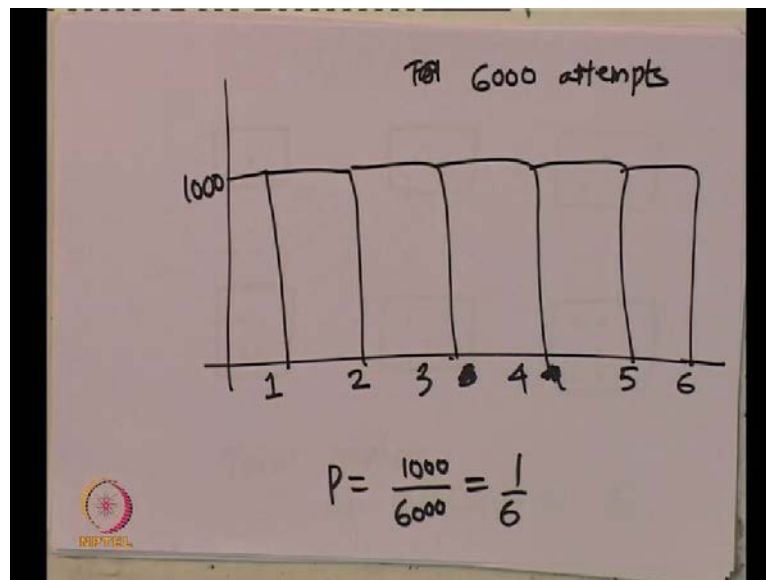


You might have seen dies like with 1 side having 1 dot, another side having 2 dots another side having 3 dots, another side having 4 dots, another side having 5 dots and you can see a side having 6 dots. This is different sides of a die

Now if you task this die if you through this die and when it falls down if you look which 1 of this is the side that is coming up you will see that there is equal probability for all of these two come so if you do it like many thousand times if you through the dies thousand times so there are 6 so let us write here total number of outcomes is 6 total number of outcome is 6.

So each of them have equal probabilities so this has same probabilities what is that mean that that means that if you have like 600 times if you throw this die or if you throw this die or if you throw die like 6 thousand times an average at least thousand times, 1 will come thousand times and 2 will come some more thousand times 3 will come thousand times 4 will come thousand times 5 will come thousand times 6 will come.

(Refer Slide Time: 07:42)

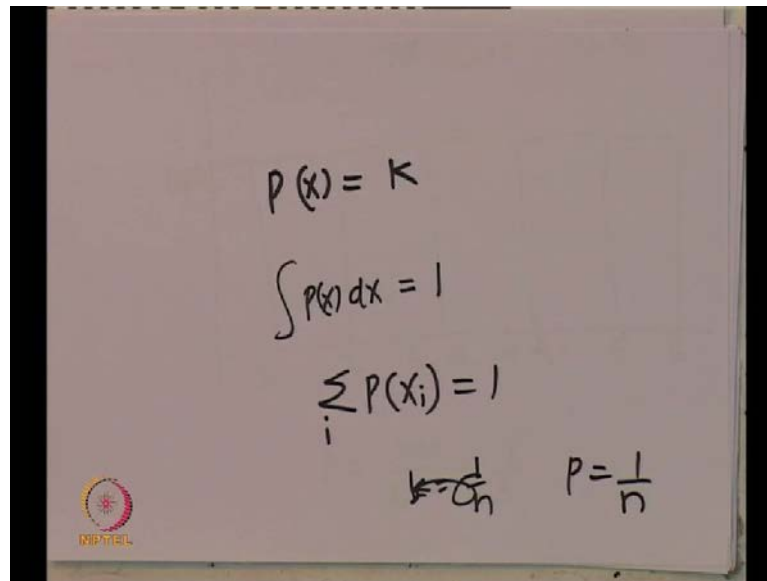


So, if you do this 6 thousand times and draw this graph, how many times you got 1? How many times you got 2? How many times you got 3, 4, 5 and 6? You would get like thousand times 1, another thousand times 2 thousand times, 3 thousand times, 4 thousand times, 5 and thousand times 6

So this is 1000; if you attempt 6 thousand attempts, if you throw the die 6 thousand times an average 6 thousand times you will get 1 thousand times, you will get 2 thousand times, you will get 3 thousand times, you will get 4 thousand times, you will get 5 and thousand times you will get 6. So, this is 1 can call it as uniform distribution. So, what is the probability of getting 1, 2, 3, or 5 or 6? Out of 6 thousand times out of 6 thousand attempts, thousand times you got 1 side; so the probability is 1000 by 6000.

Let less than 1 that is 1 by 6 or in other words, if there are 6 events possible or 6 outcomes possible, each outcome has same possibilities. So, 1 by 6 each of this 1 by 6 into 1 by 6 is 1. So, the total probability is 1; this is called uniform distribution. In other words, like you can say that  $p$  of  $x$  is some constant  $k$  that is the like it is already the same.

(Refer Slide Time: 09:03)

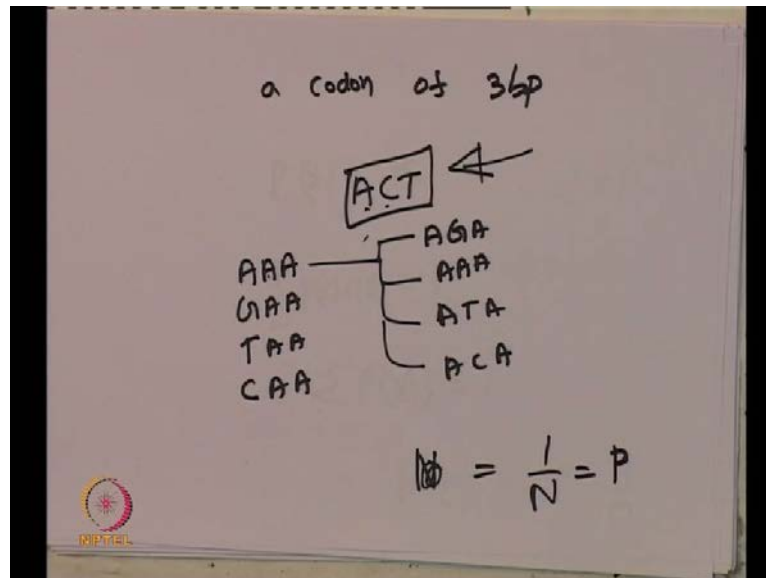


The image shows a whiteboard with handwritten mathematical equations. At the top, it says  $P(x) = k$ . Below that, the integral equation  $\int P(x) dx = 1$  is written. Underneath, the summation equation  $\sum_i P(x_i) = 1$  is shown. To the right of the summation, there are two expressions:  $k = \frac{1}{n}$  and  $P = \frac{1}{n}$ . In the bottom left corner of the whiteboard, there is a small circular logo with a star and the text 'INBTTEL' below it.

Now, what is  $k$  how do we find it so integral  $p$  of  $x$  is  $d x$  has to be 1 so the total number of probabilities are same over  $I$   $p$  of  $x$   $I$  has to be 1. So, total if you using this condition you will find that if there are  $n$  events  $k$  is  $1$  by  $n$  is there are  $n$  events  $k$  is  $1$  by  $n$  so just like before  $1$  by  $6$  or  $1$  by two there is total number of events or total number outcomes and if you divide  $1$  by the total number outcomes will be the probability so  $p$  probability will be the total number of  $1$  by total number of events

So if they distribution is uniform this will be the probability now the third example which we discussed is here there are let us say, you have a you have a random sequence so you have  $n$  base pair of a random sequence so let us say, let us think of 3 base pair let us think of random 3 base pair so like a codon of 3 base pair

(Refer Slide Time: 11:10)



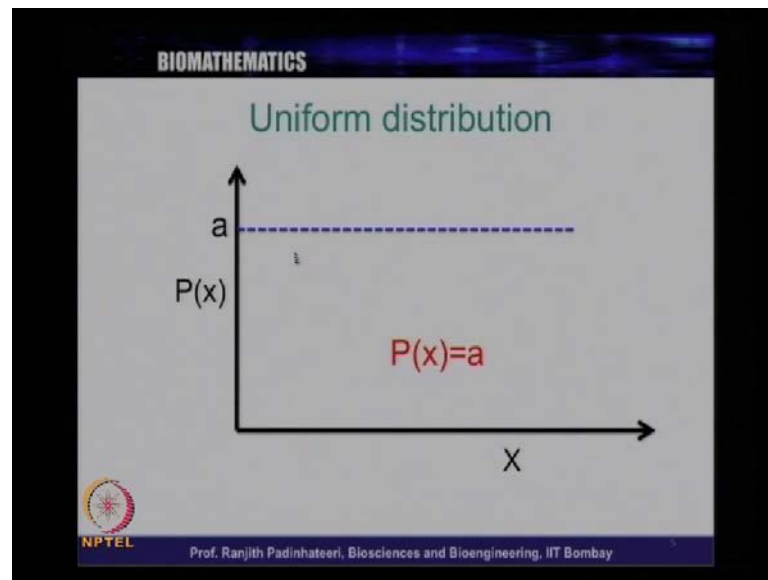
A codon of 3 base pair; what is the probability that you will find a codon of ACT? What is the probability? so to find this first of all you have to know how many kinds of codons are possible so you can now that it can have a a it can be g a it can be t a it can be c a. so for each of this later you calculate how many of them are possible right, so there is this you can change for.

Now it can each of this a can be change to so this a can be change to c g or t so a we can change this to either A G A A A A T A A C A. This can be further divided so and so on forth you count the total number of possibilities and if you find n it will be like 1 by n will be the probability of having a 3 base pair of codon A C T so taking when in mind assignment and find out what is the probability of finding a codon of A C T or all of this sequence where sequences are equally likely.

So the answer, the way to do this is write down all possible sequences with having 3 like all possible combinations of having 3 of them and then 1 over that will be the probability. If you have n sequences n different times of sequences possible for this codon, you will have 1 over n as a probability of finding any of this codon because all of them are equally likely so that is the idea of uniform distribution.



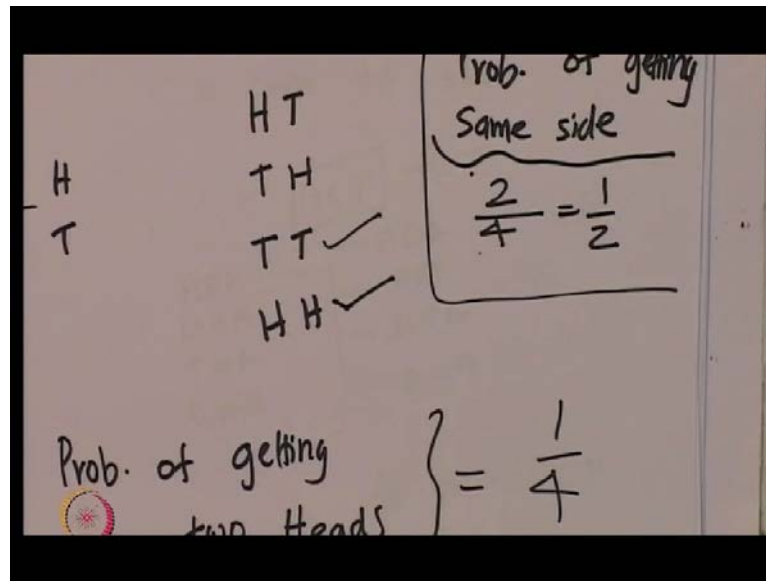
(Refer Slide Time: 13:41)



So, see here what is plotted; this is uniform distribution. The idea is plotted here  $p$  of  $x$  is  $x$  and the  $p$  of  $x$  is constant  $a$  here and as we said,  $a$  can be some number which is something which is less than 1. It will be a fraction;  $a$  is a fraction essentially between 0 and 1 **ok**. So, this is the uniform distribution.

Now, we can ask little more complicated questions; we find this like if there are  $n$  events possible,  $1$  over  $n$  is the probability. Now, let us say you are tossing a coin 2 times. If I toss a coin let us say, you toss it 2 times, what is probability that both times you get head or what is the probability that both times you get tail? The answer is that let us think, how many events are there? What do, so we are causing **I to in** coin causing  $a$ , sorry, we are tossing a coin 2 times.

(Refer Slide Time: 15:02)



So if I do 2 times of tossing this is like you might get like one head and the other one has tail. Now, one has you toss it 2 times you might get the first A C T tail and second one is the head. You ask your friend to toss it 2 times; your friend might get 2 tails. You ask somebody else to toss, you saw that person might get 2 heads; so, these are 4 possible events.

So, either in the, when there is only 1 causing there are only 2 events possible - either head or tail. Then, we found that the probability is half here; there are 4 possible events like head and tail because we are tossing the 2 times. So, if I toss it 2 times I may get head or tail A A H are followed by t. if you ask this question like, what if you do? You might get some times first one tail and second one is head; somebody else might get T T and somebody else might get H H.

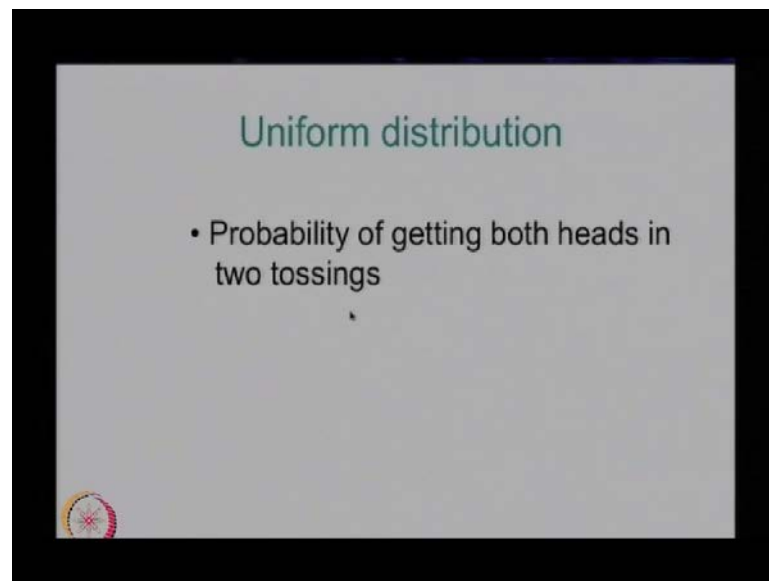
In such cases so we have 4 events and 4 of them all the 4 are equally likely so the probability of getting 2 heads or 2 tails probability of getting 2 heads 2 heads there are 4 events possible so like 1 by 4 this is just 1 event among the 4 so there are 4 events and this is just event among the 4 so just 1 by 4 you can ask the question, what is probability of getting same side, same toss; same side all the times both times?

So you are tossing it both times; how many times you will get both? Both the time you will get same side. Here, both the time you got like different sides; here also you got different sides but, here you got both the time same sides; here also you got both the time

same side. So, 2 out of 4 you got same side; so the probability of getting same sides when you toss twice, that probability is 2 by 4; that is half. This is the probability of getting same side when tossing twice is half.

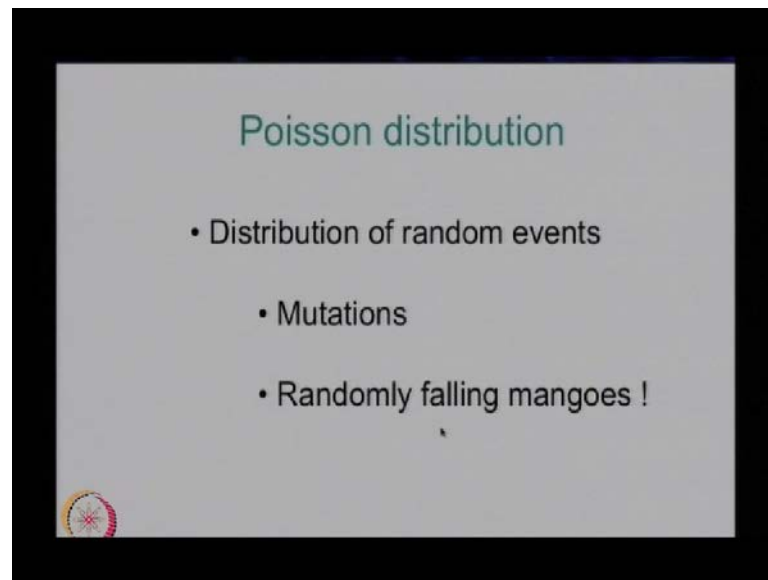
Just like counting like this, we can just learn something about these probabilities. So, I said an example of tossing a coin here in a few minutes or in a, we did not follow sometime later. We will discuss something about cancer and mutations in cancer and all that. There we will use the same idea and find out something very interesting; the same kind of counting. What is the (( )) instead of just like a instead of coin? We will talk about something else which is related to cancer. In a few moments, in something what follows, we will discuss how this idea of probabilities can be applied to understand something about cancer; so, we have uniform distribution

(Refer Slide Time: 18:55)



Now we are going to discuss another distribution called Poisson distribution.

(Refer Slide Time: 19:02)



Poisson distribution is distribution of random events when say, you say something about random events the immediate thing that comes into picture is mutations. So, mutation is something about that can immediately think of as random events. Something about mutations can be described by the Poisson distribution. Another day to day life we can think of many examples that can have happened randomly like, one example that you might, everybody might have seen just like say, you go under mango tree or something or an apple tree, whatever tree you like and you ask like randomly mangoes will fall down. So, we will discuss something about randomly falling mangoes and how can we described using Poisson distribution. First, let us go to this mutation.

(Refer Slide Time: 20:08)

**Mutations**

You have many copies of a DNA each of length  $L$ .

Let 4 be the average number of mutations per copy.

How likely that you will find a DNA with exactly 3 mutations ?

Look at here; let us say, imagine that you have many copies of a DNA each of length  $l$ ; so you have many copies. A DNA in each of having length  $l$  so let us say, 1000 base pair of DNA; so an average let us see you have 1000 base pair of DNA.

(Refer Slide Time: 20:39)

average # of mutations = 4

The diagram illustrates a DNA segment of length  $l$  with an average of 4 mutations. Below it, four horizontal lines represent individual DNA copies, each with a different number of mutations (represented by small circles) at various positions along the segment.

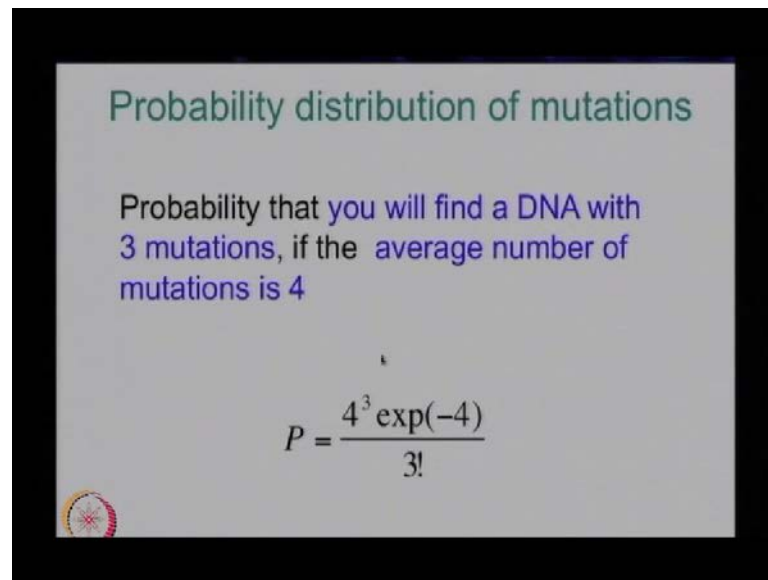
So many copies 1000 base pair of DNA, there are many copies like this. Now, each of these copies has some mutations. Let us say this has some number mutation; this has some number of mutations; this has some mutations; this has some mutations (Refer Slide time: 21:00).

So let us say, on an average there are 4 mutation average number of mutations. Let us say, there are 4 mutations an average; some of them will have 3, some of them will have 1, some of them will have 5, some of them will have 10. So, an average let us imagine there are 4 mutations. Now, you randomly pick 1 DNA piece and ask the question, what is the probability that you will find the DNA with 3 mutations? If you ask this question answer to this question is given by this **Poisson by the** distribution of Poisson; so, this is the Poisson probability distribution.

Let me make this question little more clear. you have a set of you have many pieces of DNA and each of this DNA has some number of mutations and if you find out the mean number you have average number of mutations is 4 what does that mean that means some of the DNA pieces have 3 mutations some of them have 4 mutations some of them have 5 mutations some of them have 6 mutations and some of them have two mutations. So, there are many mutations like but, an average the average number of mutations average number of mutations per filament per DNA's piece is 4.

Now, you randomly choose 1 DNA form million pieces of DNA. You have a large number of pieces of DNA you have and ask the question what is the probability that can you before doing this can we say, what is the probability that you will pick a DNA that has 3 mutations? If you ask this question, answer of this is given by something all Poisson distributions. That is what we are discussing, how likely that you will find a DNA with exactly 3 mutations? So, if the average number of mutations is 4 per copy how likely that you will find a DNA with exactly 3 mutations?

(Refer Slide Time: 23:46)



Probability distribution of mutations

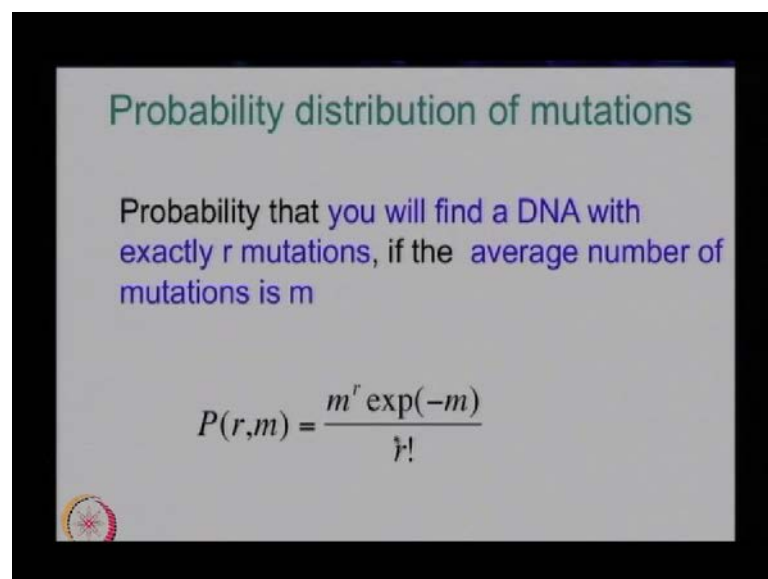
Probability that you will find a DNA with 3 mutations, if the average number of mutations is 4

$$P = \frac{4^3 \exp(-4)}{3!}$$

A small circular logo is visible in the bottom-left corner of the slide.

The answer is the probability that you will find a DNA with 3 mutations if the average number of mutation 4 is given by this calculation 4 power 3 in to exponential minus 4 by 3 factorial. So you can calculate this. This is how this comes; we will discuss later but, for the moment just realize that there is a formula and this formula is known as the Poisson formula or, the Poisson distribution; this probability distribution. So, let us generalize this; if you generalize this, what we get is that probability that you will find a DNA with exactly  $r$  mutations, if the average number of mutations is  $m$ .

(Refer Slide Time: 24:25)



Probability distribution of mutations

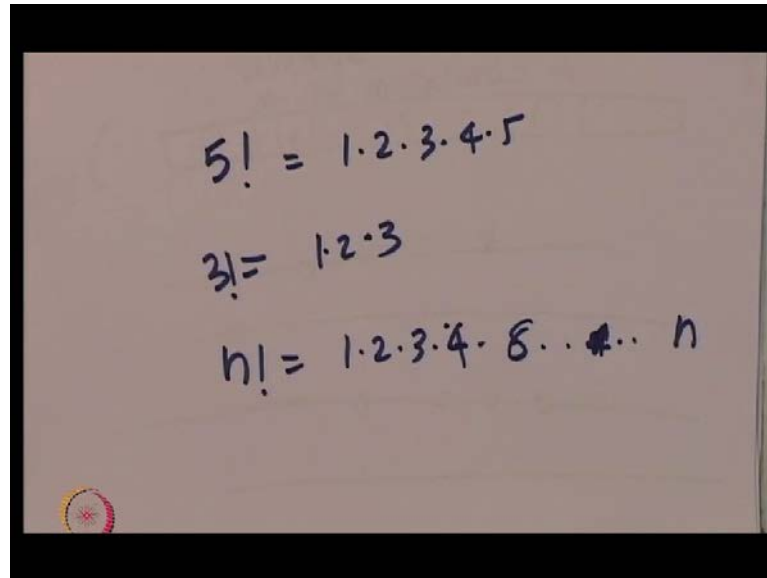
Probability that you will find a DNA with exactly  $r$  mutations, if the average number of mutations is  $m$

$$P(r,m) = \frac{m^r \exp(-m)}{r!}$$

A small circular logo is visible in the bottom-left corner of the slide.

So, probability that you will find  $r$  mutations if the average number of mutations is  $m$ ; now, you substitute  $m$  is in the previous cases;  $m$  is 4 and  $r$  is 3. You got the... you can get the previous formula we have used here  $m$  as 4 and  $r$  as 3.

(Refer Slide Time: 25:10)



A photograph of a whiteboard with handwritten mathematical formulas. The formulas are:

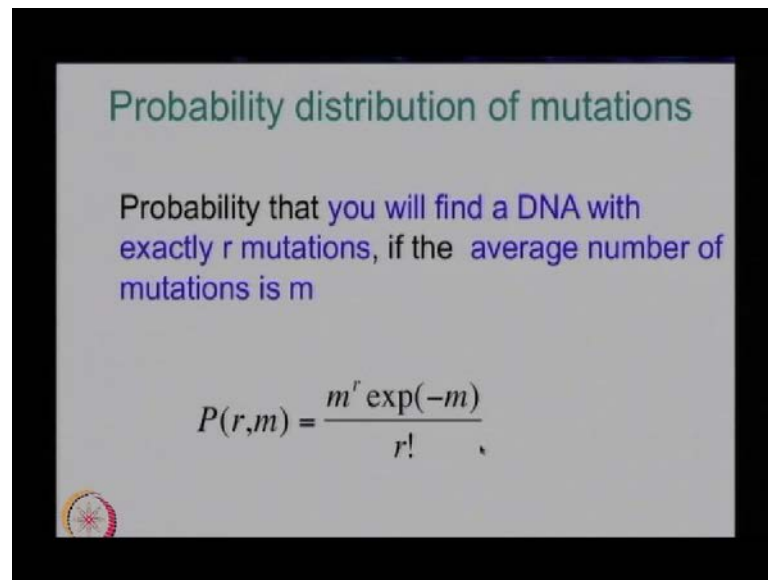
$$5! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5$$
$$3! = 1 \cdot 2 \cdot 3$$
$$n! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \dots n$$

The whiteboard has a small circular logo in the bottom left corner.

So,  $p$  of  $r$  comma  $m$  is equal to  $m$  power  $r$  e power minus  $m$  divided by  $r$  factorial. You now like 5 factorial is 1 into 2 into 3 into 4 into 5; 3 factorial is 1 into 2 into 3 like  $n$  factorial is 1 into 2 into 3 into multiply up to multiply 4, 5, 6, all those up to  $n$  1, 2, 3, 4. Like, let us say, this is 8 and so on up to  $n$ . So, you multiply  $n$  of them together you get  $n$  factorial, the definition of  $r$  factorial,  $n$  factorial.



(Refer Slide Time; 25:47)



Probability distribution of mutations

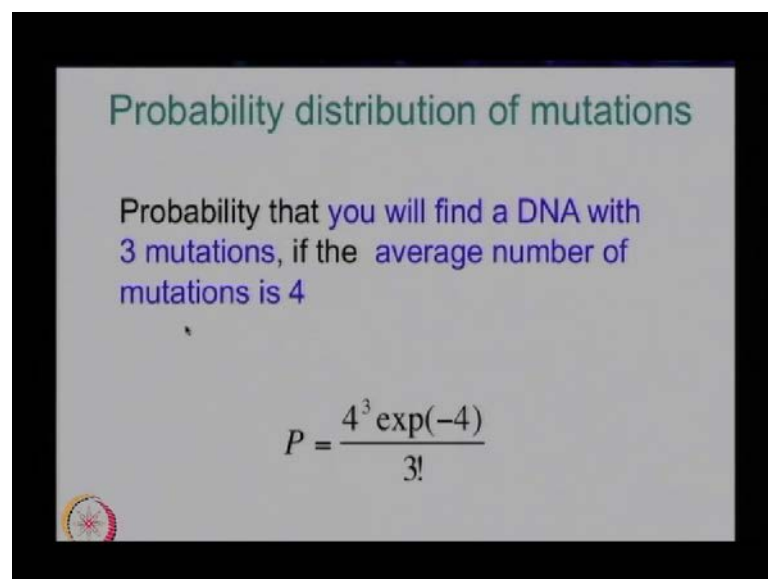
Probability that you will find a DNA with exactly  $r$  mutations, if the average number of mutations is  $m$

$$P(r,m) = \frac{m^r \exp(-m)}{r!}$$

A small circular logo is visible in the bottom-left corner of the slide.

The  $P$  of  $r$  comma  $m$  is  $m$  power  $r$   $e$  power minus  $m$  by  $r$  factorial. So, probability that you will find a DNA with exactly  $r$  mutations, this is called Poisson distribution. So, let us go back and think a bit more.

(Refer Slide Time: 26:06)



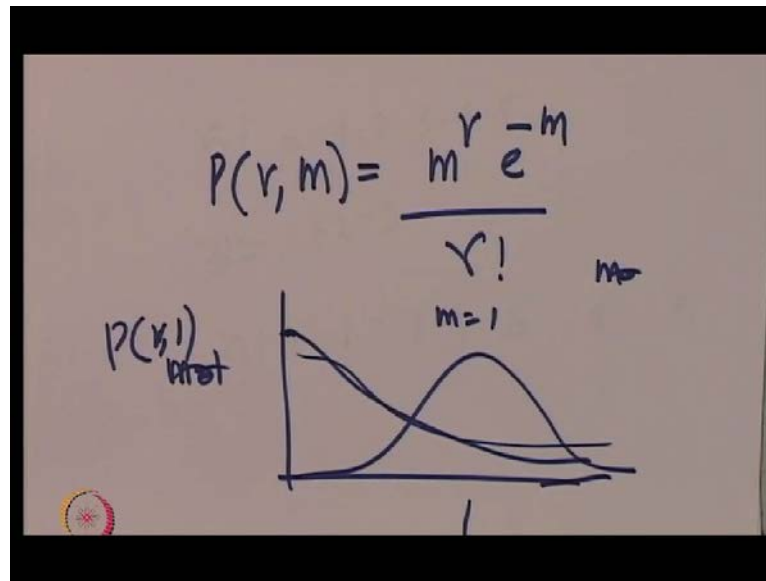
Probability distribution of mutations

Probability that you will find a DNA with 3 mutations, if the average number of mutations is 4

$$P = \frac{4^3 \exp(-4)}{3!}$$

A small circular logo is visible in the bottom-left corner of the slide.

(Refer Slide Time: 26:20)

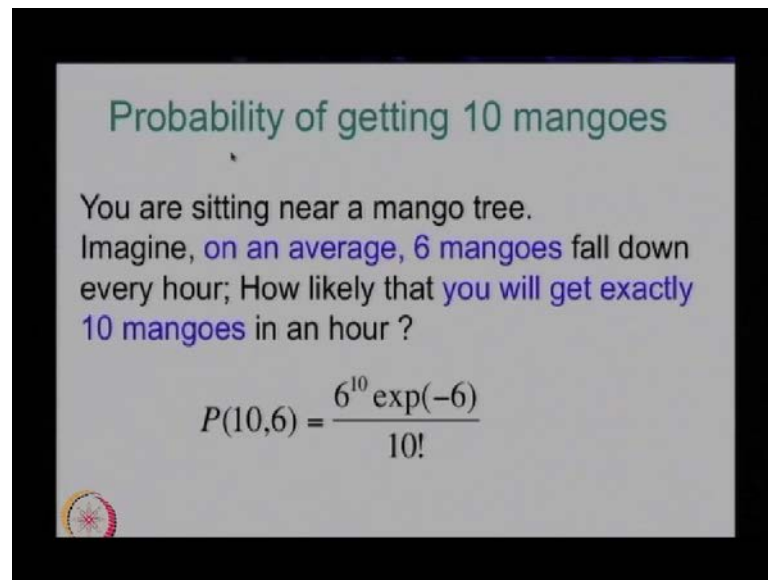


So, if you add an average 4 now we have here is  $p$  of  $R$  comma  $m$  is we said, we say this is  $m$  power  $r$   $e$  power minus  $m$  by  $r$  factorial. This is the probability of finding  $r$  mutations, if an average  $m$  mutations. So, let us say that, an average there is only 1 mutation there is only 1 mutation an average. That means, if you have some of them will have 2 mutations; some of them will have 3 mutations; some of them will have 4 mutations; but, many of will them have 0 mutations.

So, the average mutations only 3 let see the average mutations only 3 then, how do plot look like, sorry, the average  $m$  is 1 let us say,  $m$  is 1 and how do we plot  $p$  of let us say,  $m$  is 1? You want to plot  $p$  of  $r$  comma 1; so, probability of finding anything above 1 will be less because an average it is 1. So, this looks something like this; 1 you will find but, above 1 is like very likely to have very less likely to find this. So, we can plot for each  $m$ , we can plot for  $p$  of  $r$  comma  $m$  and for  $m$  is very large it can look something like this.

The shape of this curve will look like it will depend up on the value of  $m$ . So, do plot this for different values of  $m$  and see what you get for some intermediate value of  $m$  which 5 or 6; it might look something like this 3 4 and 3 4 and are all like this. So, do plot this  $p$  of  $r$  comma  $m$  for different values of  $m$  and have a look at it and get a feeling how this will look like. This is the answer for the probability of mutations; now you can ask the question, what is the probability of, you can think about probability of getting mangoes?

(Refer Slide Time: 28:38)

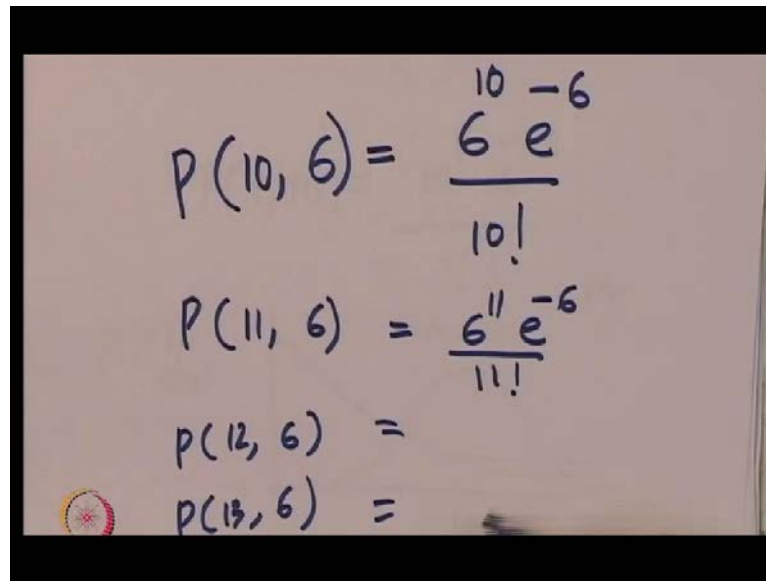


The slide has a black border and a light gray background. At the top, the title 'Probability of getting 10 mangoes' is written in a green font. Below the title, there is a paragraph of text in black font: 'You are sitting near a mango tree. Imagine, on an average, 6 mangoes fall down every hour; How likely that you will get exactly 10 mangoes in an hour?'. The words 'on an average, 6 mangoes' and '10 mangoes' are highlighted in blue. Below the text is the Poisson distribution formula: 
$$P(10,6) = \frac{6^{10} \exp(-6)}{10!}$$
. In the bottom-left corner of the slide, there is a small circular logo with a red and white pattern.

Let us say, that you are sitting under mango tree and you see that mangoes fall down. Now, wait for every day morning, you go and wait every day. Sometimes, you go and wait for 1 hour near a mango tree and you see that some number of mangoes fall down. You do this let us say, the first day you waited for 1 hour and you got 3 mangoes; the next day you waited for 1 hour near the mango tree and you got 8 mangoes; the next day you waited near the mango tree and then got 10 mangoes. All of the falling down, you do not pluck the mangoes; you only take those mangoes which are falling down. You find that, first day 3 mangoes fell down; the next day 7 mangoes fell down; the next day 10 mangoes fell down; the next day only 1 mango fell down and like.

So, you just keep counting each day how many mangoes are falling down in the 1 hour that period and that you waiting there and you find them let us say, on an average 6 mangoes per hour fall down if you wait for an hour; 6 mangoes fall down on an average. So, you find some overall days divide and by total number of days you will get average of mangoes that you get in 1 hour time and then, let that be 6. Now, you can ask this question so let us say, you did this for like about 100 days or many days; at 30 days and you count this number and then you can ask the 31st day what is the probability that I will get 5 mangoes or what is the probability that I will get 10 mangoes? So, this is again described by the Poisson distribution.

(Refer Slide Time: 31:04)



The image shows a whiteboard with handwritten mathematical formulas for a Poisson distribution. The formulas are:

$$P(10, 6) = \frac{6^{10} e^{-6}}{10!}$$
$$P(11, 6) = \frac{6^{11} e^{-6}}{11!}$$
$$P(12, 6) =$$
$$P(13, 6) =$$

What do you want? Even probability to get 10 mangoes when the average is 6; so, this is again the same formula we have to use; that we described some time ago. The answer to this is 6 power 10 into e power minus 6 divided by 10 factorial. So, that is what here the probability; if you look at here the probability that you will get exactly 10 mangoes in 1 hour, if the average number of mangoes in 1 hour is 6; that is what this tells us.

Now, this is probability of getting exactly 10 mangoes: now you can ask the question slightly differently. You can ask the question, what is the probability of getting 11 mangoes? **You can ask the question, what is the probability of getting 11 mangoes?**

So, this is 6 power 11 e power minus 6 by 11 factorial. What is the probability of getting the 12 mangoes, 13 mangoes and so on and so forth? You can calculate all of this; if you know all of this you can ask the questions, what is the probability of getting at least 10 mangoes? That means, some time it can be 11 mangoes; some time it can be 12 mangoes; some time it can be 13 mangoes; at least, 10 you should get. So it is the sum of all these probabilities because, even if you get 11 like at least, 10, the probability of getting at least 10 mangoes is given here.

(Refer Slide Time: 32:57)

Probability of getting at least 10 mangoes

On an average, 6 mangoes fall down every hour; How likely that you will get 10 or more mangoes in an hour ?

$$P = \sum_{r=10}^{\infty} \frac{6^r \exp(-6)}{r!}$$

If you look at this slide here on an average 6 mangoes fall down for every hour; how likely that you will get 10 or more mangoes in an hour so it is the sum over all this probability that get if you wrote down previously likely sum over all is equal to 10 to infinity. So you can get as many as mangoes as you want and this is 6 power x by e power minus by r factorial and sum over r. So, that is you sum over P of 10; 6 plus p of 11, 6 plus p of 12, 6 plus dot dot dot.

(Refer Slide Time: 33:32)

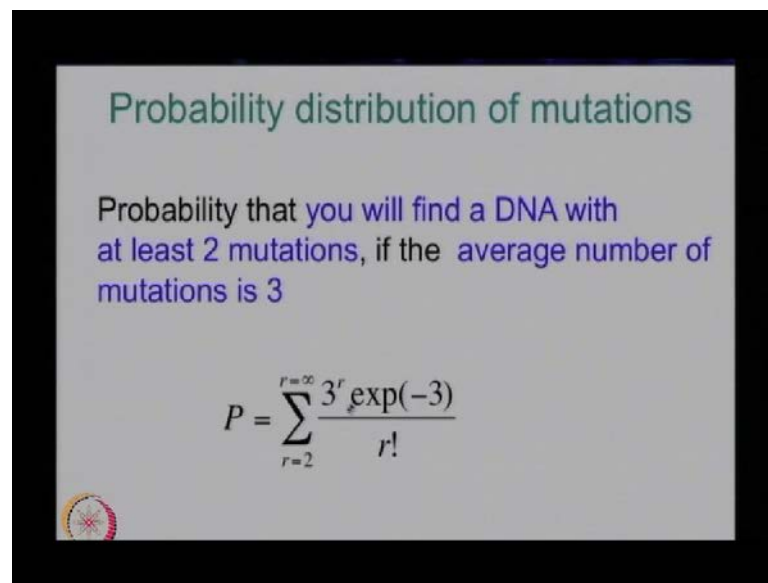
$P(10) = P(10,6) + P(11,6) + P(12,6) + \dots$

Prob. of getting 10 or more mangoes

$= P(10,6) + P(11,6) + P(12,6) + \dots$

This is the probability that you will get the at least 10 or greater than this. So, this is the probability of getting probability of getting 10 or more mangoes is this sum at least 10 is this probability; exactly 11 is this; exactly 12; is this exactly 13, exactly 14 and so and so forth. So sum up to infinity you get the answer probability of getting 10 or more mangoes due to exchange this to mutations. You can ask a question, what is the probability that you will find a DNA with at least two mutations?

(Refer Slide Time: 34:38)



Probability distribution of mutations

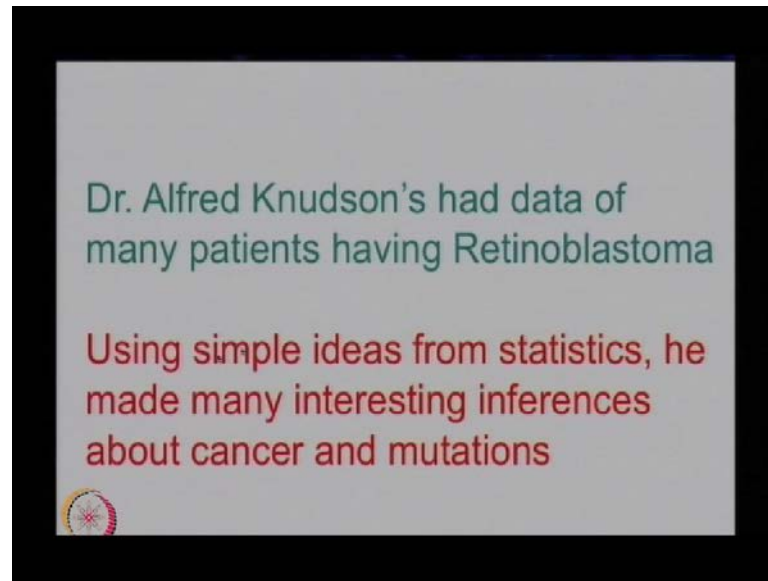
Probability that you will find a DNA with at least 2 mutations, if the average number of mutations is 3

$$P = \sum_{r=2}^{\infty} \frac{3^r e^{-3}}{r!}$$

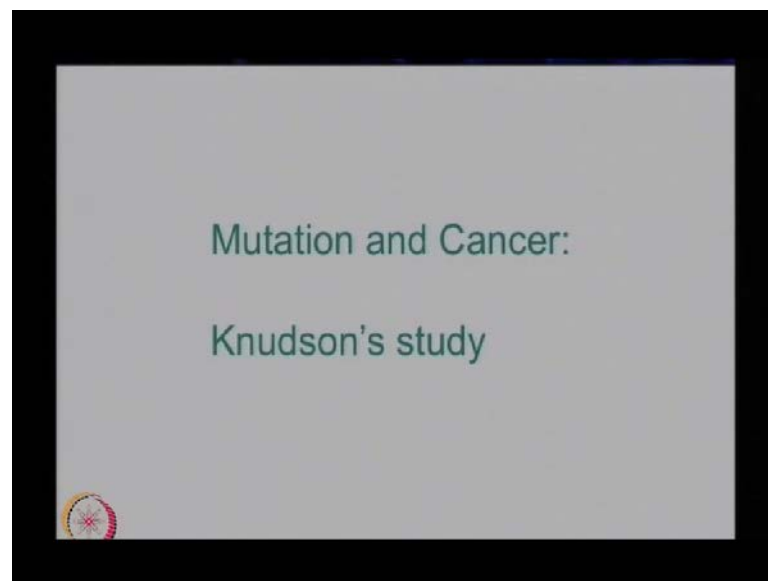
Probability that you will find a DNA with at least 2 mutations if the average number of mutation, is 3; the answer to the that is this is p is equal to summation of r is equal to 2 to infinity 3 power r e power minus 3 by r factorial; so, this is this answer **ok**.

So, now we learned a few things about Poisson distribution and now, we will take a real biological example and very famous example and discuss and find out how this idea whatever we learn so far today, can be used to understand something about cancer.

(Refer Slide Time: 35:38)



(Refer Slide Time: 35:39)



Today is mutation and cancer - Doctor Knudson's study. Doctor Alfred Knudson was a famous, was the well known doctor in ND hospital in Texas. He was the doctor treating patients with retinoblastoma. Retinoblastoma is a type of a cancer that comes for eyes lens cancer tumor in eyes that is called retinoblastoma. He had some record data of his patients having retinoblastoma; very preliminary data like he will write down just like any doctor would write down. Doctor Knudsen... the typical data from a doctor would be like what is the age and all that and then which eye you have like both the eyes, one of

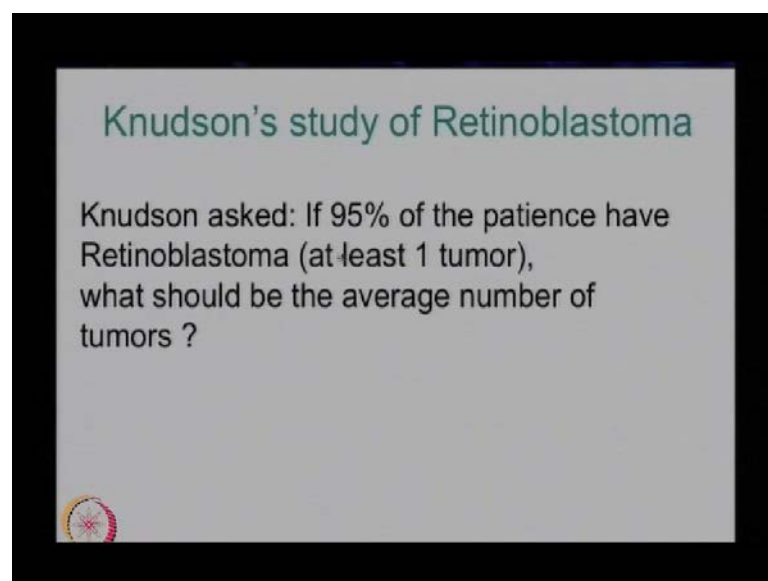
these eyes will have cancer; may be both the eyes will have of cancer, this is in retinoblastoma.

So you will write down and this is very simple information like, whether there is something like that inertia? Is it like your parents - did they have? Patients' parents did they have cancer? Or those kind of very preliminary information you will write down and he collected the data. We will see the data in a different, at a different point but, at this moment is not important; at this movement is like is a very simple things like 1 of the...

So, basically what we had to understand is that he had some data very preliminary data that any doctor will get it and using this data and some ideas of Poisson distribution and uniform distribution he found some interesting he reached some interesting conclusions and he extended this later to form something the famous to fit hypothesis cancer we could not discuss that hypothesis today but, we will discuss some interesting conclusion that can reach by just knowing Poisson distribution and uniform distribution.


So what did what did Doctor Knudsen have? Knudsen had just set of data and he asked the question the following question here if 95 percent of a patient, sorry, there is a spelling mistake here like patients, so there is typo.

(Refer Slide Time: 38:22)



Knudson's study of Retinoblastoma

Knudson asked: If 95% of the patients have Retinoblastoma (at least 1 tumor), what should be the average number of tumors ?





Here, so if 95 percent of the patients have retinoblastoma that is at least 1 tumor what should be the average number of tumors? So you found that 95 percentage of that (( )) came to this hospital had at least 1 tumor. That means we have some retinoblastoma some cancer at least 1 tumor 1 growth in a either of the eyes at least 1 tumor in tattle if this is the case, what should be the average number of tumors like you should have like 3 tumors, 4 tumors, 5 tumors are only just 1 tumor.

The answer to this can be easily found from the Poisson distribution. So, the question is he asked that let us say, on an average like you had like only 1 tumor or 2 tumors 3 tumors. So, then it can you ask the question, first you assume that for the growth, the tumor growth is the random process. So, you the because coming from mutations. So it has to be random process, if it is a random process it will have Poisson distribution. If this has the (( )) tumor probability of finding tumor is having Poisson distribution or probability of tumor growth of Poisson distribution then you can ask the question, probability that you will find a patient with exactly r tumors if the average number of tumors is m. If the probability you will find a patient with exactly r tumors if the average number of tumor is m; that is what we discussed.

(Refer Slide Time: 40:33)

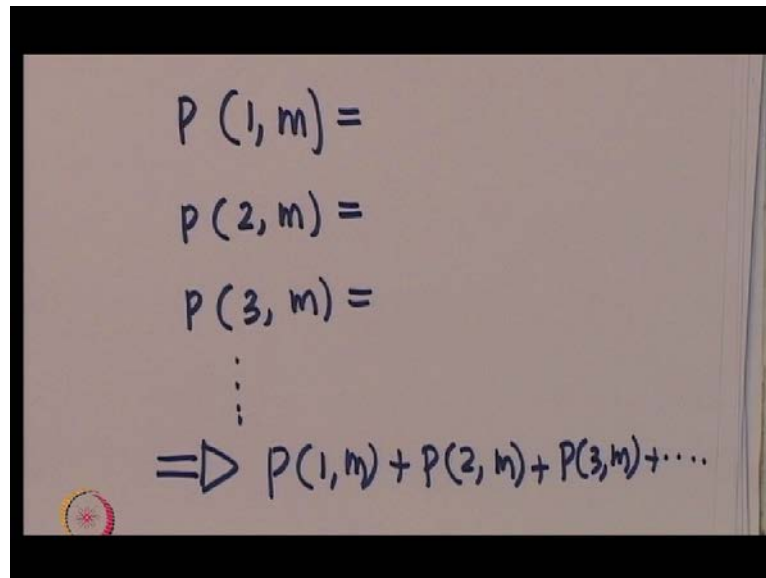
Knudson's study of cancer

Probability that you will find a patient with exactly  $r$  tumors, if the average number of tumors is  $m$

$$P(r,m) = \frac{m^r \exp(-m)}{r!}$$

So, for p of r comma m which is m power r e power minus m by r factorial but, what Knudson wanted is he does not know at least he does not want exactly r tumors or at least 1 tumor so what can we write down from this, you can write down.

(Refer Slide Time: 41:19)

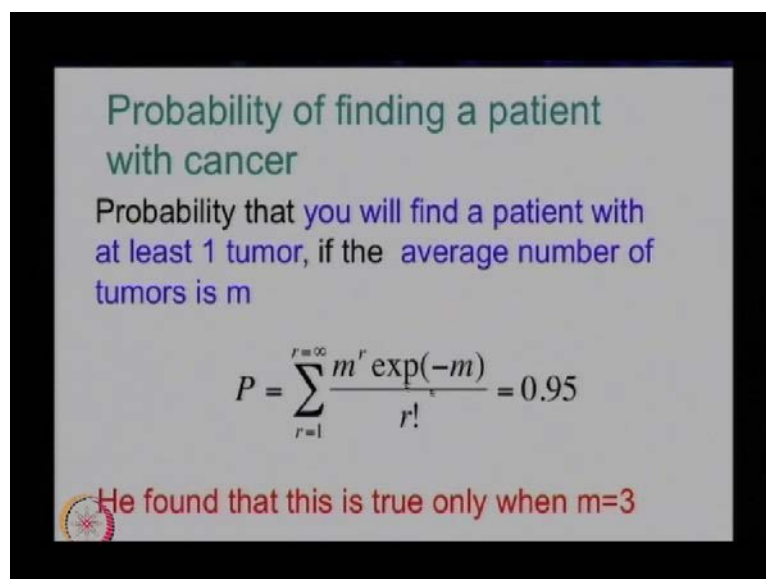


Handwritten notes on a whiteboard showing the summation of probabilities for 1, 2, 3, and more tumors:

$$P(1, m) =$$
$$P(2, m) =$$
$$P(3, m) =$$
$$\vdots$$
$$\Rightarrow P(1, m) + P(2, m) + P(3, m) + \dots$$

Probability of having 1 tumor if the average is  $m$  probability of having two tumors if the average is  $m$  probability of having 3 tumors if the average is  $m$  so on and so forth now you can sum of all these now the probability there is at least 1 tumor probability that there is a 1 tumor is given by  $p$  of 1 comma  $m$  plus  $p$  of 2 comma  $m$  plus  $p$  of 3 comma  $m$  plus dot dot dot. So, this is the probability that there is at least 1 tumor that means if a patient what is the probability that a patient will have at least you have an 1 tumor is given by this sum so that is what described here.

(Refer Slide Time: 42:18)



Probability of finding a patient with cancer

Probability that you will find a patient with at least 1 tumor, if the average number of tumors is  $m$

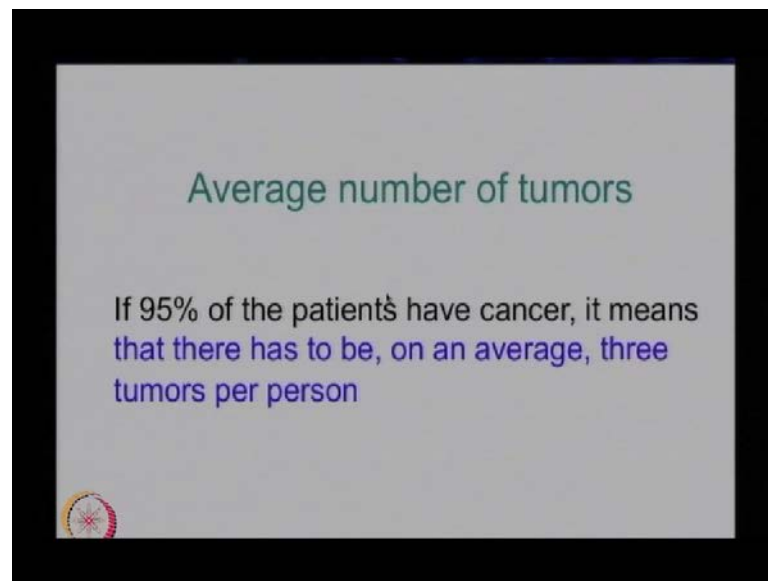
$$P = \sum_{r=1}^{\infty} \frac{m^r \exp(-m)}{r!} = 0.95$$

He found that this is true only when  $m=3$

Probability that you will find a patient with at least 1 tumor, if the average number of tumors is  $m$  is  $\sum_{r=1}^{\infty} \frac{m^r e^{-m}}{r!}$ . So, this is what we just described and Knudson found this is point 95 there is a 95 percent of the patients had at least 1 tumor; some of them have more than 1 but, all of them had at least 1.

So this is the probability of 1 finding a patient with that means if you have 100 patients 5 patients are have any tumor 95 percent had at least 1. So, if you want to get this what should be the value  $m$  for which value this sum is 95 point 95 we can ask this question. You can do for this for different values of  $m$ . So, Knudson put out tables for different values of  $m$  and  $r$  and he did sum and he found that this is true only when  $m$  is equal to 3 this equation is true only when  $m$  is equal to 3.

(Refer Slide Time: 43:35)

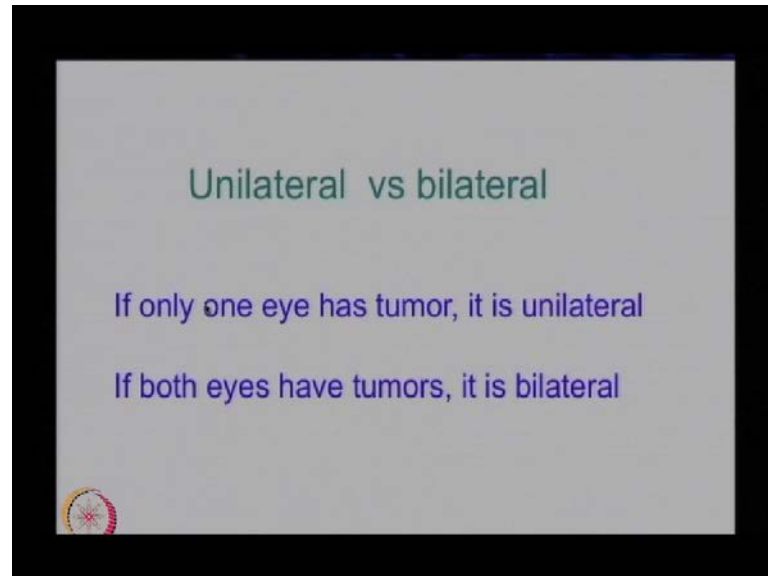


So what is that mean? That means if 95 percent of the patients have cancer it means that 3 there has to be cancer on an average 3 tumors per person. So, just by asking a simple question he could find out what is the average number of tumors slot 1, slot 2 it is 3.

If 95 percent patients have tumors at least 3 tumors have an average has to be done here. He did not measure a number of tumors in each person but, without doing a measurement from mind just from simple idea statistics  $(())$ . That there has to be an average 3 tumors this is the very interesting conclusion that he is teaching. Now, the next

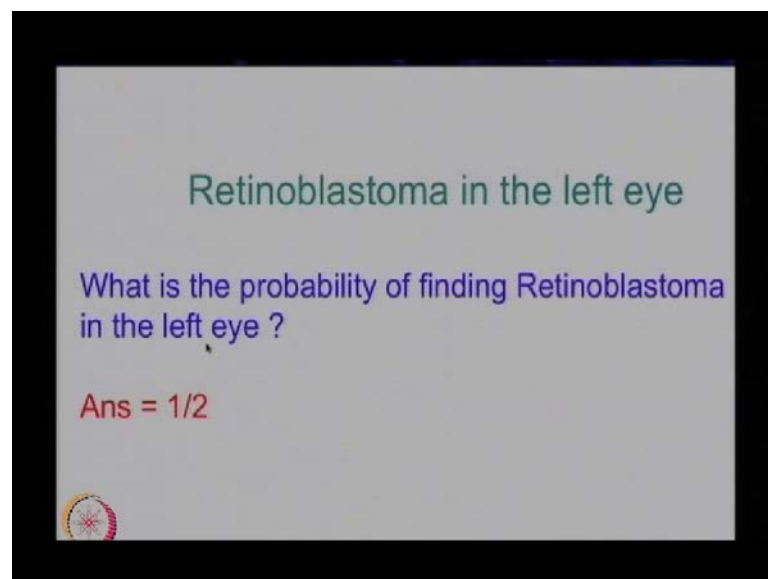
question he asked something about unilateral retinoblastoma versus S bilateral retinoblastoma why, what is this unilateral retinoblastoma?

(Refer Slide Time: 44:47)



Let look at here if only 1 eye has tumor it is called unilateral retinoblastoma; if both eye both the eyes have tumors or at least 1 tumor in both the eyes you called it a S bilateral retinoblastoma.

(Refer Slide Time: 45:09)

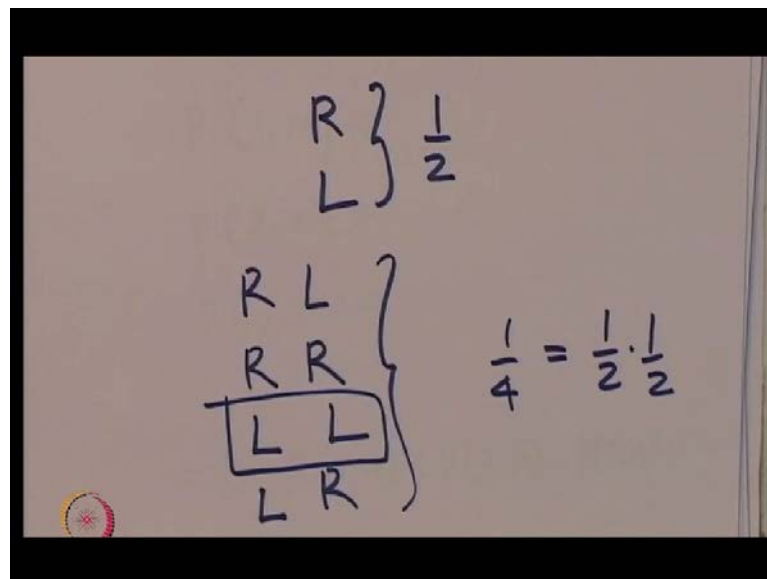


So, you can ask the question, what is the probability of finding retinoblastoma in the left eye? You can ask the question what is the probability that finding retinoblastoma in left

eye. If somebody gets an retinoblastoma as a result of some kind of a mutation then the probabilities are will be on the left eye will be half and the probability on the right eye we have equal probability you can come on the left or it can come on the right eye we do not know. So, equal probability there is no particular difference of mutation like it can, if you look a large number, of large sample of people there is no reason why it is on the left or on the right it can either.

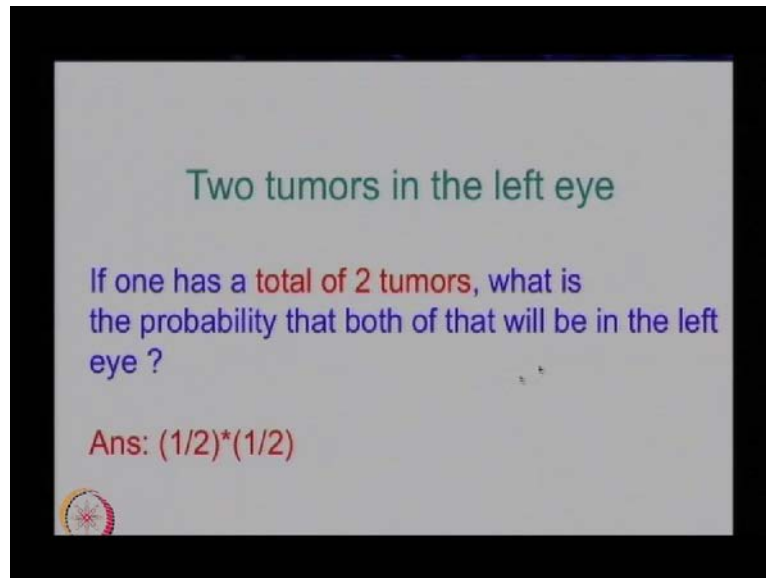
There is no particular preference; so the probability that you find the retinoblastoma on the left eye is half because, there is only two possibilities. It either come randomly or right eye or on the left eye which there are two events that we discussed earlier. There are two outcomes; retinoblastoma on the left eye or retinoblastoma on the right eye either only two outcomes possible. In these two outcomes likes this equally probable; so the probability to come on the left is half on the probability to come on the right is half **ok**. Now, just like we discussed about so what can we have we can have either retinoblastoma in the right eye or on the left eye and this is the probability of half.

(Refer Slide Time: 46:54)



Now you can ask the question if A C T a total of 2 tumors what is the probability that both of what will be in the left eye?

(Refer Slide Time: 47:07)



Two tumors in the left eye

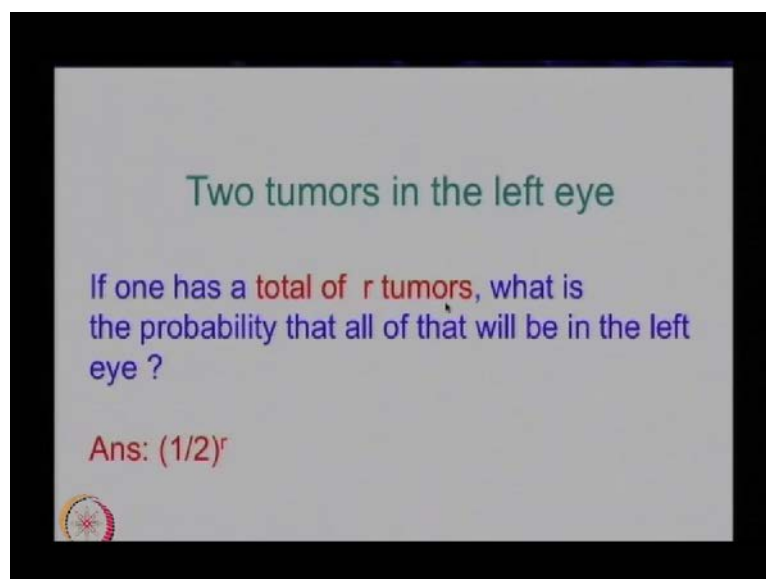
If one has a total of 2 tumors, what is the probability that both of that will be in the left eye ?

Ans:  $(1/2)*(1/2)$

A small circular logo is visible in the bottom-left corner of the slide.

So, if I, somebody if has 2 tumors in the eyes you can have what is the, what are the different possible situations is like it can be on the right 1 one of the left 1 of the right 1 on the left both on right 1 on the left and 1 on the left 1 on the left first will be on the left. The second will be on the right there are 4 possibilities, so the probability that you get on both affect on the left eye you just 1 by 4 so this is basically half into half into so the half probability get here. The next 1 is half and third, 1 so the product of this is 1 by 4 you will get there are 4 events this 1 4 on the this event so you get 1 over 4.

(Refer Slide Time: 48:42)



Two tumors in the left eye

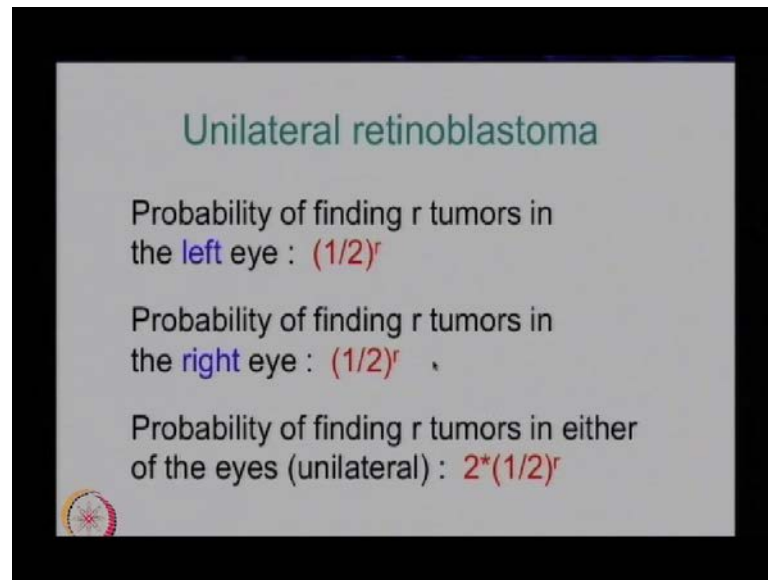
If one has a total of  $r$  tumors, what is the probability that all of that will be in the left eye ?

Ans:  $(1/2)^r$

A small circular logo is visible in the bottom-left corner of the slide.

So if A C T a total of 2 tumors what is the probability that both of that will be in the left eye? Is half into half this is 1 by 4 you can ask the question is if A C T a total of R tumors what is the probability that all of that will be in the left eye that is half power of r half into half into half that r times half power r.

(Refer Slide time: 48:58)



Unilateral retinoblastoma

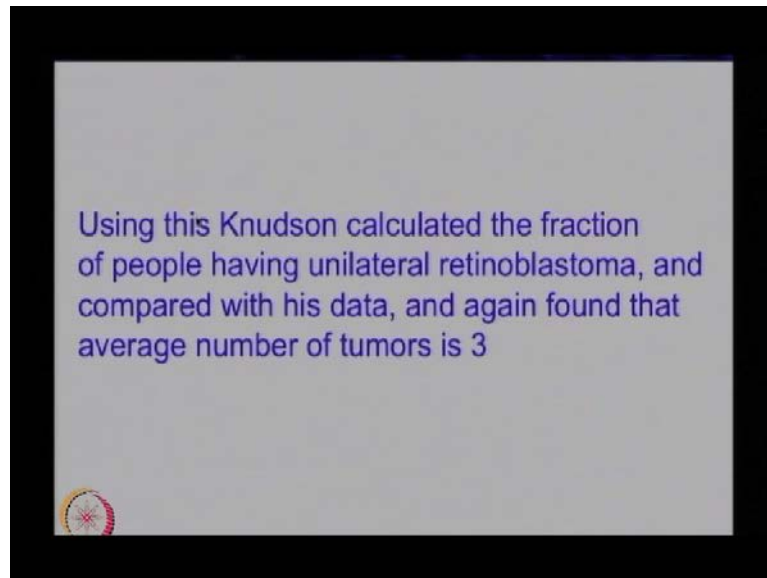
Probability of finding r tumors in the left eye :  $(1/2)^r$

Probability of finding r tumors in the right eye :  $(1/2)^r$

Probability of finding r tumors in either of the eyes (unilateral) :  $2*(1/2)^r$

So what do we have in the unilateral retinoblastoma probability of finding r tumors in the left eye is half power r probability of finding r tumors in the right eye is half power r so the total like either left eye or in the right eye is 2 times of half power r which is 2 into half; so this is 2 into half power r; **from this Knudson can find out Knudson**

(Refer Slide Time: 49:28)



Using this calculate Knutson calculated the fraction of people having unilateral retinoblastoma and compared with his data and again found that average number of tumors is 3 because he could already calculate probability that there are  $r$  tumors when the average is 3 then you could calculate what fraction of this is unilateral that is half power  $r$  into 2.

(Refer Slide Time: 49:47)

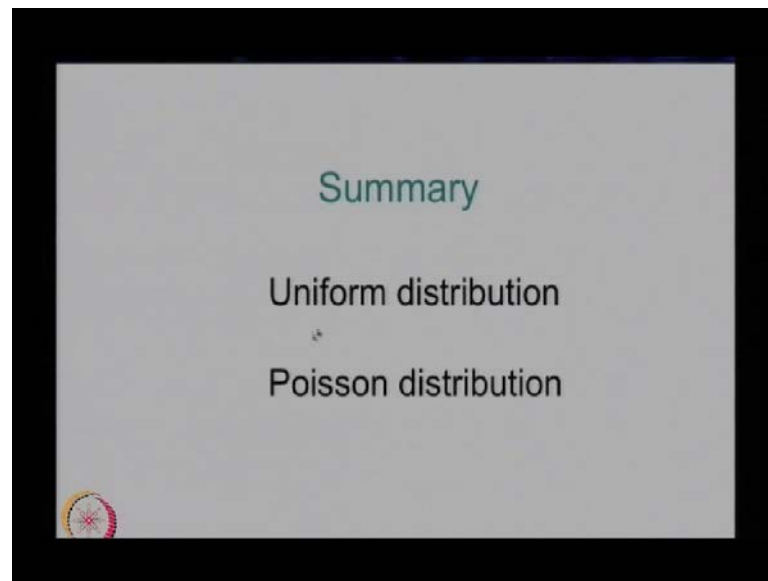
A slide with a light brown background and a black border. It shows a handwritten equation:  $P(r, m=3)$  followed by a downward arrow and  $= 2 \cdot \left(\frac{1}{2}\right)^r$ . There is a small circular logo in the bottom left corner.

So this is the fraction of this is the fraction of proposed having unilateral retinoblastoma will be 2 into half power  $r$  just like we saw. So, from this we could conclude again there



is  $m$  is equal to 3 he found that the fraction of unilateral retinoblastoma and also agreed with this  $m$  equal to 3 update **ok**. So, to summarize we discussed uniform distribution and Poisson distribution and we found that some examples of uniform distribution and Poisson distribution and we on other end discuss an example for 1 can I apply the simple ideas to get some insides about cancer. In this case particular case of retinoblastoma as doctor nelson did this in seventies.

(Refer Slide Time: 50:33)



So, this is the famous paper by Knudson and we will discuss this paper in detail about in detail process he used the some ideas statistics and found out few more things. So, that we will discuss in the coming lecturers so today we will summarize just by saying that we discuss uniform distribution, first Poisson distribution and we will discuss more about statistics later this is all today's lecture; bye.