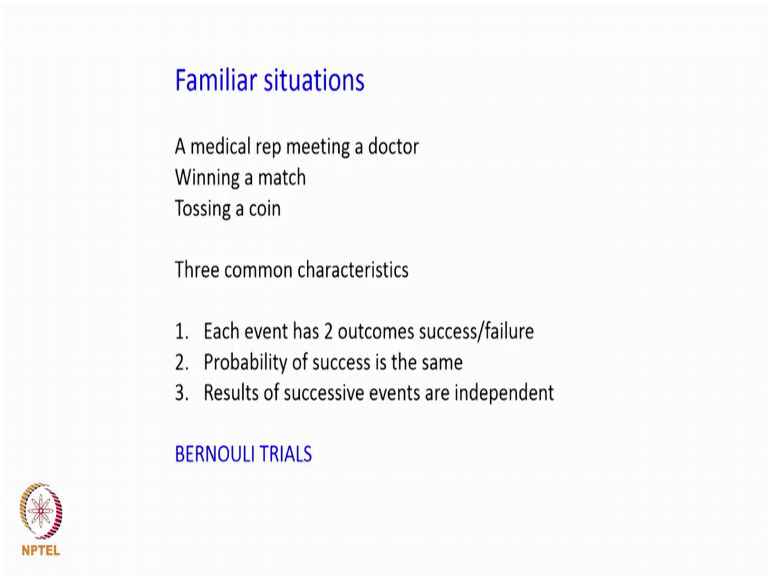
Introduction to Probability and Statistics
Prof. G. Srinivasan
Department of Management Studies
Indian Institute of Technology Madras

Lecture – 19
Binomial Distribution

In this lecture we discuss probability models for counts; we discussed binomial and Poisson distributions 2 discrete distributions. Where we actually try to count and see what is the probability of some count happening out of some possibilities.

(Refer Slide Time: 00:41)



So, let us explain this in detail familiar situations could be tossing a coin which we are very very familiar whenever we study probability. So, the simple question is a if I toss a coin 10 times what is the probability that I get 4 heads, what is the probability that I get 6 tails, what is the probability I do not get a head at all what is the probability that I get more than 1 head and so on.

Another situation that is often discussed in textbooks and literature is a medical representative trying to meet a doctor with a certain probability, representative goes and asks for a meeting and the doctor may meet the doctor may say later. So, there is a probability of meeting the doctor similar questions come out of 10 times, how many times or what is the probability that the representative is able to meet the doctor 6 times.

Just to extend it if the probability of meeting the doctor is the same, if he meets 10 different doctors what is the probability or tries to meet 10 different doctors what is the probability that he meets 6. So very similar situation; winning a match is reasonably familiar in the sense just like tossing a coin, there is a probability associated with winning a match.

All these have 3 common characteristics each is a random variable which has 2 outcomes, one of which is called a success and the other is called a failure. The moment we call one of the outcomes a success the other automatically becomes a failure. Now in the example of winning a match you could say winning a success and losing is failure, in the example of a medical representative meeting a doctor successfully having a meeting with the doctor would be called success and not being able to meet the doctor could be called failure.

Whereas, in tossing a coin we have to define what a success is and what failure is and it depends on how we define in term one could define probability of getting a head as a success and getting a tail as a failure, somebody else would define probability of getting a tail as a success and head as a failure.

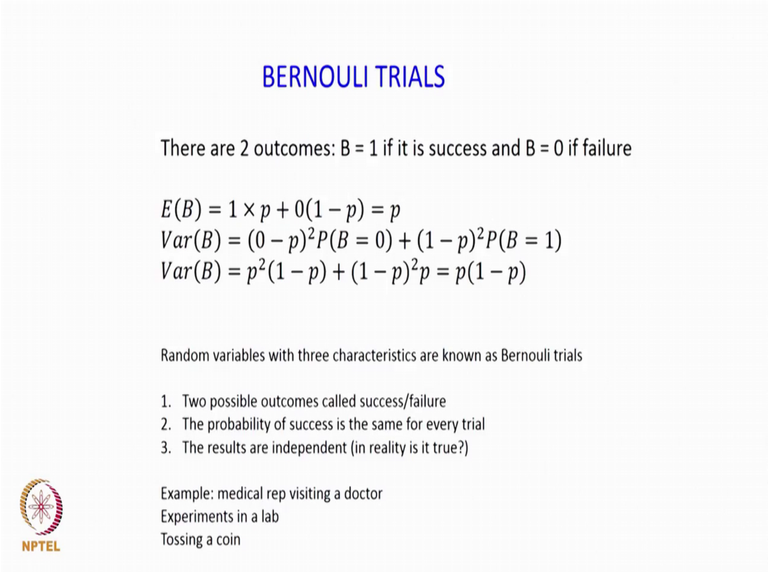
Sometimes when we tried to do inspection and try to find out defective items, success could be identifying a defective item; whereas, in reality a defective item would not mean something successful it would mean something that is not successful. So, it only depends on what we define as success and what we define as not success which becomes failure. So there are 2 outcomes which we call a success and failure.

Now, probability of success is the same irrespective of the number of times it happens tossing a coin is a very good example. So, we might have just got a head and then we toss again what is the probability of getting a head half, it neither increases nor decreases because of the earlier attempt and results of successive events are independent once again tossing a coin is a very good example of successive events being independent. It actually does not matter whether the previous toss resulted in a head or a tail, the probability of head and tail remained the same, so to that extent they are independent.

If we look at winning a match it is expected to be independent it does not matter whether you won the previous match or not, but then you play a match your probability of victory is the same. Medical representative meeting a doctor is expected to be independent at times we may question that because, maybe the last attempt we the medical representative was able to meet the doctor and therefore the doctor might possibly decline and so on.

But if we extend the same example by saying that this medical representative is trying to meet 10 different doctors then we can quickly understand that the events are independent unless the doctors talk to each other. But let us assume that these 3 common characteristics are there in this situation and such a trial is called a Bernoulli trial.

(Refer Slide Time: 05:17)

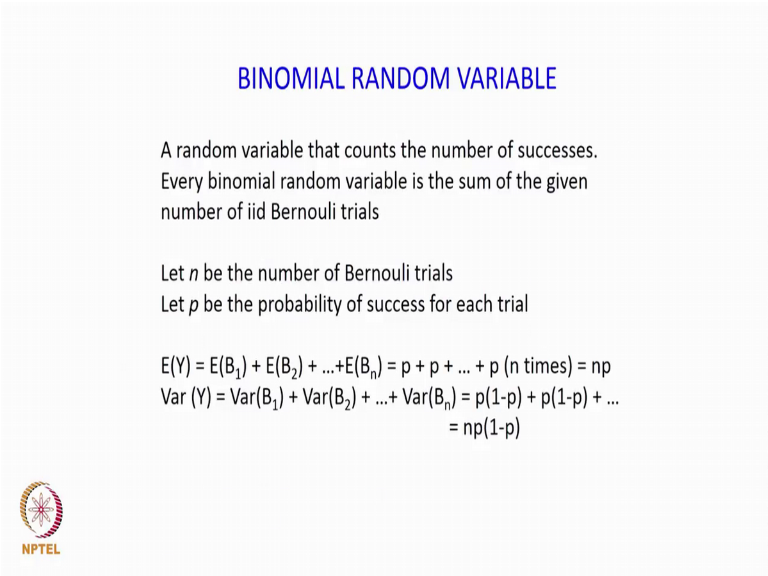


So, again we represent the same thing there are 2 outcomes B equal to 1 the trial is if it is a success and a 0 if it is a failure, success is with a probability of p and failure is a probability of 1 minus p. So, expected value of B is 1 into p plus 0 into 1 minus p which is p. The variance of B is 0 minus p the whole square into probability of B equal to 0 plus 1 minus p the whole square into probability of B equal to 1, which is p square into 1 minus p plus 1 minus p the whole square into p which is p into 1 minus p, so again to repeat random variables with 3 characteristics are known as Bernoulli trials.

So, there are only 2 possible outcomes which are called success and failure, probability of success is the same for every trial and the results are independent. So, we just ask a question in reality it is true we discuss this aspect particularly with the medical representative visiting a doctor, but then if there are 10 doctors and we want to do that, then they are independent.

Same thing is true with tossing a coin, the problem is the same whether the same individual tosses a coin 10 times and you want to find out the probability of getting 4 heads versus 10 different people tossing at the same time with the same probability of getting a head and then you want to find out; out of these 10 what is the probability that 4 got heads, so the problem is the same.

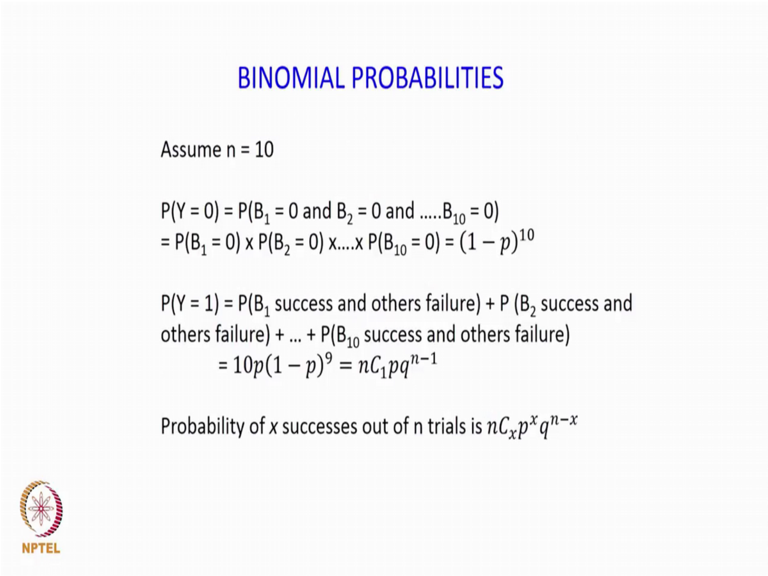
(Refer Slide Time: 06:52)



Now, we define a binomial random variable, a random variable that counts the number of success. So, every binomial random variable is the sum of the given number of iid Bernoulli trials independent identically distributed in independent Bernoulli trials. So, let n be the number of Bernoulli trials and p be the probability of success for each trial. So, expected value of Y is expected value of B 1 plus expected value of B 2 plus expected value of B n which is p plus p plus p n times.

So, when this Bernoulli trial is repeated n times expected value is n into p and the variance of y is variance of B 1 plus variance of B 2 and so on. So, it is p into 1 minus p plus p into 1 minus p n times, so n into p into 1 minus p, we consistently use p and 1 minus p to represent the probability of success and probability of failure a times we also use q equal to 1 minus p as an additional notation and then say that the variance is n into p into q where q is 1 minus p which is the probability of failure.

(Refer Slide Time: 08:13)

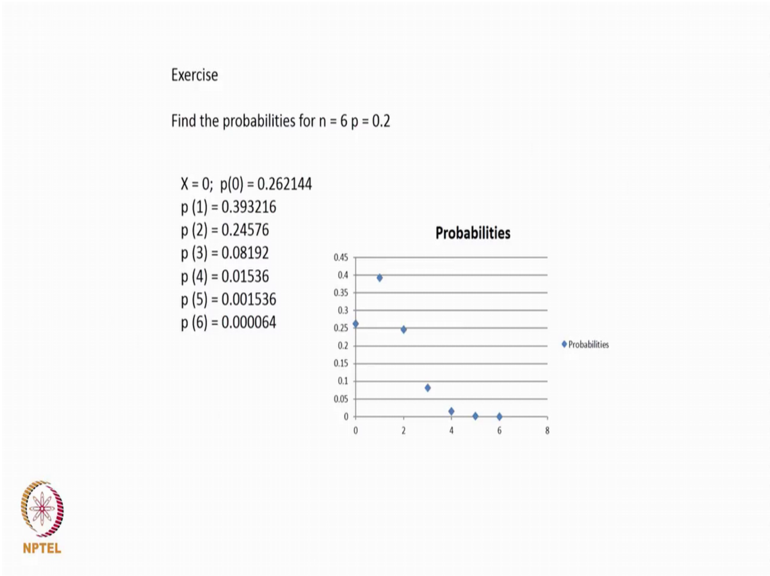


Now, assume now we define what are called binomial probabilities so assume n equal to 10. So, probability of y the random variable equal to 0 will be probability of the first one equal to 0 and second one equal to 0 and the third one equal to 0 and so on. So, each is a failure so each is 1 minus p so 1 minus p multiplied 10 times, so 1 minus p to the power 10; Y equal to 1 success. So, 1 success out of 10 is the first one being successful in the others fail the second one being successful and the others fail and so on.

So, its 10 times p into 1 minus p to the power 9 and in general we can now show that probability of x successes out of n trails is n C x p to the power x q to the power n minus x. So, there are n trials out of which x is successful so that is p to the power x the remaining n minus x our failure. So, q or 1 minus p to the power n minus x and the x successes out of n trials can happen n C x times, therefore n C x p power x q to the power n minus x. For example, if we extrapolate this as y equal to 2 then one could go ahead and say 1 and 2 being successful the rest not 1 and 3 being successful the rest not 1 and 4 being successful and so on.

So, finally it boils down to choosing 2 out of 10, 10 C 2 ways into p to the power x p square q to the power n minus x q to the power 8 or 1 minus to the power 8, so in general its n C x p power x q to the power n minus x.

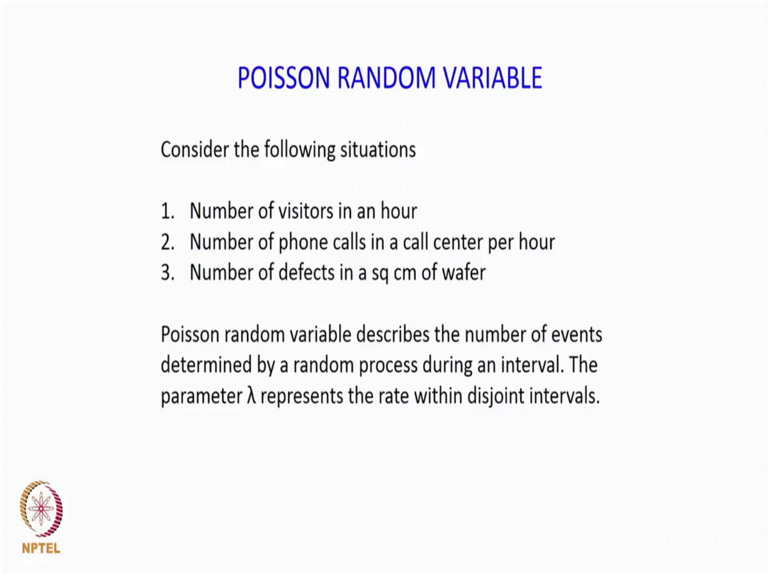
(Refer Slide Time: 10:20)



Just try to find the probabilities for n equal to 6 and p equal to 2, so x equal to 0 n C 0 p to the power 0 q to the power 6 we get 0.262144. So, probability of one success out of 6 is n C 1 6 C 1 p to the power 1 q to the power 5, so which is 6 C1 is 6 into 0.2 into 0.8 to the power 5 which is 0.393, 2 out of 6 is 245, 3 out of 6 is 081, 4 out of 6 is 015, 5 out of 6 is 001536 and all 6 out of 6 is 000064. If we try to plot these they obviously they add up to 1 we can check that 0.26, 0.39 is roughly about 0.65 this 0.25 is about 0.8 0.8889 and so on, 0.26 plus 0.39 is about 0.65 here it is about 0.25. So, 0.65 plus 0.25 is 0.9, 0.98 0.99 and the fractions add up to 1.

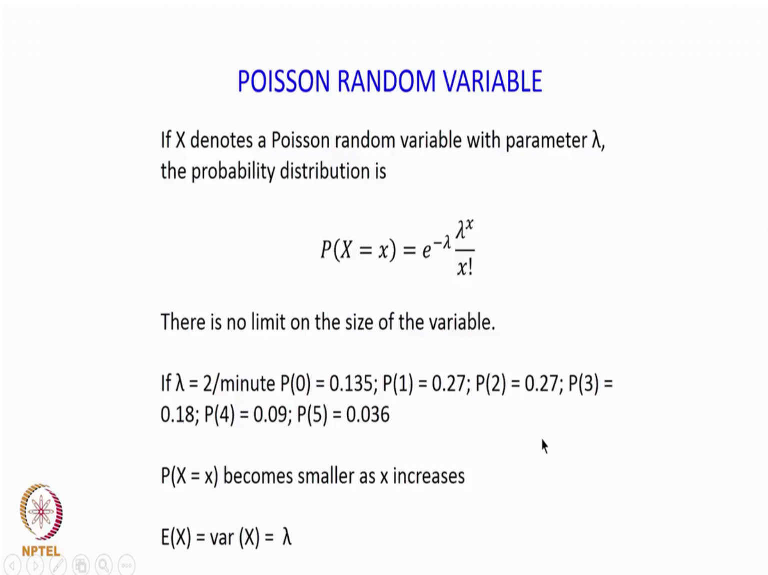
The plot also tells us something interesting that when we have n equal to 6 and depending of course on p equal to 2, since p equal to 0.2 the maximum probability happens for 1 here and so on and one can show that as p increases it moves a little bit to the right. But after some p of 4 p of 5 p of 6 etc you realize that they have very very small values and they kind of come close to 1 as we add them they come close to 1 the smaller values are closer to 0 and progressively decreasing.

(Refer Slide Time: 12:12)



Now, we try to look at Poisson random variables. So, we look at again some situations the number of visitors in an hour, the number of phone calls in a call center per hour, number of defects in a square centimeter of wafer and so on. So, Poisson random variable describes the number of events determined by a random process during an interval it is very important during an interval, the parameter lambda which is shown by this symbol here the letter the Greek letter lambda represents the rate within the disjoint intervals.

(Refer Slide Time: 12:48)



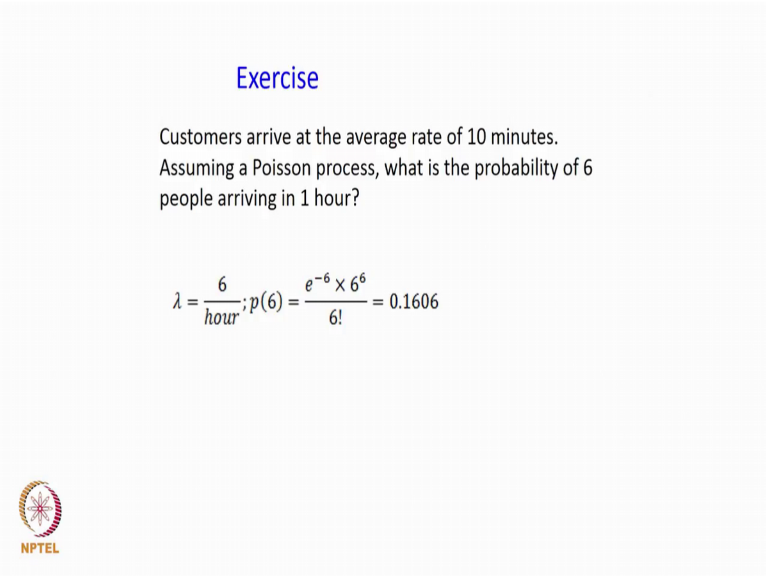
So, if x denotes a Poisson random variable with a given parameter lambda, then the probability distribution of p of X taking the value small x is equal to e to the power minus lambda lambda to the power x by x factorial. Now in the Poisson distribution there is no limit on the size of the variable X can take any value, example if lambda is 2 per minute.

For example, we say that people arrive at the rate of 2 per minute in probability of 0 people arriving is 0.135 which comes from e to the power minus lambda lambda power x by x factorial. Probability of 2 people coming in the interval is 0.135 probability of 1 person coming in the interval is 0.27, so that is got by e to the power minus lambda lambda to the power 1 by 1 factorial where lambda is equal to 2.

So, probability of 3 people coming in that interval is 0.18, 4 people is 0.09 and 5 is 0.036 and as X increases small x increases the probability of X equal to x becomes very very small. So, even here if the average is 2 per minute it is it is fairly acceptable that no person comes 13 percent of the times, 1 person comes 27 percent of the times, 2 people come 27 percent of the times and so on.

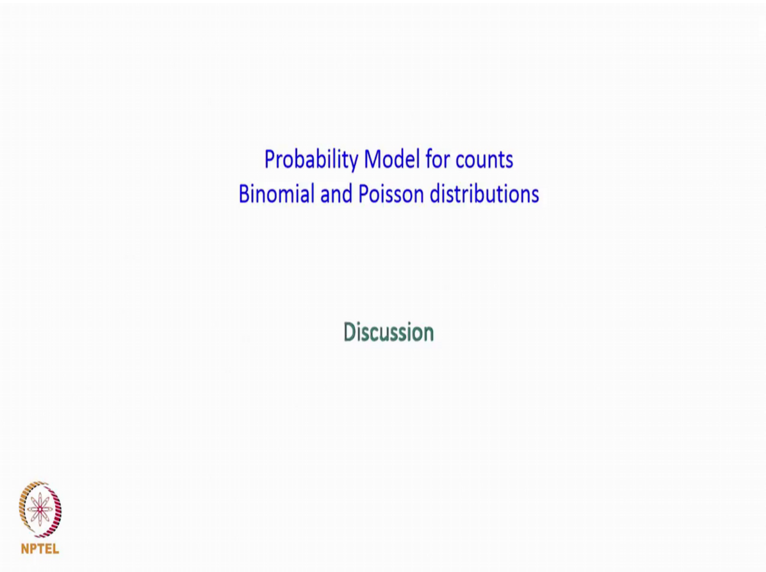
If we start adding 0.135 plus 0.27 is 0.405 plus 0.27 is 0.675 plus 0.18 is about 0.855 0.951, we realize that around with p equal probability of 5 it almost reaches 1, but then X can take any value. So, as small x becomes larger the probability becomes very very small in a Poisson random variable. Though we are not going to prove this so probability of X equal to x becomes smaller expected value of the random variable is lambda the variance is also equal to lambda.

(Refer Slide Time: 15:02)



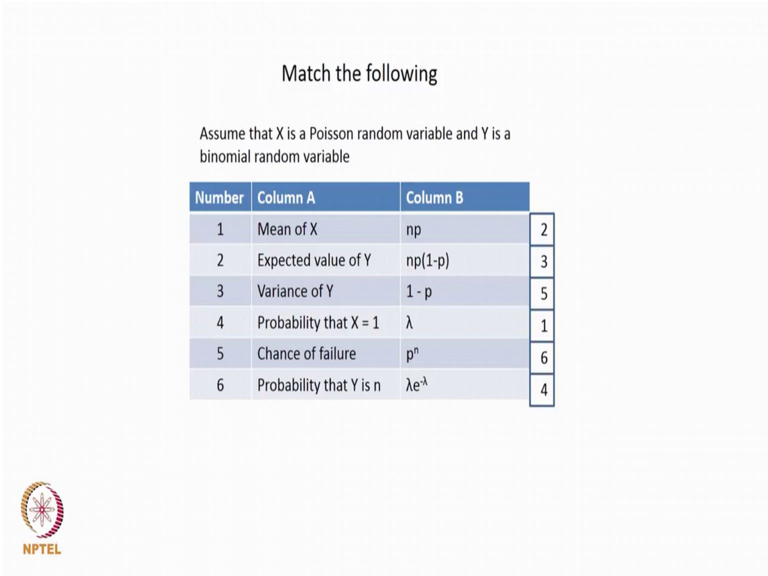
Customers arrive at an average rate of 10 minutes assume a Poisson process what is the probability of 6 people arriving in the next 1 hour. So, lambda is 6 per hour 10 minutes so 6 per hour, p of 6 is e to the power minus lambda lambda power x by x factorial e to the power minus 6 into 6 to the power 6 by 6 factorial which is 0.1606. So, even on an average 6 people arrive in an hour on an average, but then we realize their actual probability of 6 people arriving in an hour it is very small.

(Refer Slide Time: 15:41)



So, let us continue on this topic with a little bit of discussion as we have been doing in all previous topics.

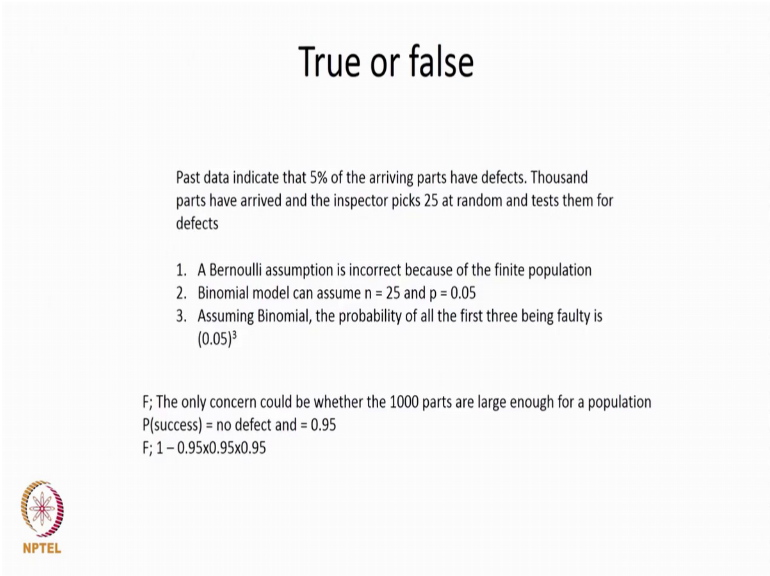
(Refer Slide Time: 15:47)



So, we now have a match the following, so assume that X is a Poisson random variable and Y is a binomial random variable. So, we try to match mean of X so mean of the Poisson variable is lambda so mean is lambda, expected value of Y so Y is a binomial random variable so expected value is equal to n p; variance of Y, Y is a Y is a binomial variable so n into p into q r n into p into 1 minus p which is shown here probability that X equal to 1 X is Poisson. So, probability that X equal to 1 is the equation is e to the power minus lambda lambda to the power X by X factorial

So, when we put X equal to 1 X factorial is 1, so e to the power minus lambda into lambda to the power 1 which is lambda e to the minus lambda chance of failure 5 binomial, so binomial we define success and failure. So, chance here is probability, so probability of success is p probability of failure is 1 minus p probability that Y is equal to n binomial. So n successes n C x p power x q to the power n minus x, so n C n p to the power n q to the power n minus n n C n is 1 q to the power n minus n is q to the power 0 which is 1 and therefore the value is n C x p power x q power n minus x, n C n p power n q to the power n minus n which is p to the power n.

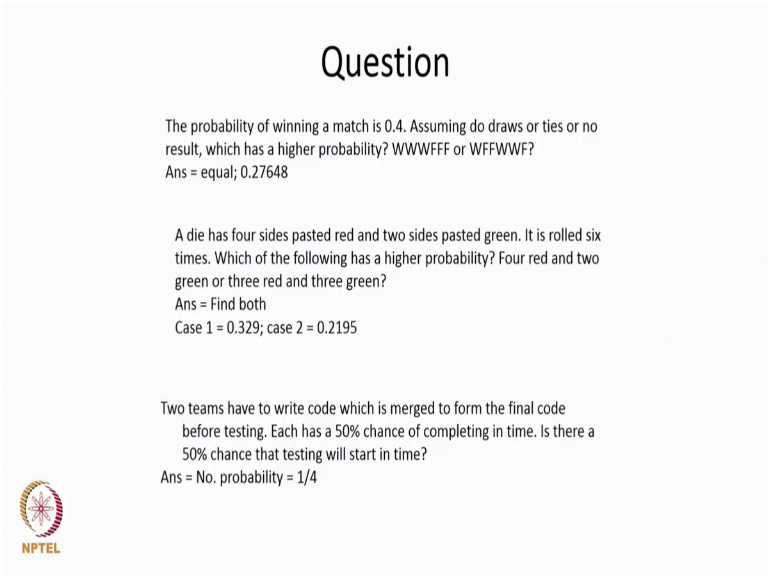
(Refer Slide Time: 17:28)



Now, let us look at some situations and try to study them past data indicate that 5 percent of the arriving parts have defects, 1000 parts have arrived and the inspector picks 25 at random and tests them for defects. A Bernoulli assumption is incorrect because of finite population one may disagree with this one can say that 1000 parts are large enough for a population. But then we could take this has to be reasonably large and continue and that is exactly how most of inspection also happens, that we take a reasonably large number and then we take a small fraction of them to do the inspection.

Binomial model can assume n is equal to 25 and p equal to 0.05. So, 5 percent so it depends on what we define a success and what we define a failure. So, if defect is a success then n equal to 20 p equal to 0.05 if not being defective is the success then p is 0.95. Assuming binomial the probability of the first 3 being faulty is 0.05 cubed it would not p. So, this will be one minus 0.95 into 0.95 into 0.95 and so on

(Refer Slide Time: 18:42)



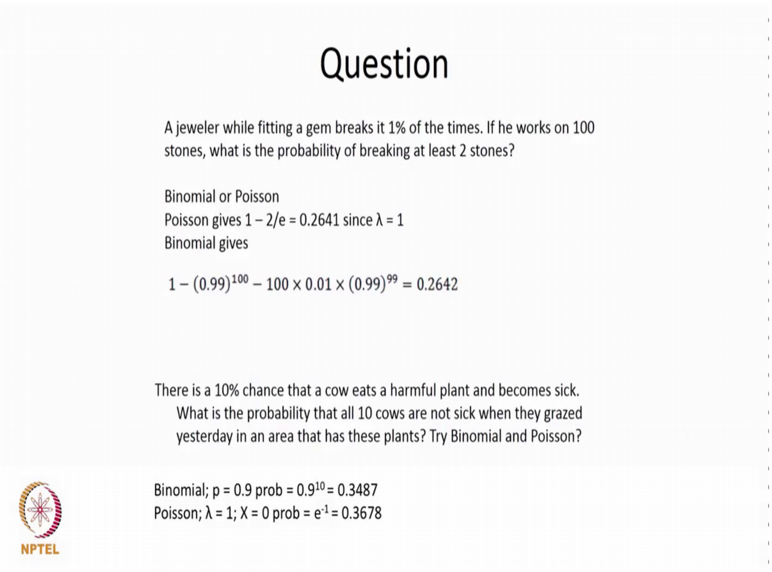
Next one probability of winning a match is 0.4 assuming that there are no draws or ties or no results and so on, which has a higher probability win win win FFF is lose or fail and win win win fail fail win win and failed. Now we look at try to model this as binomial then we realize that out of 6 matches 3 victories and 3 defeats is the probability that we are looking at. So, the sequence does not matter I think that is that that is a big learning from this the sequence does not matter. So, probability of 3 wins irrespective of the order in which they arrive is the same. So, this will be n C x p power x q to the power n minus x. So 6 and 3 so we could do this 6 C 3 0.4 to the power 3 0.6 to the power 3 which works out to be 0.276.

A die has 4 sides pasted red and 2 sides pasted green it is rolled 6 times which has a higher probability 4 red and 2 green or 3 red and 3 green. Even though this question is about a die so it is not about the numbers 1 to 6, therefore we should not use the probability of 1 by 6 and so on.

Now, this has 4 sides pasted red and 2 sides pasted green, so if we define red as a success then probability of success is 4 by 6 which is 2 by 3 and probability of failure is 1 by 3. Now we have to find out probability of 4 red and 2 green which is given by n C x p power x q to the power n minus x. So, we would have 6 times it is rolled so 4 red, so 6 C 4 2 by 3 to the power 4 1 by 3 to the power 2 which is 0.329 and the other one is 6 C 3 2 by 3 to the power 3 1 by 3 to the power 3 which is 0.2195 and therefore 4 red and 2 green has a higher probability than 3 red and 3 green.

Now, 2 separate teams have to write code which is merged to form the final code before testing, each has a 50 percent chance of completing in time. Is there a 50 percent chance that the testing will start in time no, it would be one could take 1 by 2 as success and 1 by 2 as failure because of the 50 percent and then we realize the answer is actually when it started in time will be both will be successful. So, 2 C 2 half to the power 2 half to the power 2 minus 2 which is 1 by 4, another way of doing it is probability that term team a successful is 0.5 team B successful 0.5 both being successful is 0.5 into 0.5 which is 0.25 and therefore we do not have a 50 percent chance of starting the testing in time.

(Refer Slide Time: 21:45)

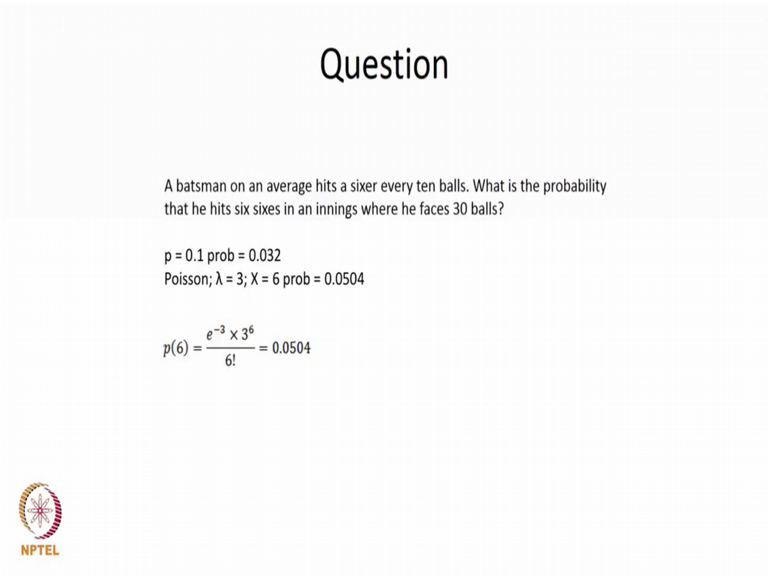


Now, a jeweler while fitting a gem into an ornament breaks at 1 percent of the times, if he works on 100 stones what is the probability of breaking at least 2 stones. So, we could model this as binomial or Poisson. Poisson would give us a lambda, so 1 percent of the times he breaks out of 100 times. So, we can take lambda equal to 100 into 1 percent which is 1 and therefore Poisson breaking at least 2 stones is 1 minus probability of breaking no stone less probability of breaking 1 stone. So, each would become 1 by e therefore, the answer is 1 minus 1 by e minus 1 by e which is 1 minus 2 by e which is 0.2641.

Now, if we use binomial then we would have 1 minus probability of 0 break and 1 breaking. So, 1 minus 0.99 to the power 100 minus 100 C 1 which is 100 into 0.01 in to 0.99 to the power 99 which on simplification gives us 0.2642. So, we also observe that in this instance either a binomial way of approaching it or approaching it as Poisson gives us the same probability. There is a 10 percent chance that a cow eats a harmful plant and becomes sick, what is the probability that all 10 cows are not sick when they graced yesterday in an area that has these plants try binomial and Poisson. So, p is 0.9 because there is a 10 percent chance that the cow can become sick, therefore probability that all the cows are not sick is 0.9 to the power 10 which is 0.3487.

So, if we look at poison 10 percent chance there are 10 cows, so lambda is 1 X equal to 0. So e to the power minus 1 e to the power minus lambda lambda to the power X by X factorial so e to the power minus 1 1 to the power 0 by factorial 0 so e to the power minus 1 which is 0.3678

(Refer Slide Time: 23:54)



Batsman on an average hits a 6 every 10 balls what is the probability that he hit 6 sixes in an innings where he faces 30 balls. So, every 10 balls he hits one 6, so p equal to 0.1 q equal to 0.9 and then we have to do out of 30 what is the probability of hitting 6 sixes. So, 30 C 6 0.1 to the power 6 0.9 to the power 24 which is 0.032, when we do a Poisson so he hits a 6 every 10 balls so 30 balls so lambda is equal to 3 and x is equal to 6 sixes so probability is 0.0504. So, p of 6 is equal to e to the power minus 3 into 3 to the power 6 by 6 factorial 0.0504.

Poisson in binomial we have used alternately for some problems you have actually used both, it is also possible to show that binomial approaches Poisson when n is large and p is small and it approaches Poisson distribution and therefore we would find that in some cases the answers are close, while in some cases the answers are slightly different. So, with this we complete our discussion on binomial and Poisson models and in the next lecture we would look at the normal distribution and with that we would summarize the course and wind up the course after we study normal distribution.