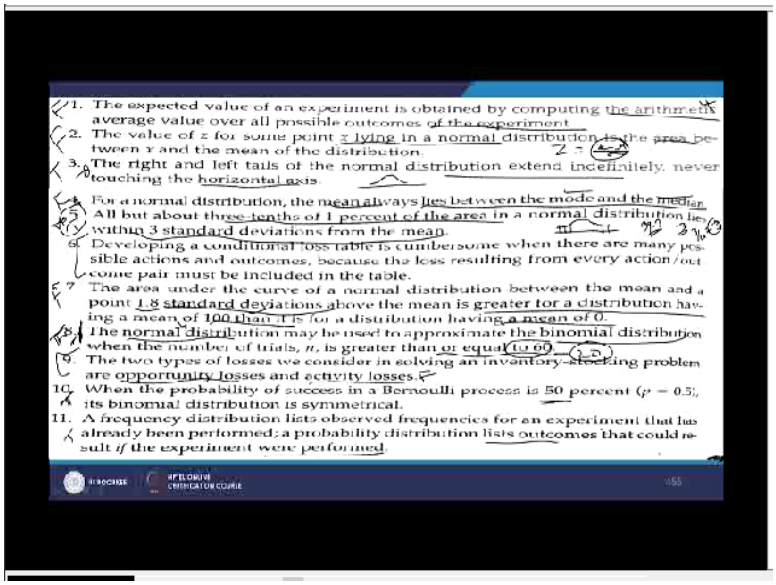
Business Statistics
Prof. M. K. Barua
Department of Management Studies
Indian Institute of Technology-Roorkee

Lecture-21
Chapter Concepts–Probability Distributions

Good morning friends, I welcome you all in this session as you are aware in couple of last few sessions we have discussed about probability and probability distributions. We have seen in probability distribution basically two types of distributions discrete and continuous. In discrete we have seen binomial, Poisson. In continuous distribution we have seen normal you have uniform and we have seen exponential distribution.

In today's session I would like to take up couple of questions related to probability distributions and these questions will make your fundamentals clear. So, let us look at first question and these are different statements and you have to identify whether these statements are true or false.

(Refer Slide Time: 01:23)



So, the first one is the expected value of an experiment is obtained by computing the arithmetic average value over possible outcomes of experiment. So that there is something called expected value which is nothing but a weighted mean, so this not arithmetic mean right. So this statement is false. Now if you look at second statement the value of Z for some point X lying in normal

distribution is the area between X and the mean of the distribution, is it like that? What is Z, Z is standardized normal distribution right.

And how do we calculate Z, Z was it was is equal to $x-\mu/\sigma$ is not it, so what about second statement is that true or false. It is again false right because Z is equal to this right. Now let us look at statement number 3 the right and left tails of normal distribution extend indefinitely never touching the horizontal axis is that true. We have several characteristics of normal distribution and one of them was that the normal distribution this never touches horizontal axis right.

So, this statement is true ok, the next one is for a binomial distribution the mean always lies between mode and for normal distribution the mean always lies between mode and median is it true what are the characteristics of normal distribution let us how one of them is this is not it third one. And there are several characteristics for example it is a bell shaped curve, it is symmetric curve is not it, and mean mode and median are at the same point right.

So, mean does not lie between mode and median, so this is false right. Statement number 4 is false. Let us look at statement number 5th, all but about 3 10th of 1% of the area in normal distribution lies within 3 standard deviation from the mean. So, what is this, let me draw a curve for this question, so if you look at this you have got this is your normal distribution for example right. So, you have got 1 sigma limit, 2 sigma limits and 3 sigma limits right.

So, within 3 sigma what is the area under the curve it is 99.7 is not it. So, all but 3 10th of 1 = what is that value this point 13 right 3 of 10th is 3/10 is 0.3 is not it. So, 99.7 + 0.3 it becomes 100% right. So, this statement is true right statement number 5 is true, let us look at the next one developing a conditional loss table is cumbersome when there are many possible actions and outcomes.

Because the loss resulting from every action since this is on expected value and we did not discuss much on expected values. So, you can skip this for the time being right question number 6 ok, let us look at the next one the area under the curve of a normal distribution between mean

and mean and then point 1.8 of standard deviation above the mean is greater than for a distribution having mean this, then it is for the distribution having mean 0.

So, there are 2 distributions one has got mean 100 the other one has got mean 0 and if there is a point which is 1.8 standard deviation above the mean would that be greater than for this mean, correct or incorrect question number 7 is actually a false statement, is not a correct statement ok. Now look at the next one, 8th one, the normal distribution may be used to approximately binomial distribution when the number of trials n is greater than or equal to 60 is that correct.

We have seen that we can always use normal distribution for binomial distribution when the number of trials in is greater than or equal to 60 is that correct, no this not correct statement. Because it is np and nq both have to be greater than or equal to 5 right, so statement number 8 is false. So, this is false right statement number 8 is false, now let me tell you one more thing here we can approximate binomial distribution.

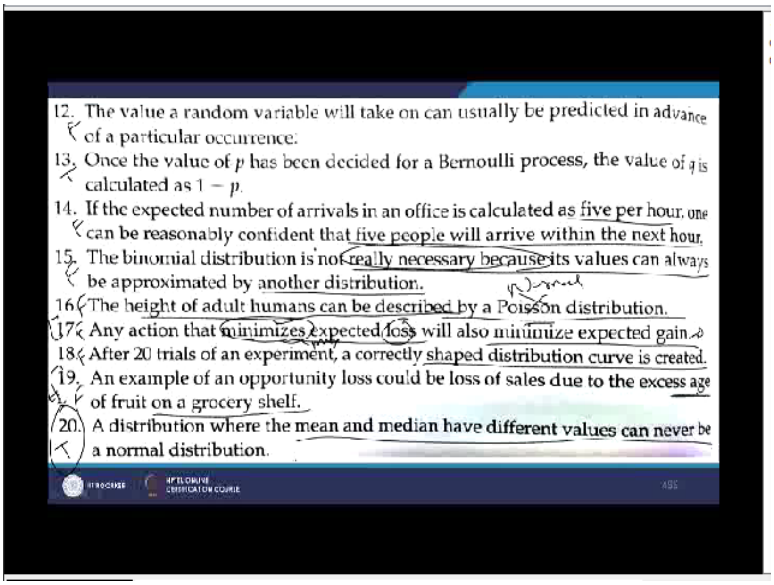
As a Poisson distribution or in other words we can use Poisson distribution instead of binomial distribution if N is greater than or equal to 20. So, had it been 20 over here and had been Poisson distribution here it would be this statement would have become true right but right now this one is false. Let us look at the next one the 2 types of losses be considered in solving an inventory stock in problem or opportunity losses in fact this again the question on expected value so we can skip it but for the but this statement is false.

Because in expected value we always considered 2 types of losses opportunity loss and obsolescence loss right, so this statement is false but you can escape for the time being. Let us look at next question when the probability of success in Bernoulli process is 50% or P is equal to 0.5 it is binomial distribution is symmetrical is it correct, yes this statement is true right. Because if it is 0.5 then it is symmetric if less than 0.5 or more than 0.5 then it becomes asymmetric.

So for statement number 10 this is true right, next let us look at next one a frequency distribution lists observed frequencies for an experiment that is already been performed. A probability distribution list outcomes that good result if the experiment were performed its or incorrect true

or false. This statement is true ok, so frequency distribution is lists observed list observed frequencies for an experiment that is already been performed while a probability distribution list outcomes right.

(Refer Slide Time: 10:00)



Let us look at the next one, value of random variable will take on can usually be predicted in advance of a particular occurrence is it possible can you predict the random variable value, no this is false right. Now let us look at the next one statement number 13, once the value of p has been decided for Bernoulli process the value of q is calculated 1-p is it true, yes this is 100% true right, because p is known you can calculate Q if Q is known you can calculate P right.

Because p + q=1, let us look at the next one the expected number of arrivals in an office is calculated as 5 per hour. One can be reasonably confident that 5 people will arrive within the next 1 hour or within next hour is it true. So, you know from past data that the 5 people arrive office in an hour time, so what you say that in next 1 hours' time 5 people will arrive as well no this false right, 14 number statement is false.

Let us look at the next one, the binomial distribution is not really necessary because its value can always be approximated by another distribution. We have seen in one of the session that you can use Poisson distribution instead of binomial distribution you can use even normal distribution instead of binomial distribution. So can we say that it is not really necessary, no it is necessary

because we can approximate binomial distribution as Poisson and as normal distribution under certain conditions right not always.

So this statement is false ok, let us look at the next one, the height of adult humans can be described by Poisson distribution is it possible. We have seen the characteristics of Poisson distribution and binomial distribution normal distribution. So what do you think, is it possible the height of adult humans can be described by Poisson distribution no, not by Poisson but normal distribution ok. So this statement is false 16 number is false.

Let us look at the next one any action that minimizes expected loss will also minimize expected gain. Though again this is a question on expected value we can skip this but this statement is false. Because when you maximize when you minimize losses your gains will automatically increase right, so both cannot be minimize right. So, minimize expected orders will also will maximize the expected gains right.

If let say maximizes expected losses will minimize expected gain, so here it been like this maximize any action that maximizes expected losses will also will minimize expected gain. In that case it would have been true statement right but right now this one is false right. Let us look at the next one of a 20 trials of an experiment a correctly shaped distribution curve is created is it possible just by having 20 trials.

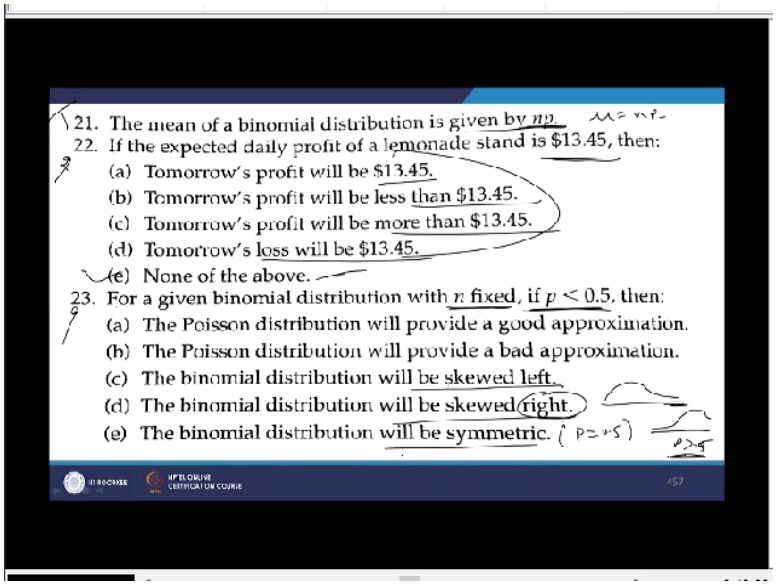
You can have a correctly shaped distribution, no it is not possible it is false right is not it. So, you need to have large number of trials right. let us look at the next one an example of an opportunity loss could be loss of cells due to the excess age of fruit on grocery sale. In fact this again equation on expected value but let me tell you what happens in when we calculate expected value of a given question.

Then generally we have got two types of losses either opportunity loss or obsolescence loss. Now opportunity loss is basically a loss when there is a demand of an item and you do not keep stock of that item, so that is kind of opportunity loss. Now obsolescence loss is a loss wherein

you have the item but you are not able to sell it. So, **so** this is a case of obsolescence loss not the opportunity loss, so this is false question number 19 is false right.

A distribution where the mean and median have different values can never be normal distribution is it yes is correct a distribution where mean and median have different values can never be a normal distribution yes it is correct right point number 20 is correct.

(Refer Slide Time: 16:17)

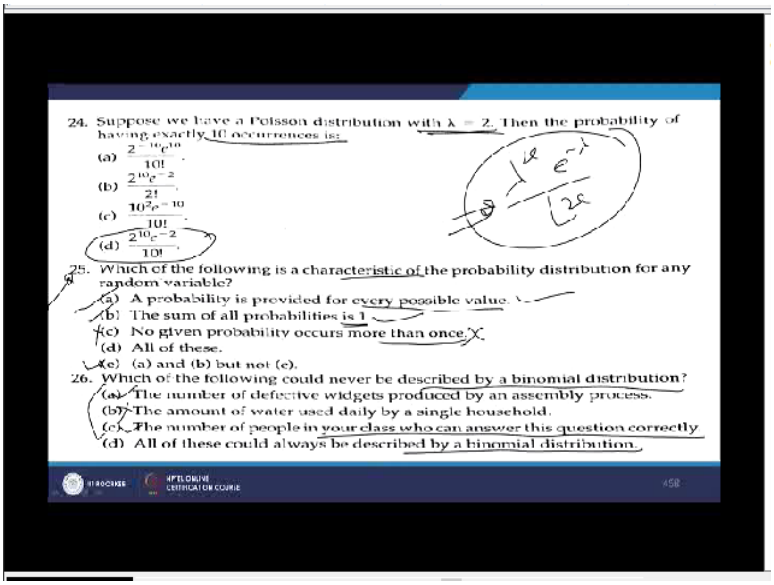


Let us look at the next one, the mean of binomial distribution is given by np yes correct or not correct yes it is correct, so μ=np right ok, what about next one so 21st is true, the expected daily profit of lemonade stand is dollar 13.45. Then tomorrow's profit will be this much tomorrow's profit will be less than this much more than this much or tomorrow's loss will be this much or none of the above. If the expected daily profit of lemonade stand is 13.45.

So what would be the tomorrow's profit or tomorrow's loss you cannot identify right, so this none of the above so E is the right answer for question number 22nd. Let us look at next question 23rd for a given binomial distribution with n fixed, p is less than 0.05, the Poisson distribution will provide a good approximation. The Poisson distribution will provide a bad approximation the binomial distribution will be skewed left or it will skewed right or the it will be symmetric.

Let me tell you that the binomial distribution will would be symmetric when p=0.5 is not it but it is less than 0.5. So, what would be the answer, it will be right skewed it would be like this ok. So, for p, p less than 0.5 right skewed and p more than 0.5, more than 0.5 would be left skew right. So, just remember these 3 in case of it is 50% it would be symmetric, less than 50% it would be right skewed otherwise left skewed.

(Refer Slide Time: 19:01)



Now suppose we have a Poisson distribution with λ=2 then the probability of having exactly 10 occurrences is what is 2 to the power -10, e to the power 10, factorial 10, 2 to the power 10, e to the power -10 factorial 2, 10 to the power 2 e to the power -10 factorial 10 2 to the power 10 e to the power -2 factorial 10 which is correct. In fact you need to remember formula for Poisson distribution right, so this λ to the power x, e to the power - λ is the formula right for Poison distribution.

So we know that λ = 2 right, so 2 to the power 10, right and this one is the right. So this is a quite a simple question because you know this formula right. Now look at the next one which of the following is a characteristics of the probability distribution for any random variable, which of the following is a characteristics of probability distribution for any random variable.

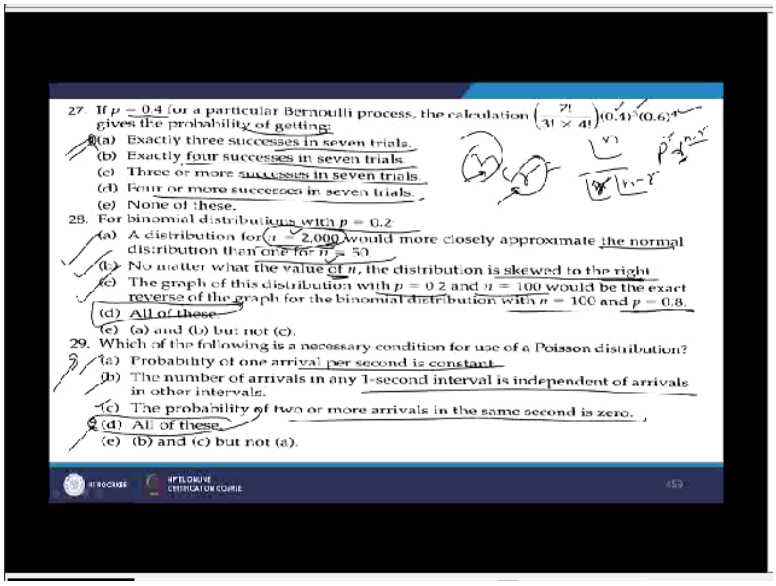
A probabilities provided for every possible value, the sum of all probabilities is 1, no given probability occurs more than once, all of this a and b but not c what do you think which of the

following is a characteristics of probability distribution for a given random variable. So, a probability is provided for every possible values yes it we provide probability right. So, this correct, this sum has to be won this also correct no given problem occurs more than one this in correct right.

So, d cannot be the answer because c is incorrect, so e is the right answer, a and b are correct but c is incorrect right a and b are correct c is incorrect ok. Let us look at the next one which of the following could never be described by a binomial distribution, the number of defective widgets produced by an assembly process, the amount of Water use daily by a single household, the number of people in your class who can answer this question correctly.

All these could always be described by a binomial distribution, is it possible, can you describe all these 3 points by binomial distribution, no we can describe the first part and third part not the second part right.

(Refer Slide Time: 22:31)



So, let us move on to next one if p=0.4 for a particular Bernoulli process the calculation this gives the probability of getting what so ^nC_r is equal to what? What is the formula factorial n factor n factorial, n-r is factorial r right, e to the power right. So, ^nC_r number of trials and number of success is right. So, what is this if p=0.4. so, p is 0.4, r is 3 right q is 0.6, n-r is 6 right.

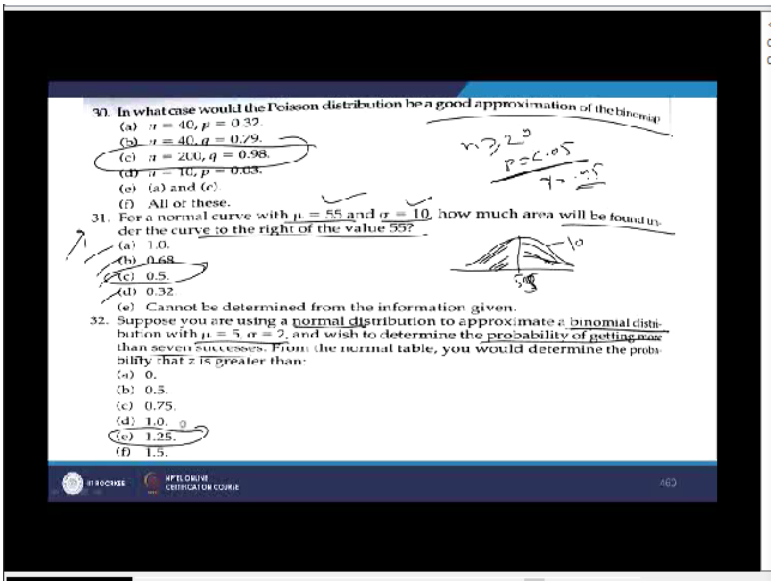
So, what is that value it is exactly 3 success in 7 trials or 4 success in 7 trials, 3 or more successes in 7 trials, 4 or more successes in 7 trials, none of this which one is correct ${}^{n}C_{r}$, number of successes and number of trials. So, for this question the answer a is correct, 3 success in 7 trials right, so here n is 7 right. So, these are trials and r are successes right. So, this is number of successes number of trials, for binomial distribution with p= 0.2.

A distribution with n is equal to this would more closely approximate the normal distribution than more than one for n is equal to, so the difference between these two is here sample size is 2000 here it is let us n is 2000 n is 50. So, which one would be more closely approximate the normal distribution see when n is higher than that would when that would approximate the normal distribution closely right.

So, a is correct for question number 28 right a is correct what about this second part no matter what the value of n the distribution is skewed to right. We have seen in one of the questions in previous slide where we have seen that when p is less than 0.5 their distribution is rightly skewed right. So whatever is the value of n the distribution would be skewed to right. So, this also correct ok, the graph of this distribution p equal to this and n=100 would be exactly reverse the graph for binomial distribution n is equal to this.

And p is equal to 0.8 yes is this also correct, so, the answer to this question is b, so a b and c all these are correct now let look at the next question which of the following is necessary condition for use of Poisson distribution. So, probability of 1 arrival per second is constant yes we know this the number of arrivals in any one second and the other is independent of arrive was yes this is also true. The probability of 2 or more arrivals in same second is 0 yes it is not 0 but it is very small value. So, but we take it equal to 0, so all of this right. So, the answer to question number 29th is all of this.

(Refer Slide Time: 26:57)

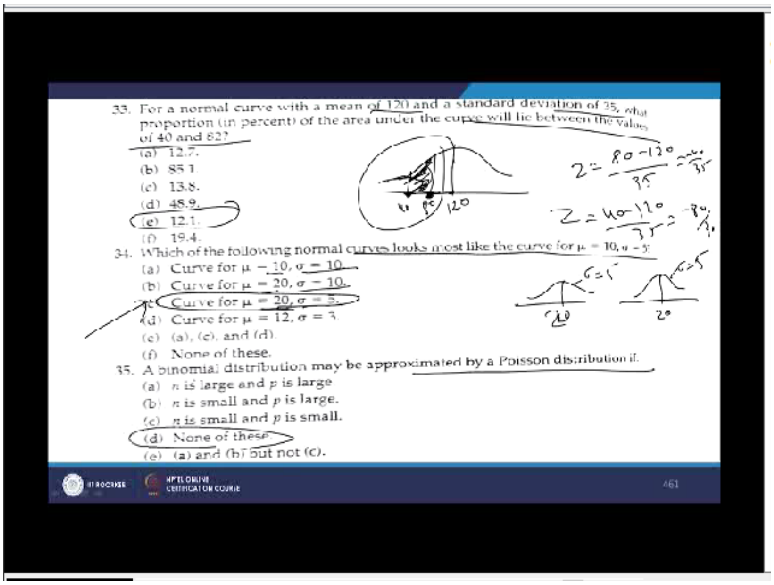


Let us look at the next one in what case would the Poisson distribution be good approximation of binomial in what case Poisson distribution is good approximation of binomial. So, n has to be more than or equal to 20 and p has to be less than or equal to point 0.05 is not it. So, when we said this is 0.05 it means q has to be more than 0.95 is not it. So, n has to be more than 20 and q more than 0.98, so this is the correct answer.

Let us look at next one for normal curve with mean this and standard deviation this how much area will be found under the curve to right of the value 55. So, it is 1.68, 0.5, 0.38 and so this is the answer to question number is 15 for normal curve mean this how much area will be found under the curve to the right of the value 55 is 50% right. So, this is mean 55 sorry 55 is standard deviation 10, so this is 50% right.

Suppose you are using a normal distribution to approximate the binomial distribution, normal distribution which is a continuous distribution to approximate of binomial distribution which is a discrete distribution with mean and standard deviation equal to 5 and 2 respectively. And which to determine the probability of getting more than 7 successes from the normal table you would determine the probability of that Z is greater than. So, you can solve this question and the answer would be 1.25 ok.

(Refer Slide Time: 29:12)

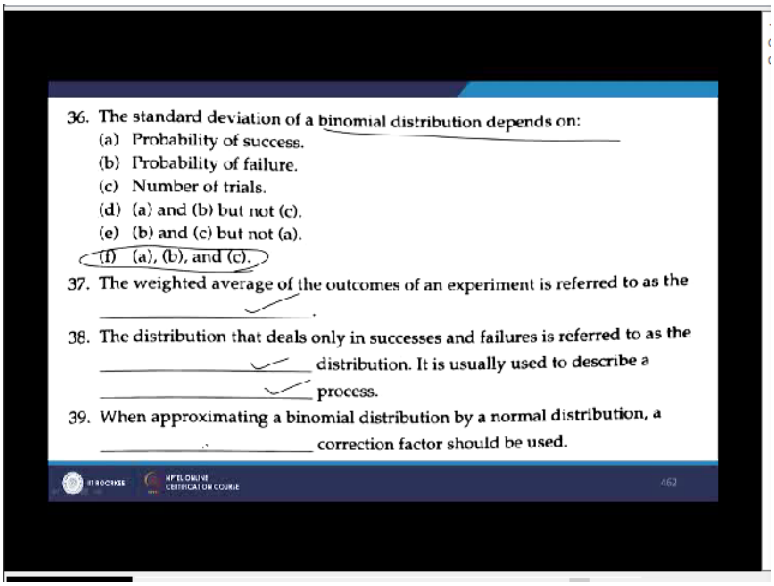


For a normal curve with mean equal to this and standard deviation=35, what proportion of the area under curve will be between 40 and 82. This is actually mean is 120 we have to find out area between 40 let say 40 and this is 82 right, so we are interested in this area. So, what you should do calculate Z values, so (x-μ) this 80-120/35you will get some Z value in that would be let say -40/35 you get some Z value.

So, that would be the area from here to here and Z is equal to 40-120/32, so -80/35 you will get some another value of Z, so this. Now you know the area from here to here from table you will also know area from here to here right. And you know this area is 50% right. So, when you do the calculation this would be the answer for this question right, which of the following normal curves look most like the curve when μ=10 and σ=5.

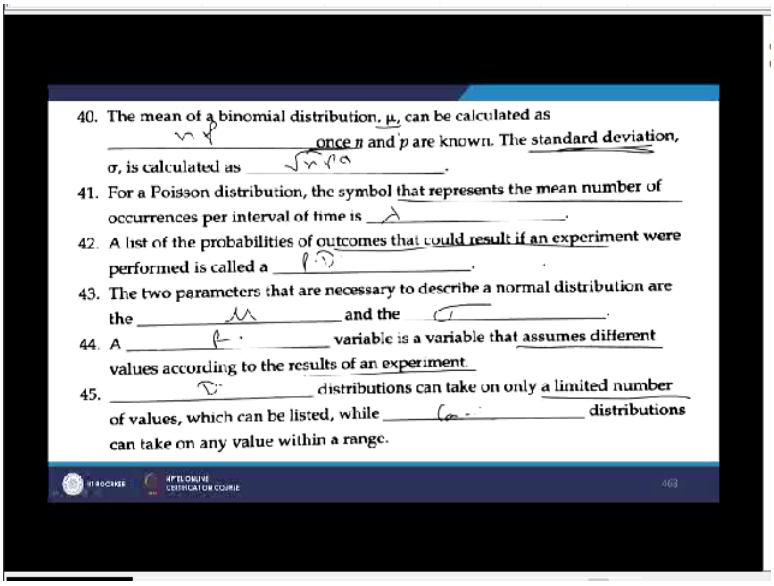
So, mu=10, standard deviation 10; mu = 20, standard deviation 10; then mu 20, standard deviation 5. So, the answer would be this because only this value will change instead of this will be 5 is not it. So, instead of 10 you just write 20, so the answer is c. Let us look at the next one binomial distribution may be approximated by Poisson distribution. If n is large, when p is small right, n is small p is which of the binomial distribution may be approximated by Poisson distribution. When n is large and p is small is that option is here, no that option is not there, so d is the correct one.

(Refer Slide Time: 32:04)



So, this is the last slide last but one slide the standard deviation of binomial distribution depends on probability of success probability of failures and number of trials yes and all these 3. The weighted average outcome of an experiment is referred as expected value. The distribution that deals only in success and failures is referred to as is binomial distribution it is usually used to describe a Bernoulli process right. When approximating a binomial distribution by normal distribution continuity correction factor there is something that a continuity correction factor.

(Refer Slide Time: 32:54)



This is the last slide, the mean of the binomial distribution is μ can be calculated as the mean of binomial distribution can be calculated as np. Once np are known the standard deviation is calculated as npq right is not it, so this is the answer for this question root npq is the standard

deviation of binomial distribution. For a Poisson distribution the symbol that represents the mean number of occurrence per interval is known as lambda is not it.

A list of the probabilities of outcomes that could result if an experiment work performed is called a list of probabilities of outcomes is called probability distribution right, probability distribution. The two parameters that are necessary to describe a normal distributions are mean and standard deviation right, this mean and standard deviation. Now variable is variable that has assumes different values according to the results of an experiment a random variable.

Distribution can take only limited value which can be so discrete distribution and continuous distribution right. So with this we come to the end of this particular session, today we have seen several questions which would have made your fundamentals clear are related to probability and probability distribution, thank you very much.