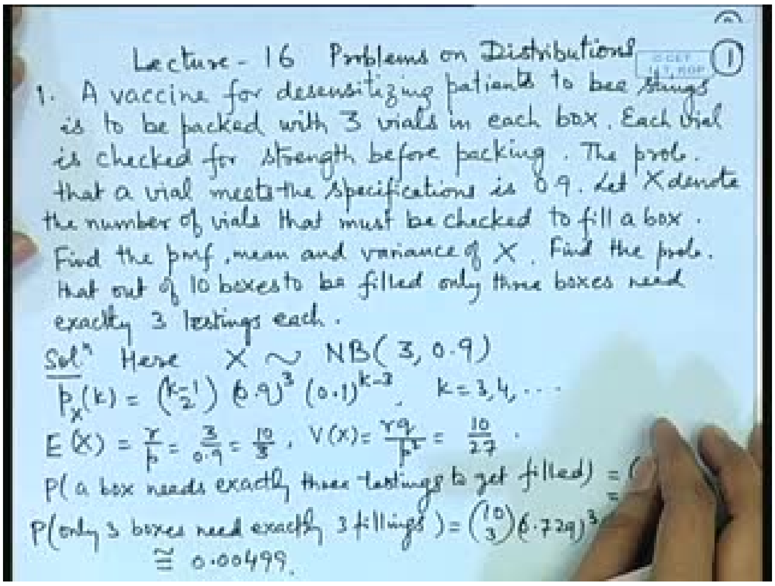
Probability and Statistics

Prof.Dr.Somesh Kumar
Department of Mathematics
Indian Institute of Technology Kharagpur

Lecture No.#16
Special Distributions – VII

(Refer Slide Time: 00:28)



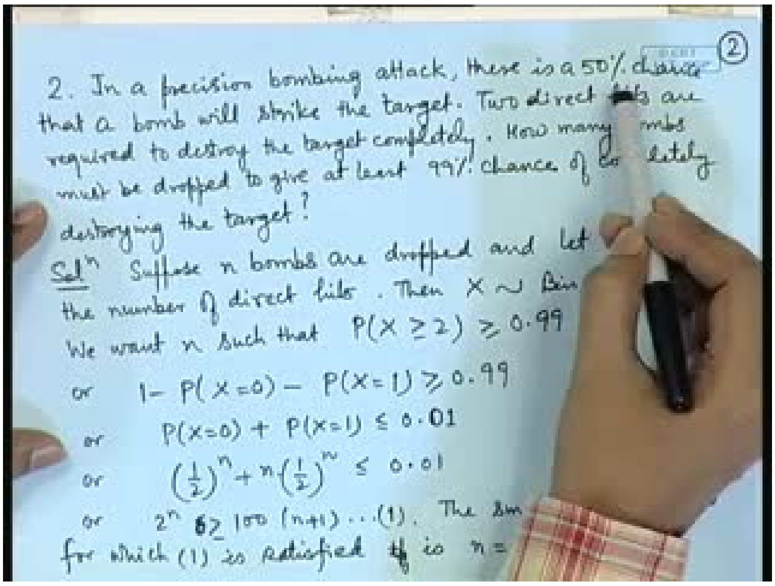
We have discussedvarious special discrete and continuous distributions, which arise frequently in practice.Today, we willlook at various applications of these distributions and this will be explained through certain problems.

Let us look at the first problem, a vaccine for desensitizing patients to bee stings is to be packed with 3 vials in each box;each vial is checked for strength before packing.The probability that a vial meets the specifications is 0.9.Let X denotes the number of vialsthat must be checked to fill a box.Find the probability mass function, mean and variance of X, find the probability that out of 10 boxes to be filled, only 3 boxes need exactly 3 testing is each. Now, let us look at the setup of this problem; so, in each box, we are packing 3 vials, but the vial has to be checked, so it may meet the specification or it may not meet the specification; we are assuming that, all the vials are having identical probability of meeting the specification and each checking is done independently.Under these assumptions, the vial meeting a specification or not becomes a Bernoullian trial; so, this is a sequence of Bernoullian trials,now in, we keep on checking,till 3 vials meet the specification and then we pack it in a box; so, this is negative anomialsampling, and therefore,if we consider X as the number of vials, which are needed to fill a box, that means, the first time 3 vials are correctly meeting the specification, then the distribution of X is negative anomialwith r is equal to 3 and P is equal to 0.9, that means, the probability mass function of this will be k minus 1 c 20.9 cube 0.1 to the power k minus 3, which is the probability mass function of a Binomial negative anomialdistribution with parameter r is equal to 3 and p is equal to 0.9, the values of k are 3, 4 and so on.

As the mean of a negative binomial distribution is r by p, so that is 3 by 0.9, that is 10 by 3and variance is r q by p square, which after a simplification becomes 10 by 27.Find the probability that out of 10 boxes to be filled, only 3boxes need exactly 3 testing is,now what is a probability that one box needs exactly 3 testing is, that means, the first 3 vials, which are checked, all of them meet the specification; so, this is corresponding to k equal to 3 term here, which is giving 0.9 cube that is p cube.

Now, the second part of this problem is that, each box may need exactly 3 testing or it may not need exactly 3 testing; so, if total number of boxes are 10, a particular may need 3 testing or may not need 3 testing; so, it again becomes likea Bernoullian trial with probability of success p as equal to 0.9 cube, that is, 0.729; so, out of 10 boxes,3 boxes will need 3testing is, it is the binomial probability of X is equal to 3, where n is equal to 10 and p is equal to 0.729;so, by applying the formula n c x p to the power X q to the power n minus X, we get 10 c 30.729 cube 0.271 to the power 7, which is approximately 0.00499.You can see here, we have to make the assumption of independence, and identical nature of the trials, so that, we can apply the concept of binomial or negative binomial distribution here.

(Refer Slide Time: 04:59)

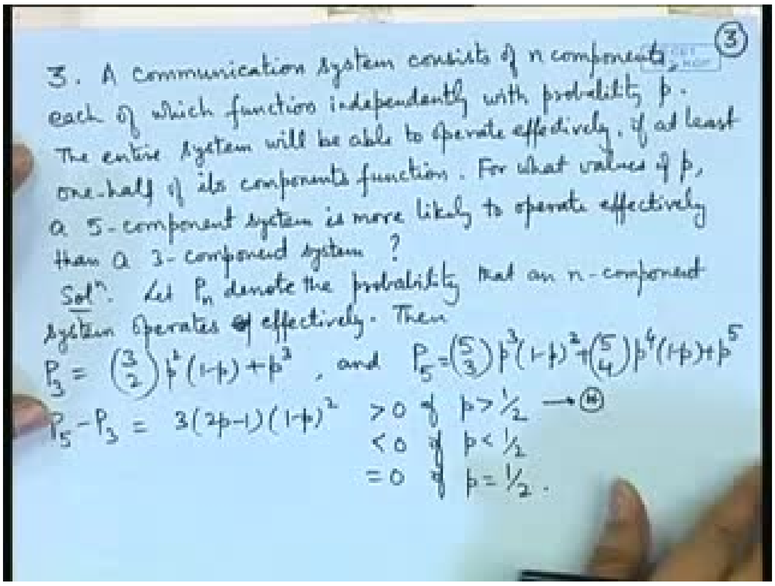


Let us look at another application; in a precision bombing attack, there is a 50 percent chance that a bomb will strike the target.Two direct hits are required to destroy the target completely.How many bombs must be dropped to give at least 99 percent chance of completely destroying the target?Now, here you can see, a particular bomb may hit the target or it may not hit the target; so, each trial can be considered as, that means,hitting of the,throwing of a bomb or attacked by a bomb, that be can be considered as a Bernoullian trial, so it may strike the target or it may not strike the target; we assume that the attack by the bombs is independently done, and it is identical in nature, that is a probability of striking is same for all.

So,you will have,if I consider out of n bombs,X bombs are the direct hits, then the distribution of X will be binomial n and here the probability of striking is half, because there is a 50 percent chance; so, it becomes binomial n half.So, we want that, there is at least 99 percent chance of destroying the target; since, we need at least two hitting is, that means, what is the probability that X is greater than or equal to 2.So, we want this probability to be greater than or equal to 0.99; that means, what should be the value of n for which this condition is satisfied.So, we apply the formula of the binomial probability mass function here, the probability of X greater than or equal to 2 is having many terms, so we consider the complementation of this event that is, probability X equal to 0 and probability X equal to 1, so 1 minus this must be greater than equal to 0.99

So, after simplification, it becomes probability X equal to 0 plus probability X equal to 1 is less than or equal to 0.01 andnow here P is equal to half; so, this is n c 0 P to the power 01 minus P to the power n which is half to the power n and this is n c 1P into 1 minus P to the power n minus 1, since P n,1 minus P, both are half, so it is again half to the power n, n c 1 is n, so the term is n plus 1 divided by 2 to the power n; so, after simplification, this condition is equivalent to that 2 to the power n greater than or equal to 100 into n plus 1, what is the smallest value of n for which this is satisfied.So, we can check it and it turns out that, n is equal to 11, first time satisfies this condition.So, at least, we have to drop 11 bombs, so that the target is completely destroyed with probability greater than or equal to 0.99.

(Refer Slide Time: 08:35)



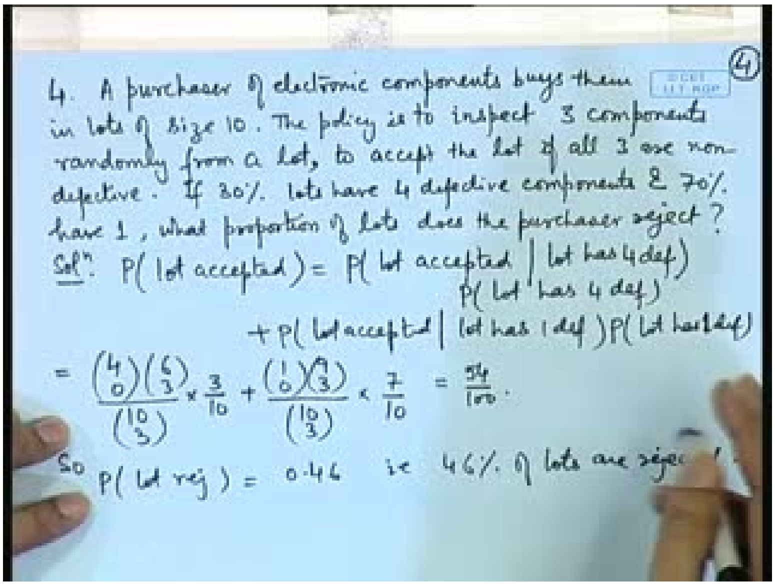
So, here if you see once again, we have made the assumption that the trials, that means, the dropping of the bombs are striking of the target etcetera is considered as Bernoullian trial, that is independence and identical nature of the trials has been considered here.A communication system consists of n components, each of which function independently with probability P.So, once again functioning of components is a Bernoulliantrial, because the component may fail or it may not fail; so, if it is working, the working probability is p, the entire system will be able to operate effectively, if at leastone half of it is components function, that means, suppose there are 10 components, then at least 5 should function, then the system will be operating effectively.

For what values of p a 5 component system is more likely to operate effectively than a 3 component system.So, we have to calculate the probability of a 5 component system working effectively and a 3 component system operating effectively; so, let us use a notation P n, let it be the probability that an n component system operates effectively.

So,P3 will denote the probability that a 3 component system is working effectively, that means, at least 2 or 3, that means, either 2 or 3 of the components are working correctly.So, out of three, two are working, so 3 c 2p square 1 minus p and all the three are working, so p cube, so the probability of p 3 is given by this.

In a similar way, a 5 component system will be operating effectively, if either 3 or 4 or all 5 components are working properly.So, the probabilities of them are given by these, so we add up, so p 5 is given by this.We have to check that whether a 5 component system is more effective or 3components, so we consider the difference P5 minus P3, which after certain simplification is equal to 3 into 2 p minus 1 into 1 minus p square.Now, here you observe 1 minus p square is a positive term, because p is a number between 0 and 1; so, this term is positive, if p is greater than half, that means, a 5 component system is moreeffective, if the probability of each system working effectively is more than 50 percent, so this answers the question here.

(Refer Slide Time: 11:48)



We can also see that this is less than 0, if p is less than half, that means,if each component fails with a probability which is less than 50 percent, then if we add more components in thesystem, it is actually reducing the probability of operating effectively, and if p is equal to half, then both the systems have the same probability of operating effectively; so, these are some of the applications of the binomial and negative anomial distribution etcetera.

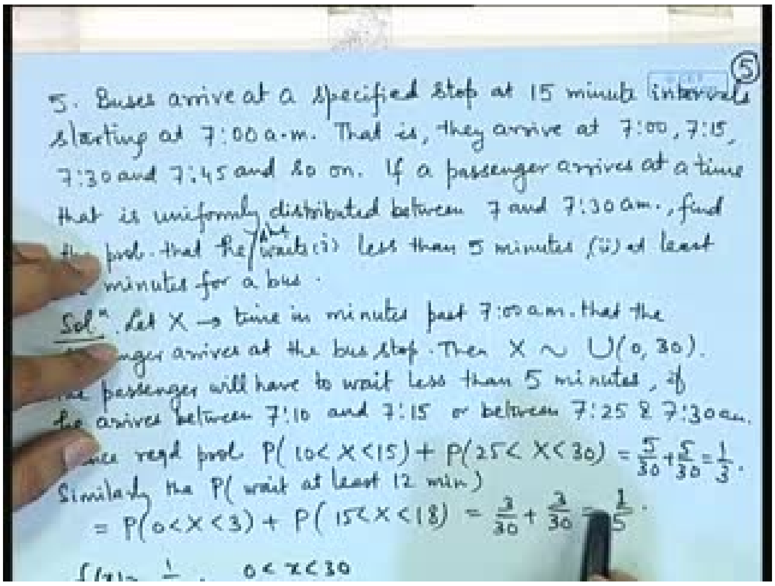
Let us look an application of hyper geometric distribution; a purchaser of electronic components, buys them in lots of size 10; so, the policy is to inspect three components randomlyfrom a lot, so a lot each lot is having 10 components; so, what the purchaser will do, he randomly selects three components or three parts from that lot, and he inspects them, if all the three arenon-defective, he accepts the lot, otherwise he rejects the lot; now, it is known to us, that 30 percent of the lots have 4 defective components and 70 percent of the lots have 1 defective component; so, in general what proportion of lots the purchaser will be rejecting.So, in place of probability of rejecting, we can calculate the probability of lot getting accepted, so the lot is accepted.Now, this is conditional upon two types of possibilities, that the lot has come fromthe set, where four defective components are there or where one defective component is there; so, we can apply a theorem of total probability the probability, that to the lot is accepted, given that the lot has 4 defectives into the probability that the lot has 4 defective plus probability that lot is accepted given that lot has one defective multiplied by probability that lot has one defective term.

Now, we evaluate these probabilities, so, probability that the lot has 4 defectives is 0.3, because 30 percent of the lots have 4 defective components, and similarly the probability that the lot has one defective, that is 0.7.

Now what is the probability of the lot getting accepted, if the lot has 4 defective items?So, total number of items in the lot is 4is 10, we are selecting randomly 3.So, the lot will be accepted, if all the 3 checkingis, which have been done from here are for the good ones.So, since the lot which is having 4 defective,6 will be good; so, the 3, which the purchaser has selected must be all good; so, it is 6 c 3, and out of the bad ones, none of them is selected, so 4 c 0 into 6 c 3 divided by 10 c 3, which is the hyper geometric probability.

In the second case, if the lot has one defective, so out of ten, nine are good one, is defective, and our selection all the three must be from good; so, it is 1 c 09 c 3 divided by 10 c 3; so, after simplification this turns out to be 0.54,so probability of the lot getting rejected will be 1 minus this that is 0.46; so,46 percent of the lots get rejected, which is quite high which is understandable, because the person checks, and each of the components which are checked must be alright, then only we will accept; so, the condition is relatively tough, and therefore,almost 50 percent of the lots are actually getting rejected under these given conditions.

(Refer Slide Time: 15:36)



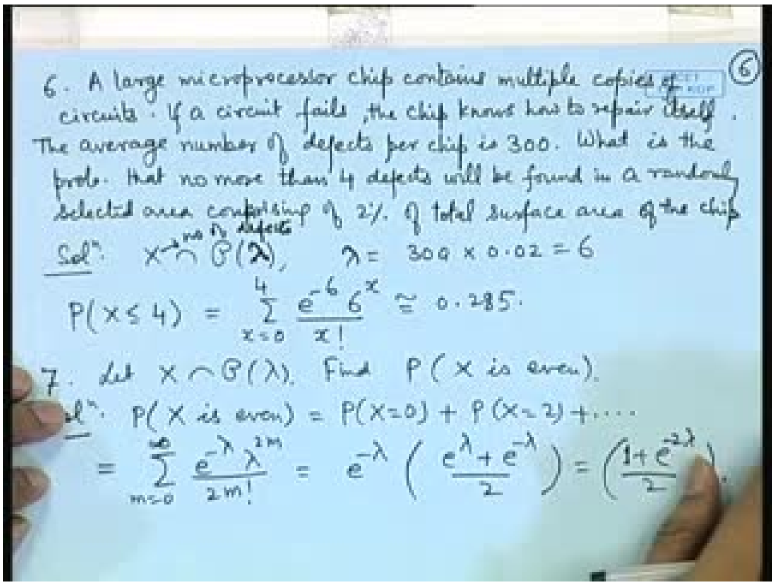
Let us look at a application of continuous uniform distribution; buses arrive at a specified stop at 15 minute intervals starting at 7 am, that means, bus will come at 7, it will come at 7:15,7: 30,7:45 and so on; so, if a passenger arrives at a time uniformly distributedbetween 7 to 7: 30 am, find the probability, that he has to wait for less than 5 minutes or least 12 minutes for a bus, that means, when he comes, then he will get a bus in less than 5 minutes; so, let us consider the arrival time of the passenger; so, since the time of the passenger is between 7 to 7: 30 am, if I say X is the time in minutes past 7 am, that is the arrival time of the passenger at the bus stop, then we can say that under the given conditions X follows uniform 0 to 30, where the measurement is done in the minutes, because it is between 7 to 7: 30, but we are considering the starting point as0, so between 0 to 30.

The passenger will have to waitless than 5 minutes, if he arrives between 7:10 to 7:15, because the next bus is at 7:15, so if he arrives say 7:05, then he will have to wait for 10 minutes, if he comes at 7:07, then he has to wait for 8 minutes and so on; so, if he arrives between 7:10 and 7:15, he will get a bus in less than 5 minutes.Similarly,after this bus departs, then the next bus comes at 7: 30, so if the passenger arrives between 7:25 to 7: 30, then he will have to wait less than 5 minutes; so, the required probability, that the passenger has to wait for less than 5 minutes is probability that X lies between 10 and 15 or X lies between 25 to 30, since,it is a uniform distribution on the interval 0 to 30, the density of X is 1 by 30, 0 less than X less than 30;so, probability of 10 less than X less than 15 will become 15 minus 10 by 30, that is 5 by 30, and similarly,probability of 25 less than X less 30, that will become 30 minus 25 by 30, that is 5 by 30; so, the total probability is 1 by 3.

That means, the probability that he has to wait for less than 5 minutes is one-third of the time.In a similar way, the probability that he has to wait for at least 12 minutes, so he will have to wait for at least 12 minutes, if he arrives between 7 to 7:03, because if he arrives after 7:03, then he has to wait less than 12 minutes, because the next bus is at 7:15.

Similarly, if he arrives between 7:15 to 7:18, then he has to wait for more than 12 minutes.So, the required probability of waiting at least 12minutes is that,X lies between 0 to 3 or X liesbetween 15 to 18.So, once again using the uniform density, it is 3 by 30 plus 3 by 30, that is 1 by 5, that means,20 percent of the times, he has to wait for more than 12 minutes.

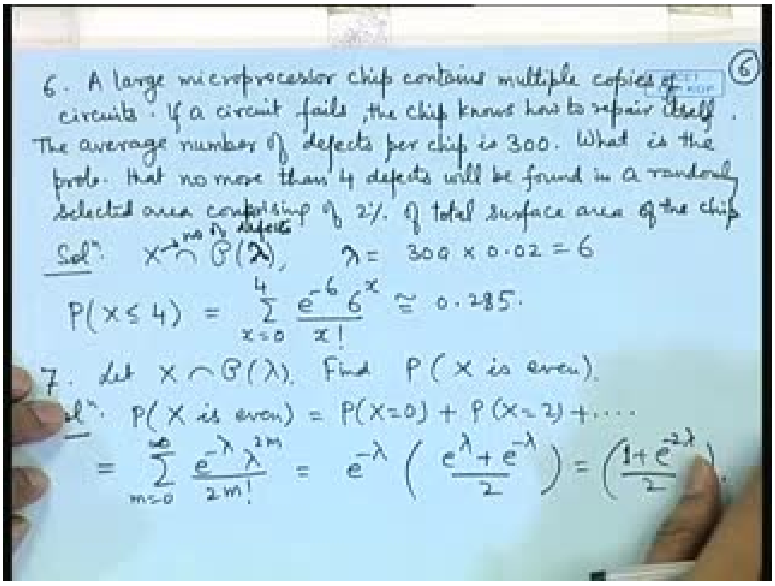
(Refer Slide Time: 19:21)



Let us look an application of Poissondistribution; a large microprocessor chip contains multiple copies of circuits, if a circuit fails, the chip knows how to repair itself.The average number of defects per chip is 300; so, this is the condition that is the rate kind of thing, that is, one chip has nearly 300 defects, what is the probability that no more than 4 defects will be found in a randomly selected area comprising of 2 percent of the total surface area of the chip.

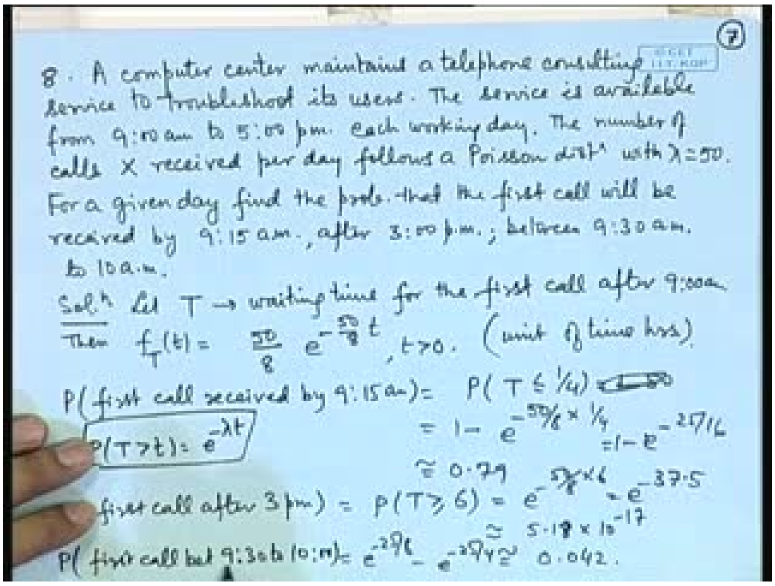
So, here, if I consider X as the number of defects in the 2 percent of the surface area, then if I assume it to be Poisson distribution with parameter lambda, then lambda here will be because 300 is for the full area; so,300 into 0.02, that means, in 0.2 percent of the surface area, what is the number of defectives, for that the rate will be 300 into point 02, that is 6, that is the rate of detecting the defects.Here,we want, what is the probability that not more than 4 defects will be found; so, it is probability of X less than or equal to 4, based on the Poisson probability mass function, so, it is e to the power minus lambda lambda to the power X by X factorial X isequal to 0 to 4, so by putting lambda is equal to 6, this can be evaluated either by direct or by Poisson tables, it is approximately 0.285.

(Refer Slide Time: 21:19)



Let us look at one more application of the Poisson distribution; suppose,X is a Poisson lambda distribution, what is the probability that X is an even number, as in the Poisson distribution X can take values 0,1,2, 3 and so on, that means, any non-negative integer value; so, the probability that X is even is probability X equal to 0, probability X equal to 2 etcetera, that means, e to the power minus lambda lambda to the power j by j factorial, where j is of the form 2 m, so we can write it in this from m is equal to 0 to infinity.So, if we take out e to the power minus lambda, the infinite series is actually the sum ofe to the power lambda plus e to the power minus lambda divided by 2, so it is equal to 1 plus e to the power minus 2lambda by 2.

(Refer Slide Time: 22:16)



A computer center maintains a telephone consulting service to troubleshoot it is users.The service is available from 9 am to 5 pm, each working day.The number of calls X received per day follows a Poisson distribution with lambda is equal to 50.For a given day find the probability, that the first call will be received by 9:15 am after 3 pm between 9:30 am to 10 am.

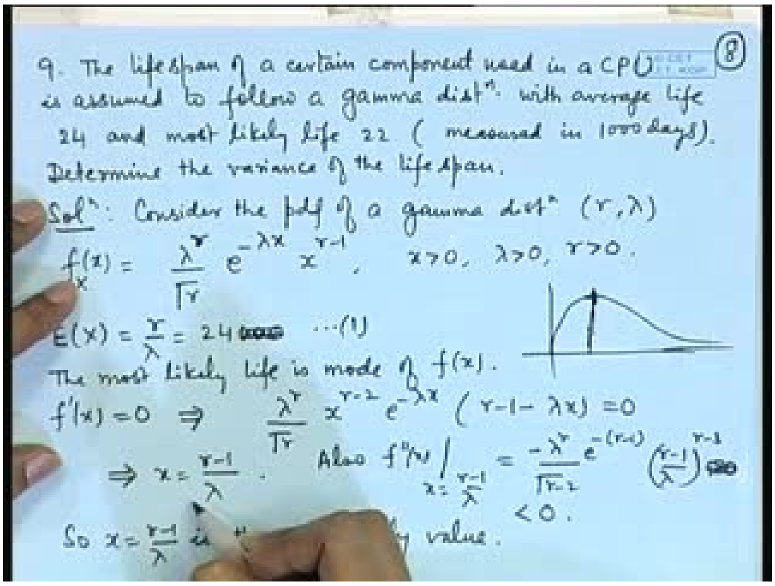
So, let us look at the modeling of this problem; since, here time is hours, so we can consider T as the waiting time for the first call after 9 am; now,here we are considering the day as an 8 hour period, and in 8 hour period, the number of calls received is rate lambda is equal to 50; so, in 1 hour period, it will be 50 by 8; now, if I consider t as the waiting time for the first call after 9 am, then the distribution will be negative exponential with lambda is equal to 50 by 8, so lambda e to the power minus lambda t.

What is the probability that the first call is received by 9:15, that means, in quarter of an hour, that means, what is the probability of T less than or equal to 1 by 4, this unit of measurement is in the hours;now, we apply the formula for the exponential distribution, probability of T greater than some small t is equal to e to the power minus lambda t; so, if we use thisformula probability of capital T less than or equal to 1 by 4 is 1 minus e to the power minus lambda t, so lambda is minus 50 by 8 and time is 1 by 4.

So, it is after simplification 0.79, which looks surprisingly quite high, that is almost four-fifth of the probability, the reason is that, in a day, we are receiving roughly 50 calls, so within first 15 minutes, a call will be received with a substantially high probability, likewise if we calculate what is the probability that the first call is after 3 pm, now from 9 am to 3 pm, it is 6 hours, so that means, what is the probability of t greater than or equal to 6; so, if we utilize this formula, it is e to the power minus lambda t, that is 50 by 8 into 6, which is extremely small probability,which it must be, because the rate is quite highof receiving the complaints, and here we are saying that from morning 9 to 3 pm,there is no call, so the probability of that event must be pretty small.

Similarly,what is the probability that the first call is between 9: 30 to 10; since, the rate is high, the first call is after 9: 30, the probability must be small, and then, we are saying between 9: 30 to 10 which is further small, so it is after calculation using this formula, it is turning out to be 0.042.So, these aresome of the applications of the Poisson distribution or the Poisson process, you can notice here, that in order to apply the Poisson distribution, we have to look at the rate in the appropriate time interval or the area, that is lambda t, we have to calculate.

(Refer Slide Time: 25:59)



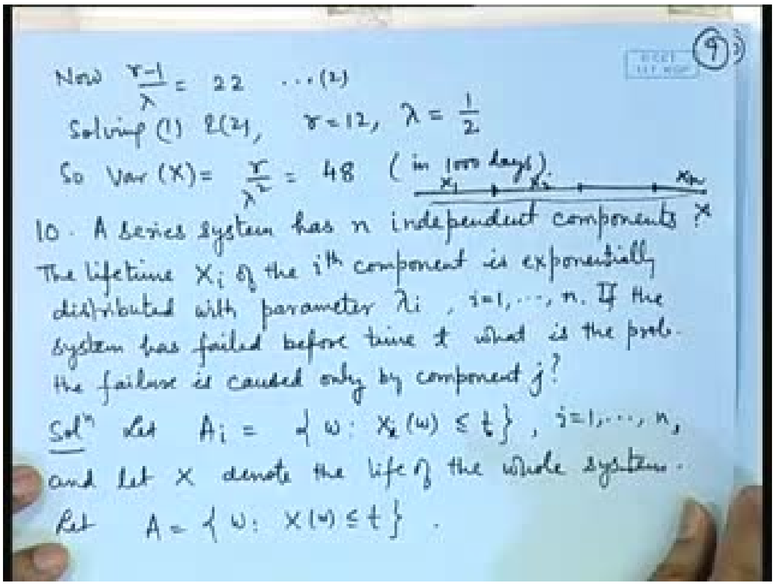
Let us look at the lifespan of a certain component used in a CPU is assumed to follow a gamma distribution with average life 24, and most likely life 22, it is measured in 1000days; so that means, average life is 24000 days, and most likely life is 22000 days, find the variance of the life span.

So, the standard formof a gamma distribution, we have considered as a waiting time for the rth occurrence in a Poisson process; so, we had the parameters r and lambda, and the form of the probability density function was given by lambda to the power r by gamma r e to the power minus lambda x xto the power r minus 1; so, if we take the mean of this distribution, it was r by lambda, that is given to be 24, the most likely life, so by most likely life, we mean the maximum value of the density function.In the discrete case, it would mean that, the point corresponding to the maximum probability mass function, and if we take analogous value of that, in the continuous case, it means the mode of the distribution.

So, if we have the density function like this, then the maximum value is attained at this point.Now, for a gamma distribution, the maximum value can be calculated, so here the density function is given by this, we can use the ordinary calculus by looking at the derivative, so f prime x is 0 for this, that is x is equal to r minus 1 by lambda, and at this point, we can check the second derivative, it is actually negative.So, this is xis equal to r minus 1lambda is the mode of the distribution that is the maximizing point.

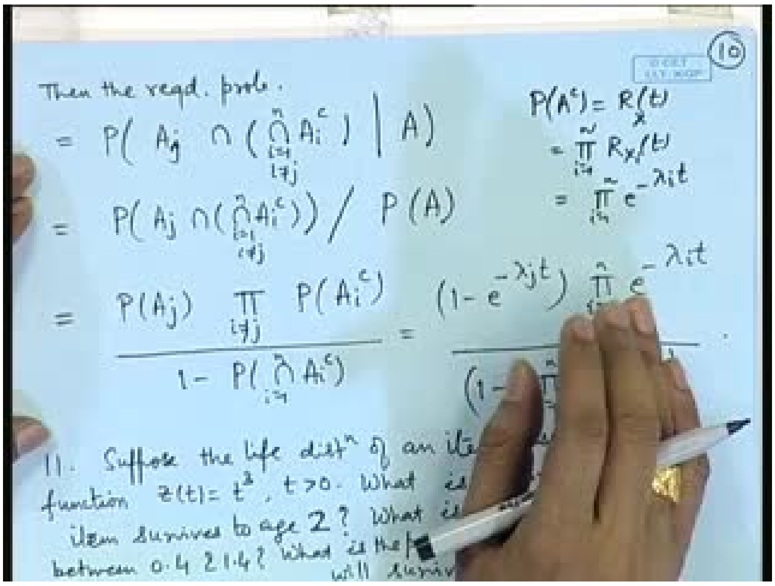
It is given here that r minus 1 by lambda is equal to 22; so, we have two equations 1 is r by lambda is equal to 24, and another is r minus 1 by lambda is equal to 22; so, if we simplify this, we get r is equal to 12 and lambda is equal to half.So, the distribution of the life span ofCPU of a certain component, using CPU is specified as a gamma 12 and half;now, if we want the variance of this life span, variance of a gamma distribution is r by lambda square; so, after substitution, it becomes 48 and in 1000, so 48000 days that is in the squared units.

(Refer Slide Time: 28:57)



Let us consider a series system; a series system has n independent components, the life time, so the components are attached and the life c areX 1, X2, X n, suppose the total life of the system is X; so, individual life times are assumed to be exponentially distributed with parameters lambda i, so lambda 1, lambda 2, lambda n, if the system has failed, before time t what is the probability that the failure is caused only by component j, basically we consider it as a weak link.

(Refer Slide Time: 30:13)



So, let Ai denote the event that the ith component fails before time t, so Xi omega less than or equal to t; consider X to be the life of the whole system, so A is the event that the entire system fails before time t, we are asked to find out the probability that the system has actually failed, so what is the probability that the failure is caused by the component j; so,it is actually the conditional probability, that the jth component fails before time t, because Ai denotes the event that ith component fails before time t; so, here we are interested in that the jth component fails, and all components other than the jth component do not fail, this is A1 complement,A2 complement etcetera, except Aj complement here; so, it is the simultaneous occurrence, that is jth component fails and all the components other than the jth component are working at time t; so, what is theconditional probability of this event given that actually the system has failed.

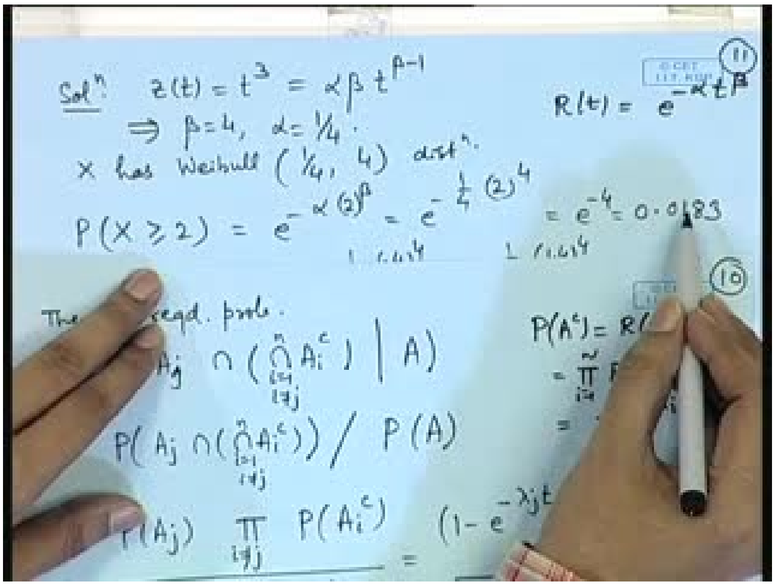
So, we apply the formula for the conditional probability, so probability of some event e given an event f, so it is equal to probability of e intersection f divided by probability of f; now, if you look at this event, here it means, one of the component fails, and intersection with the event that the system fails; since,it is a series system, one of the component fails necessarily implies that the system has failed.

So, this probability, this event is a subset of the event a;therefore,the intersection will give me only the event, which is described here, that is a j intersection with the intersection of Ai complement, where i is notequal to j; now, at this stage, we make use of the assumption, that the components are independently working, that means, failure or not failure of a individual component, does not affect the failure or not failure of any other component; so, here we can apply the formula for the probability of the intersection of the events, for independent events; so, it becomes probability of Aj into probability of these events, which again can be split as the product of the probabilities divided by probability of A.

Now, what is the probability of A, this means, that system has failed before time t, so it is equal to 1 minus the reliability of the system at time t; now, reliability of a series system is nothing but the probability of theproduct of the reliabilities of the individualterms.So, here, probability of a complement that is equal to reliability of the system at time t which is equal to product of the reliabilities of individual components.Now, each of the lives is exponentially distributed, so reliability of the ith component is e to the power minus lambda i t product i is equal to1to n, so that is the term coming here corresponding to probability of A complement.

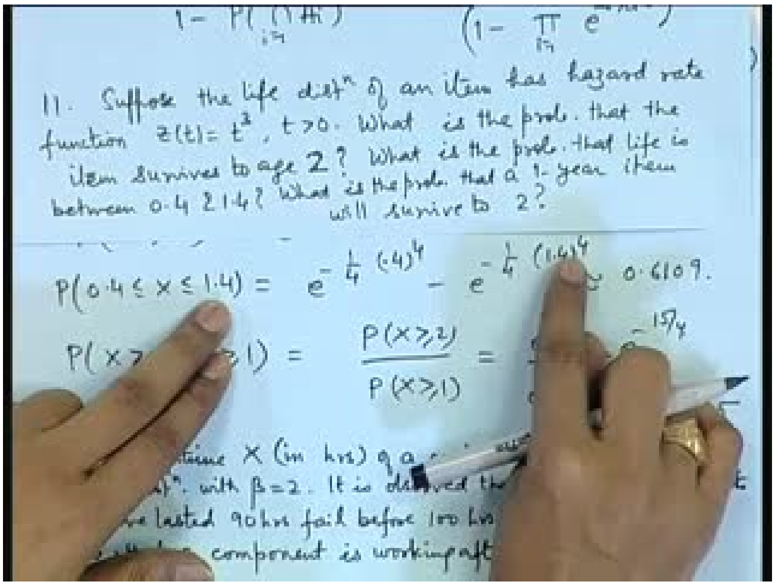
In the numerator probability of Aj, that is the jth component fails before time t, it is 1 minus e to the power minus lambda j t, and here it is reliabilities of the all other components except the jth component, so it is product of e to the power minus lambda i t, where i is not equal to j; so, this is denoting the conditional probability, that thefailure is cast by the component j alone, that means, the system has failed that what is the conditional probability that the failure was cast by the jth component.

(Refer Slide Time: 35:30)



Let us look at applications ofWeibull distribution; so, suppose the life of an item has a hazard rate function Z t is equal to t cube, so if you recallwe considered that the hazard rate function is of polynomial type, that is alpha beta t to the power beta minus 1, if and only if the distribution of the life is Weibull distribution with parameters alpha and beta.So, here if the life distribution is given to be Z t is equal to t cube, it is exactly of the form of a hazard rate function of a Weibull distribution. So, we can determine what are the parameters of the Weibull distribution here; so, if we compare it with alpha beta t to the power beta minus 1, it gives alpha is equal to 1 by 4, and beta is equal to 4, that means, the life of the item has Weibull distribution with parameter alpha is equal to 1 by 4, and beta is equal to 4; so, what is the probability that the item survives to age 2; so, probability of X greater than or equal to 2, the reliability function of the Weibull distribution at t is e to the power minus alpha t to the power beta; so, here if we substitute the values of alpha and beta and t is equal to 2, then the value turns out to be e to the power minus 4, that is 0.0183 which is quite small.

(Refer Slide Time: 36:03)

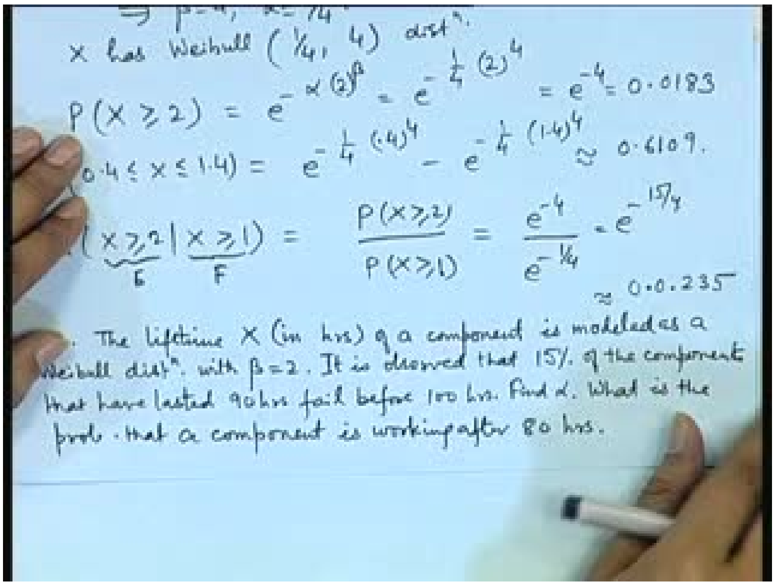


What is the probability that the life is between 0.4 and 1.4; so, we are interested in finding out X lying between 0.4 to 1.4; so, naturally this can be written as probability, that is the reliability at 0.4, and the reliability at 1.4, the difference of the reliabilities, so we make use of this formula, that is at time t, the reliability here is e to the power minus alpha t to the power beta; so, after substituting the values of alpha, beta and t, t has 0.4 and 1.4, after simplification it turns out to be 0.61

What is the probability that a one year item will survive to age 2, that means, it is given that the item is working at age 1, what is the probability that it willwork till age 2, so it is the conditional probability of X greater than or equal to 2, given that X is greater than or equal to 1.Once again we make use of the formula for the conditional probability, this is event e this is event f; so, probability of e given f is equal to probability of e intersection f divided by probability of f; once again notice that, here e is a subset of f, so in the numerator, we will have probability of event e divided by probability of f.

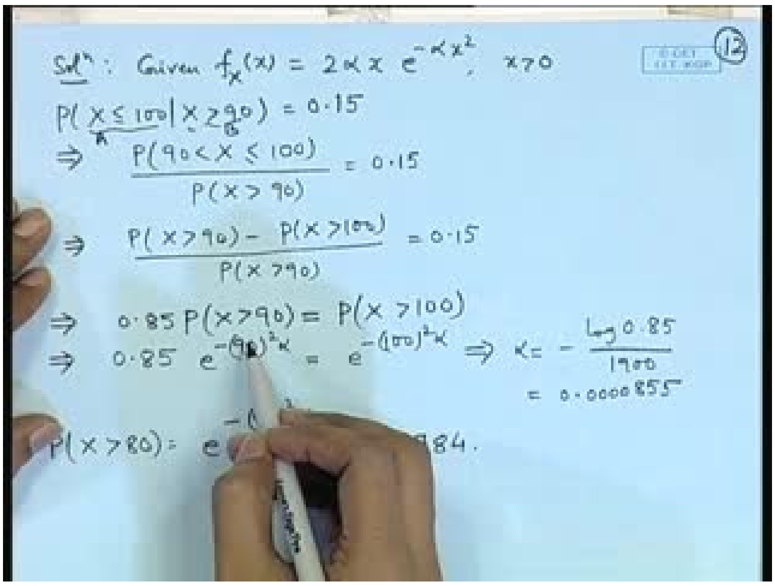
So, here this is turning out to be the reliability of the system at time 2, and this is the reliability of the system at time 1; so, in the formula for the reliability, we substitute the values of alpha, beta and t, and we get e to the power minus 4 and e to the power minus 1 by 4, and after simplification this 0.0235.

(Refer Slide Time: 38:17)



Let us look at another application of Weibull distribution; so, life in hours, we denote by the random variable X of a component is modeled as a Weibull distribution with beta is equal to 2, it is observed that 15 percent of the components, that have lasted 90 hours, fail before 100 hours, find the value of alpha.What is the probability that a component is working after 80 hours.

(Refer Slide Time: 39:07)

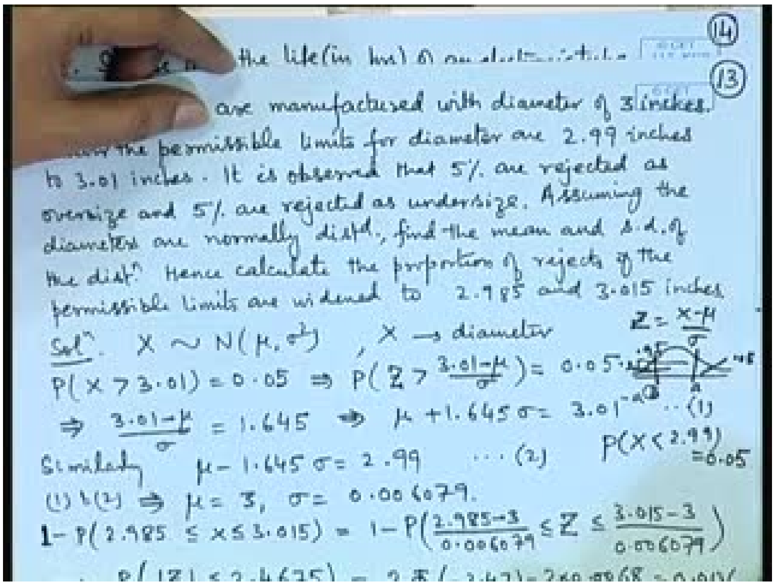


Now, here notice that for Weibull distribution, the parameter beta has been specified the parameter alpha has not been specified, but certain condition is given; so, we can use this conditionto determine the value of alpha.So, the Weibull distributions density is alpha beta x to the power beta minus 1 e to the power minus alpha x to the power beta, if we substitute the value of beta is equal to 2, then the density reduces to the form given here, that is 2 alpha x e to the power minus alpha x square.

Now, it is given that,15 percent of the components which last90 hours, fail before 100 hours, that means they fail before 100 hours, that is the X less than or equal to 100 given, that they work till 90 hours, that is X is greater than 90 this proportion is 0.15; so, consider these events, this is event say A,and this is the event B,so probability of AintersectionB becomes, that X lies between 19 to 100, andprobability of B is X greater than 90; so, the numerator is the difference of the reliabilities at 90 and 100 hours divided by the reliability at 90 hours; so, we can simplify this terms, so it is 0.85 probability X greater than 90 is equal to probability X greater than 100, substitute the values of the reliability function of the Weibull distribution as e to the power minus alpha t to the power beta.So, if we use this, then the reliability here at X greater than 90 is e to the power minus alpha t to the power beta; so, beta is equal to 2 here, and here it is e to the power minus 100 square alpha,that is t is equal to 100. Now, after certain simplification this value of alpha turns out to be quite small 0.000855.

So, what is the probability that a component is working after 80 hours, that means, the reliability of the component at 80times; so, by applying this formula that is e to the power minus alpha t to the power beta, the value turns out to be 0.5784.

(Refer Slide Time: 41:52)



So, these are some of the applications of the Weibull distribution, we have already seen the applications of exponentialgamma etcetera also.Now, we look at application of the normal distribution; so, consider steel rods which are manufactured, so with the diameter of 3 inches, now in any industrial process,the specifications are given; so, here the specification is the diameter should be 3 inches for the rods, however in the actual manufacturing, there will be some deviation, that means, it may be 2.99 9 inches, it may be 3.001 inches etcetera; so, in any industrial production, the producer or manufacturer or the customer, he specifies, that what should be his desired specifications, rather than telling exactly 3, because that may never be met; so, the permissible limits for diameter are taken for example from 2.99 inches to 3.01 inches, that means, if a rod is manufactured, and if it is diameter turns out to be less than 2.99, it is not considered to be meeting the specification.

If it is more than 3.01 inches, then also it isnot meeting the specifications.Now, it is observed that 5 percent of thesteel rods, which are produced by this process, they are rejected as oversize, that means, they are having diameter more than 3.01inches and 5 percent are rejected as undersize, that means, they are having the diameter less than 2.99 inches.

So, assume that the diameters are normally distributed, find the mean and the standard deviation of the distribution; hence, calculate the proportion of rejects, if the permissible limits are widened to 2.985 and 3.015 inches.

So, let us look at X as the diameter of the rod produced; so, it is given that, it is normally distributed; so, let us assume that,it is normally distributed with certain mean, mu and certain variance sigma square; so, it is given that 5 percent of the rods have diameters more than the upper limit, they are oversize; so, probability that X is greater than 3.01 is 0.05.

So, now we have seen that the probabilities related to any general normal distribution can be transformed to probabilities related to standard normal random variable; so, here the standard normal random variable can be obtained by looking at Z as X minus mu by sigma; so, this is reducing to X minus mu by sigma greater than 30.01 minus mu by sigma, so this is Z; so, this is nothing but the point on thenormal distribution, such that so this 3.01 minus mu by sigma, let us call it say a, this is the point such that beyond this you have 0.05 probability or before that you have 0.95 probability.

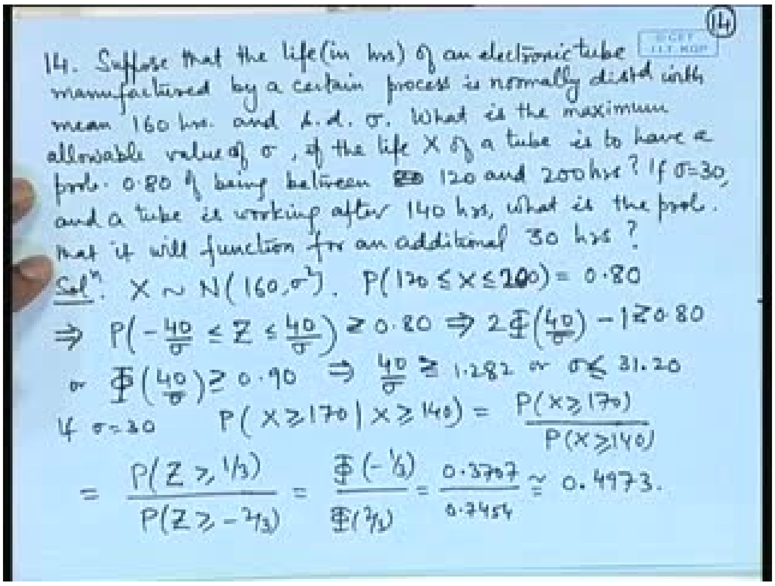
So, see the tables of the normal distribution, this point a is 1.645, so we are getting an equation 3.01minus mu by sigma is equal to 1.645; so, after simplification, the equation reduces to a linear equation in variables mu and sigma.

In a similar way, if we consider probability of X less than 2.99 is equal to 0.05, then after transformation probability of Z less than2.99 minus mu by sigma is 0.05, that means, what is the point, say B such that this probability is 0.05;now, by the symmetric property of the normal distribution, this point will be actually minus a, so we will get 2.99 minus mu by sigma is equal to minus 1.645; so, after simplification, it leads to the equation mu minus 1.645 sigma is equal to 2.99.So, if we solve this two linear equations in two unknowns mu and sigma, we get mu is equal to 3, which isbecause it is given that the distribution is symmetric, and here the assumptions of the problem, here we are assuming that it is normally distributed, so mu must be 3, because that is the target here, and sigma turns out to be a pretty small value 0.006.

Now, under these values of mu and sigma, if we extend the permissible limits by little more that is 0.015, that is from 2.985 inches to 3.015 inches, then how many or what is the proportion of rejecting of the steel rods.So, we calculate the probability of accepting the rod, that means,X lies between 2.985 to 3.015 and 1 minus that if we take this is the probability of rejection.So, this probability we can transform to standard normal distribution by subtracting 3 and dividing by the sigma; so, after certain simplification, actually this term turns out to be 2.4675, and thisterm is a minus of the same term, so it is 1 minus probability of modulus Z less than or equal to 2.4675; so, it is 2 times the CDF of the standard normal variable at the 0.2 point minus 2.4675. So, from the tables of the normal distribution, we can see and this value turns out to be 0.0136; so, the probability of rejecting is nearly 0.01, that is one in a 100 will be rejected.

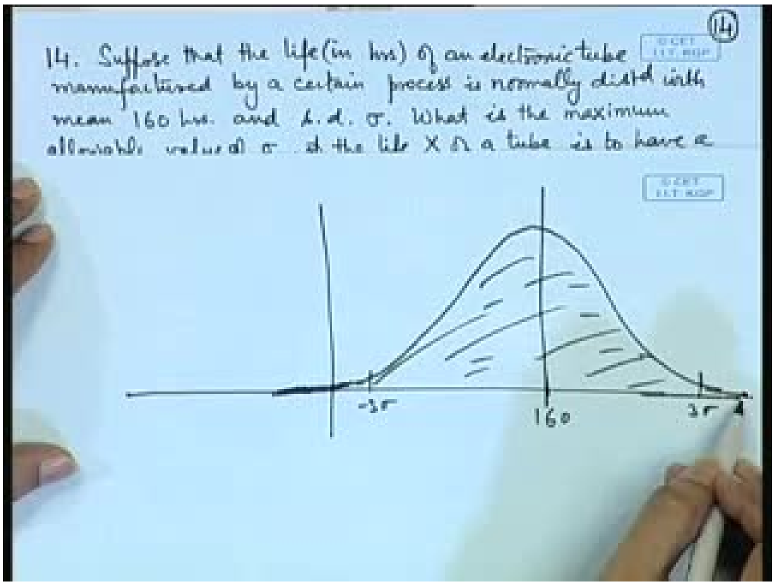
So, you can see here, that initially 10 percent are rejected, if we are having 0.01,as the, that is on either side of 3, we are having 0.01 as the acceptable limit, if we widen little bit more then only 1 percent are getting rejected; so, this is pretty fast. This point we had seen earlier also, that in the normal distribution a large probability is concentrated around the mean.

(Refer Slide Time: 49:14)



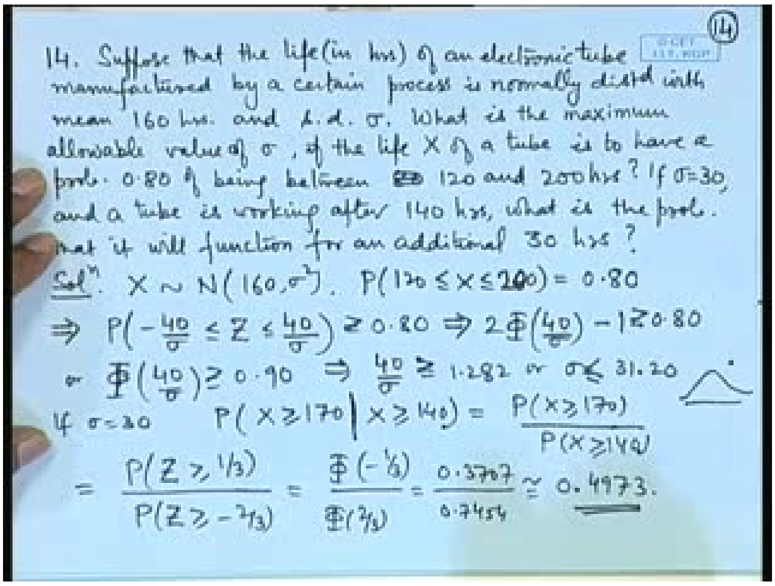
Suppose that the life in hours of an electronic tube manufactured by a certain process is normally distributed with mean 160 hours, and standard deviation sigma.What is the maximum allowable value of sigma, if the life of a tube is to have probability point 8 of being between 120 and 200hours?If sigma is equal to 30 and a tube is working after 140 hours, what is a probability that it will function for an additional 30 hours.

(Refer Slide Time: 50:17)



Here one natural question, one may ask is that, here we are looking at the life of the tube, so life is a non-negative number or a positive real number; so, how it can be normally distributed, because normal distribution is from minus infinity to infinity, that means, it takes any real value.Now, this can be explained like this, that the life may be positive, but suppose you are having a here mean is 160 hours; so, since that the normal distribution most of the probability is concentrated within 3 sigma limits, that is minus 3 sigma to plus 3 sigma actually more than 99.9 percent of the probability is concentrated here.

(Refer Slide Time: 51:12)



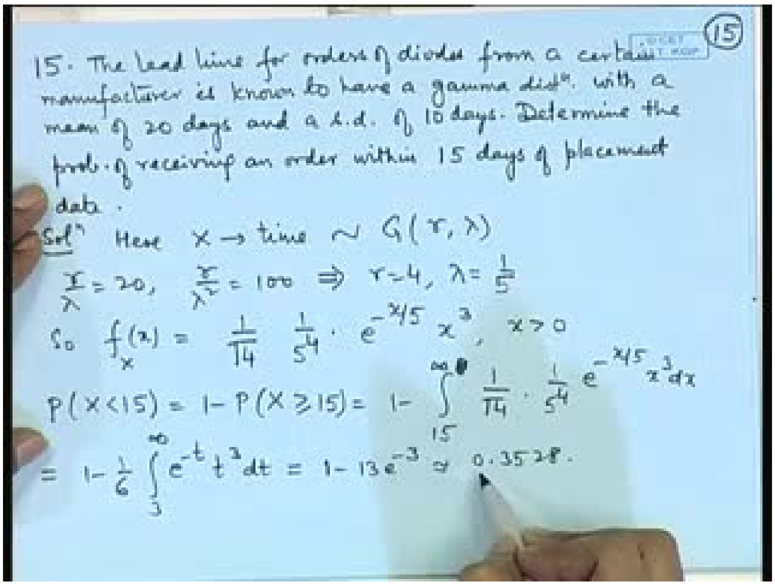
So, actually the values before 0 will not be coming into picture, so this is only a theoretical approximation to the practical situation.Theoretical normal distribution is having values from minus infinityto plus infinity, but in practice, the values will concentrated in 3 sigma limits around mu.So, in this particular case, if X is denoting the life, then X follows normal distribution with mean 160and variance sigma square; now, it is given that, the life is between 120 to 200 hours with probability point 8, so if you transform X to standard normal by subtracting 160 and dividing by sigma, then this is reducing to probability of Z lying between minus 40 by sigma to 40 by sigma.

We want this probability to be at least point 80; now, here we can simplify this term, it is phi of 40 by sigma minus phi of minus 40 by sigma, we make use of the property that phi of t plus phi of minus t is equal to 1; so, this is reducing to twice capital phi of 40 by sigma minus 1 greater than or equal to point 80.So, from the tables of the normal distribution, we see the point, that is the probability is more than 0.90, such that phi of phi 40 by sigma is greater than or equal to this.So, we can calculate this and sigma turns out to be less than or equal to 31.2.

Now, if sigma is equal to 30, what is the probability of an item working till additional 30 hours, which has already worked upto 140 hours; so, it is probability X greater than or equal to 170 divided by X greater than or given that X is greater than or equal to 140; so, it is the conditional probability and it will beturning out as the ratio of probability X greater than or equal to 170 divided by probability X greater than or equal to 140.

So, after transforming to the standard normal probabilityfunction, this value can be evaluated, and it is approximately 50 percent of the probability, that means, if theitem has already worked for140 hours, the probability that it will work for another 30 hours is nearly half.

(Refer Slide Time: 53:41)



The lead time for orders of diodes from a certain manufacturer is known to have a gamma distribution with a mean 20 days, and a standard deviation 10 days.Determine the probability of receiving an order within15 days of placement date.

So, let X denote the time; so, this is following a gamma distribution with parameter r and lambda, then the mean of a gamma distribution is r by lambda, that is given to be 20, and the variance r by lambda square is given to be 10 square, that is 100; so, after the solving these two equations, we get r is equal to 4 and lambda is equal to 1 by 5; so, the probability density function of the time is 1 by gamma r lambda to the power re to the power minus lambda x x to the power r minus 1.

After substituting the value of r and lambda, we get the form of the probability density function as this.So, probability that X is less than 15 is 1 minus probability X greater than or equal to 15, so that is 1 minus 15 to infinity, the probability densityfunction; here, we can make the transformation X by 5 is equal to t, then it is reducing to a simple integral 3 to infinity e to the power minus t t cube d t, this can be evaluated using integration by parts and the term is equal to 1 minus 13 e to the power minus 3, which is approximately0.35; so, under these conditions the probability of receiving an order within 15 days is nearly one third.

So, today we have seen various applications ofdiscrete and continuous distributions.In the next lecture, we will consider the distributions of the functions of random variables, so we will stop here.