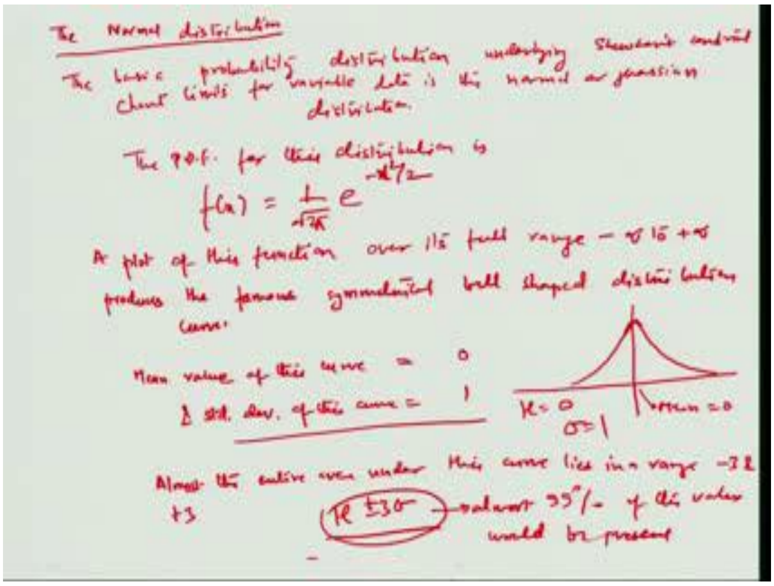
Manufacturing System Technology - II

Prof. Shantanu Bhattacharya
Department of Mechanical Engineering or
Industrial and Production Engineering
Indian Institute of Technology, Kanpur

Lecture – 27

Hello and welcome to this manufacturing systems technology part 2 module 27. We learn about the process distribution and the in this particular module will be talking more about the normal distribution.

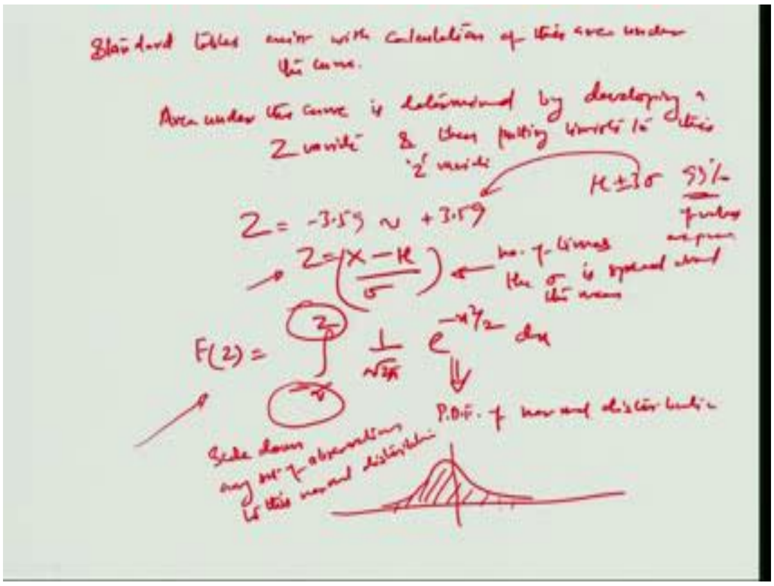
(Refer Slide Time: 00:26)



So, the basic probability distribution underlying the Shewhart's control charts, as I think I had illustrated earlier limits for variable data is the normal or Gaussian distribution. The probability density function for this distribution is given by fx equal to 1 by root of 2 pi e to be power of minus x square by 2, and a plot of this function over its full range minus infinity to plus infinity produces the same as symmetrical by shades distribution curve. So it is actually continues curves over its full range and has a mean a value look around a 0 and a standard deviation of about one. So, that is how you can calculate the way that distribution are spread of actually.

So almost of the… So, giving this initial input that the mean value of this curve 0 and standard deviation of this curve equals 1. So, almost the entire area under this curve something as indicated here this is mean around 0, and there is standard deviation of associated with this curve. So, almost the entire are under this curve lies in the range minus 3 and plus 3. So, basically typically within new plus minus 3 sigma, sigma be mu equal to 1 and mu is 0 in this particular case almost you can say 99 percent of the values would be present.

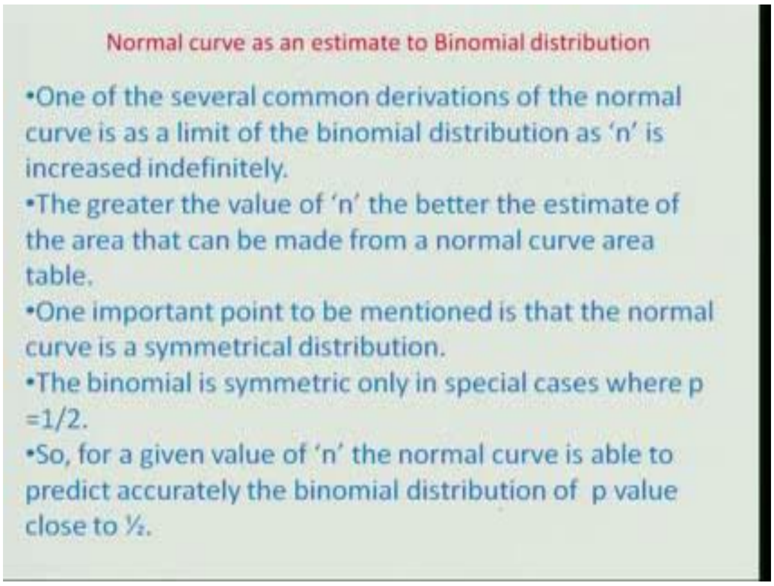
(Refer Slide Time: 04:35)



So, therefore obviously, there are standard tables, reflects these values. And the area under the curve is determined by developing a z variate, and then putting limits to this z variable. Typically the z value is varying between minus 3.59 to plus 3.39 and the z is actually equal to whatever the value x minus mu by sigma. So, it is actually the number times the sigma is spread around the mean, you already know that within mu plus minus three sigma almost 99 percent of the values are present. In fact, if you just extend this from 3 to 3.59, they will cover almost hundred percent of the values. So, fz therefore, it is nothing but the area under the scarves between minus infinity to 2 this value z 1 by root of pi e to the power of minus x square by 2 dx, which was earlier the value for the probability distribution function, of normal distribution.

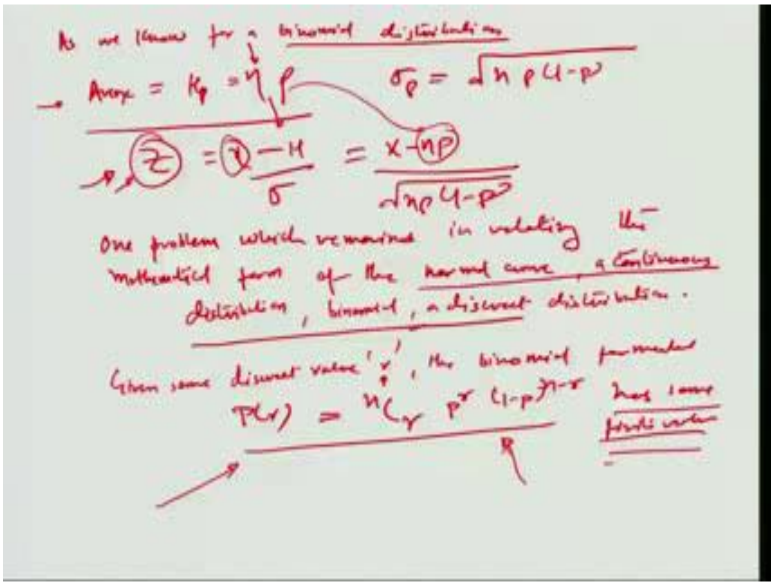
So, this is whatever common knowledge is needed about this normal distribution. So obviously, if I can the idea is, but I can scale down any set of observations to this normal distribution in the occurrence of need event within the are under the curve becomes very, very easy which is actually important for statistical quality control.

(Refer Slide Time: 07:35)



So, we would like to sort of you know go head, and try to see what is the derivation of this normal distribution curve or normal curve as the limit to the binomial distributions same way as we did for as n is increase in definitely the greater the value of n the better the estimate of the area that can be made from the normal curves area table. And one important very important point that distribution is actually symmetrical distribution. So, binomial distribution become symmetrical only in a special case, where you have exactly probability of 50 percent to be fraction defective or fraction good. So far a given value of n normal curve is able to predict accurately the binomial distribution of P value close to half, but it does not mean that cannot. So, that is how the limit of this particular thing can be determine.

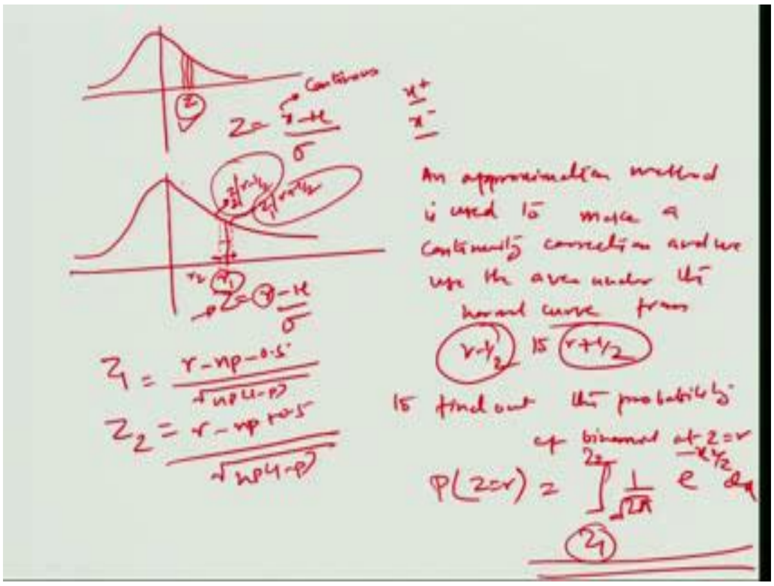
(Refer Slide Time: 08:46)



So, in light of this as we know that for a binomial distribution, the average value mu p, if I would look at the binomial distribution is given by np as proved earlier, and the standard deviation was given by n times of p times of 1 minus p were, p was the fraction defective or fraction good 1 minus p was the fraction good. So, we want to substitute these values in the z variate. So, the z earlier from this normal distribution was x minus mu by sigma. So, how many times the x is spread from the mu value or the mean value of the distribution in terms of number of sigma's. So, I just simply substitute these values from the binomial. So, x minus np divided roof of np times of 1 minus p. So, although we can try to calculate z variate in this manner, but one problem which comes or 1 problems which remains in relating the mathematical form of the normal curve, which is a continuous distribution and the binomial distribution which is discrete distribution function.

So, therefore how to sort of correlate a continuous to a discrete is major question that is being asked. So, just by putting the substituting the mu p value does help us much because you know the z variate is the continues domain, you know variate where as binomial distribution is normally for discrete functions, you have discrete value of p and you have counts 1, 2, 3, 4, so on so forth. So, there are completely disguised distribution. So, given some discrete value r, where r is the number of times rejection happens probably. The binomial formula predicts the probability of this r to occur as ncr p to the power of r 1 minus p to the power of n minus r and this has a final value. So, how do we predict, you know this particular probability function with normal distribution.

(Refer Slide Time: 12:20)



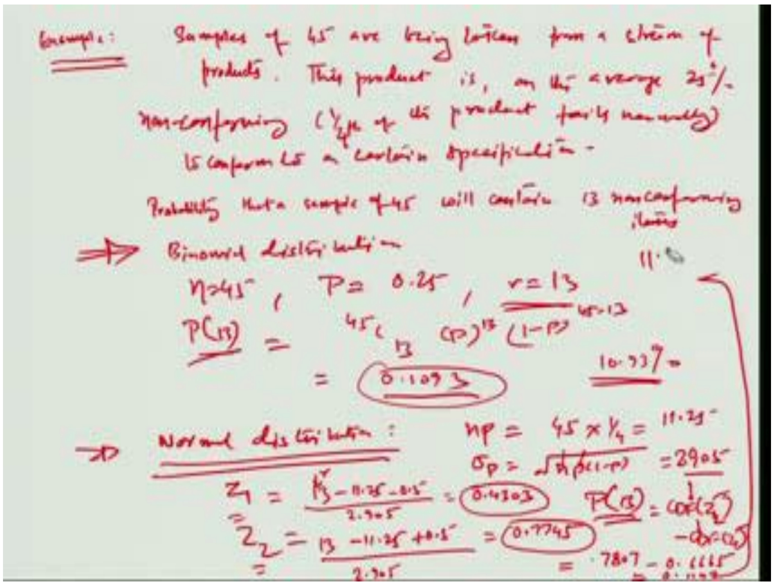
So, typically the idea is that when you have a fit to this normal distribution function and you are looking at the probability of a value z. So, basically try to look around this value and since in this the z variant is nothing but x minus mu by sigma, and if it is continuous variable then we have no problem, because there would exist you know something in the near vicinity of x plus and x minus where there would be still a existing a value, but in this case if I wanted force fit a binomial distribution, where this z is made of r minus mu by sigma really where r is basically the number of rejects.

So, around the value r; there are as such no other continues domain values on both sides and everything is discritise. So, corresponding to r 1 let us say and corresponding to r 2 we have 2 different discrete values. So, the area under in this particular case add a certain level of let us say z value is really not calculable unless you make certain corrections and try to use the approximation method. So, an approximation method is use to make a, you can say this can be called as a continuity correction you know, and we use the area under the normal curve from r minus half to r plus half. So, that this way I can try to establish values near this r which is exactly the number of times you are you know doing the trial etcetera. So, is a discrete values. So, around this r I am establishing a domain, so that you have local continuity at r to find out the probability of binomial add z equal to r.

I am going to be more clearly representing it when we actually come to discussing a numerical problem, but here the idea would be that you will be trying to establish the z variate functions has 2 variables rather than 1 single variable at r in one case you will try to associate z r minus np minus 0.5 divided by root of np times of 1 minus p, and the other case you will try to associate r minus np plus 0.5 divided by root of np into 1 minus p in a manner. So, that the probability corresponding to a z equal to r can be found out as the difference in the regions between z 1 variate and z 2 variate 1 by 2 pi root e to the power of minus x square by 2 dx which in any event is the cumulative area.

So, this is basically a case where you are assuming a z variate corresponding to r minus half. So, let us write this down as z variate corresponding to r minus half another z variate corresponding to r plus half let us called this z 1 and this z 2. You basically trying to find out the area under the curve up to the point z 1, and delete or deduct the are up to the point z 2 to find local probability in that particular domain. So, you are trying to force fit this into a local continuity. So, that the values can be calculated in accordance to what is they are in the normal distribution on normal curve.

(Refer Slide Time: 16:29)



So, I would like to discuss a small example problem here. So, let us say samples of 45 are being taken from a stream of products, this product is on the average 25 percent nonconforming. So, the p value here is 25 by 100 or 2 and 25. So, about a fourth of the product fails in normally to confirm to a certain specification. So, we would like to find out the probability that is a sample of 45 will contain 13 nonconforming items. So, the first goal here is to represent this by a binomial distribution. So, we will calculate using a binomial distribution. So, let us say n equal to 45 and the probability of acceptance or a, sorry rejection here is about 25 percent quarter of the products are rejected, r here is 13.

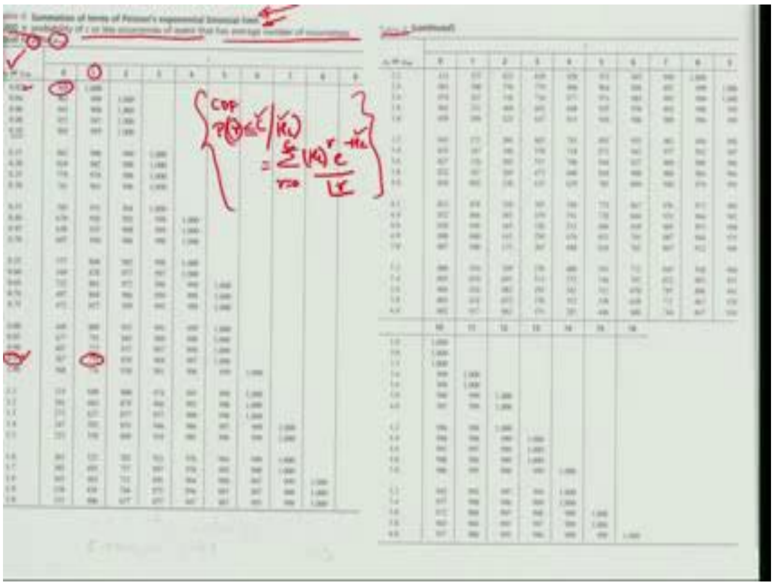
So obviously, the probability of having exactly r rejects the sample of 45 is given by 45 c

13 p to the power of r. So, p to the power of 13 times of 1 minus p to the power of 45 minus 13 in the binomial manner, and this can be this is coming out to be 0.1093. If you want to use the normal distribution to calculate the same thing, so therefore the probability is about 11 percent or may be 10.93 percent if you do it with the binomial distribution while using normal distribution will again use the similar kind of limits as talk about the making a local continuity. So, we want to create 2 variate is a z 1 and z 2 here, where z 1 is basically 13 that is the r value minus np, so obviously np in this particular case, because the sample size is 45 it is about a quarter reject is about close to 11.25, and similarly the sigma p from the normal distribution here is root of np times 1 minus p. So, this comes out to be equal to 2.905, when you calculate substitute the p value and n value as 1 fourth and 25 or 45 respectively.

So, here we can calculate this is r minus np that is minus 11.25 minus of 0.5 making a local continuity around the value r divided by 2.905, it comes out to be 0.4303. Similarly z 2 variant here would be in terms of 13 minus 11.25 in the plus side of the r we make a local continuity by putting r plus 5 of r plus 0.5. So, this divided by 1 2.905 which is actually 0.7745. So, the probability of having exactly 13 rejects here, and this case would actually be the area under the curve up to the point z 2. So, let us I just called the probability distribution function or cumulative distribution function CDF of z 2 minus of cdf of z 1 and from the normal tables these values can be obtained as 0.7807 minus 0.6665 corresponding to this z 2 and z 1 value and this comes out to be equal to about 0 0.1142. So, in this case with the normal distribution and assuming a local continuity, if fairly close you are getting about 11.42 percent which is very, very close to the value express earlier about 10.93 percent. So, you see that you know it does not give much impact if you make a local continuity discreet value of r, so that you can implement the normal distribution function suitably.

So, in this whole module we have sort of analyze various things related to the zone distribution, how that can be fitted and also the normal distribution and how this can be aligned to the binomial distribution, and binomial distribution earlier has been giving summery in a very nice manner has to how this can be a line to accepting sampling plan. So obviously, all these things fall in line that all this three distribution can be use to appropriate an affluently has the case may suggest; for example, in ozone case it is a case of large sample with very small probability or in this case normal distribution case it is you know just to simplified the calculation that would otherwise generate out of binomial your are making a local continuity. So, that the normal can be force fit to the binomial the idea is in this all this manner you can predict the probability of occurrence of the certain number of reject which is very, very important. So, only other thing that I would like to share now is a what are the real value which are plotted in the tables which approximates both the ozones in the normal distribution case.

(Refer Slide Time: 23:05)



You may recall the ozone probability distribution had a cumulative density function CDF which was provided by the probability of the value r to be less than or equal to some value c given a certain mu c rate of occurrence, and which is the average and which was represented as sigma r wearing between 0 and c mu c to the power of r e to the power of minus mu c by factorial r. So, have in said this has the probability of having exactly rejects less than equal to c and for a particular mu c case or a particular mu np or whatever you call the mean of the distribution.

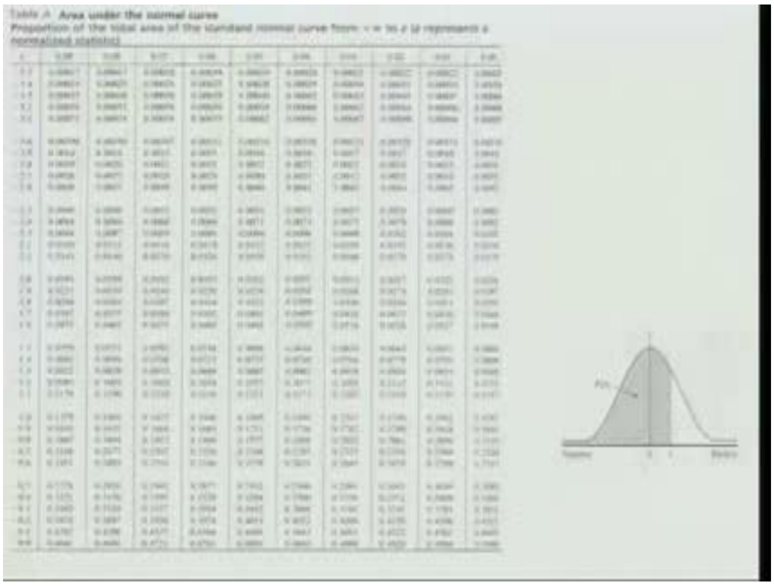
The table g here shows the summation of these terms which is ozone sort of you know exponential binomial limit. So, this table g is basically 1000 times the probability that is being reflected here.

So, 980 exactly means 0.98 of c or less occurrences of event that has an average occurrence which is given equal to mu c or mu npn. So, if you see if the mu c value write about here is let us say 0.02, and the number occurrence is are exactly 0 times or 1 time the probability is 0.98 and 1 respectively. Similarly the mu c is around the 0.95 and the you know occurrences r exactly let us say 1 time. So, corresponding to 0.95 and number of occurrences equal to 1 or less the total cumulative distribution is 0.754. So, this is you read this table g and you can actually applied ozone case to find out, what is the probability of occurrence of the certain event in the whole distribution.

So, that is how you can sums these are tables, and these tables as you can see correspond to a mu c value from 0.02 all the way to almost about 22 of 25 or so. And the c value varies all the vary from 0 to about 43; obviously, as you know there are certain terms which are redundant, so far an example for a case of 10.5 mu c. If I go about 23 the probability all going to be 1, we have just made that redundant. So, whenever there is a limit to the full probability of 1, you sort of does not repeat the values here and` truncate the table in this manner. So, this write here is up to forty three occurrences; that means, for a 43 or a less c values, you can plot or you can read from these tables all mu's all the way about 25 words. So, that is how this calculation of can be perform.

You can keep on changing and increasing in this table by looking at the distribution function, the cumulative distribution function that has been given to you earlier and discussed in some details, but the idea is that this is the way to apply the ozone to calculate the probability of occurrence of some particular event, in such complex distributions were the probability percentage is very low and the numbers which are involved of the number of data; data sets of very, very large.

(Refer Slide Time: 26:30)


Area under the normal curve

Similarly that the thing mentioned about the normal tables this right here shows the area and the normal curve corresponding to certain z variate; the z variate again varies between minus 3.52 almost about positive 3.5; obviously, this means to accommodate almost 99.99 percent values of the occurrences, and this corresponds again 2 a, z value of all the way from 0.00 to 0.09 and so on so far as given in table a. So, you can use this tables for the purpose of you know you can again have, you look at this, this is part a of the table again part b of the table. In a similar manner to like just in the cases of fozones and you can use this values for and these are well able in any other you know source including even online.

So, you can look at this values and try to estimate the probability of occurrence of a certain event r; for example, r could be the number of draws you know in as we just executed earlier in case of defining the binomial distribution. So, that is how you calculate using the normal table, the advantage here is the calculation becomes much more simplified in this particular case, just because you are force fitting a function which as having a, which as entered about 0 and a standard deviation of plus or minus 1. So it becomes lot more easier to handle this way numerically. So having said that I would like to close this module here and the next module would start to plotted the some of these acceptance sampling plans, and how p chart can be used to govern such a plotting and what could be the interpretations coming out of the systems.

So, as of now thank you so much.