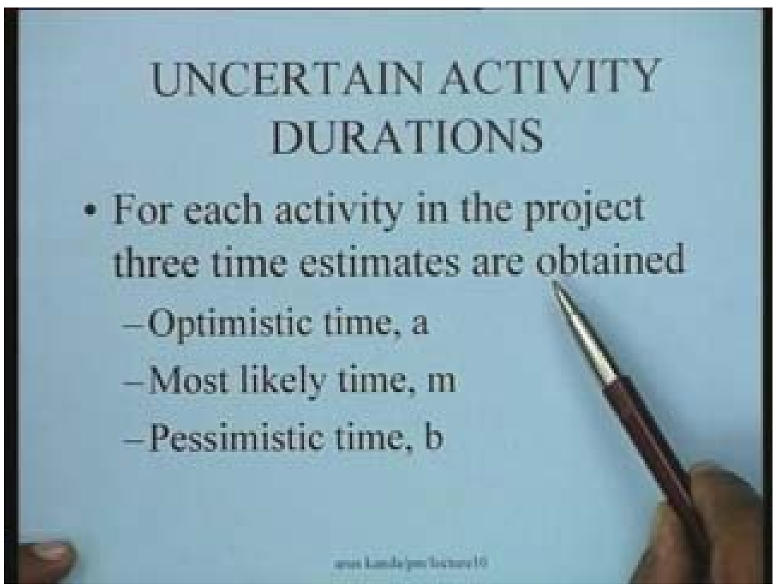
Project and Production Management
Prof. Arun Kanda
Department of Mechanical Engineering
Indian Institute of Technology, Delhi

Lecture - 11
Project Scheduling with Probabilistic Activity Times

So far we have been talking about basic project scheduling in which the activity times were deterministic. Today we are going to talk about project scheduling with probabilistic activity times. In fact we shall be dealing with the basic PERT model and we shall see how uncertain activity times or random activity durations can be accommodated in the project scheduling framework using the basic PERT methodology. The essential problem is that activity durations are uncertain and what is basically done in PERT is that for each activity in the project we assume that there are 3 time estimates.

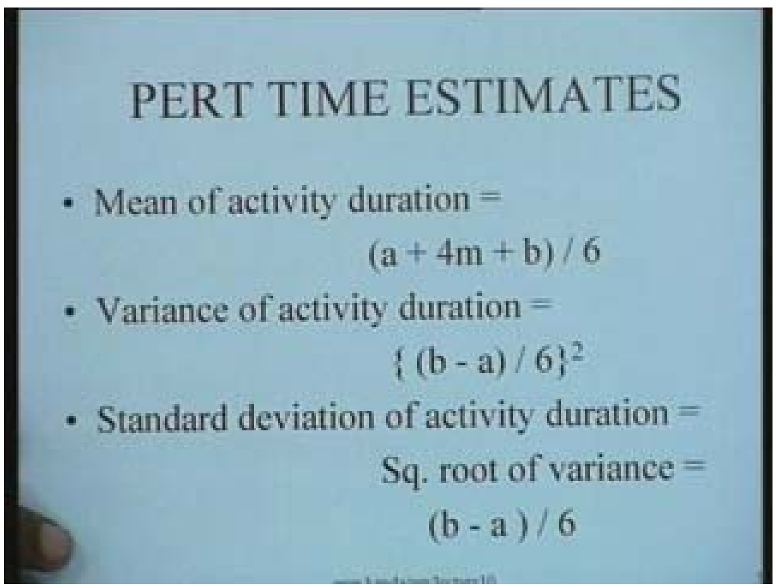
(Refer Slide Time: 1:55)



These 3 time estimates are typically referred to as the optimistic time. The optimistic time is that time which is probably the shortest possible time in which that particular activity can be accomplished and it would therefore be a situation where everything goes favorably and you can perform the activity in the minimum possible time. Then we can at the other extreme talk about a pessimistic time b. The pessimistic time would be an estimate of the time that a particular activity would take under the worst possible circumstances which means that this would be the longest likely time for that particular activity and between these two extremes we would have an estimate of time which we refer to as the most likely time m which would in fact be the mode of the distribution of the activity duration in that particular set. This is how we can capture the uncertainties in activity durations by making 3 time estimates instead of a single time estimate as we do in the basic scheduling computations that we did in the c p m analysis.

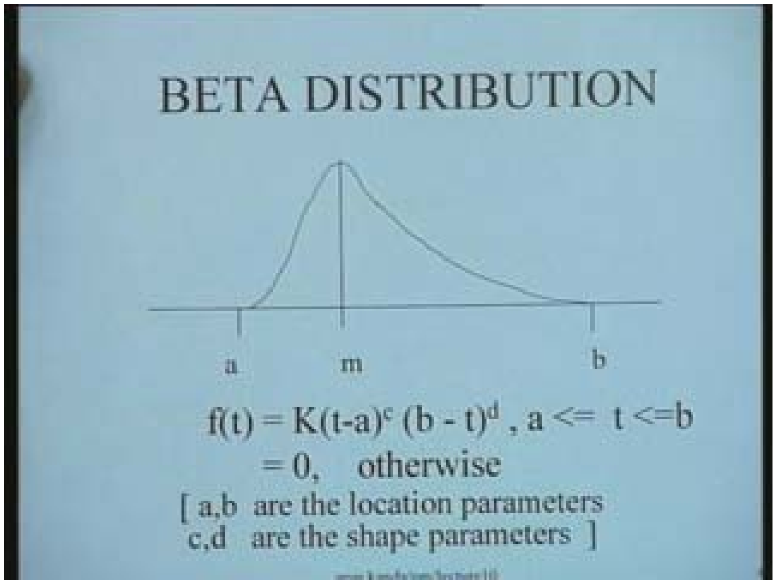
These three time estimates that we talk about could in fact be easily combined to obtain the various parameters of the distribution. The first parameter which is of interest to us is the mean of the activity duration and under standard PERT computations we estimate the mean to be a + 4m+ b by 6. It's like saying that we give 4 times the weightage to the most likely time as we give to the optimistic and the most likely times and we divide the whole thing by 6. It's like a weighted average of the optimistic, the pessimistic and the most likely times and we get this particular value of the mean of the activity duration and then in standard PERT computations the variance of the activity duration is simply computed as b-a by 6 whole squared. Here b-a would in fact be the range of the distribution. The range divided by 6 is an estimate of the standard deviation and therefore we can say that the standard deviation of the activity duration which is the square root of the variance is nothing but b-a by 6 in that particular way.

(Refer Slide Time: 4:58)



In standard PERT analysis the distribution assumed for the activity times is in fact a beta distribution and the beta distribution looks something like this.

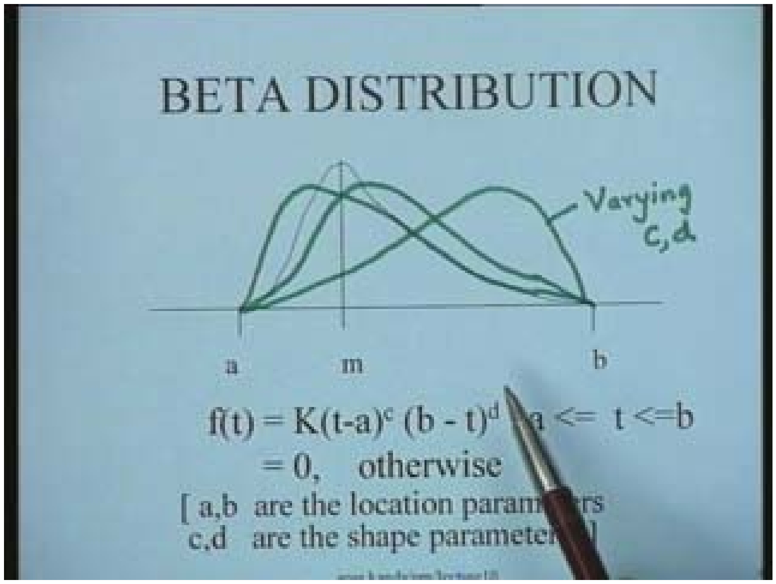
(Refer Slide Time: 5:24)



The beta distribution is a distribution which has a minimum value at a and a maximum value at b and it has a mode, a maximum peak at this particular value of m and in fact the probability density function of the beta distribution is given by this particular function where K is a constant and this is multiplied with t minus a to the power c and b minus t to the power d where t lies between a and b and it's zero otherwise. Notice that in the beta distribution that we have here a and b are referred to as the location parameters. The reason for doing this is that these two parameters determine where the distribution is located on the time axis. It's between a and b. That's why these are very legitimately known as the location parameters and c and d are the shape parameters. In fact by changing the values of c and d we can get different distributions wounded between the same particular values. For instance I would get a distribution that's something like this. Then I could get a distribution something like this. Then I could get may be a distribution something like this.

These are by varying the values of c and d which are the two shape parameters for this particular problem. By choosing c and d we can vary the shape of the distribution and by choosing a and b we can identify the location of the distribution on the entire axis, the time axis. That's why a and b are called the location parameters and c and d are called the shape parameters. You would see very easily from this example that the beta distribution is in fact a 4 parameter distribution. You compare it for instance with normal distribution. A normal distribution is a 2 parameter distribution the two parameters being mu and sigma, the mean and the standard deviation or the mean and the variance of the distribution. Notice that K is not a parameter. Why is K not a parameter? It's simply not a parameter because the total area under the curve must be equal to 1 and if you do that you would get the value of K in terms of the various parameters a, b, c and d and therefore K is really not a parameter of the distribution.

(Refer Slide Time: 8:18)

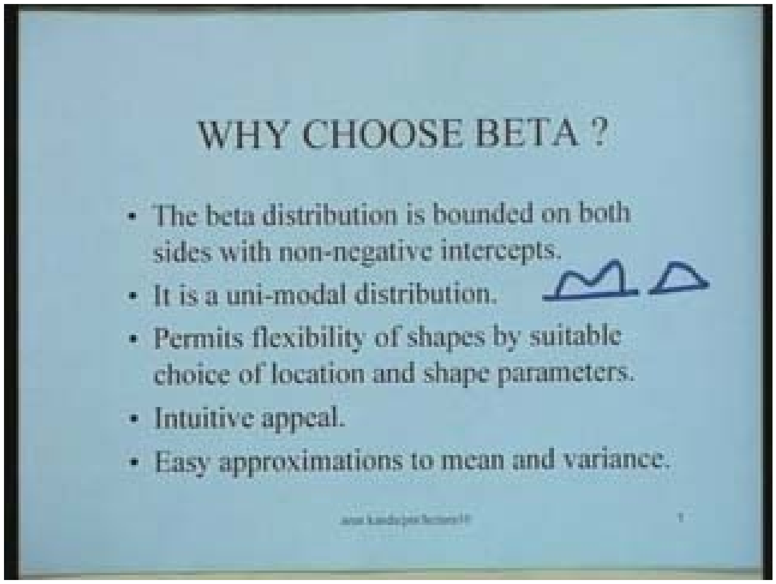


This is a beta distribution which is chosen by the originators of PERT to represent the activity times.

A question that arises at this stage is why did the originators of PERT choose a beta distribution? Why didn’t they choose any other distribution? This is a question. Why chose beta? When there is a plethora of distributions available why did they choose beta? Beta has certain properties which we can see and we see that it favors the representation of the distribution of time. First thing is the beta distribution is bounded on both sides with non-negative intercepts unlike for instance a normal distribution which extends from minus infinity to plus infinity on both sides. The 2 non-negative intercepts can be very conveniently given the interpretation of the optimistic and the pessimistic times, the minimum possible time and the maximum possible time. This is one property of the beta distribution.

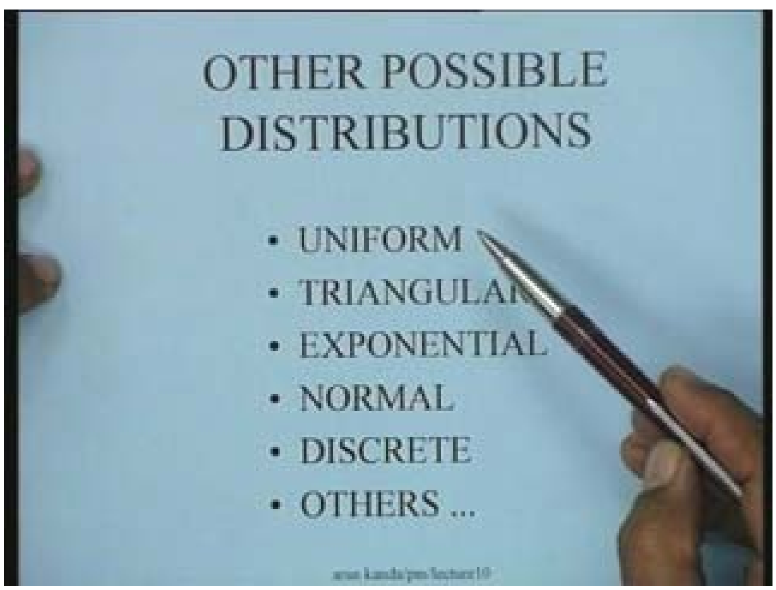
Second thing is it’s a uni-modal distribution. It’s a distribution that has a single mode. It doesn’t have multiple modes. What would be a multiple modal distribution like? This is a bimodal distribution if we had a distribution like this. The beta distribution is in fact a distribution which has a single mode. That is a single particular peak in the distribution. There is this advantage of then referring to this particular value as the most frequently occurring value of time because it’s the peak of this frequency distribution. Then the beta distribution permits flexibility of shapes by suitable choice of location and shape parameters. This we just saw that by choosing different values of the shape parameters you can give a different shape to the distribution. It has an intuitive appeal because of the factors that we just mentioned and it permits easy approximations to computations of the mean and the variance. The two formulas that we just mentioned for computing the mean and computing the variance can actually be derived as very simple approximations of the true values of the mean and the variance from the properties of the beta distribution and these formulas are very convenient to use or adopt in practice.

(Refer Slide Time: 11:04)



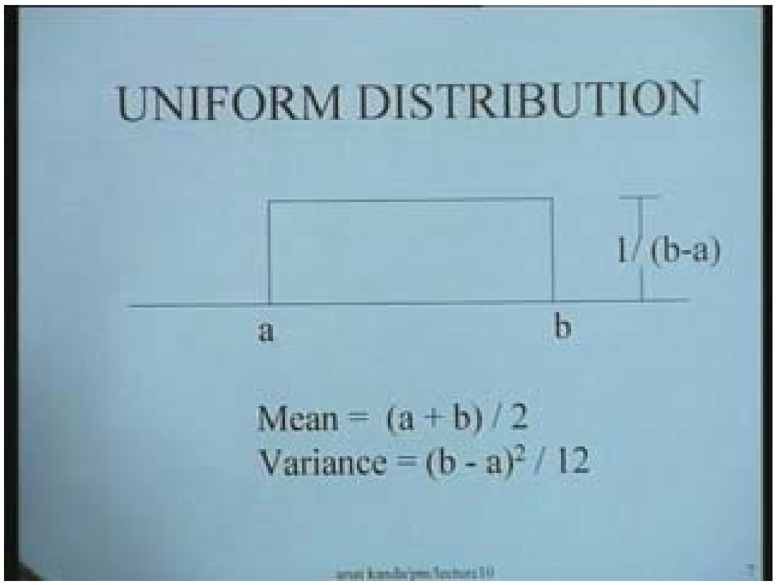
However I would like to point out here that there are other distributions and in fact we could use any possible distribution for activity times. Here are some instances of other possible distributions for activity times.

(Refer Slide Time: 11:21)



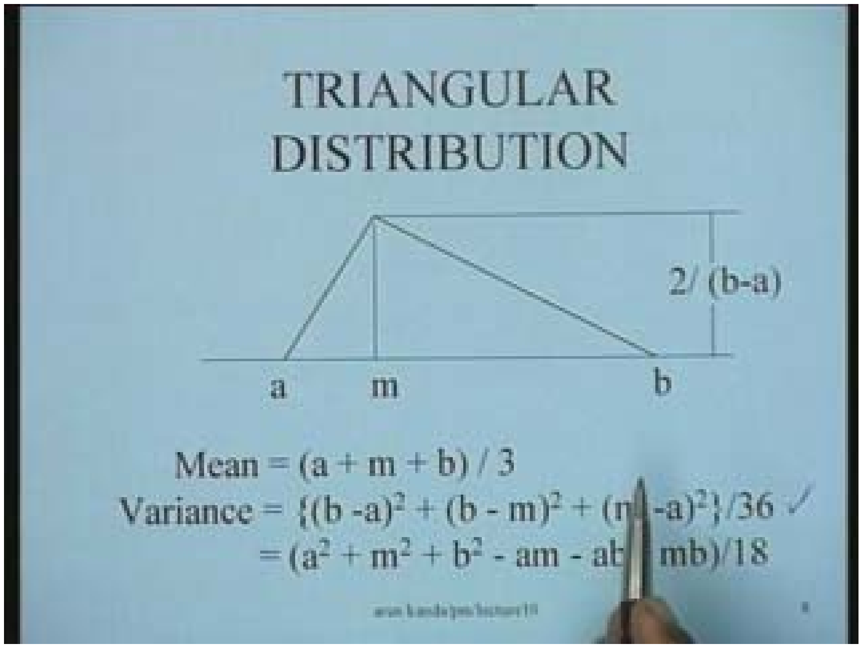
One could use a uniform distribution or a triangular distribution or an exponential distribution or a normal distribution or a discrete distribution or from the many of the other distributions which are available you could select any distribution and use that as a distribution for the activity times. In fact this can be done very easily. For instance the typical results that we would require for carrying out a PERT analysis are the uniform distribution would show that any time between two limits a and b is equally likely. The probability density function of this distribution would be a simple rectangle of this shape and you can see that because the area is 1, this particular height of the rectangle would be 1 upon b-a and therefore the p d f of this distribution is completely defined and the mean of this distribution is simply a + b by 2 and the variance is b-a whole square by 12 and these results could be utilized if we were to use a uniform distribution instead of a beta distribution.

(Refer Slide Time: 12:38)



Similarly one could work with a triangular distribution. A triangular distribution is also an intuitively appealing distribution because it has 3 time estimates the optimistic, the most likely and the pessimistic time. But what we are assuming is that the variations from a to m are linear in this particular fashion. Then the pdf of this distribution would look like this. This height would be equal to 2 over b-a because this entire a area of the rectangle must be 1 and consequently the mean of this distribution is a very convenient simple formula a + m + b by 3. Just the simple average of the 3 values and the variance of this distribution can be computed for purposes as just the square of b minus a, square of b minus m and square of m minus a. Basically these intervals 1, 2 and 3, squares of these divided by 36 or you can simplify this expression if you like to write a square plus m square plus b square minus am minus ab minus mb divided by 18. This particular form is much more convenient for computation because all one has to do is get this interval, this interval and the total interval, square it up and get the values.

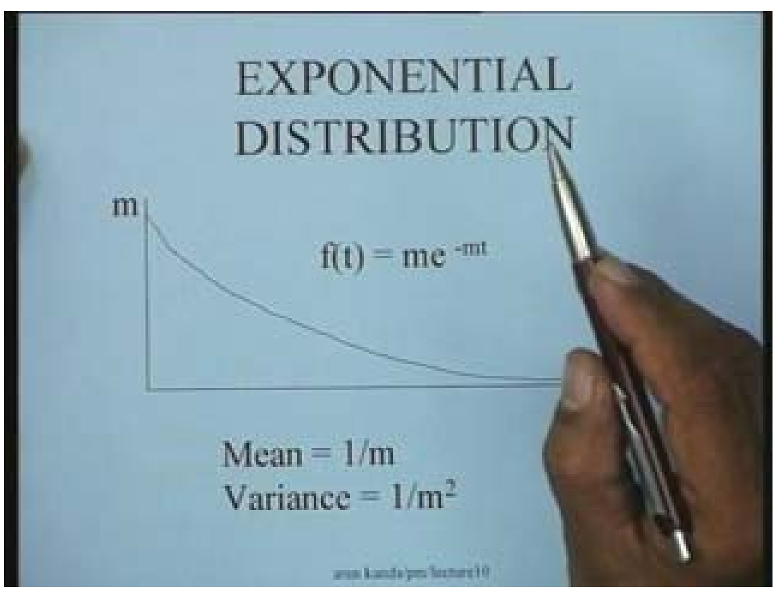
(Refer Slide Time: 14:09)



It’s not a kind of beta distribution. It’s a different distribution. It’s a distribution in its own right. Very true it looks like a beta distribution but a beta distribution has a probability density function which we saw was governed by the particular expression K into t minus a to the power c and so on. Here these are straight lines. They are similar in the sense they have these 3 time estimates but the distributions are entirely different.

Another commonly used distribution could be an exponential distribution.

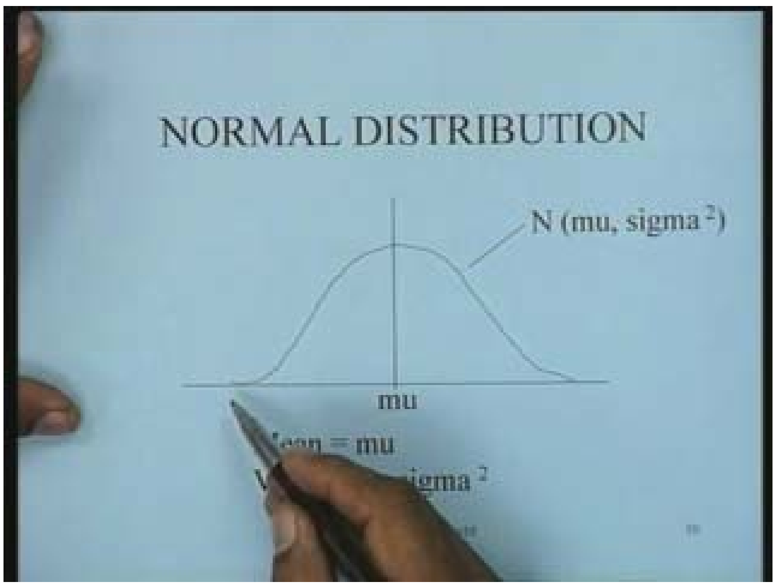
(Refer Slide Time: 14:52)



An exponential distribution as you all know is a distribution which reads f(t) is equal to m e to the power –mt. If this is the intercept m, then it gradually drops and goes on till infinity and the mean of an exponential distribution is just 1 by m and the variance is just 1 by m square. Very convenient expressions for the mean and the variance of the distribution in case you want to use the exponential distribution as a distribution for the activity times.

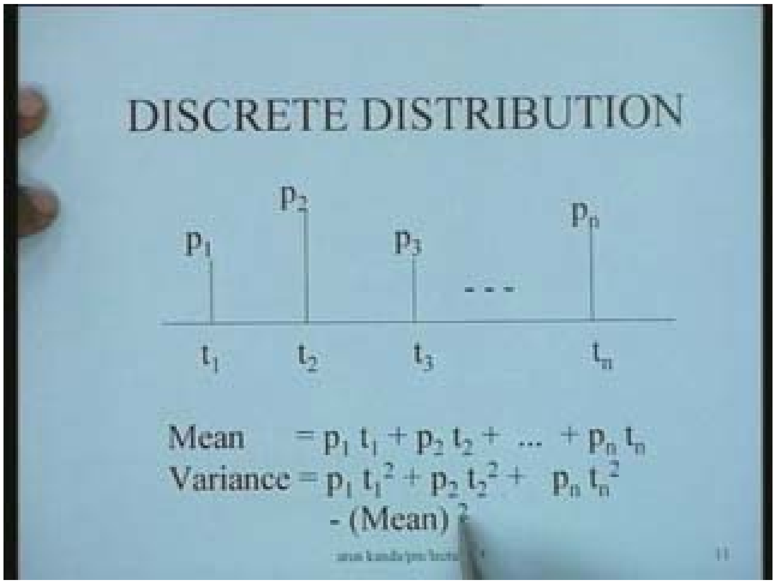
Then we can talk about a normal distribution. You are all familiar with a normal distribution. Normal distribution is a bell shaped symmetric distribution which looks something like this.

(Refer Slide Time: 15:46)



It's characterized by 2 parameters the mu, which is the mean of the distribution and the sigma square which is the variance of the distribution and this particular distribution in fact is one of the most popular distributions in statistics. Apart from this one might want to use a discrete distribution. A discrete distribution would be something like this. When the activity could take the time either t_1 or t_2 or t_3 and so on up to t_n with probabilities of p_1, p_2, p_3 and so on up to p_n. This may happen for instance if you have the historical record of the time that a particular activity took. It might have taken either 10 days or 12 days or 15 days or 20 days in the past and if you have done it for say 100 times you will know how many times it took 10 days and how many times it took 20 days and how many time it took 30 days and so on and these estimates would be then estimates of the probability that you have for this distribution. In a situation like this the mean of the distribution is nothing but p_1t_1 plus p_2t_2 and so on up to p_nt_n and the variance of the distribution can then be computed as p_1t_1 square plus p_2t_2 square plus so on up to p_nt_n square minus the mean whole square.

(Refer Slide Time: 17:24)

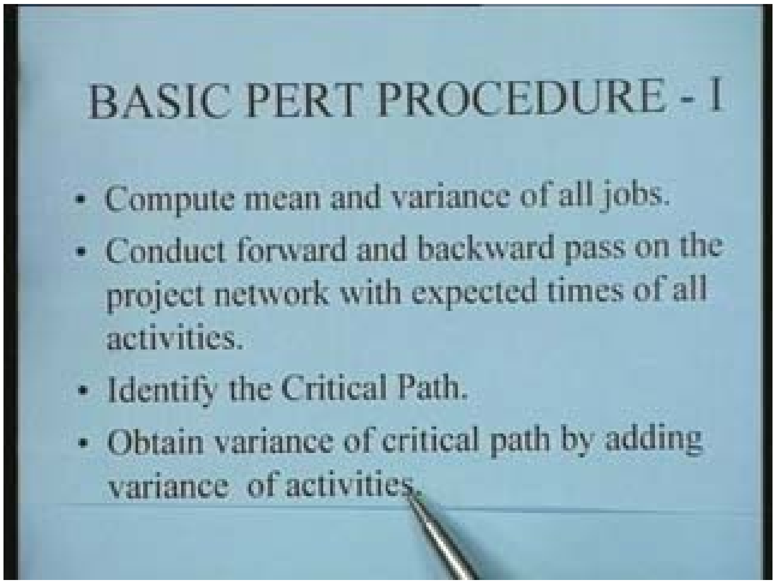


We are in fact here using that standard formula for the expected value of the expected value or the variance of x is nothing but the variance of x square minus the variance of the expected value of x whole square. You get these 2 formulas from that particular thing. The point is that the activity duration can have any distribution. Though the originators of PERT used a beta distribution there is nothing to prevent us from using any other distribution which we think is appropriate for the PERT computations.

Let us now try to look at the basic PERT procedure. When you try to use PERT analysis the first step is generally, after you have obtained information about the activities and their times and all that, to compute the mean and variance of all the jobs. So you compute the mean and variance of all the jobs. Thereafter we conduct a forward and a backward pass on the project network with the expected times for all activities. What is being done here is that since we have computed the expected time for each job we sort of convert the problem into a deterministic problem and do basic scheduling in much the same manner that we determine the critical path in a deterministic network. The only thing that we are doing is we are using the expected times for all the activities and are now computing a forward and backward pass.

The next thing we do is to identify the critical path. This you can do after you have done the backward and forward pass. Then you can identify the critical path of this particular situation. The next step is to obtain the variance of the critical path by adding the variance of activities which lie on the critical path.

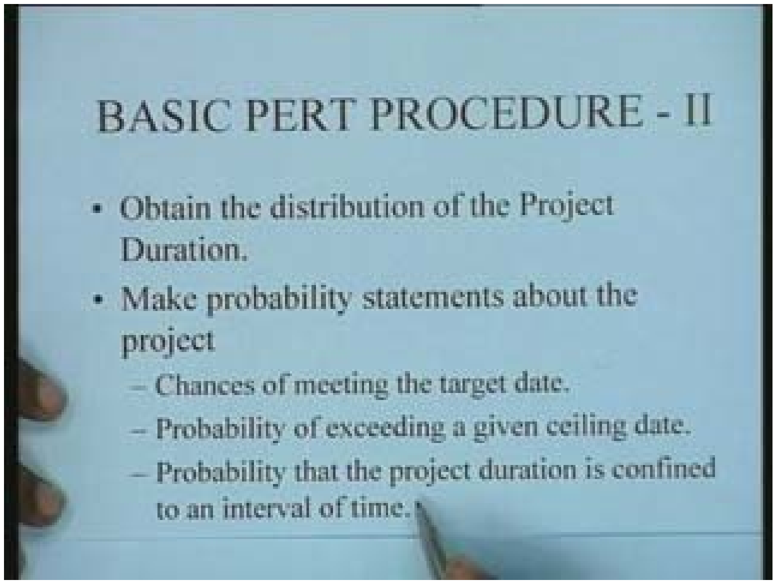
(Refer Slide Time: 19:39)



Having done this we are now in a position to compute the distribution of the project completion time because that is the primary output from the whole exercise of PERT. We then obtain the distribution of the PERT of the project duration and the project duration is in fact normally distributed with mean equal to the sum of the means and variance equal to the sum of the variances of the activities that lie on the critical path. Can somebody identify the reason for this? Well the reason is fairly simple. The reason is that the critical path or the project duration is the sum of the activity times. The activity times are random variables and the sum of the random variables if the number of random variables is fairly large would by the central limit theorem be approximately normally distributed with the mean equal to the sum of the means and the variances equal to the sum of the variances. It's basically an application of the central limit theorem that leads to the distribution of the project duration.

We will see when we do the computations as to how this is obtained but this is the underlying logic behind the whole thing and once we have obtained the distribution of the project duration we can make probability statements about the project and these project statements or the probability statements are what are the chances of meeting the target date. This is the question that the project manager would be very much concerned about. Suppose he has negotiated a project whose target date is 3 months from now and you know in all projects there are uncertainties in activity times. He would be concerned about finding out what are the chances of meeting the target date of 3 months that he had originally negotiated because if he doesn't meet the target date then there are likely to be penalties. He should be concerned about those things. Then he would also like to find out the probability of exceeding a given ceiling date. What are the chances that the project will exceed a certain value? This kind of information can be obtained from the distribution that we have just obtained and then you can be interested in finding out the probability that the project duration is confined to a certain interval of time.

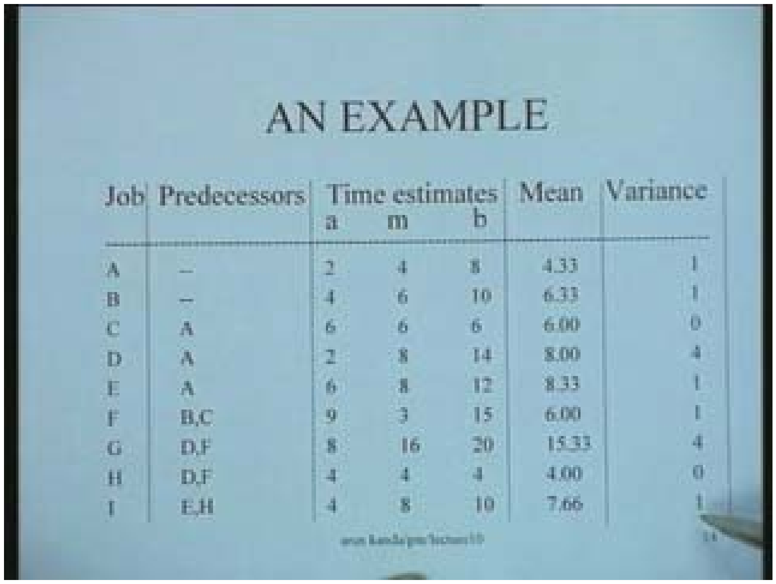
(Refer Slide Time: 22:29)



That means what are the chances that the project would be completed between 3 months and 4 months? If you are interested in answering this question you could answer these questions very exactly by making appropriate probability statements from the distribution of the project duration. This is essentially the basic use of PERT. We are able to capture the uncertainties in the environment and we are able to incorporate those uncertainties into the modeling process and make firm estimates of how long the project is likely to take, what are the probabilities that it's likely to be delayed and so on.

Let us now illustrate the basic PERT procedure with an example. Let us assume that we have a project which is something like this. The basic starting data for us would be the list of jobs, the list of predecessors for the jobs and the time estimates. As I indicated to you earlier there would be 3 time estimates: the optimistic, the most likely and the pessimistic time estimates for each job. In this case for instance we feel that job a can be done in 2 days or it can go up to 8 days. But the most likely time for completing job a is 4 days and similarly for job b we have these estimates. However we are fairly sure of the time of activity c. We are quite sure that activity c will take 6 days. We indicate 6 on all the 3 values of a, m and b and similarly activity h also we are fairly sure that it will take only 4 days. It will take only 4 days and we have these estimates and there are certain activities which have greater variance depending upon the nature of the activity and the kinds of factors which affect the activity. This is the starting point of obtaining the time estimates. Once these time estimates are obtained we compute the mean and variance time for each activity by utilizing the two formulas that I had enunciated a short while ago. The mean is computed by simply a + 4m + b by 6. The mean for instance for the activity 1 is a + 4m + b by 6 which is going to be 16 + 10 = 26 divided by 6 and this value is 4.33. In a similar manner all the means for all these activities are computed. We can look at this particular activity, probably a mistake in the typing of the activity. I think what needs to be done is 3 and 9 need to be reversed. You get the means and you get the variances for each activity.

(Refer Slide Time: 26:13)

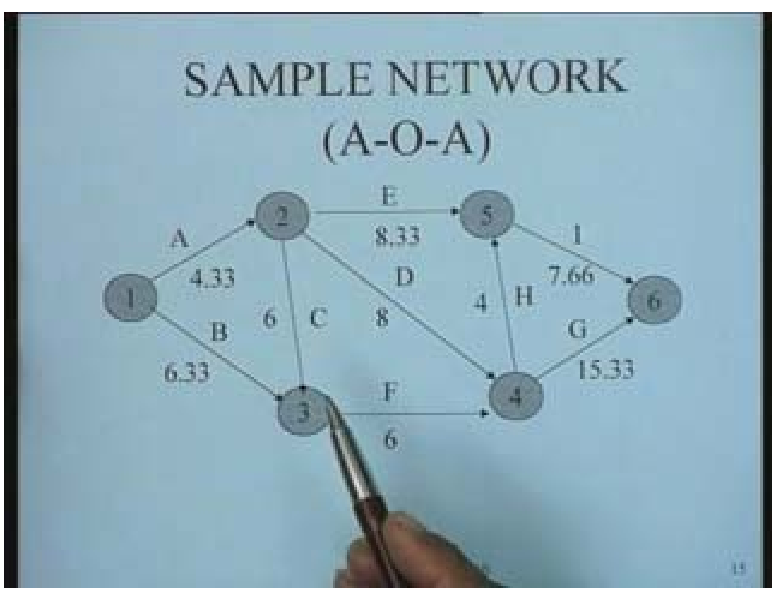


AN EXAMPLE

Job	Predecessors	Time estimates a	m	b	Mean	Variance
A	–	2	4	8	4.33	1
B	–	4	6	10	6.33	1
C	A	6	6	6	6.00	0
D	A	2	8	14	8.00	4
E	A	6	8	12	8.33	1
F	B,C	9	3	15	6.00	1
G	D,F	8	16	20	15.33	4
H	D,F	4	4	4	4.00	0
I	E,H	4	8	10	7.66	1

The variance for each activity for instance for a is nothing but b-a and b-a divided by 6 would be the standard deviation. Square of that is 1 which is the variance. Similarly the variances for the various activities in this particular example could be simply worked out as shown in this table here. Step 1 is to compute the means and the variances of the activities which can be very easily done. Then to start the computations what we need to do is draw a network and the network for this example is shown here.

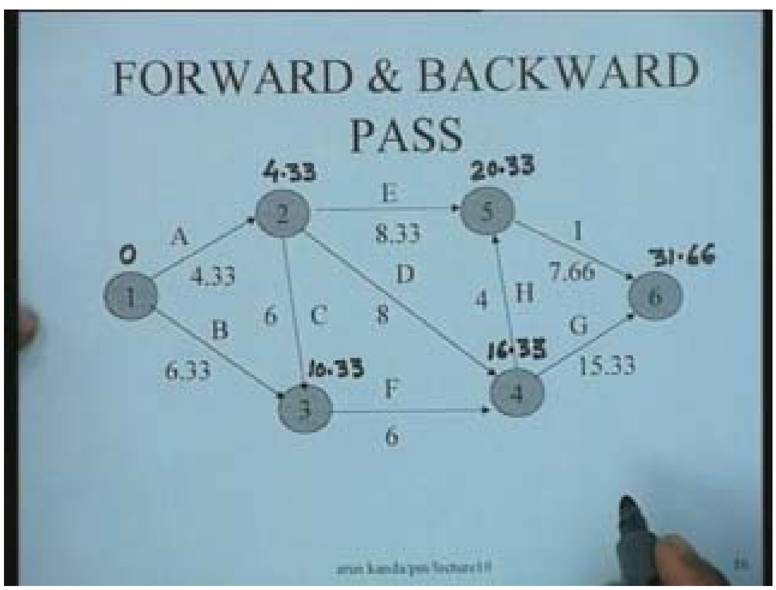
(Refer Slide Time: 27:00)



We have drawn the network of the activity on arc mode and when you are performing PERT computations it’s the A-O-A mode which is used. The times that are shown adjacent to each activity are the expected values of the times for each of the activities and we have in fact therefore the entire network which is shown in this particular diagram.

Now what we will do is we will do a forward pass and a backward pass on this particular network and identify the critical path just as we had done here. Let us do a forward and backward pass. We are going to determine the earliest occurrence times and the latest occurrence times for all the nodes. Starting from node zero, the earliest occurrence time of this node is zero. For node two the earliest occurrence time is going to be 4.33 because the duration of this particular activity is 4.33. When you come to node three it's going to be the maximum of 6.33 and 10.33. So this value is going to be simply 10.33. Then when you come to node 4, node 4 has 2 incoming arcs. In the usual manner when we are doing the forward pass this value is going to be 12.33 if you come via activity d and 16.33. So 16.33 being the larger of the two values will in fact be the earliest occurrence time of the node 4 and then for node 5 from this side we would get a value of 12.66 whereas when we go from this side we would get a value of 20.33; 4+16.33 which is 20.33 and subsequently when you come to node 6, the values that you will get from this side are something like 28 whereas if you go from here you would get a value of 31.66.

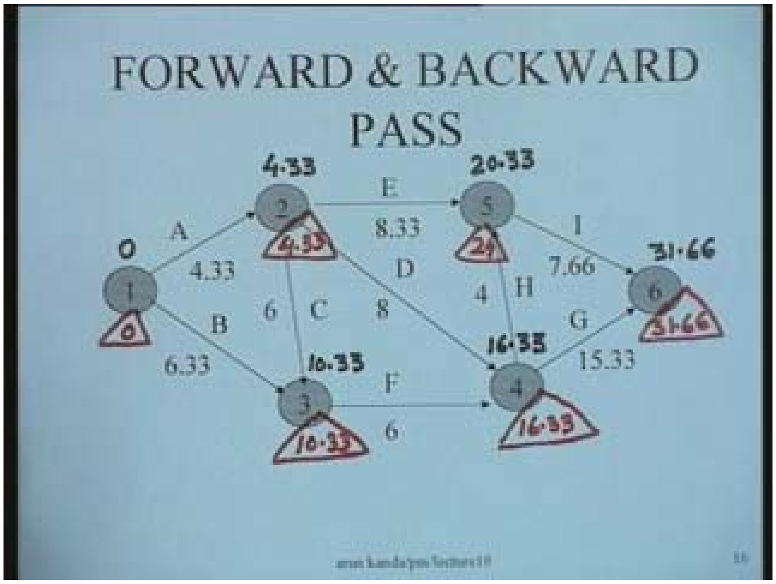
(Refer Slide Time: 29:24)



The project expected completion time is 31.66. That is one thing and in fact we can do a backward pass also. Let us do a backward pass. Let me use a different color for the backward pass. Let's say this value is 31.66. This is in fact the value of the latest occurrence time of node 6. Subtracting this value from node g we get 16.33 from here-, sorry We first have to compute the value of node 5. Because there are 2 outgoing arcs we cannot yet compute node 4. So 31.66-7.66 is going to be 24. Having computed 24 we will then have to say 24-4 which is 20 and 31.66-15.33 which is 16.33. We can take the value which is lower and the lower value in this case is 16.33 and then we come to node 3. 16.33-6 is going to be 10.33 for this particular node. We compute the latest occurrence time of this node and when you come to node 2 you will have to take the minimand over

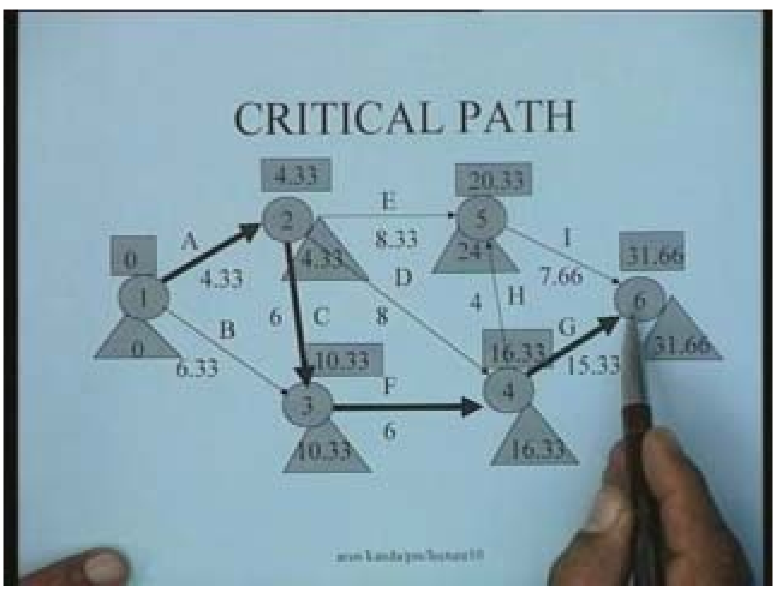
10.33-6 which is 4.33; 16.33-8 eight which is 8.33 and 24-8 which is going to be 16.33 and the least value is 4.33. This would be the value of latest occurrence time for this particular node and finally for this node you can see very easily that this value comes out to be one - but written as zero pl check. We carry out a forward pass and a backward pass in the usual manner and then we can identify exactly the critical path in the usual fashion.

(Refer Slide Time: 31:31)



For instance in the identification of the critical path for this example what you find is that the critical path happens to be 1, 2, 3, 4 and 6.

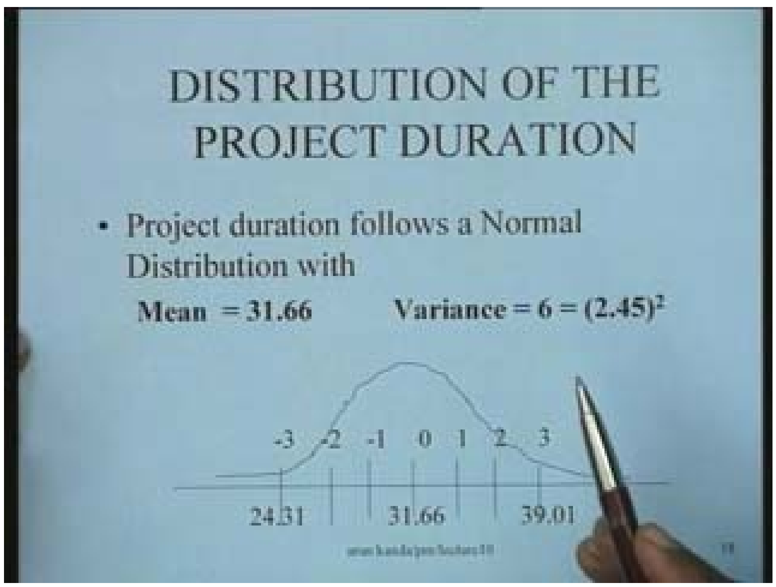
(Refer Slide Time: 31:43)



That's the critical path and the earliest occurrence times and the latest occurrence times are the values that we have just computed for each node through a forward pass and a backward pass respectively. The assumption that is made in PERT is that once you identify the critical path in this fashion the critical path has a mean duration of 31.66. Critical path is the sum of the durations of the activities lying on the critical path and what we do is we now take into account the variances of these individual activities A, C, F and G. We had computed their variances. The variance of a plus the variance of c plus the variance of f plus the variance of g is in fact equal to 6 in this particular situation from the table of variances that we have.

We can identify the distribution of the project duration. The project duration follows a normal distribution with a mean equal to 31.66 and a variance equal to 6. That means the standard deviation is 2.45 for this particular project.

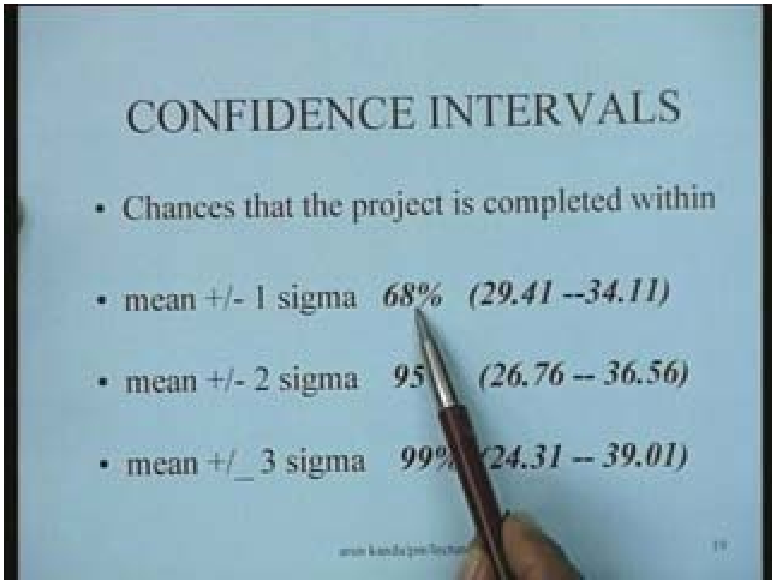
(Refer Slide Time: 33:07)



The normal distribution would typically be a distribution which is something like this and you notice here that the mean completion time of the project is something like 31.66 and we can make use of certain properties of the normal distribution. The normal distribution is contained almost entirely within +/- 3 sigma limits. The +/- 3 sigma means 3 into 2.45 +/- that gives you values of 24.31 to 39.01. We can in fact see that this particular project is with all probability likely to be completed in a minimum time of 24.3 days and a maximum time of about 39 days. We get this estimate of the worst case situation or you can say that there is more than 99% probability that the project would be complete within this interval. In fact the distribution of the project completion time, which was a normal distribution, has come out by virtue of the central limit theorem by adding up the durations of activities on the critical path and by ignoring all the non-critical activities we get this distribution. This distribution can answer a variety of questions, questions of managerial significance.

For instance we can determine the confidence levels. What do we mean by confidence level? That means you can find out the chances that the project is completed within ……. If you take mean +/- 1 sigma the chances are about 68% because within +/- 1 sigma about 68% of the area of the normal distribution lies. We can say that between 29.4 to 34.1 days, within this interval there are 68% chances that the project will be completed.

(Refer Slide Time: 35:24)

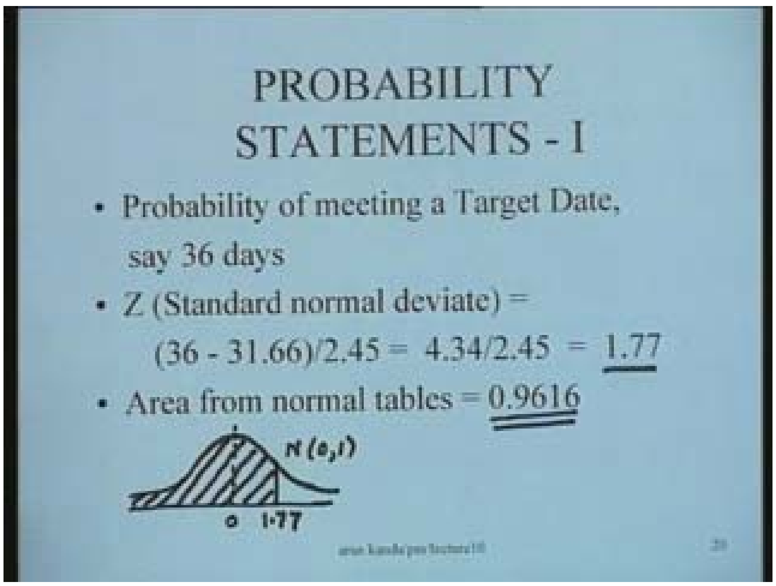


Similarly if we take mean +/- 2 sigma here we say that the chances are approximately 95% and the corresponding limits can be worked out in this example as something like 26.76 to 36.56. These limits are obviously wider than this. These are 2 sigma limits and if you look at the 3 sigma limits which we had seen before in the 3 sigma limits approximately more than 99.76% of the area lies within 3 sigma limits. But for practical purposes we can say that 99% of the area lies between 24.31and 39.01 days. It's quite clear that this particular project has a degree of uncertainty which can be captured by this information which says minimum of 24 days maximum of 39 days is what this project is likely to take based on the information pertaining to the uncertainties of the individual activities. This is one of the things which we can very easily say.

Apart from this analysis which is in fact very straight forward we can make some very interesting statements. Let us look at the statements. One kind of statement could be that what is the probability of meeting a target date? This is the kind of information that a project manager would be very interested, very much keen to find out. Let us say that the target date for our example is 36 days. It's like trying to say if you are working 6 days a week after 6 working weeks what is the probability that the project will be completed after 36 days? This is the point that I want to find out. What you can do is for this particular example you can compute the standard normal deviate which is nothing but x minus mu by sigma. x minus mu by sigma would be that I am interested in finding out for 36 days. So 36 minus the mean value of the distribution which is 31.66 divided by the standard deviation which is 2.45 and this particular value works out to 4.34 divided by 2.45 which is 1.77. Then you can refer to the standard normal tables.

What we are actually looking for is that the normal distribution is actually a distribution which is something like this. Standard normal table means it says normal distribution with mean zero and variance 1 and we have tables available which specify the area. So this is zero. This particular value of 1.77 that we computed let's say would be somewhere here. This is 1.77 and what we are interested in computing is basically this area under the curve to the left of this particular value of 1.77. On a standard normal variant we are interested in this area and this area from the tables actually works out 0.9616.

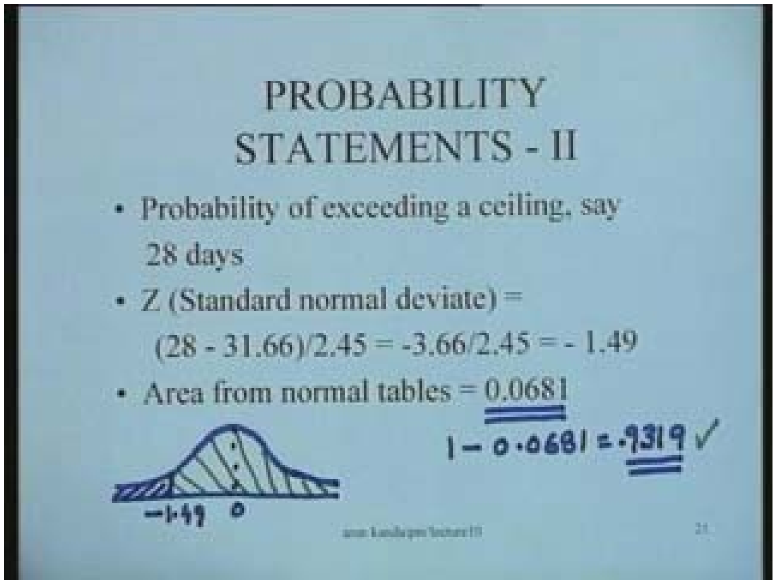
(Refer Slide Time: 39:04)



We can say that the chances of meeting the target date in this particular case with the kinds of uncertainties that you have is about 96% percent and you can be fairly confident that you can meet the target date and this would be the area for this particular problem.

For instance we try to work out the problem of exceeding a ceiling. Suppose I want to find out what is the probability that I exceed the ceiling of 28 days that I want to complete the project? Again I can follow the same procedure, compute the standard normal deviate which is x minus mu by sigma. x is less than the mean here. So 28-31.66 divided by 2.45 which is -3.66 divided by 2.45 and this is -1.49. Let's see the interpretation of this particular value on the normal curve. What we have in this particular situation is we have a standard normal distribution and this is the zero value. -1.49 would be somewhere here and we are actually looking for this area to the left of -1.49 which is barely 6.81% of the entire area. This is the probability that this is less than this. You can work out both. The area that we have worked out here is this particular area here. If you are interested in the probability of exceeding this value you are actually interested in 1-0.0681 and the value in this particular situation is going to be 0.9319. So 0.9319 would be in fact the probability that you exceed that particular value. That means it's this particular area which is the probability that you are actually interested in computing, the one that is shown in green. It's here. The area shown in blue here is 0.0681 but the area shown in green here is the probability of exceeding a ceiling. The probability that the project will take more than 28 days is in fact given by 93.19%.

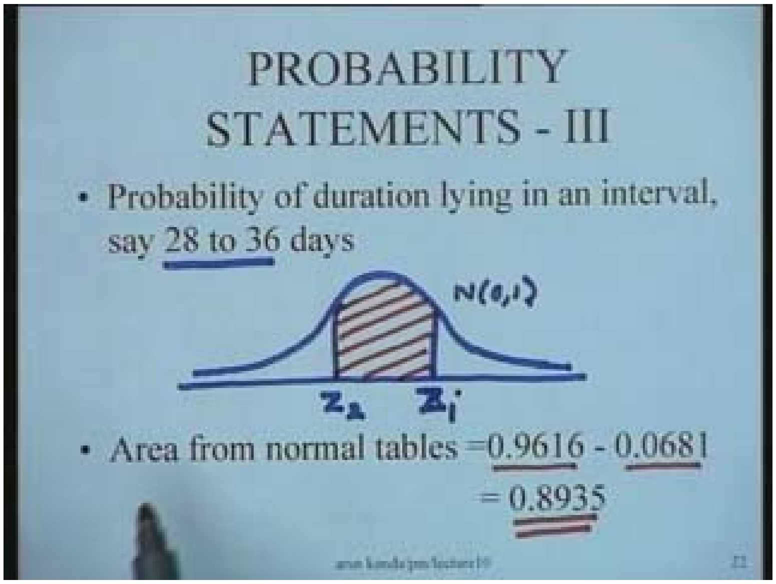
(Refer Slide Time: 41:50)



Depending upon the interpretation that you have, you have to suitably interpret the values that you obtain from the tables. This is the kind of computation that you can perform to find out the probability of exceeding a ceiling rather than talking about this particular value.

On the same lines, suppose our objective was to find out the probability that the project would be completed within a certain interval of time. The standard normal distribution would be something like this and we are interested in finding out the probability that the project would be completed in between 28 to 36 days. Again you would compute the Z value for each one of these. 28 we had already computed and 36 we had already computed. For each case you would have got a value of the areas which are there. Let us say the 36 days would be here. For the 36 days you would calculate the corresponding Z. So we write Z_1 here and we write Z_2 here and we would have got the areas. If you take the difference between the two you are then talking about this particular probability. This is the probability that is exactly the probability that the project would be accomplished between the time Z_1 and Z_2. That is in between these two times when you are talking about an interval from 28 to 36. From this case the corresponding areas for Z_1 which is the same value that we computed earlier and for Z_2 which is 0.06. Ultimately we get 0.8935 as the value of the probability corresponding to this.

(Refer Slide Time: 43:55)

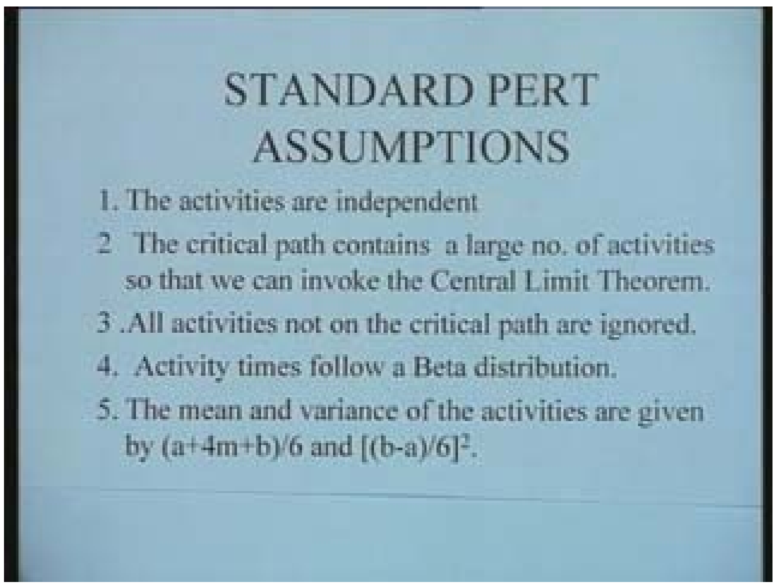


The point that really emerges is that we can make 3 different kinds of probability statements of managerial interest. We can talk about the probability of meeting a target date, we can talk about probability of exceeding a ceiling or taking more time than a certain date and we can find out the probability that the project is accomplished within a certain time interval. All these three kinds of computations can be made very conveniently by using the distribution of the project completion time and in fact this is the greatest accomplishment of PERT. PERT analysis is basically useful only because you can make these kinds of probability statements about the project completion time.

Now let us look at the analysis that we have done. What we have done so far is referred to as the basic PERT analysis and this basic PERT analysis actually suffers from a variety of errors and these errors are there primarily because of the assumptions which have been made in the whole PERT analysis. We will now try to do an evaluation of PERT. based on what kinds of analysis what kind of First thing is let's see the assumptions that are made in PERT. It's very important to understand these assumptions if you want to do any serious work on trying to refine or improve the basic PERT procedure. The activities are assumed to be independent. This is one thing so that the activity times corresponding to each activity are independent random variables. This may not be so. The activities could be co-related. That means if there are conditions favorable in real life they could be favorable for all activities or they could be favorable for some subset of activities which means that there is a co-relation. between various.

For instance the market is favorable. It's quite likely that all the activities pertaining to sales promotion and so on will go on faster. There could be co-relation between various activities but nevertheless for analysis we assume that the activities are independent. Second and the most crucial assumption which is made here is that the critical path contains a large number of activities so that we can invoke the Central Limit Theorem. I will explain this because normally in practice for the central limit theorem to be valid the number of random variables should be greater than at least 4 or 5. In real life the number of activities in the project could be very many, generally much more than 4 or 5 and therefore this assumption is generally valid. Another interesting thing is that an assumption that has been made in PERT is all activities not on the critical path are ignored. At no stage did we consider those activities which were not occurring on the critical path and yet because of the randomness of activity times it can happen that a path which is near critical or not critical at the moment could become critical at a subsequent time because the activities could stretch. Activity times are assumed to follow a beta distribution. This we have discussed and the mean and the variance of the activities are given by a + 4m +b by 6 and b-a by six whole square. This again is an assumption or an approximation from the beta distribution.

(Refer Slide Time: 47:43)

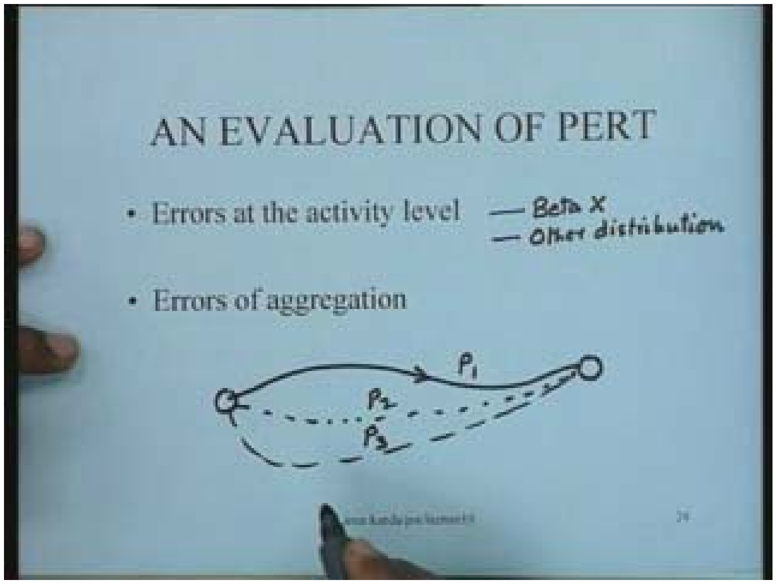


These are the 5 major standard assumptions which are made and all the errors in PERT analysis occur because of these assumptions.

Now let us try to do an evaluation of PERT and what I would like to say here is that errors in PERT can occur primarily because of these two factors. Errors can occur at the activity level. What this means is that we have made some assumptions about the activity distributions. We have said that the activities follow a beta distribution. They may not be following a beta distribution. This might not be valid. It might be following some other distribution which we have probably not even identified. This could be some other distribution. As a consequence of the distribution of the activity times you could have certain errors and by virtue of the fact that you assume them to be beta there could be some errors here.

Then there are errors of aggregation. These are perhaps much more serious. What happens in errors of aggregation is that we assume essentially that there is a critical path and the critical path is the path which runs from the source to the terminal. We identify this critical path and we sort of forget about the other paths which are going from the source to the sink. There could be many paths in the project. Isn't it? This is path P_1. This could be path P_2 and this could be path P_3.

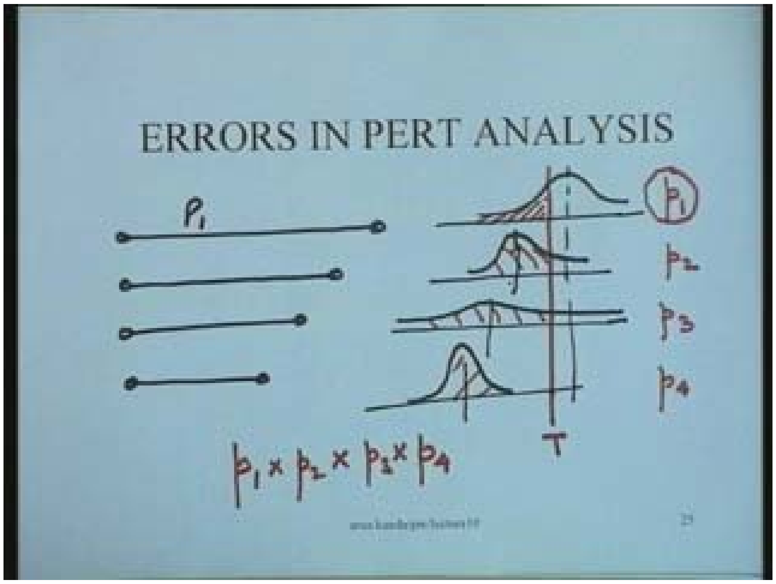
(Refer Slide Time: 49:30)



If we ignore all these paths then things are simple because what happens is by virtue of the Central Limit Theorem each path will be normally distributed with a distribution with a mean equal to the sum of the means and the variance equal to the sum of the variances. If we look at these various paths what would these paths' distribution look like? For instance they would look like this. You have the critical path which is P_1. You have a path which is slightly less than this. You have other paths. You have other paths. You have the paths in the descending order of their lengths and these are all paths going from source to sink. Critical path is the longest path. What we are simply trying to say is that this distribution will be a normal distribution with mean equal to this. This is the mean of the critical path. This particular path will have a smaller mean, may be it has this mean. It could still have a normal distribution which will probably be something like this but its mean is smaller. This particular distribution will have a path whose mean is probably still smaller and it would have a distribution which could probably be very dispersed like this. This particular path could also have a distribution which will be something like this. It would not be important for us to take it in.

By using the PERT analysis that we have computed we say what is the probability that the project duration will be equal to some particular value? Let us say what is the project duration at less than this particular value? I mark this particular line here. What is the probability that the project duration is equal to this time t? In the standard PERT analysis all that we say is that this distribution is equal to the probability p_1 which we compute by a standard normal distribution which is given by this area. If you take this path corresponding to this path also there would be a certain probability. This would be p_2 which is actually this area. This would be p_3 and this entire area, whatever is the entire area it's p_4. If we assume to begin that all these paths are independent then the total probability is going to be p_1 into p_2 into p_3 into p_4 and since the probabilities are less than 1 this probability is the probability that all these paths are of length less than or equal to t. That means the probability that the project duration is t or less is given by this probability and not by this probability which actually means that this probability will be smaller than this probability which goes to prove that the estimate that you obtain by PERT are going to be optimistic. It's going to give you a higher probability.

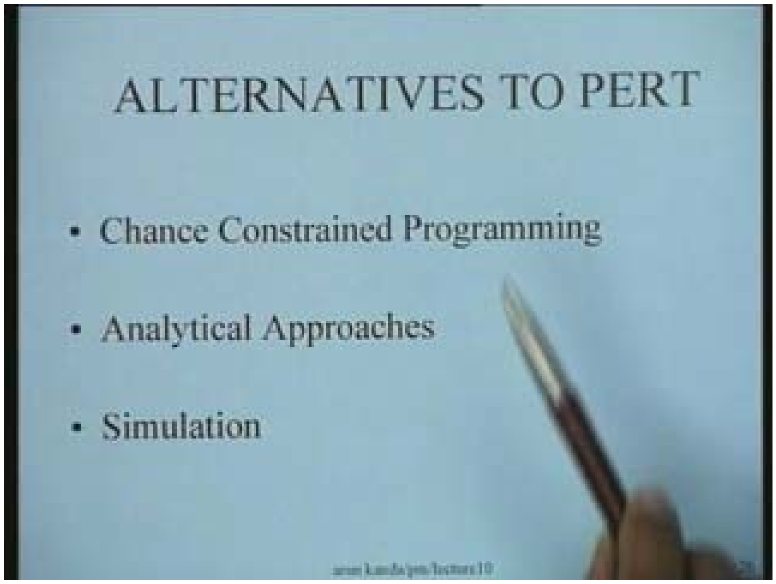
(Refer Slide Time: 52:47)



It's going to be basically higher probability means that it's going to give you more optimistic estimates. It's giving you higher probabilities. Actually the probability is lower just as I mentioned here but then these paths are not independent. There would be some co-relations coming in which would make this but this is like a lower bound of the probability. The real probability would be between these and these values in real life situation. What are the alternatives to PERT?

One approach for dealing with this kind of situation is chance constrained programming where we develop a model based on constraining the chances that the particular event times talk about certain ……….

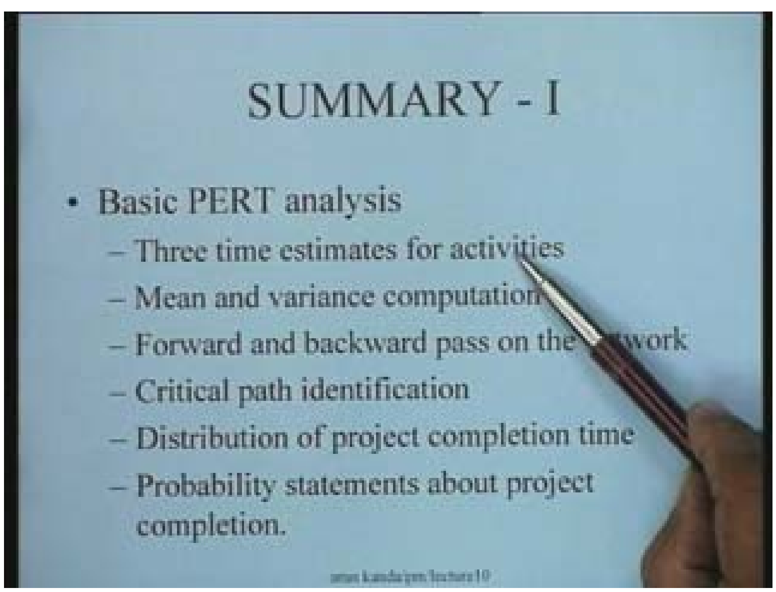
(Refer Slide Time: 53:31)



Then there are analytical approaches which we can explore and finally simulation is a very versatile and a powerful approach which can deal with this entire option. We shall be exploring some of these cases, some of these alternatives in our subsequent lectures.

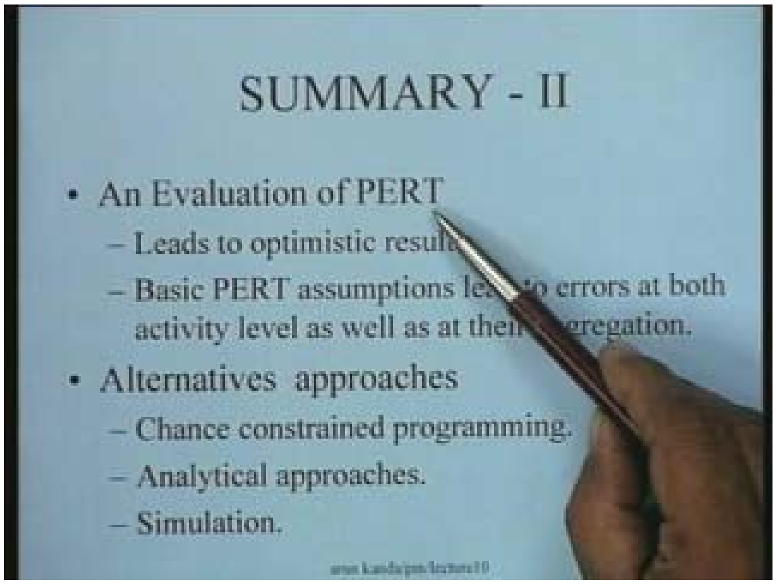
To summarize our discussion, in today's lecture we have tried to look at basic PERT analysis. We have looked at the 3 time estimates for activities.

(Refer Slide Time: 53:57)



We have looked at how to calculate the mean and the variances? We have seen how to do a forward and a backward pass on the network. We have done the critical path identification. We have tried to determine the distribution of the project completion time and make probability statements about the project completion. Then we have tried to do an evaluation of PERT.

(Refer Slide Time: 54:21)



We have seen that PERT leads to optimistic results. Basic PERT assumptions lead to errors at both activity level as well as their aggregation and finally the alternatives to PERT are chance constrained programming, analytical approaches and simulation and we will try to look at some of these techniques in some of our future lectures. Thank you!