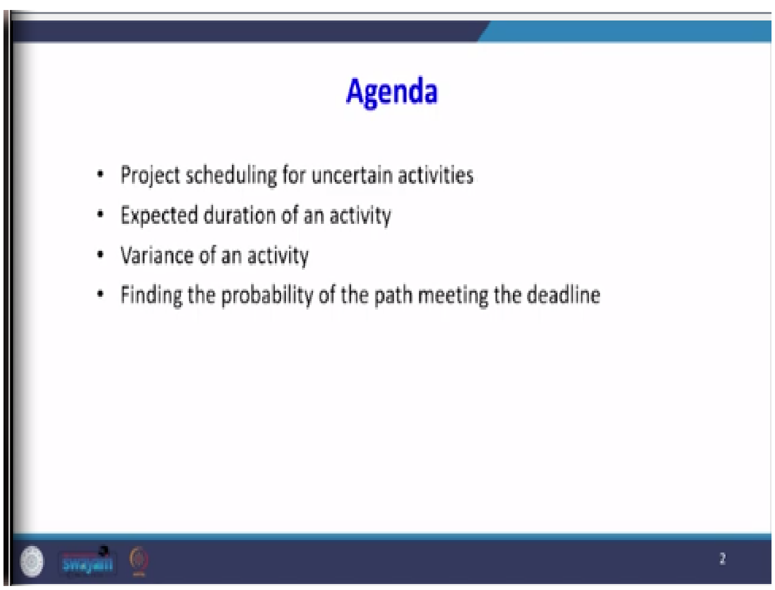
**Decision Making With Spreadsheet**
**Prof. Ramesh Anbanandam**
**Department of Management Studies**
**Indian Institute of Technology-Roorkee**

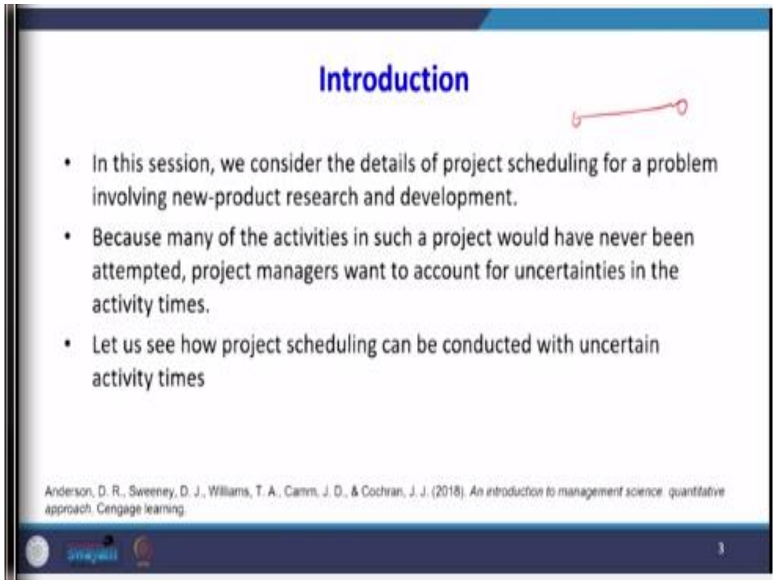
**Lecture-31**
**Project Scheduling for Uncertain Activity Duration: PERT/CPM-III**

Dear students, in the previous class, we estimated the earliest starting time, earliest finishing time, and the latest starting time and latest finishing time for an activity, which has a one-time estimate. What is the meaning of a one-time estimate? Only one time is given; that is when you say activity A, for example, 9 weeks, it is precisely 9 weeks. But in this lecture, we are going to find the schedule of the whole project where the 3-time estimate is given. What is the 3-time estimate?

I will explain in the lecture another way: instead of giving a 3-time estimate, we can say the activity duration is uncertain. Whenever the activity duration is uncertain there are different ways to find out the expected duration of an activity. So, I will cover these details in this lecture.

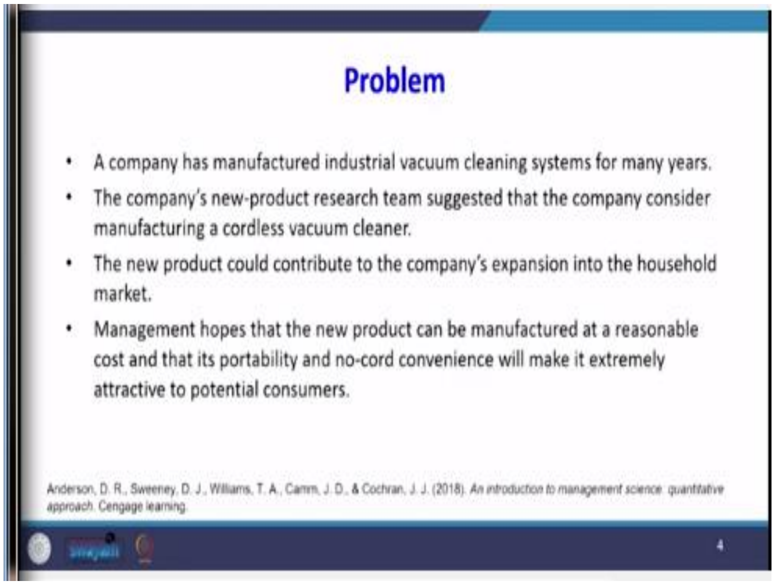



So, the agenda for this lecture is project scheduling for uncertain activities, then finding the expected duration of an activity, then finding the variance of an activity, and at last, finding the probability of the path meeting the deadline. So, these topics will be covered in this lecture.

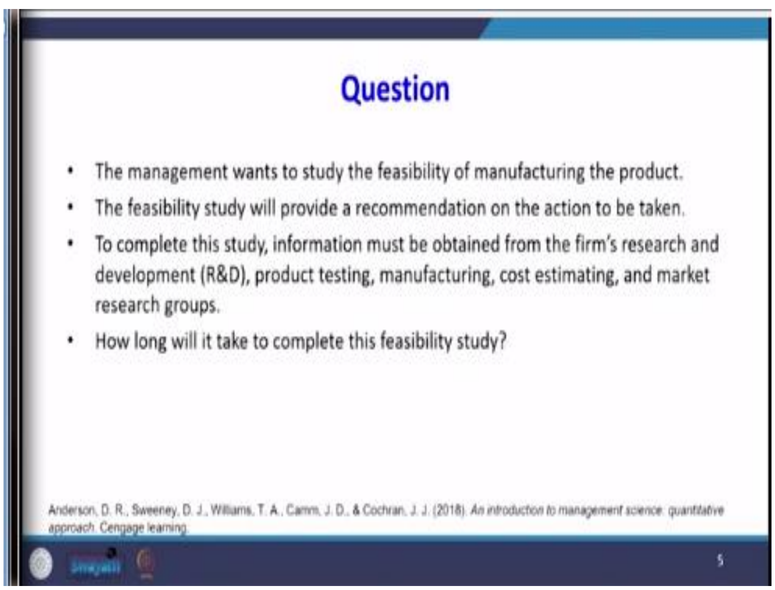



The reference for this lecture is from the book by Anderson et al. So, in this session, we consider the details of the project scheduling for a problem involving new product research and development. Whenever the project is something innovative, something new, the project activity time cannot be estimated perfectly; that is the reason this example has been taken for this session. So, here is what we are going to do.

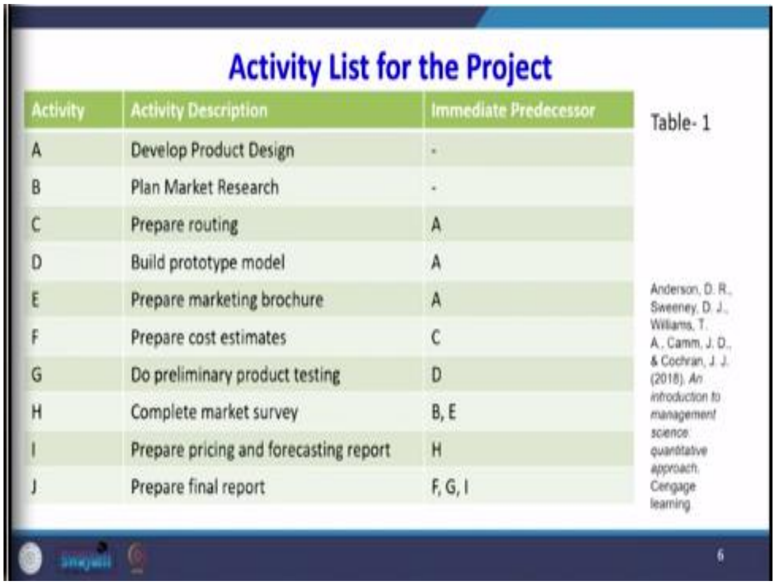
There will be an activity; there will be a 3-time estimate that will be given: optimistic time, most likely time, and pessimistic time, which will be covered in this lecture. Because many activities in such a project would have never been attempted, the project manager wants to account for uncertainties in the activity times. Let us see how project scheduling can be conducted with the uncertain activity times.




The problem that I have taken for this session is a company that has manufactured industrial vacuum cleaning systems for many years. The company's new product research team suggested that the company consider manufacturing a cordless vacuum cleaner without wire. The new product could contribute to the company's expansion into the household market. Management hopes that the new product can be manufactured at a reasonable cost, and that its portability and no-cord convenience will make it extremely attractive to potential consumers.




So, what questions is management asking? The management wants to study the feasibility of manufacturing the product, and for that, they are going for a feasibility study. The feasibility study will provide a recommendation on the action to be taken. To complete this activity, information must be obtained from the firm's research and development department, product testing department, manufacturing and cost estimating, and market research groups. The question that needs to be answered is how long will it take to complete this feasibility study.



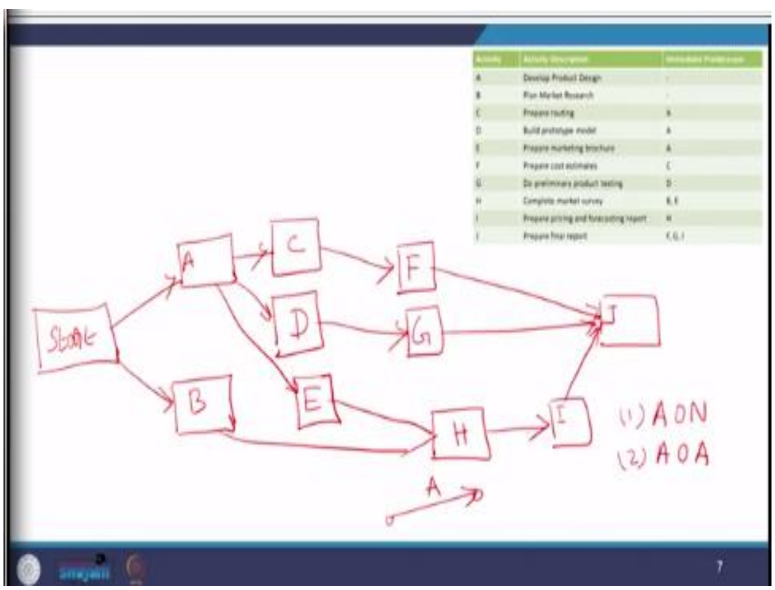
## Activity List for the Project

| Activity | Activity Description | Immediate Predecessor |
|---|---|---|
| A | Develop Product Design | - |
| B | Plan Market Research | - |
| C | Prepare routing | A |
| D | Build prototype model | A |
| E | Prepare marketing brochure | A |
| F | Prepare cost estimates | C |
| G | Do preliminary product testing | D |
| H | Complete market survey | B, E |
| I | Prepare pricing and forecasting report | H |
| J | Prepare final report | F, G, I |

Table- 1

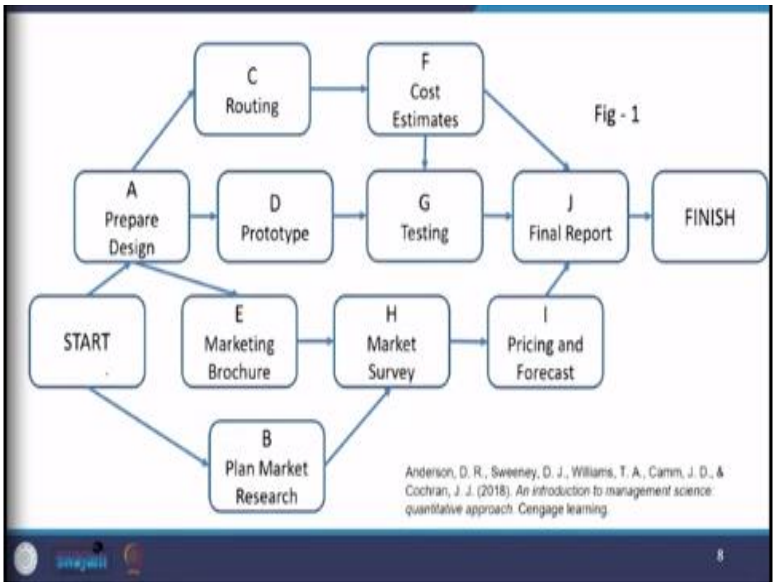
Anderson, D. R., Sweeney, D. J., Williams, T. A., Camm, J. D., & Cochran, J. J. (2018). *An introduction to management science: quantitative approach*. Cengage learning

So, to complete the feasibility study, the activity list is given activities A, B, C, D up to J 1, 2, 3, 4, 5, 6, 7, 8, 9, there is a 10 activity. Activity A and B do not have any precedence; C, D, and E have precedence A like there are other activities that have preceding activities. So, the precedence constraint needs to be considered while drawing the network.

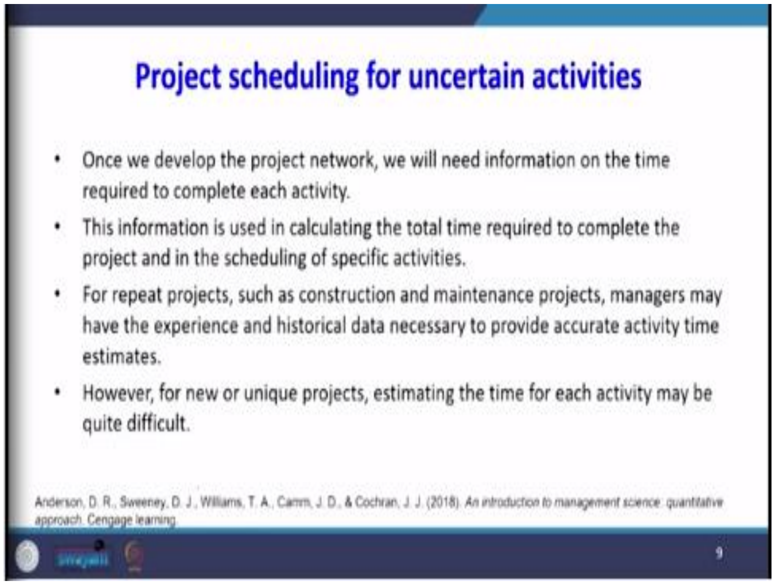



So, what am I going to do? I am going to draw the network for this. So, for activity, I will start from here start, activity A and B do not have any predecessor, so I am writing this is A, this is B for activities C and D, and E also the predecessor is A, so for C, for D, similarly for E, for activity F predecessor is C, for activity G predecessor is D, for activity H predecessor is B and E, so for activity I predecessor is, for activity J the predecessor is F, G and I.

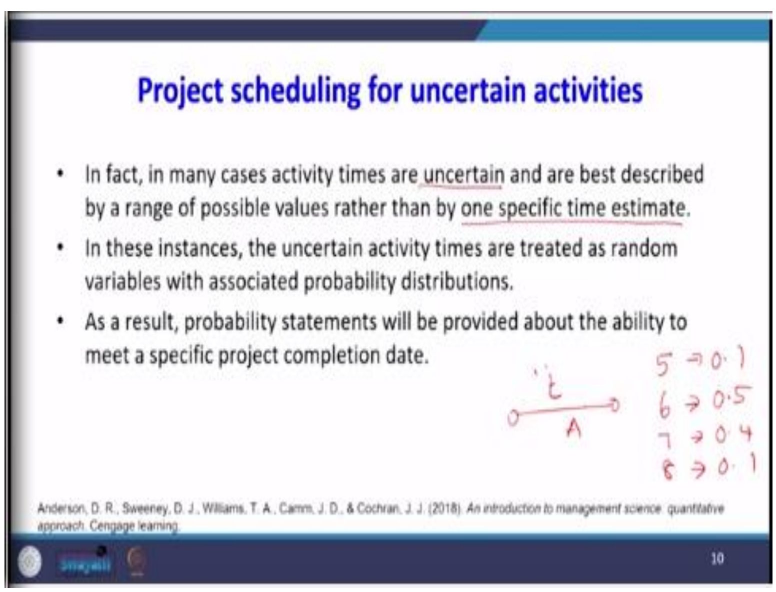
There are different ways of drawing the network; one method is called activity on the node, and another method is called activity on the arc. Here is the activity I have written in node; many software packages follow this activity or node. The way I have drawn this activity I have mentioned on the node. There are some networks in which we will mention activity, for example, activity A, which is activity on arc, but we are following activity on node. So, this is the way there should be an arrow; this is the network by considering all the precedence constraints.




So, I have brought it, so we are starting from here; I have written all the activities, and finally, there is a finish. Start and finish only for my understanding, the start and finish are not the activities; A and B are the first activities.



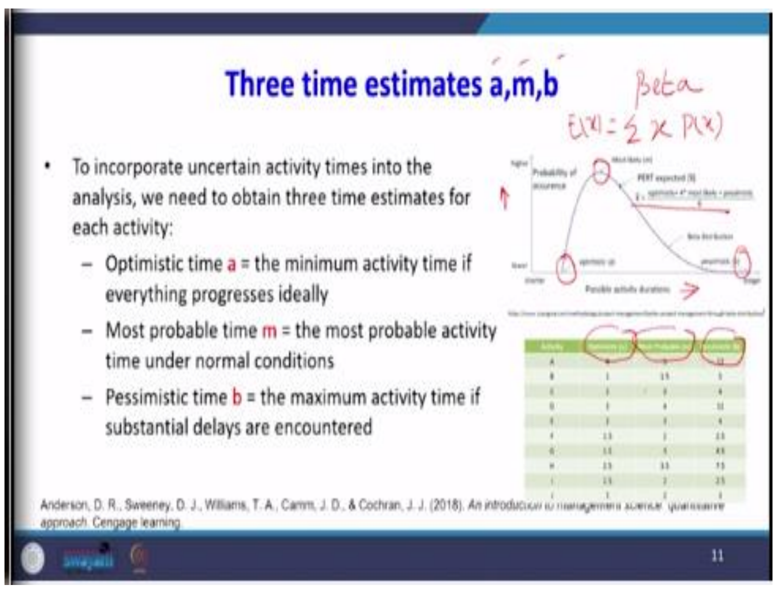

Now, regarding the project scheduling for uncertain activities, once we develop the project network, as I have discussed in the previous slide, we will need information on the time required to complete each activity. This information is used to calculate the total time required to complete the project and schedule specific activities. For repeat projects, such as construction and maintenance projects, the manager may have the experience and historical data necessary to provide accurate activity time estimates. However, for a new or unique project estimating the time for each activity may be quite tricky.




In fact, in many cases, activity times are uncertain and are best described by a range of possible values rather than by one specific time estimate. In the previous lecture, we found the schedule of each activity by having one specific time estimate, but in this lecture, we are going to estimate the project schedule by considering an activity which is having a range of possible values.

So, in these instances, the uncertain activity times are treated as random variables with associated probability distributions. For example, if this is an activity, the activity completion time may follow any distribution. So, the probability of completing, say, 5 days maybe 0.1, and the probability of completing the same activity for completing 6 days maybe 0.5. So, the probability of completing this activity is 0.4, so 8 completing this activity is 0.1. So, what is happening?

This time estimate t follows an empirical distribution; it can follow any distribution, this activity duration time may follow a normal distribution, or it can follow a uniform distribution, or it can follow a beta distribution. That is what we are going to discuss in this class. So, in this class, we are going to assume that the activity completion time follows beta distribution.

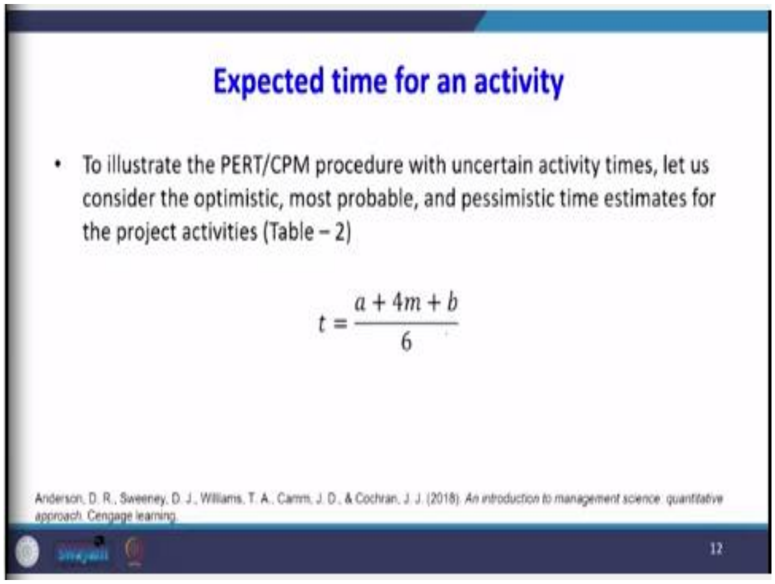



So, now we are going to discuss the 3-time estimate for an activity. So, we are going to call it a, m, b. To incorporate uncertain activity times into the analysis, we need to obtain 3-time estimates for each activity. So, the first time estimate is called optimistic time a, what is the minimum activity time if everything progresses ideally, if there is no interruption for completing, so the optimistic time is a minimum activity time. The second one is the most probable time m, which is the most probable activity time under normal conditions.

The third time estimate is the pessimistic time, which is the maximum activity time if substantial delays are encountered. So, we are assuming this 3-time estimate follows a distribution called the beta distribution. What have I written here? This picture is beta distribution time; look at this: the minimum time is optimistic time a, and look at the b, right; the longest time is called pessimistic time. So, in between, there will be a time called most likely time. So, on the x-axis, there is a possible activity duration, and on the y-axis, probability.

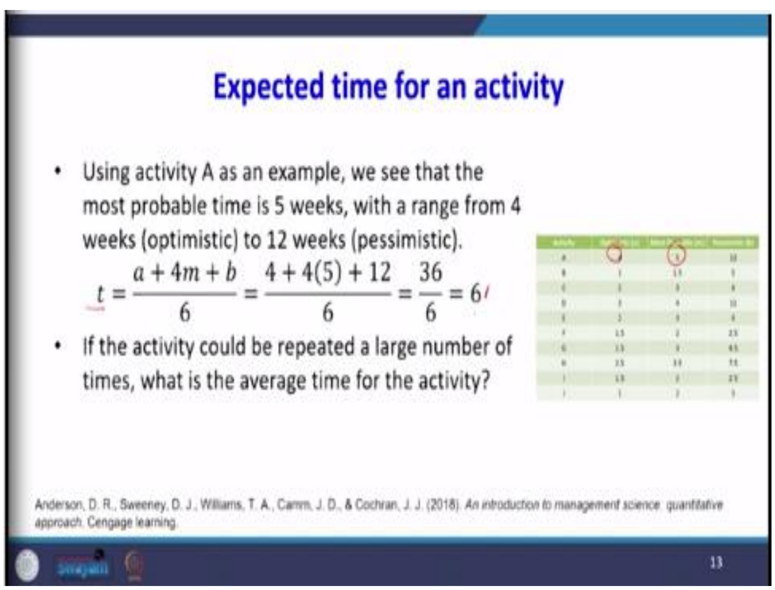
So, the mean of this beta distribution is optimistic time + 4 times, most likely time + pessimistic time divided by 6. So, this mean formula for beta distribution, if the project completion time follows, say, uniform distribution, then the mean will be (a + b)/2. If it follows a normal distribution, the mean will be the mean itself. So, since we assumed that this 3-time estimate follows beta distribution, the formula for mean is this one optimistic time plus 4 times most likely time + pessimistic time divided by 6. Instead of this, if the activity completion time follows simply an empirical distribution.

So, for that, how will you find the meantime? So, E of x = summation of x into p of x, x is the project completion time, and p of x is the probability. So, you can see what we saw in the previous slide I wrote here. So, here, the mean is five multiplied by 0.1 + 6 multiplied by 0.5 + 7 multiplied by 0.4 + 8 multiplied by 0.1; this is the way to get the mean of an empirical distribution. You see that for each activity, we have the optimistic time, pessimistic time, and most likely time. Most probable time, the optimistic time is the minimum time, the pessimistic time is the maximum time, and the most probable time is in between.



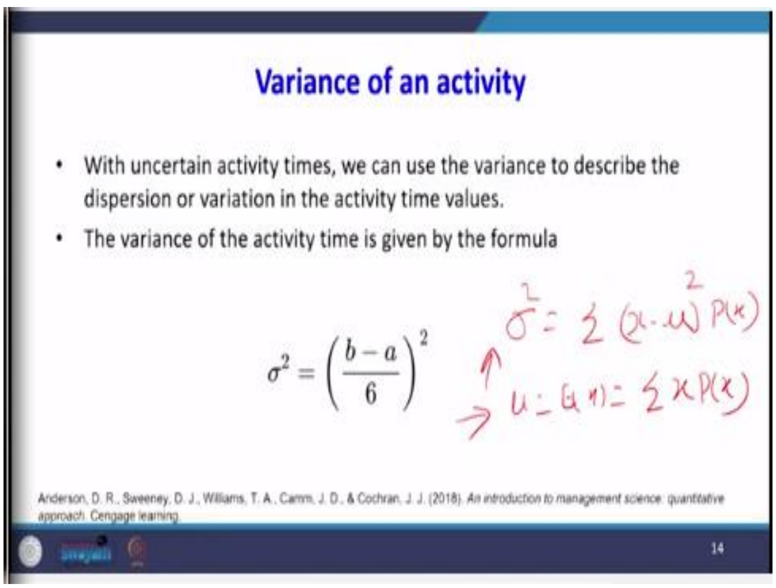

So, how do we find out the mean or expected time for an activity? To illustrate the PERT/CPM procedure with uncertain activity times, let us consider the optimistic, most probable, and pessimistic time estimates for the project activities, which are given in the previous table. So, the formula for finding the mean of a beta distribution is optimistic time + 4 times most probable time + pessimistic time upon 6.

$$t = \frac{a + 4m + b}{6}$$



Now, I will tell you how to find out the expected time for each activity, for example, activity A. For activity A, we see that the most probable time is 5 weeks, with a range from 4 optimistic times + to 12 weeks of pessimistic time. So, the formula for finding the mean is (a + 4 m + b)/ 6 is a 4+, so this four comes from the formula four multiplied by 5 + 12 upon 6. So, what is the meaning of this 6? If the activity could be repeated a large number of times, the average time for the activity is nothing but 6, the average time is 6.

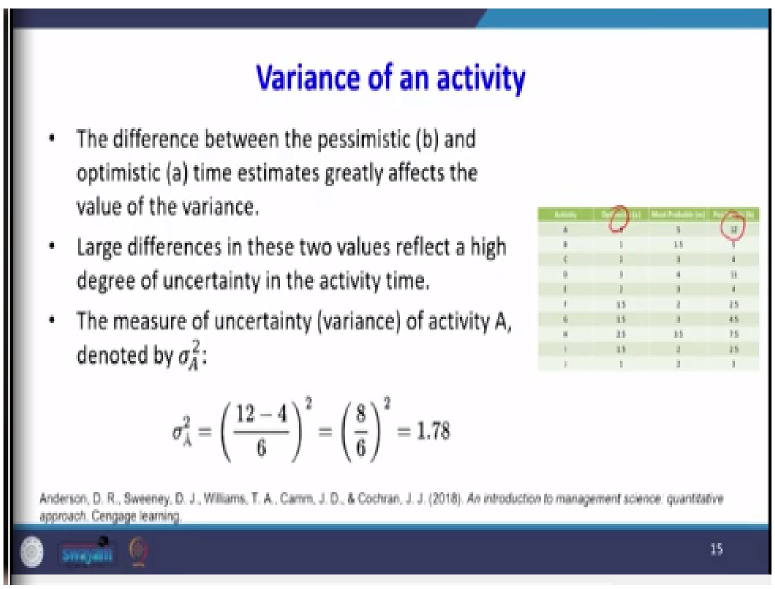
$$t = \frac{a + 4m + b}{6} = \frac{4 + 4(5) + 12}{6} = \frac{36}{6} = 6$$




Once we know the average then we have to find out the variance of that activity also. So, when uncertain activity times, we can use the variance to describe the dispersion or variation in the activity time values. So, the variance of the activity time is given by the formula for beta distribution, which is that the sigma square is equal to b – an optimistic time upon six whole squares. If it is an empirical distribution, what is the variance formula?

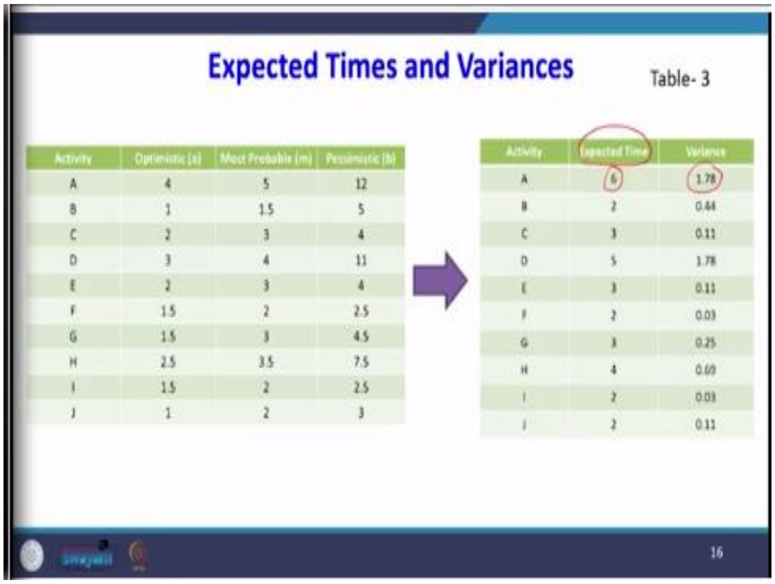
$$\sigma^2 = \left(\frac{b-a}{6}\right)^2$$

Variance is the sigma of x - mu whole square p of x; here, mu is the mean. So, before finding the variance, you should know the mean value, I already told you the mean is E of x and sigma of x into p of x. First, you have to find the mu, and then you have to find the variance if the activity completion time follows empirical distribution. But here, we assumed that the activity time follows the beta distribution. So, the variance formula for beta distribution is b - a upon six whole squares.

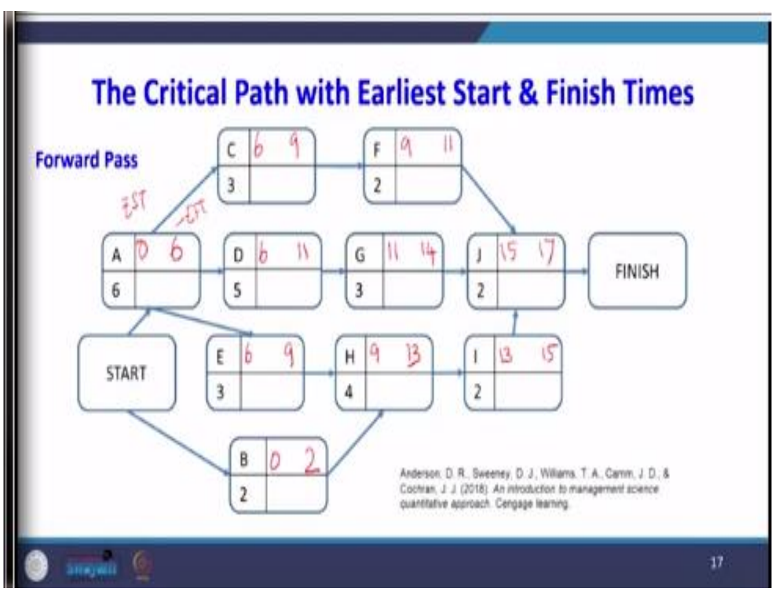



So, the difference between the pessimistic and optimistic time estimates greatly affects the value of the variance. So, a large difference in these 2 values reflects a higher degree of uncertainty in the activity time. So, the measure of uncertainty that is the variance of an activity A denoted by sigma A square is nothing but 12 - this 12 - 4 whole square 4 upon 6 whole square, so 8 upon 6 whole square this is 1.78 if you want to know the standard deviation you have to take the square root of this.

$$\sigma_{Å}^{2} = \left(\frac{12-4}{6}\right)^{2} = \left(\frac{8}{6}\right)^{2} = 1.78$$

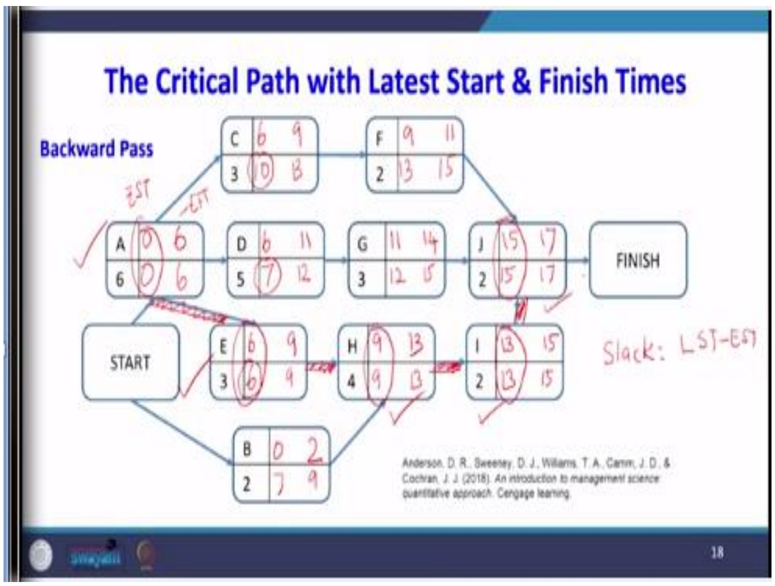



So, what have we done? By using the mean and variance formula, so this 3-time estimate is converted into a single-time estimate for each activity. For example, for activity A, we have the 3 time estimates 4, 5, and 12 that are converted into a single time estimate of 6; for that activity, we found a variance of 1.78. Like that I have found the expected time and variance for all the activities.



Now, we will find the earliest start and earliest finishing time using the forward pass method that I discussed in the previous lecture. So, we start from 0 earliest starting time is 0, so 6, so here 6 + 3, 9, so 6 + 5 = 11, so here 6 + 3 = 9, so here 0 + 2 = 2, here 9 + 2 = 11, here 11 + 3 = 14. Here you see there are 2 ways to reach activity H 9 and 2, so we have to consider the largest value between 9 and 2, so 9 + 4 is 13, so this 13 + 2 is 15.

Now you see, for activity J, there are 3 ways we can come from activities F, G, and I. Which is the largest one? The largest one is 15, so 15 + 2 is 17. So, for this project, the total time taken to complete all the activities is 17 weeks. As we know, this represents the earliest starting time, and this represents the earliest finishing time.




Now, using the backward pass technique, I am going to find out the latest starting and finishing times. So, I start from the last activity, which is J, so the latest finishing time is 17, so the latest starting time will be 17 - 2, 15. So, here it is a 15, 15 - 2 is 13, so 13, 13 - 4 is 9, so it is 9, 9 - 3 is 6, so here it is 9, 9 - 2 is 7. Now we will come to activity J, which is 15, 15 - 3 is 12, so here for F, it is 15, 15 - 2 is 13, so here also 13, 13 - 3 is 10, so for activity D, 12, 12 - 5 is 7.

Now I have to write the latest finishing time for activity A. So, what are the succeeding activities C, D, and E? So, here, the earliest starting time is 10, the earliest starting time is 7, and the earliest starting time is 6, so what do I have to do? I must find out the minimum earliest starting time so that the minimum earliest starting time will be the latest finishing time of the preceding activity. So, the 6 will be the latest finishing time, so 6 - 6 is 0.

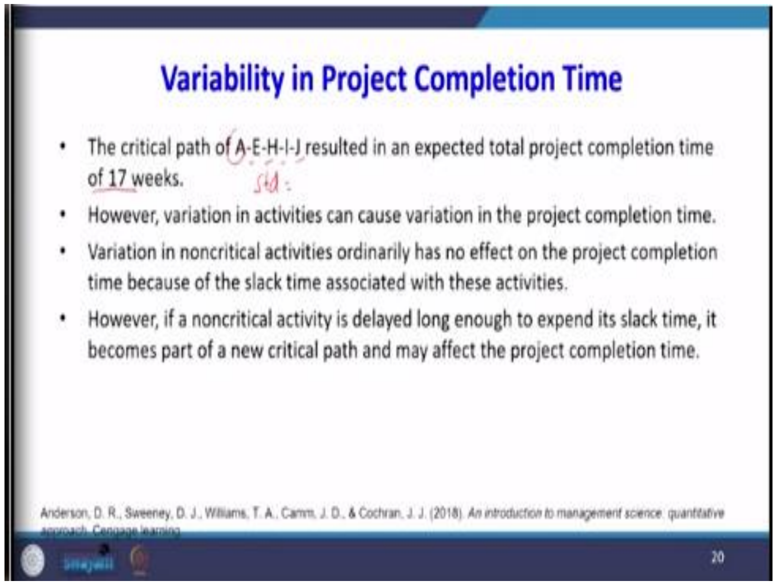
Now we have to find out the critical path, so wherever the flow to 0, that path is the critical path. What is the formula for a float? Float means slack the previous class, that is latest starting time - earliest starting time. So, for this, when you look at this, here it is 15, so here slack is 0, here also it is 0, 9, 9 0, 6, 6 is 0, here also 0, so the critical path is this one, this, this, this. So, using different colors, this, this, this, this.

So, all the activities in the critical path are called critical activities because the float or the slack is 0, which means these activities cannot be delayed at all. If we delay that activity the project duration will be extended to 17 weeks.

Table- 4

| Activity | ES | LS | EF | LF | Slack (LS-ES) | Critical Path? |
|---|---|---|---|---|---|---|
| A | 0 | 0 | 6 | 6 | 0 | YES |
| B | 0 | 7 | 2 | 9 | 7 | |
| C | 6 | 10 | 9 | 13 | 4 | |
| D | 6 | 7 | 11 | 12 | 1 | |
| E | 6 | 6 | 9 | 9 | 0 | YES |
| F | 9 | 13 | 11 | 15 | 4 | |
| G | 11 | 12 | 14 | 15 | 1 | |
| H | 9 | 9 | 13 | 13 | 0 | YES |
| I | 13 | 13 | 15 | 15 | 0 | YES |
| J | 15 | 15 | 17 | 17 | 0 | YES |

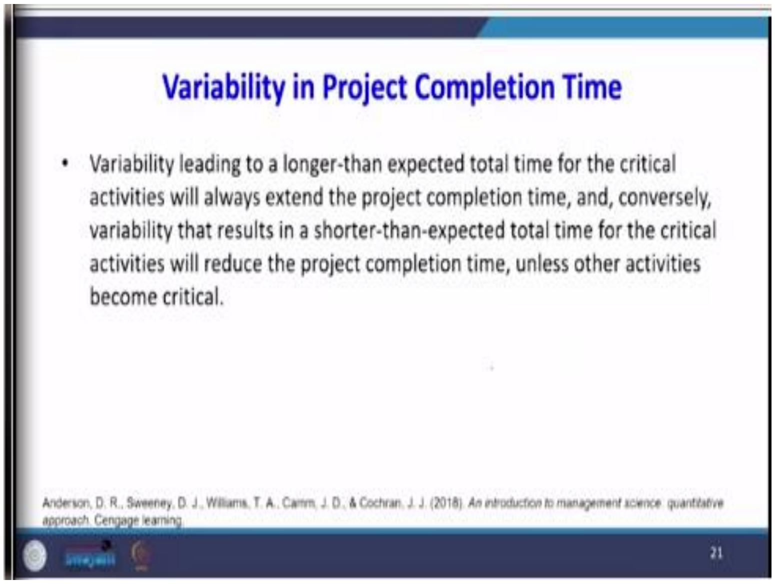
Now what are the critical activities? Activity A I go back yes, this is activity A second one, so when you find float activity E yes, activity E, the next one is activity H will go back activity H, next one is I and J, so I and J, these are the critical activities. Apart from this critical path and critical activities, we have the schedule for the earliest starting time and latest starting time, earliest finishing time, and latest finishing time.



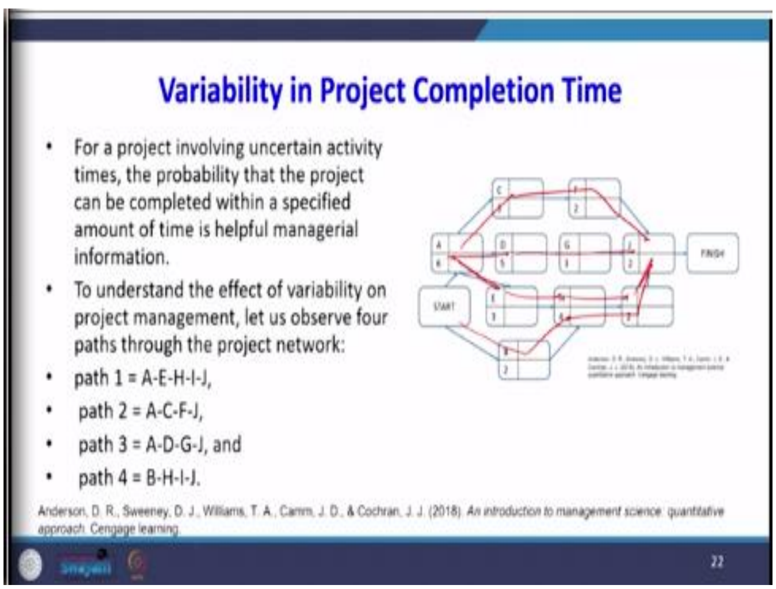

So far, we have found the expected total completion time. Now, we are going to consider another element apart from the meaning. We are going to find out the variability in the project completion time. So, the critical path A-E-H-I-J resulted in an expected total project completion time of 17 weeks. So, the mean completion time is 17 weeks but we have to supply the variance also apart from the 17.

However, the variation in activities can cause variation in the project completion time; the 17 is there, so what do we have to do? We have to find out the variance, we have to sum the variance of all the activities which are in the critical path. That is, we have to find out the variance of activity A, E, H, I, J, we have to sum the variance, then we have to take the square root of that variance, so that will be the standard deviation of that critical path.

So, variation in noncritical activities ordinarily has no effect on the project completion time because of the slack time associated with these activities. Because there will be a positive slack time, that will not affect our project completion time. However, if a noncritical activity is delayed long enough to expand, it is slack time. It becomes part of a new critical path and may affect the project completion time. But generally, we are concerned about the activities which are in the critical path.

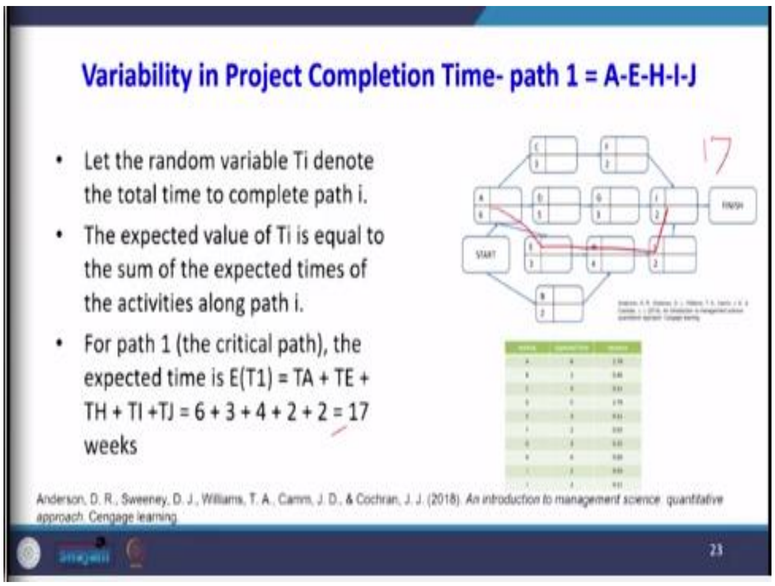



So, the variability leading to a longer-than-expected total time for critical activities will always extend the project completion time. Conversely, variability that results in a shorter than expected total time for the critical activities will reduce the project completion time unless other activities become critical. So, what is the point here is whenever we provide the mean total completion time, we have to provide the variance of that activity also.

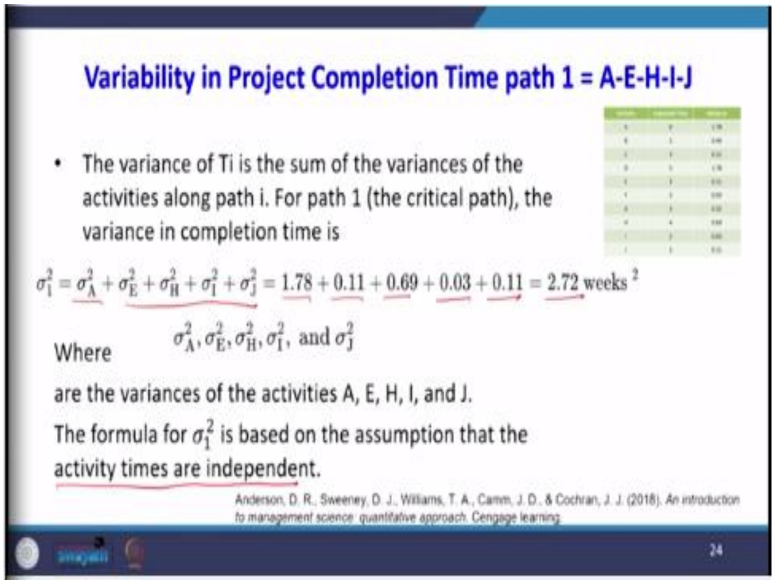



Now we are going to find out the variability in the project completion time. For a project involving uncertain activity times the probability that the project can be completed within the specified amount of time is helpful for managerial information. To understand the effect of variability on project management let us observe 4 paths through the project network. For example path 1 A-E, this one A-E-H-I-J, this is one path.

Here is another path, A-C-F-J; the third path we can consider A-D-G-J; the fourth path we can consider B-H starting from here, B-H-I-J. So, we have considered 4 paths, and there will be a different total completion time for all these 4 paths. So, what are we going to do? We are going to find out the total completion time for all these 4 paths and we are going to find out the variance of these each path, that is what we are going to do.

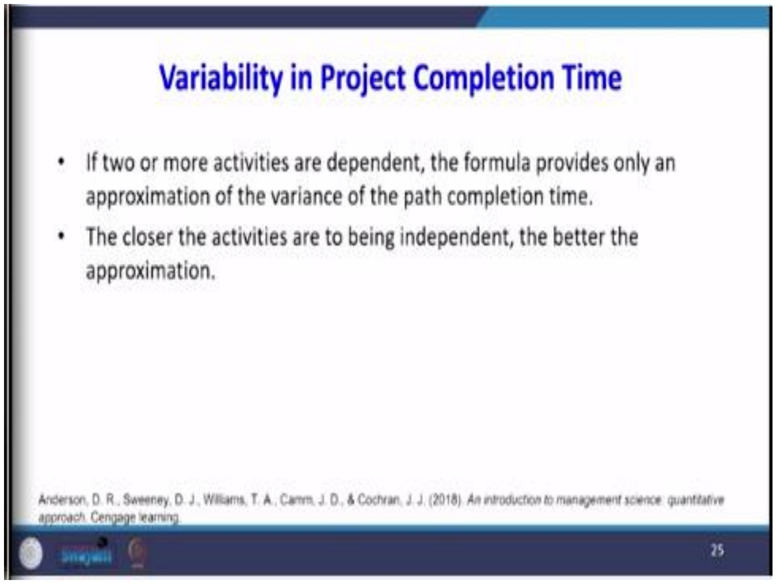



So, first, you will take path 1 A-E-H-I-J. This one A-E-H-I-J is the critical path that we all found 17 weeks ago. So, let the random variable Ti denote the total time to complete path i. So, the expected value of Ti = the sum of the expected times of the activities along the path, so what is that? Here 6 + 3, 9 + 4, 13 + 2, 5, 15 + 2, so it is 17, so we got 17 weeks. So, what do we have to do now?

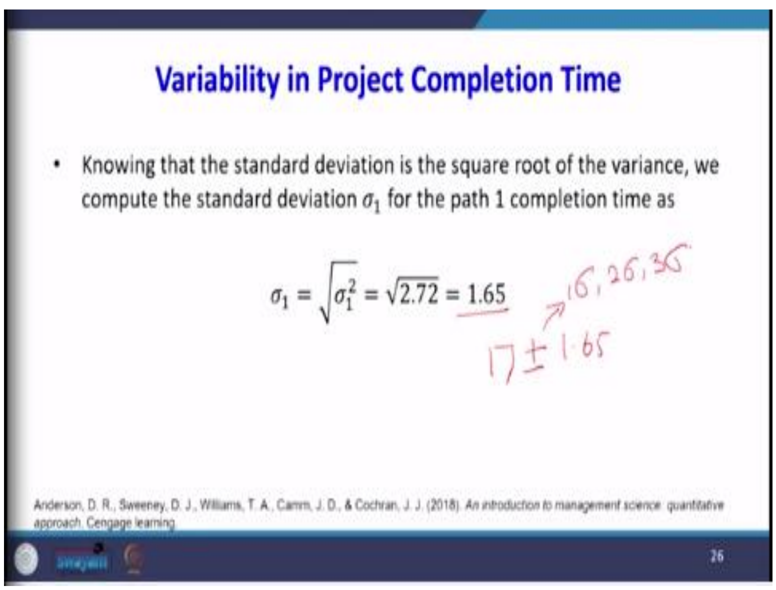



We have to find out the variance of this path to complete this project within 17 weeks. So, the variance of this Ti is the sum of the variance of the activities along the path i. So, what do we have to do? So, we have to sum the variance. We should be very careful. We should not add the standard deviation. We have to sum the variance. So, because we already have the variance data with this, 1.78, 0.11, 0.69, 0.03, 0.11, so we will get 2.72.

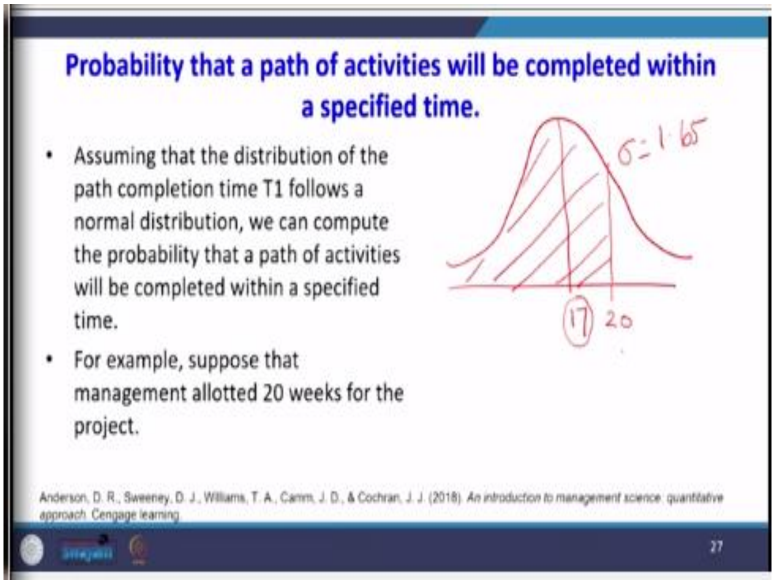
So, where this sigma A square, sigma E square, sigma H square, sigma I square and sigma J square are the variance of activities which are in that path. The formula for sigma 1 square, which means variability of path 1, is based on the assumption that activity times are independent; this assumption is important. So, these activity times do not have any dependency; these are independent.




If 2 or more activities are dependent the formula provides only an approximation of the variance of the path completion time. The closer the activities are to being independent, the better the approximation; here the assumption is all the activities are independent.



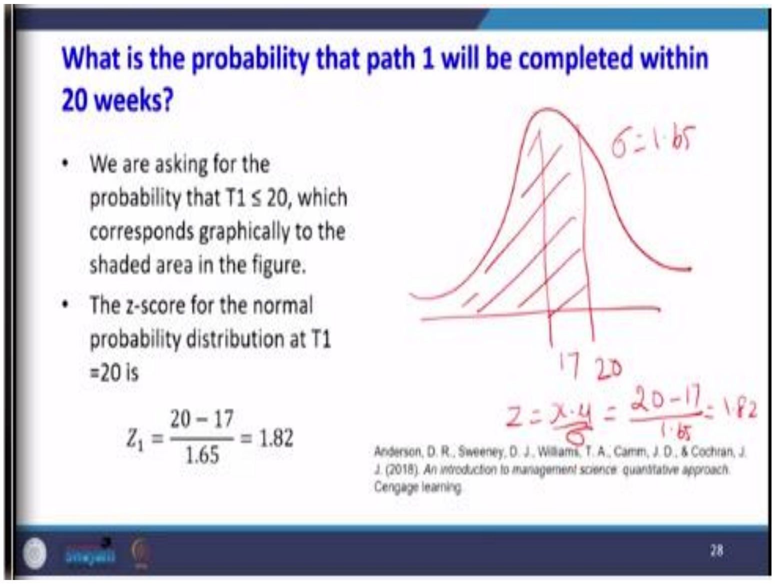

Knowing that the standard deviation is the square root of the variance, we compute the standard deviation of path 1, and that value is 1.65. Already, we know the mean is 17, so plus or minus, if you want to have 1 sigma, 2 sigma completion time here, you can write 1.65. Here, if you want, you can write 1 sigma, sometimes 2 sigma, sometimes 3 sigma.




Now we are going to find out the probability that a path of activities will be completed within a specified time. Assuming that the distribution of the path completion time T1 follows a normal distribution, we can compute the probability that a path of activities will be completed within a specified time. So, what point here I am discussing? You see that the project's completion time follows normal distribution.

So, for path 1, the mean is 17, and the standard deviation is 1.65. Now, suppose that the management allotted 20 weeks for the project; suppose the management says what is the probability that the project will be completed within 20 weeks, what do we have to do? We have to find out the area of this shaded portion, that is what we are going to do. So, this area will give you the answer for the probability of completing this path within 20 days.

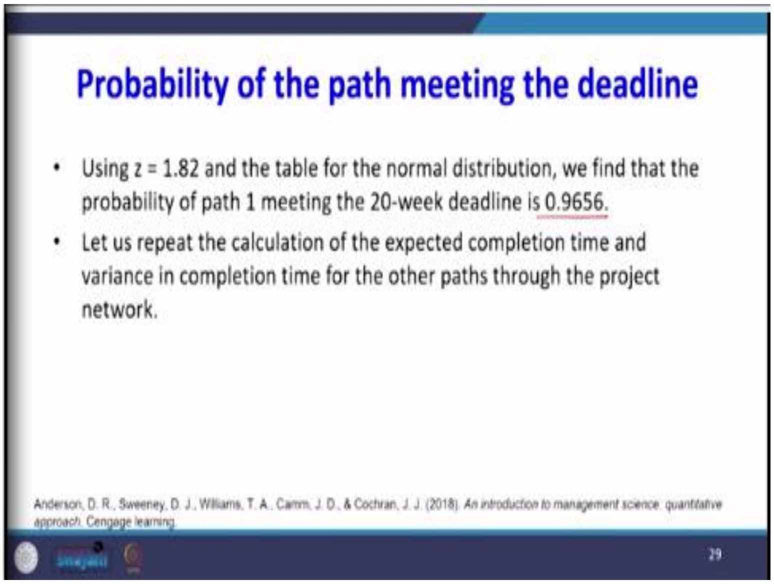
But you see, what is the probability of completing these activities in 17 days 0.5 because we found the average and because 17 is the average time, what will be the probability? It is 0.5, but if you want to complete this project in 20 days, obviously, the probability will be more than 0.5.



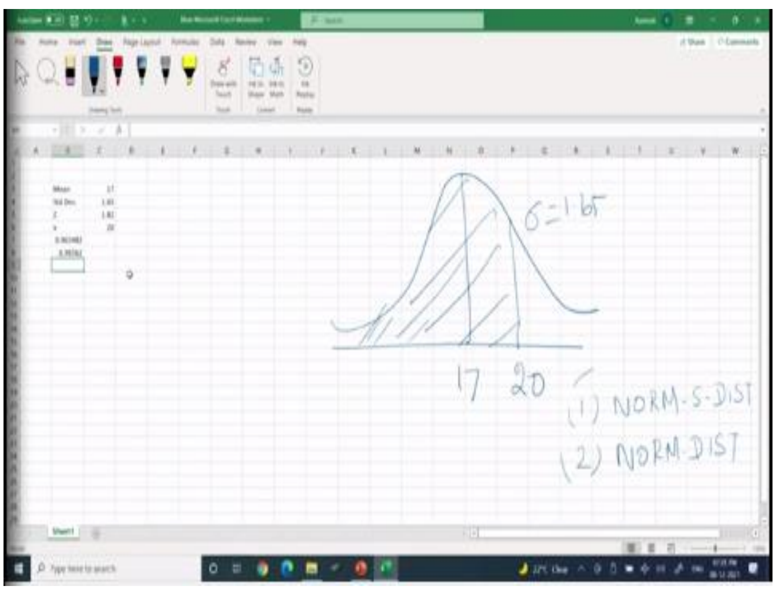

We are asking for the probability that the T1 is less than or equal to 20, which corresponds graphically to the shaded area in the figure. So, what do I mean by this? This is 17; this is 1.65, so this is 20, so we must find out the area of this shaded figure. So, to find the area, there are 2 ways: one is directly we can find the area of this distribution with the help of Excel; otherwise, we can find the area with the help of a table and a normal table.

So, if you are using a normal table, this 20 has to be converted to a z scale, so what is the formula for z? x - mu by sigma, so here x is 20, the mean is 17, sigma is 1.65, so when you simplify this, you will get 1.82. So, in a table where the z value is 1.82 the left side area will give you the probability of completing this path within 20 days.

$$Z_1 = \frac{20 - 17}{1.65} = 1.82$$




Using z = 1.82 and the table for the normal distribution, we find that the probability of path 1 meeting these 20 weeks is 0.9656; this probability we can find with the help of Excel also. So, now I am going to open excel. Now, I am going to find out the area of a normal distribution using Excel; I know the mean = 17, the standard deviation is 1.65, and I have already found that the value of z is 1.82. So, here is the value of x. For example, x = in B6, you write x = tab 20. Now, what I am going to find out when x is 20, what will be the area?




So, what I mean here is suppose there is a normal distribution, the mean is 17, standard deviation is 1.65. If x is 20, that is what will be the area on the left-hand side. So, in our problem context, what is the probability of completing this project within 20 weeks? So, we need to find out the area. In excel, there are 2 ways we can find the area of a normal distribution; one is NORM.S.DIST.

So, directly, we can find out the area of a standardized normal distribution; otherwise, there is another function NORM.DIST, you need not standardize directly. You can find the value of the probability of the normal distribution. First, I will use the second one, NORM.DIST. So, my cursor is in B 7 = = NORM.DIST, see what I need x, x is my 20, select 20, the mean is 17, the standard deviation is 1.65,

You see that this is a continuous distribution. We need the cumulative value, so click true and close the bracket. So, the area is 0.96 what is the meaning of this 0.96? There is a 96% chance that the project can be completed within 20 days. Then, next, I will use the other formula, NORM.S.DIST, so it is equal to NORM.S.DIST. Now we have to substitute the value of z value, so the value of z is 1.82, cumulative, close the bracket.

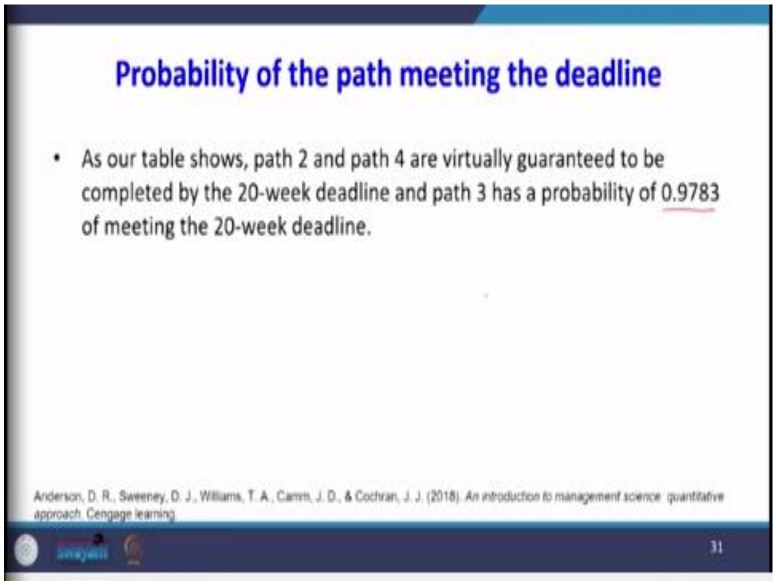
So, this is the way you see that both the values are the same; this is the way to use Excel to find the area of a normal distribution. Now I will go back to the lecture. Now you see that the area for completing within 20 weeks is 0.9656. Let us repeat the calculation of the expected completion time and variance in completion time for other paths through the project work.

Table- 5

| Expected Path Completion Time | S. D. of Path Completion Time | Z-Score | Probability of Meeting Deadline |
|---|---|---|---|
| E(T1) = 6+3+4+2+2 = 17 | $\sigma_1^2 = 1.78 + 0.11 + 0.69 + 0.03 + 0.11 = 2.72$ | $Z_1 = \frac{20-17}{\sqrt{2.72}} = 1.82$ | 0.9656 |
| E(T2) = 6+3+2+2 = 13 | $\sigma_2^2 = 1.78 + 0.11 + 0.3 + 0.11 = 2.03$ | $Z_2 = \frac{20-13}{\sqrt{2.03}} = 4.91$ | >0.9999 |
| E(T3) = 6+5+3+2 = 16 | $\sigma_3^2 = 1.78 + 1.78 + 0.25 + 0.11 = 3.92$ | $Z_3 = \frac{20-16}{\sqrt{3.92}} = 2.02$ | 0.9783 |
| E(T4) = 2+4+2+2 = 10 | $\sigma_4^2 = 0.44 + 0.69 + 0.03 + 0.11 = 1.27$ | $Z_4 = \frac{20-10}{\sqrt{1.27}} = 7.02$ | >0.9999 |

30

So, I have created a table that we have found already. Similarly, for path 2, I found the expected completion time is 13, and the standard deviation is 2.03. Then, if you find the area, it is 0.999. With the help of Excel, you can find out the area. Then path 3, the mean is 16, and the probability of completing this project in 20 weeks is 0.97. In path 4, the expected completion time is 10, the standard deviation is 1.27, and the probability of completing in 20 days is 0.99. So, what are we infer from this table?




Path 2 and path 4, see this path 2 and path 4 the probability is 0.99. That is, paths 2 and 4 are virtually guaranteed to be completed by the 20-week deadline, and path 3 has a probability of 0.9783 of meeting the 20-week deadline. So, in this lecture, I discussed project scheduling for uncertain activities, and then we found the expected duration of an activity. Assuming that the activity duration follows beta distributions, then I have found the variance of an activity.

Here, the assumption is that activity duration follows beta distribution. You should remember one point that activity duration follows beta distribution, but the project completion time follows a normal distribution, that point you should remember. Then, at last, what have I done? I found the probability of the path meeting the deadline. So, here I want to insist on the point that the probability of completing the deadline, that is, project completion time, follows normal distribution; thank you very much.