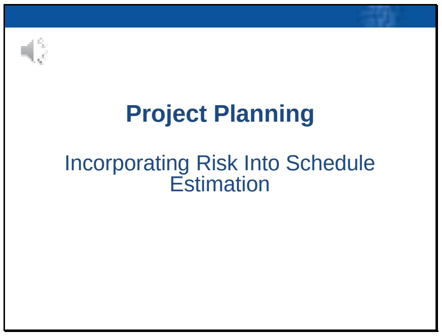
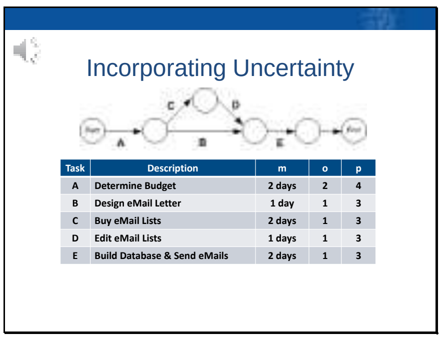
1



In this lecture we'll discuss some ways to incorporate risk into project schedule estimation. In particular, we'll discuss how to associate probabilities with schedule estimates.

2



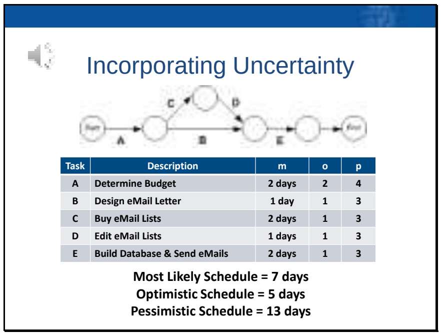
Task	Description	m	o	p
A	Determine Budget	2 days	2	4
B	Design eMail Letter	1 day	1	3
C	Buy eMail Lists	2 days	1	3
D	Edit eMail Lists	1 days	1	3
E	Build Database & Send eMails	2 days	1	3

I'll continue to use the seminar marketing campaign project to illustrate how to incorporate uncertainty using probabilities. For this example, I'm using duration estimates for the project tasks. For most software projects, I'd use effort estimates…which would automatically be converted into calendar durations by a project management tool.

As we've already seen, one way to incorporate uncertainty into schedule estimates is to use the three-point estimating technique that was introduced in an earlier lecture…and calculate three separate estimates for the schedule. I've included estimates for most likely, optimistic, and pessimistic task durations in the table.

As a quick exercise, I'd like you calculate schedule estimates for the most likely, optimistic, and pessimistic task estimates. Please pause this lecture now, and continue it when you have completed the exercise.

3

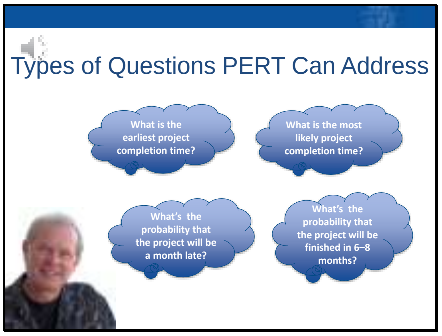


The most likely schedule is 7 days, the same as our deterministic schedule estimate in an earlier lecture. The optimistic schedule is 5 days and the pessimistic schedule estimate is 13 days. The difference between the optimistic and pessimistic estimates is more than 100 percent…that's quite a difference, but if the assumptions that went into what determined each of our task estimates are valid, then that much variation is a reality. If you didn't get the same results as these you may want to revisit the earlier lecture that illustrated how to calculate a project's critical path.

The next issue is how would one present such variation to management and customer stakeholders. In my experience, if three estimates like this were presented, there would be a lot of pressure to commit to the optimistic estimate. This is why I recommended documenting assumptions earlier. If the assumptions turn out not to be true, then the validity of the estimate is impacted.

Now…for the sake of discussion, suppose the powers that be gravitated to a schedule that was somewhere in between the optimistic and the most likely. What would be the probability that the project could actually be completed within that estimate? Hard to say…right? Well…here's where the PERT technique I mentioned earlier can help out. Let's see how.

4

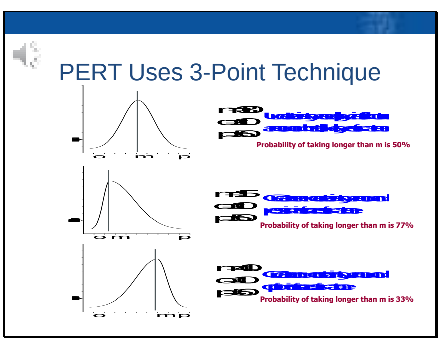


First a little bit of history. PERT is an estimating and scheduling technique that was developed for use in planning projects that have never been done before and for which there exist numerous schedule risks. It is often used in projects for which schedule, rather than cost, is the most important factor. PERT was developed by the U.S. Navy in the 1950s to assist with the planning and management of the Navy's Polaris missile project.

Some of the types of questions PERT can help answer are questions like "what's the earliest project completion time?", "What's the most likely project completion time?", "What's the probability that the project will be finished a month late?", and "What's the probability that the project can be completed in 6-8 months?". Notice the repeated use of the word "probability" in these questions. You can actually calculate these probabilities using PERT, though not everyone does.

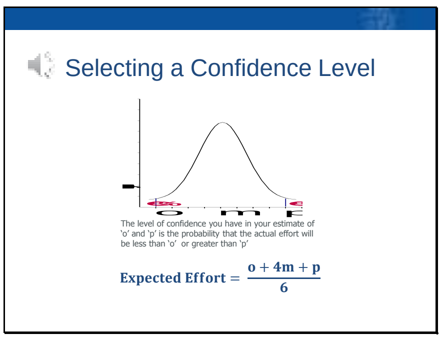
Whether or not the probabilistic aspects of PERT are used, in its basic form it can be used to incorporate risk into effort and schedule estimates.

5



In an earlier lecture we discussed using the three-point technique to estimate effort or cost…and I used it again at the beginning of this lecture to illustrate possibly calculating separate schedules. Recall that one estimate was the optimistic…or best case…estimate, another was a pessimistic…or worst case…estimate, and the third estimate was the one we felt was the most likely to be the actual. As we saw in that lecture, depending upon the nature of uncertainty associated with the three estimates the probabilities of effort being more or less than the most likely estimate can be equally distributed or skewed in one direction or the other. PERT uses that same three-point technique…but in a more rigorous way…by incorporating probability and statistics.

6



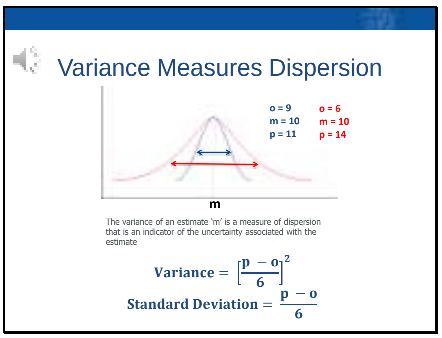
When we use PERT, we factor probabilities into the estimates of 'o' and 'p'…the optimistic and pessimistic estimates…by choosing what's called a confidence level. I've used the letter 'c' in this diagram to illustrate the confidence level. We choose an optimistic estimate so that we feel there is only a 'c percent' chance that the actual effort will be less than our optimistic value…and we choose a pessimistic estimate so that we feel there is only a 'c percent' chance that the actual effort will be greater than it. Common values for the confidence level are 1 percent, 5 percent, and 10 percent. This is really a guess on our part, but it allows us to make some statistical calculations and statements. And…in practice, we would document any assumptions behind our choices.

In PERT, we then estimate the project schedule by using a weighted average for each task estimate. The formula for the weighted average is to sum the optimistic estimate, the pessimistic estimate, and 4 times the most likely estimate…and divide by 6…to get the estimate of expected effort.

Now, that formula for calculating the average or expected effort is certainly not intuitive. It is based on the average for a type of statistical distribution called a beta distribution.

The expected effort estimates are then used to calculate the project schedule. If we are using a project management tool, we would use this weighted average for the task efforts. Some tools will actually let you enter all three estimates and calculate four different schedules…an optimistic schedule, a pessimistic schedule, a most likely schedule, and an expected schedule.

7



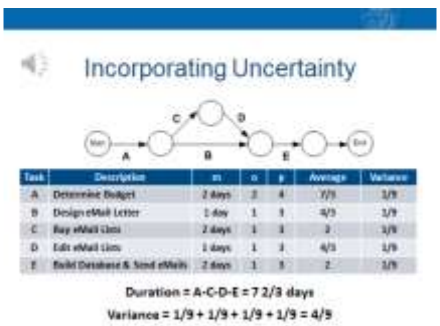
To use PERT we need to know a little bit about statistics. A common measure used in statistics is the variance associated with an estimate. The variance measures the dispersion around an estimate…and in the PERT technique it reflects the amount of uncertainty around an estimate.

For example, suppose we made three estimates for a task in which the most likely estimate is 10, the optimistic is 9 and the pessimistic is 11. If we plotted the probability that the actual value would fall between the optimistic and pessimistic, the curve might look something like the one illustrated in blue. If the optimistic was 6 and the pessimistic was 14, the probability distribution would look like the one illustrated in red. Simplistically, the width of the red curve is greater than the width of the blue curve…and it's a visual illustration of variance. It has a larger variance than the blue curve.

In this example there are bell-shaped curves around an estimate, because the optimistic and pessimistic estimates are equidistant from the most likely. The earlier showed examples where the probability distributions were skewed towards either the optimistic or pessimistic.

When we use PERT, the variance of an estimate is calculated according to the formula the square of the pessimistic minus optimistic divided by six. The standard deviation is the square root of the variance, or pessimistic minus optimistic divided by 6. Both of these measures are used when we calculate probabilities. Like the weighted average formula, these formulas are not intuitive. They are also associated with a beta distribution…but that's not really important in order to use them.

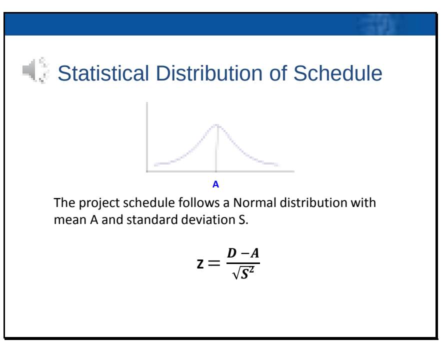
8



For the email campaign project, the weighted task averages are shown in this table. As it turns out, the estimated project duration is the same as in the earlier example, where the three-point estimates were not used. In practice, it could very well be different when the weighted values are used.

I also calculated the variance of each task estimate, as well as the variance of the critical path estimate. The variance of the critical path estimate is the sum of the variances of the critical path tasks. In this case, the value is 4/9. The variance and the weighted average will be used in calculating probabilities.

9



So…how do we actually calculate probabilities? It's really pretty mechanical…but first…a little more statistics.

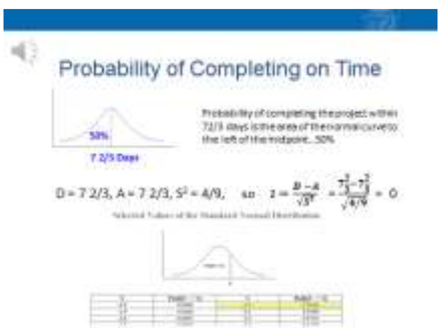
Using the formula for the weighted average task durations…it is assumed that the project duration is distributed statistically according to a Normal distribution. A normal distribution is a bell-shaped curve. The mean, or average, of that distribution is the critical path duration, which we'll call A in this illustration, and the standard deviation of the distribution is the standard deviation of the critical path tasks…calculated by taking the square root of the sum of the critical path task variances…which we'll call S^2.

The use of a Normal distribution has been shown to be a reasonable approximation of the Beta distribution in practice.

From the Normal distribution, we can calculate probabilities from a table of pre-calculated values by using something called a standard normal deviate and the formula illustrated here. Let's assume we want to calculate the probability that the project will be completed within D days. Z is the standard normal deviate, A is the estimated schedule value, and S^2 is the sum of the critical path task variances. We calculate the value of Z and then look up the probability value from a probability table. That's all there is to it.

Let's take an example.

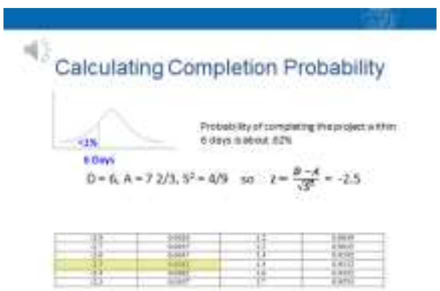
10



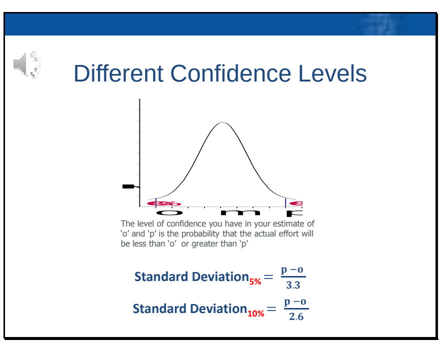
For our email campaign project, the expected completion time, A, is seven and two-third days. What's the probability that the project will be completed within seven and two-third days? Well…since we're using a Normal distribution with mean value of seven and two-thirds, we can see the answer just by looking at the illustration. The area under the entire curve is equal to one…or 100 percent…by definition. Since the mean…seven…is the halfway point, the area to the left of the mean is the probability that the project completes within seven and two-third days. Since the distribution is symmetric, half the area is to the left, so the probability of completing within 7 and two-third days is 50 percent.

We can also use the standard normal deviate calculation and look up the probability. The value of the standard deviate is zero, so we look up the area corresponding to a standard deviate value of zero in a probability table…highlighted in yellow…and see that the result is 50 percent.

We can also use the standard normal deviate calculation and look up the probability. The value of the standard deviate is zero, so we look up the area corresponding to a standard deviate value of zero in a probability table…highlighted in yellow…and see that the result is 50 percent.

11 

Let’s take another example. What’s the probability of completing the project within 6 days? In this case, the value of the standard normal deviate is -2.5, so we look up the probability associated with -2.5…which is .0062… a little less than one percent.
For your reference, a copy of the probability tables is included on the course website as part of the materials for this unit.

12 

The formula for standard deviation that was used in my examples assume that the confidence level of our pessimistic and optimistic estimates were one percent. That is, there was a one percent chance that the actual optimistic duration could be less than the optimistic estimate and a one percent chance that the actual on the pessimistic side could exceed our pessimistic estimate.

Other values for confidence level that are often used in practice are 5 percent and 10 percent. To use these confidence levels with our Normal probability distribution we need to adjust the standard deviation formulas as indicated here. Note that these formulas result in larger values of standard deviation than the one I used, because there is more margin of error in or estimates…but these estimates may be more realistic and easier to come up with than for a one percent confidence level.