Modeling Schedule Uncertainty without Monte Carlo Methods

For small projects (or small companies), a simple estimation procedure for schedule uncertainty follows.

PERT

Recall that the PERT method models schedule (and cost) uncertainty by using a three-point estimate. The project team provides optimistic (minimum, Min), pessimistic (maximum, Max), and most likely (Mode) estimates of duration for each activity. Only activities on the critical path are used to compute the schedule baseline. The general observations - and most important points - of the PERT method are that project personnel are optimistic in estimating activity durations and actual schedules finish later than estimates. (A RAND Report on, and analysis of, cost risk for 70 major DoD procurement programs indicated both a systematic bias toward underestimating costs as well as a substantial uncertainty in estimating final costs, reaffirming these observations, see Reference 1).

A Beta Distribution models these observations quite well. The Beta Distribution is asymmetric – the mode value does not always coincide with the mean value. The mean could be in advance of, or later than, the mode. For any activity, the distribution may be calculated from the Min, Max, Mode, and standard deviation. The difficulty is what to use for, or how to estimate, the standard deviation. According to the PERT method, it is estimated as (Max-Min)/6. The advantage to the method is that the Min/Max/Mode are easily estimated, but the disadvantage is only three of the four parameters needed to compute the Beta Distribution are independent.

Many texts suggest using the computed PERT mean value as the schedule basis, the computed PERT standard deviation for the spread, and then assuming a *Normal Distribution* to calculate a cumulative probability. Furthermore, the overall schedule is computed from the sum of the individual (critical path) activity mean values while the variance is computed from the sum of the individual variances. This argument invokes the "central limit theorem" and "statistical convergence." Once the latter are invoked, a Normal Distribution results. (This is also called the Moments Method). This shifts away from the fundamental PERT assertion – which is the schedule estimates tend to be *asymmetric*, not symmetric as is the Normal Distribution function. It is possible that the variance of an individual activity might dominate the sum of the variances, which then makes the resultant distribution non-normal. Also, Normal Distributions do not model finite start and finish times.

The best way of handing schedule uncertainty is to use Monte Carlo simulation methods. Each critical path activity could be modeled with an appropriate distribution function and Min/Max/Mode inputs, and the project simulated in its entirety. Software is available for use with Microsoft Project to aid the simulation, but not all projects need this level of computation (or expense).

An alternative is to provide the Min/Max/Mode inputs to all activities, find the critical path manually (or with Microsoft Project), and then use the three overall schedule values for the critical path - optimistic, pessimistic, and most likely - as inputs to a Beta Distribution model. This would serve as an overall critical path schedule model from which a Cumulative Probability may be determined using Excel. One could then easily estimate, for example, the time the project should be at 50% complete, 95% complete, etc. from the built-in Excel functions.

Triangular Distribution

An alternative to the classical PERT model is the use of the Triangular Distribution. Both the probability distribution and Cumulative Probability are completely determined from the Min/Max/Mode estimates.

$$Mean = \frac{(Max + Mode + Min)}{3}$$

$$Mode = Max$$

$$Median = Min + \frac{\sqrt{(Max - Min) \times (Mode - Min)}}{\sqrt{2}} \quad \text{for} \quad Mode \geq \frac{Max - Min}{2}$$

$$Median = Max - \frac{\sqrt{(Max - Min) \times (Max - Mode)}}{\sqrt{2}} \quad \text{for} \quad Mode \leq \frac{Max - Min}{2}$$

$$StdDev = SQRT\left(\frac{Min^2 + Max^2 + Mode^2 - Min \times Max - Min \times Mode - Max \times Mode}{18}\right)$$

The Cumulative Distribution Function is:

$$\frac{(X - Min)^2}{(Max - Min) \times (Mode - Min)} \quad \text{for} \quad Min \leq X \leq Mode$$

$$1 - \frac{(Max - X)^2}{(Max - Min) \times (Max - Mode)} \quad \text{for} \quad Mode < X \leq Max$$

The difficulty with the Triangular Distribution is that it models a “hard stops” schedule – that is, there can be no activity before the optimistic estimate or after the pessimistic estimate. While this might be true for an optimistic schedule – usually nothing starts earlier than the optimistic estimate – this is unrealistic for the pessimistic schedule – most often the project/activity finishes later. [How many times have project managers seen an activity get to 90% complete and never get to 100% complete until the end of the project?]. Nonetheless, because it is easy to compute, the Triangular Distribution is often used in calculations.

I maintain that at this level of uncertainty, as long as Min/Max/Mode estimates are sought from the project team as inputs, and the inputs for the individual activities are statistically independent, it is just as easy to use the Beta Distribution as the Triangular Distribution and the Beta Distribution models reality better than the Normal Distribution. The exception would be when there is specific information to suggest using a different probability model ab initio for any activity.

Sample Calculation

As an illustration, assume four activities, *all on the critical path*, with estimates for each activity for a minimum, maximum, and mode value as indicated in Table 1. The PERT calculated mean = (Min + Max + 4xMode)/6 and the PERT calculated standard deviation = (Max-Min)/6. The averages and standard deviations assuming a normally distributed random variable are shown in the last columns for each activity, respectively.

Activity	PERT VALUES					Normal Distribution Calculations	
	Inputs			Calculations			
	Min	Mode	Max	Mean	Std Dev	Avg	Std Dev
A	2.0	3.0	5.0	3.2	0.5	3.3	1.5
B	4.0	5.0	7.0	5.2	0.5	5.3	1.5
C	1.0	2.0	3.0	2.0	0.3	2.0	1.0
D	4.0	5.0	15.0	6.5	1.8	8.0	6.1
Total	11.0	15.0	30.0	16.8	2.0	18.7	6.5

Table 1 Sample Schedule Activity Calculations

The total standard deviation is the root sum square of the individual standard deviations for each activity. Notice the PERT total standard deviation is 2.0 and for the Normal Distribution it is 6.5. Most single point estimates use the PERT mean value for the critical path (here, 16.8) and add 2 PERT standard deviations to the mean to obtain "about a 95% probability of finishing" (according to many texts) at 20.8. This approximation tends to underestimate the schedule finish.

As Figure 1 shows, both the Beta and Triangular Distributions have different mean values though the same mode value. Both are skewed toward shorter finish times. The Normal Distribution would of course be symmetric about 18.7 and extend to infinity on either side. The Triangular Distribution was computed from the equations in the preceding section above and the Beta Distribution calculated as described in the following paper,

Model PERT Project Schedules with the BETA Distribution Using EXCEL available for download at: http://www.laserlightnetworks.com

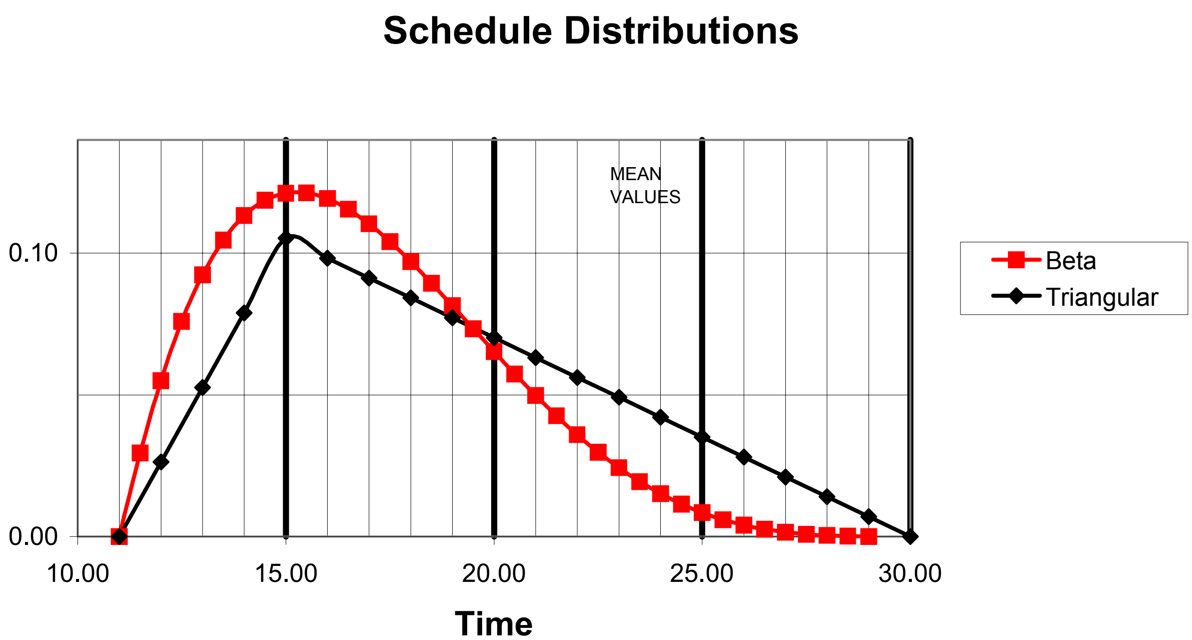

Figure 1 Beta and Triangular Model Schedules

Figure 2 shows all three cumulative probability distributions plotted together. Table 2 summarizes results from the cumulative probabilities for the three distributions for six percentile levels.

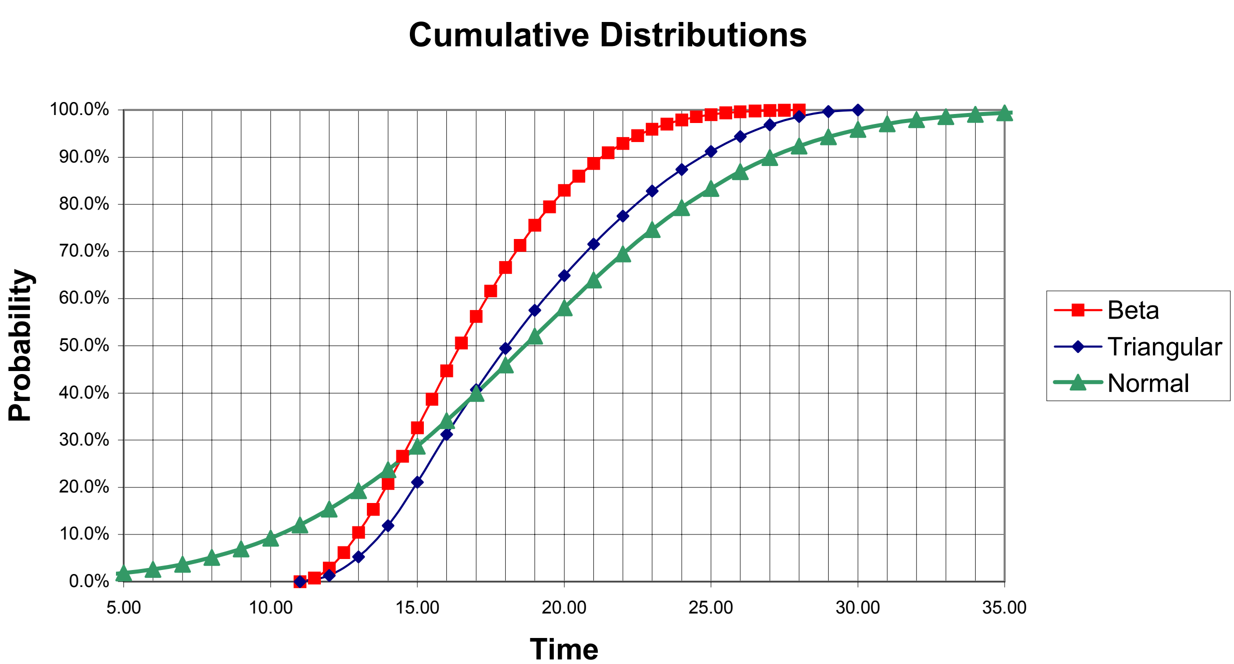

Figure 2 Beta, Triangular, And Normal Cumulative Probabilities

Distribution	Min	Max	Mode	Mean	Std. Dev	PERCENTILES 25%	50%	75%	90%	95%	99.90%
Beta	11.0	30.0	15.0	16.8	3.2	14.4	16.4	18.9	21.3	**22.7**	**27.0**
Triangular	11.0	30.0	15.0	18.7	4.1	15.4	18.1	21.6	24.7	**26.2**	**29.5**
Normal			18.7	18.7	6.5	14.3	18.7	23.1	27.0	**29.4**	**38.9**

Table 2 Summary Cumulative Probabilities for 3 Models

Discussion

The Normal Distribution spreads more over the time period of interest and predicts a later finish date than the maximum estimate (of course, theoretically it concludes at infinite time). The Triangular Distribution finishes abruptly at the maximum estimate. The Beta Distribution predicts a steeper descent to a finish than the Normal Distribution, but not as steep as the Triangular Distribution.

The last percentile figure was deliberately calculated at “three nines” accuracy for each distribution. Using more significant figures would calculate later maximum finish times for the Beta and Normal Distributions but would be an incorrect use of estimation accuracy level. The point is to compare the 99.90% level for each distribution to the estimated maximum finish time, 30.0. The Triangular Distribution finishes at, the Beta before, and the Normal after, the maximum estimated finish time. This highlights the resulting differences when selecting one of these model distribution functions for schedule purposes.

A mode value estimated closer to the center of the spread between the Max and Min would make both the Triangular and Beta Distributions more symmetric and results for the Normal and Beta Distributions would be closer. (The Beta Distribution can be made symmetric by making the shape factors equal; thus, it can resemble a Normal Distribution but with finite end points).

Conclusions

Monte Carlo methods are generally superior for estimating schedule uncertainty. However, if such software is unavailable or simply overkill to use for some small projects, Excel is an alternative for modeling schedules. Recalling the PERT methodology, a single activity or the entire schedule can be modeled as a Beta Distribution using a three-point estimate scheme.

Observations on many, many realistic schedules indicate a trend toward asymmetric distribution functions. The asymmetric Beta and Triangular Distributions are better than the Normal Distribution since the Normal strictly has no finite limits and is symmetric. [Note that a Truncated Normal is still symmetric.]

The Beta Distribution is better than the Triangular Distribution. The Triangular Distribution is a hard start/hard finish model estimate, while the Beta models schedule with a more gradual descent to the finish – again, closer to observations.

Ref 1: **Impossible Uncertainty: Cost Risk Analysis for Air Force Systems**, M. Arena, O. Younossi, L. Galway, et al, RAND Project Air Force, RAND Corporation, 2006, pp. xix, xx

To comment, ask questions, or to suggest changes/additions, send an E-Mail:
mailto: dwonica@laserlightnetworks.com