Adding Probability to Your "Swiss Army Knife"

John C. Goodpasture, Vice President, Legal Services Operations, Lanier Professional Services, Inc.

Introduction

The project schedule is perhaps the "Swiss army knife" of tools available to help manage the triple constraint. Look at what all it contains: first, the project calendar is there, so the time line in an absolute sense is available. Good schedules are organized by WBS, so all of the project deliverables contained within the approved scope are represented, and all of the tasks necessary to achieve the deliverables are there. Good scheduling includes an examination of relationships among project tasks, so the relationships are there also. The schedule contains, or at least is based upon, the project resource plan for needed skills, units required, and when they are required. If the project is one provided on contract, then the contract Statement of Work is decomposed into the schedule tasks, so from the schedule, the SOW could be synthesized; and, of course, the project risks are represented by the slack available to absorb unknowns. With such robust content, every effort should be made to leverage the schedule for insight and problem avoidance. One way to do this is to apply statistical methods to better understand the uncertainties of the plan and to get out in front of trouble before it impacts performance.

Risk as a Schedule Tool

Let's work from the following proposition: business leaders are really investors, and they make investment decisions when they charter projects. They have an expected outcome that is usually couched in business improvement terms. Project managers have a mission to deliver on the chartered scope, taking measured risks to achieve the value proposition of the project. Consider this: the executive's assignment of resources is precise and deterministic. It is not probabilistic, approximate, or statistical. But the project situation is different. Forecasting the results of projects can only be made approximately, at best. So, how is alignment achieved between management expectation and the estimates of the project team? Enter probability. Probabilistic estimation provides the project manager a means to reach agreement with the management team on a "point solution" for the project completion date or budget. Such a "point solution" is actually a data point, which has been picked from a range of possibilities that satisfies the organization's level of risk aversion established with agreement between the Project Manager and the executive. Once it becomes the project estimate, the project manager must then manage the project in a manner to defeat the possibility that this estimate may, in fact, be overrun!

> The task of the project manager in risk management is to defeat the quantitative prediction of a risk analysis and bring the project in for the value assigned by management.

Schedules and Statistics

Schedule Risk Analysis

Because schedule has such a high leverage on project success, usually much more than cost, let's focus on the schedule risk as the primary balance sheet risk to manage (Hulett, 1996). There are three main purposes of a schedule risk analysis:

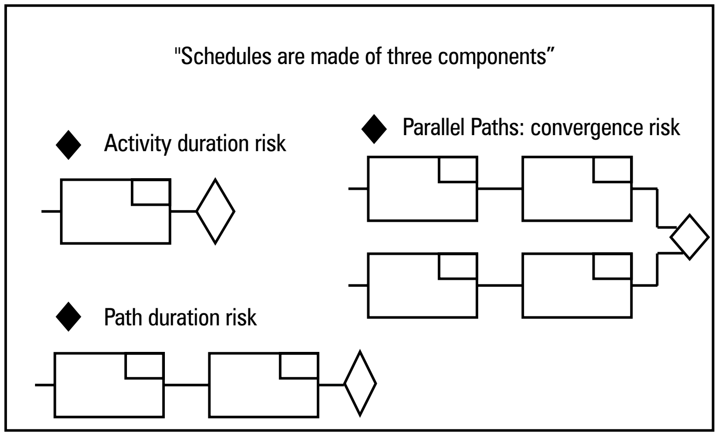
1. Determine the likelihood of overrunning the schedule
2. Estimate the likely exposure or contingency needed to drive the remaining risk of overrun down to a level acceptable to the company
3. Identify the locus of the key risks in the schedule to guide risk management efforts.

Quantitative schedule risk management is statistically based. The basic terminology can be found in any statistics text. In the PMBOK, in Chapter 11 on Project Risk Management, can be found the probability density functions most used in schedule analysis. They are the Triangular Distribution and the Beta distribution. The equations for mean, variance, and standard deviation for these distributions are likewise found in there.

In developing a risk-manageable schedule, an understanding of these scheduling fundamentals is essential:

- *Critical Path schedules are deterministic*: That is, a critical path method [CPM] schedule has at least one path through its network, called the Critical Path, that determines the length of the project. The duration of critical path schedule is computed by adding single-point estimates of activity durations along the longest contiguous

Exhibit 1. Three Basic Components of Schedule Risk Analysis



path through the network. The critical path predicts the completion date only if everything goes *according to plan.*. This method is limited because durations really are not certain. Using probabilistic analysis, we find that there is some likelihood that the durations will conspire so that a path which is not the critical path will really become critical. Thus, we define the "near critical path": a path that is not deterministically critical, but statistically may become critical, thereby delaying the project.

- *Activity duration estimates are really probabilistic assessments about what can happen in the future* . That is, when the task leader schedules the tasks, the duration, although a single number, is in fact an estimate of the statistically "most likely" duration of the task.
- "*Well, at least if you use the most likely estimates of duration you'll come up with the most likely completion date, won't you?*" No, actually you will not, as will be made clear. *The completion date is not even the most likely date*, much less a date with acceptable level of risk!

Three fundamentals of schedules

To understand how useful statistics can be to risk management of the schedule, consider that schedules are really made up of three basic primitives as shown in Exhibit 1:

- The first component is the single task with an activity duration risk. Each activity appears in the schedule with a duration that represents the most probable number of work periods necessary to accomplish the task.
- The second schedule risk component is the path. A path is composed of at least two activities linked by a relationship, most commonly finish-to-start when using Precedence Diagramming Methods [PDM].
- The third component is the merge point where n-paths join at a milestone. In this case, the milestone is not complete until all paths complete. Schedule risk often increases at a merge point where any of the paths can delay the project. If each path is statistically independent, then the likelihood of completing the milestone is their probabilities multiplied together.

Quantitative Analysis

Now, let's put schedule elements and statistics together to get a view of the uncertainties in the schedule and a quantitative estimate of what they will do to the CPM dates. We've already made the point that it is common practice to use fairly simple probability distributions to model the uncertainty of the durations caused by various risk factors. Because they model real task behavior quite well, either the Beta distribution or the Triangular distribution is often used. Their asymmetry is key to modeling the real world which rarely has perfect symmetry.

First, we can deal with the uncertainty in the most fundamental schedule element, the single task. The schedule error in one long task can be "managed" in the following sense. In an example shown in Exhibit 2, "Managing Duration Risk", the single baseline task is labeled WBS Activity 1.0. It is assumed that its probability distribution can be modeled with a Triangular Probability Distribution. The metrics of interest are the Average duration and the Standard Deviation. The latter is a measure of the dispersion of duration errors, and generally for this statistic, the smaller the better. To examine the risk of the baseline we

Exhibit 2. Managing Duration Risk

Structure of Scheduled Activities in Days	Triangle Probability Distribution of Duration: Minimum (-10%)	Most Likely	Maximum (+30%)	Average	Variance (Days-Squared)	Standard Deviation (Days)
Baseline	54	60	78	64.00	26.00	5.10
Baseline Restructured into Four Sub-Tasks and a Summary Task						
WBS Activity 1.1	13.5	15	19.5	16.00	1.63	1.27
WBS Activity 1.2	13.5	15	19.5	16.00	1.63	1.27
WBS Activity 1.3	18	20	26	21.33	2.89	1.70
WBS Activity 1.4	9	10	13	10.67	0.72	0.85
Summary (New Baseling)	Distribution Unknown			64.00	6.86	2.62
Average=(min+max+most likely)/3 Variance=[(max-min)](max-min)+ (most likely-min)(most likey-max)]/18 Standard Deviation=sq root (Variance)				No Change from Baseline	74% Improved from Baseline	49% Improved from Baseline

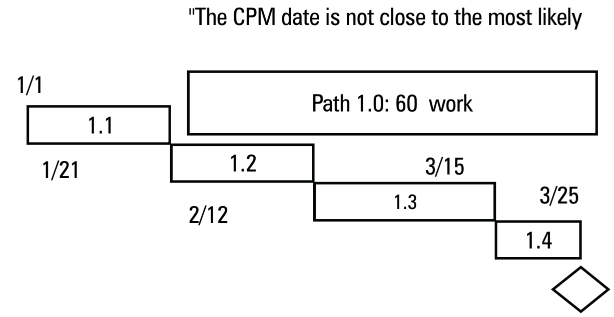
go to the detailed statistics where the risks can better be identified and quantified. (Raz and Globerson 1998), In Exhibit 2, four subtasks are shown as an example. Notice that the overall variance and standard deviation are improved for the group of four subtasks as compared to the singular baseline task.

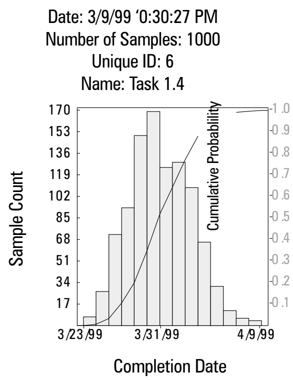
Monte Carlo Simulation

But what about our confidence of making the most probable or the average date, or any other date the Project Manager might have picked from the distribution? To avoid the complexity of the mathematics involved in looking at all the statistics of all the activities, a process known as a "Monte Carlo simulation" is used to statistically estimate the overall schedule and set confidence limits. This is simply a procedure whereby many iterations of the schedule are calculated, or "run". Each iteration uses a different combination of possible durations for the uncertain activities. These durations are chosen at random from the full range of possible durations. The full range is specified by the risk analyst, usually after interviews of knowledgeable experts, and expressed as the probability distribution. We see the results in Exhibit 3, 4-Task Project. There is shown the probability distribution and the cumulative distribution for 4-task one path project. Parameters were selected as -10% from mean for the most optimistic and +25% from the mean for most pessimistic. The average duration date, March 30th, is about 50-60% likely, but other dates, particularly the CPM "most likely" date of March 25th, is about 5% likely. In fact, the CPM date is not really close to being the most likely date! The risk analysis shows that to be about March 30th.

Next, we want to deal with the whole schedule network and see what insights statistics bring. Exhibit 4, "Parallel Paths", illustrates a Monte Carlo analysis of all three schedule primitives. Triangular distributions are used consistently with the parameters of -10%, +25% for minimum and maximum from the most likely, respectively. We've already seen the impact of the simple activity and the activity string. Lets look at the example of the two tasks merging. To do this, we examine a new project with two identical paths. Each is made up of risky activities and the project ends in a finish milestone. We can see that confidence of meeting a date for the merge milestone is approximately the product of the confidences of each path, and for a given confidence, the date is "shifted right". For instance, in the single path, as previously shown in Exhibit 3, recall that the confidence in meeting March 30th is about 50%. At the merge point, as shown in Exhibit 4, Parallel Paths, the confidence at March 30th is approximately 50%*50%, or 25%. Note that the CPM date of March 25th which was about 5% likely with one path is

Exhibit 3. 4 Task Project





Completion Std Deviation:2.4d
95% Confidence Interval:0.1d
Eash bar represents 1d

Completion Probability Table

Prob	Date	Prob	Date
0.05	3/25/99	0.55	3/31/99
0.10	3/25/99	0.60	3/31/99
0.15	3/26/99	0.65	4/1/99
0.20	3/26/99	0.70	4/1/99
0.25	3/29/99	0.75	4/1/99
0.30	3/29/99	0.80	4/2/99
0.35	3/29/99	0.85	4/2/99
0.40	3/30/99	0.90	4/5/99
0.45	3/30/99	0.95	4/6/99
0.50	3/30/99	1.00	4/9/99

just barely "on the map" with two paths. Finally, we should pause to remind ourselves that the CPM says each of these projects finishes on March 25th, which is further evidence of the need for risk analysis. What is the most likely date? It is about April 1st or 2nd with two paths, not March 30th as it is for one path, or certainly not March 25th as predicted by the CPM!

Two very important observations emerge from the schedule risk analysis:

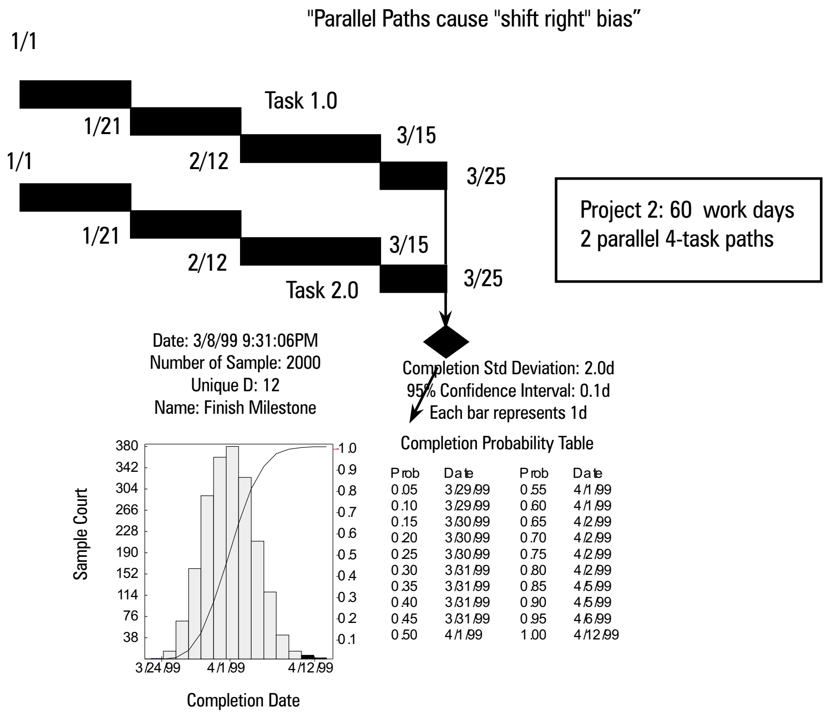
- The most likely duration is never the sum of the most likely durations of each schedule activity. That is, *"most likely's" [the modes of the distributions] can not be added!* However, averages can be added along a single path project, so the sum of all the average task durations is the average duration of the schedule path. But, the worst problem is, as seen in Exhibits 3 and 4, picking the CPM date as the most likely date is rarely a good bet. Even with a single path project, the CPM date is almost always far too optimistic!
- *And, perhaps most importantly for risk management purposes, there is a phenomenon at the merge point called "merge bias".* Merge bias means that at a merge point, the probability that the milestone will finish later than planned is nearly a certainty. (Archibald and Villoria 1967). This is so because the *likelihood of meeting the merge date is the product of the likelihood's of each of the joining paths,* if the paths are independent of each other. Schedule merge points stand out almost automatically as "shift-right" risk. If the merge point is on the critical path, then any shift-right problems translate directly into a lengthened project schedule! With complex schedules and many parallel paths and merge points, this effect is pronounced. Project managers ignore risk at their peril.

Summary and Conclusion

Summarizing the important points presented:

- Risk analysis, at least quantitative risk analysis, uses the concepts of statistics. This means, as a practical matter, that schedule elements and cost elements should be statistically estimated. The simplifying tool is the Monte Carlo simulation. The simulation results give the project team and the project manager the information necessary to construct a workable Project Balance Sheet.
- The highest risk architectural element in a schedule is the "merge point" of multiple tasks. There is a bias, the merge bias, toward a "shift right" of the schedule at merge points.

Exhibit 4. Parallel Plans



Completion Probability Table

Prob	Date	Prob	Date
0.05	3/29/99	0.55	4/1/99
0.10	3/29/99	0.60	4/1/99
0.15	3/30/99	0.65	4/2/99
0.20	3/30/99	0.70	4/2/99
0.25	3/30/99	0.75	4/2/99
0.30	3/31/99	0.80	4/2/99
0.35	3/31/99	0.85	4/5/99
0.40	3/31/99	0.90	4/5/99
0.45	3/31/99	0.95	4/6/99
0.50	4/1/99	1.00	4/12/99

- Risk can be "quantitatively managed" for both cost and schedule elements of the WBS by using the concepts of statistics to minimize the probability of extended durations and overrun cost.

In conclusion, risk management is an effective tool to manage projects outcomes, especially when combined with statistics in order to get quantitative measures. The main task of the project manager, once the quantitative information is available, is to defeat the predictions and bring the project to a successful conclusion for the parameters established in the project charter.

Acknowledgment

The author is indebted to Dr. David T. Hulett, President Hulett Associates, for his collaboration on a manuscript entitled: "The Project Balance Sheet: A guide to Risk-based Value" which as been accepted for future publication in PM Network. This paper is derived from that manuscript

References

Hulett, David, "Schedule Risk Analysis Simplified", PM Network, July 1996, pp 23-30. This article provides additional information on schedule risk analysis.

Raz, T. and Globerson, S.(1998), A similar concept to size WBS cost elements is described in: "Effective Sizing and Content Definition of Work Packages", Project Management Journal, Volume 29, Number 4, December, pg. 17

Archibald, Russell and Villoria, Richard (1967). Network-Based Management Systems (PERT/CPM), (Wiley), Appendix C pp. 447—461. The original description of this bias was by Kenneth MacCrimmon and Charles Ryavec while they were at The Rand Corporation.