Risk vs Uncertainty
What's the Difference?

Melvin R. Etheridge, Jr.

Decades after the recognition that cost estimates were in fact sums of random variables that could be analyzed using Monte Carlo simulation, there still remains some confusion within the cost estimating profession regarding the terms *risk* and *uncertainty*. Perhaps this is due, in part, to differences in viewpoint and terminology between different communities within program management and the maturation of analytical techniques with the profession. This paper seeks to clarify the distinction between the two terms and in the process align their meaning within the cost profession with other program management professions. Additionally, analytical techniques for use in *Excel*™-based cost models will be presented.

One of the more ambiguous terms in program management is *cost risk*. To the program manager, cost risk is the probability that he/she will run out of program funds before he/she runs out of program scope. But this can be the result of many causes. If a cost estimate that provides the basis for a program's funding includes all of the program scope at the time the estimate was developed and includes analysis of program risks and uncertainties, lack of sufficient funding should not be attributed to the cost estimate.

First, we should define the two terms in order to provide a framework for what follows.

<u>*Risk*</u>. The American Heritage World Dictionary, Second College Edition, defines risk as *the possibility of suffering harm or damage.* For the purposes of this discussion in the context of program management, we shall define it as a*n event, which <u>may</u> occur, that will have an adverse effect on a project's execution in terms of time or money.* Note, that in program management, events which have a positive effect are called *opportunities*.

<u>*Uncertainty*</u>. Here, the American Heritage World Dictionary, Second College Edition, defines uncertainty as *the condition of being in doubt*. We shall define it as *a quantitative input to a cost model which is not known to degree of accuracy sufficient for the estimate*. While, for example, we may not need an estimate of software source lines of code (SLOC) to a precision greater than the nearest thousand SLOC, in fact we may not know the SLOC count to accuracy within more than several thousand SLOC.

So, to distinguish between the terms, we should point out that:

- A risk is a <u>discrete event</u> with a probability of occurrence. The risk effect (impact) is only felt if / when the event occurs.
- There is no probability of occurrence with an uncertainty – you know that you don't know the actual value of the input variable.

Well managed programs identify those risks that may impact the execution of those programs. These *program risks* are identified, categorized, tracked, and reported by a risk management team (board / committee) within the program management office. The risk management team is typically multidisciplinary with stakeholder representation from within the program. Risks are categorized in two dimensions, both usually with five categories. Each category has a range of values.

- Probability of occurrence or likelihood, ranging from low probability (designated *A:* e.g., 0%-20%) to high probability (designated *E*; e.g., 80%-100%)
- Effect given occurrence or impact, Often impacts are expressed as percentages of total program costs. Impact categories range from low impact (designated *1*; e.g., 0%-0.1%) to high impact (designated *5*; e.g., >10%).[1]

Each program risk has a rating that is a combination of likelihood and impact; e.g., *A2*, *C4*. It is important to keep in mind that, typically, likelihood and impact categories are often the result of "expert" opinion as opposed to analysis.

Once program risks have been identified and cataloged, program management must make decisions regarding managing these risks. If a risk cannot be transferred to another entity, it may be possible to improve the risk rating by either reducing the likelihood or impact, or both. Potential risk reduction approaches for each risk are identified. One approach, especially for risks with low ratings, is always to accept the risk; i.e., do nothing. Each risk reduction approach will have a cost and/or schedule impact and would result in a reduction in impact and/or likelihood. It is a management decision to adopt a particular risk reduction approach or accept the risk. If a risk reduction approach is selected, the cost and schedule impact of the strategy become part of the point cost estimate and the reduced impact and likelihood is used in the cost estimate uncertainty analysis.

A closed-form solution of the effects of both program risks and input uncertainties is only possible in the most trivial of cases. Therefore both are analyzed simultaneously using Monte Carlo simulation.

To analyze the effects of program risks, the analyst must identify specific input variables or cost elements impacted should the risk occur. Since the risk impacts are generally expressed in terms of the effect on total program costs, they can be translated into a dollar amount by multiplying the upper and lower impact percentages by the point estimate, then allocating this total to different affected cost elements. This should be done in consultation with the risk management team and appropriate technical subject matter experts. If more than one variable or element is affected, the impacts would probably be different for each one.

Next, an expected value of the risk impact on each cost element or variable should be calculated:

$$EV = P_O * E|O$$

Where:

P_O = Probability of occurrence (likelihood)

[1] In Risk Management, a risk's probability of occurrence is termed *likelihood* and its effect given occurrence *impact*. The terms will be used interchangeably in this paper.

E|O = Effect given occurrence (impact)

Note that both the probability and the effect are random variables, so that the expected value is the product of random variables and hence is, itself, a random variable.

When setting up the Monte Carlo simulation, the analyst must select an appropriate distribution with accompanying parameters for both the probability and the effect for each risk. The distribution for the probability will be the same for all affected cost elements. But just as the effects may be different for each risk and element combination, it is conceivable that the effect distributions themselves could be different.

As with the allocation of risks to cost elements and input variables, the analyst should consult with the risk management team and subject matter experts when selecting distributions and parameters selection. When considering the selection the analyst consider two questions:

- Is there a central tendency to the particular probability or effect; i.e., is there any value within the range assigned that is more likely than another?
- Is there a possibility that the value could lie outside of the specified range?

Generally, there is an assumption that the answer to both of these questions is "no." In such a case, a uniform distribution with the two parameters (minimum and maximum) the end points of the range. However, the potential for an affirmative to one of the other of these questions would require selection of a distribution other than uniform.

Input uncertainties are often expressed as simple values, when in reality they are random variables. Again, estimate credibility requires that the analyst work with the subject matter experts that provided the inputs, if available, to determine the appropriate distributions and parameters. Where the input is an assumption (including expert opinion) one would expect the degree of uncertainty; i.e. the standard deviation, to be somewhat greater than that for inputs with a basis such as prior experience.

Unlike the program risks where uniform distributions are often assigned, input variables will generally have a central tendency (the original value assigned) and the distribution will likely be skewed. In working with the sources of the input variables the analyst should realize that many of these subject matter experts do not think probabilistically and may tend to assign unreasonably narrow ranges and symmetrical distributions when, in fact, there is a reasonable expectation of skewness.

Distributions that are often assigned to input variables include:

- *Triangular*. The probability density function for this distribution looks like a triangle. The parameters are minimum, most likely (mode), and maximum. The mean is (min + most likely + max)/3. It is a simple distribution and the parameters are easily understood by those unfamiliar with statistics. The probability of a value less than the minimum parameter and more than the maximum is zero.
- *Normal*. This is the familiar "bell" curve, so called because the probability density function looks like a bell. The parameters are mean and standard distribution. It is continuous and

symmetrical, hence the mean = the most likely. While most people have a concept of the normal distribution, they do necessarily understand the standard deviation. In consulting with these persons, it is sometimes useful to point out that about 2/3 of the observations will fall within one standard deviation of the most likely value. Note that, in theory, the normal distribution extends to ±∞, in practice; however, the probability of a value more than three standard deviations away from the most likely is about 1 in 1,000.

- *Log-Normal*. The natural logarithm of the log-normal distribution is the normal distribution; i.e., if the log-normal were to have a mean of 1, the normal would have a mean of 0 (ln(1) = 0). The parameters are the mean and standard deviation. The log-normal is right (positively) skewed and cannot have a value less than zero. Hence it is intuitively attractive to the analyst since unlike the triangular distribution, the log-normal distribution does not have a theoretical maximum and cannot extend below zero. However, since it is right skewed, the most-likely value will be less than the mean. How much less depends on the standard deviation. Therefore, translating simple parameters such as a most-likely, minimum, and maximum, which are easily understood by most sources of input values, into the mean and standard deviation of a log-normal is not easy.
- *Uniform*. This is a continuous distribution with no most-likely value. The parameters are the minimum and maximum and all values in between have the same probability. The probability density function looks like a rectangle.
- *Discrete*. This is a distribution which can only have certain, specific values. It is appropriate where the input variable is not continuous, such as travel days or people. It is impossible to have fractional people. In these cases the discrete values are integers. Typically, the analyst and subject matter experts would determine the probabilities assigned to each value.

Example

The following is not intended to be a complete or realistic cost estimate, but simply to illustrate the application of the concepts discussed above. We have the cost estimate for a information technology system with hardware and software. Costs are estimated for a single year. There are only three cost elements:

- Hardware procurement,
- Software development, and
- Program management contractor support.

Input variables are:

Variable	Value	Units
Average Unit Cost	$25.0	BY $K
Total Quantity	500	Units
Cost per contractor FTE	$200.0	BY $K
Contractor FTEs, PM support	5	FTEs
Software development labor hours	25,000	labor hours
Risk reduction software development labor hours	5,000	labor hours
Labor hours per FTE	2,000	labor hours

There are two program risks identified by the Risk Management Team:

	Program Risks	**Costs affected**	**Risk Rating**	**Likelihood**		**Impact** ***(Increase in Total Program F&E)***	
				Min	**Max**	**Min**	**Max**
				Unmitigated			
1	The system may require a more powerful processor	Hardware production	C4	40%	60%	5%	10%
	Risk Reduction Strategy: Write additional software: 5,000 labor hours			**Mitigated**			
			B2	20%	40%	0.10%	1.00%
2	The program schedule may slip	*Not a discrete event: treat as an uncertainty*					

The Program Manager decided to reduce the risk for Risk #1, *that the system may require a more powerful processor*, by writing additional software which will require an additional 5,000 labor hours for software development and result in a reduction of the risk rating from C4 to B2.

Risk #2, *the program schedule may slip*, is not a discrete event since it does not specify the period of the possible slip. Therefore, it is treated as an uncertainty in program duration, not a program risk.

The point estimate is:

Cost Element	**Methodology**	**Calculation**	**BY $K 20XX**
Hardware Production	AUC * Quantity	$25.0 * 500	$12,500
Software Development	(Dev Hrs./Hrs. per FTE) * Cost per FTE	{(25,000+5,000)/2000}*$200	$3,000
Contractor Program Management Support	FTEs * Cost per FTE	5 * $200	$1,000
Total			$16,500

The mitigated risk dollars are calculated as shown in the table below:

				Likelihood		**Impact** ***(Increase in Total Program F&E)***	
	Program Risks	**Costs affected**	**Risk Rating**	**Min**	**Max**	**Min**	**Max**
				Unmitigated			
1	The system may require a more powerful processor	Hardware production	C4	40%	60%	5%	10%
	Risk Reduction Strategy: Write additional software: 5,000 labor hours			**Mitigated**			
			B2	20%	40%	0.10%	1.00%
			Total F&E from point estimate (BY $K)			$16,500	$16,500
			Cost impact (BY $K)			$17	$165
			Increase in Hardware Production			0.13%	1.32%
2	The program schedule may slip	*Not a discrete event: treat as an uncertainty*					

The cost uncertainty analysis is set up in a series of tables containing the distribution parameters and cells containing the random variables. In the Crystal Ball™ Monte Carlo simulation add-in to Excel™ these are called "assumptions" and are shaded green.

The Processor risk is tabulated below. The likelihood and impact are assumed to have uniform distributions.

	Min	Max	Assumption Value
Risk	**Mitigated Likelihood**		
Processor Risk	20.0%	40.0%	30.0%
	Mitigated Impact		
	0.13%	1.32%	0.7%

The expected value of the cost increase is:

Expected Value AUC Increase = 1 + (30% * 0.7%) =100.2%

Since the only cost element affected by this risk is Hardware Production, there is only a single expected value. Note that since this is the product of two random variables, it is also a random variable and although the displayed value is 100.2%, the actual value will change with each trial of the simulation.

The input uncertainties are tabulated below:

Uncertainties	Distribution	Minimum	Most Likely	Maximum	Assumption Value
Cost per FTE (BY $K)	Triangular	$175.0	$200.0	$250.0	$200.0
Contractor FTEs, PM support*	Triangular	5.0	5.0	7.0	5.0
			Mean	**Std. Dev.**	
Software Development Hours	Normal		30,000	5,000	30,000

* Due to possible schedule slip

Note that the program risk #2, *the program schedule may slip* is treated as an uncertainty in the number of Contractor FTEs for program management support, with the assumption that this would be the only element affected by a slip. Also, in this case, since the minimum is equal to the most likely value, this assumes the program for whatever reason could not complete early.

The Monte Carlo simulation is run and results in the following tabular cumulative distribution function:

Percentiles	Contractor Program Management Support (BY $K)	Hardware Production (BY $K)	Software Development (BY $K)
0%	$879	$12,503	$1,183

10%	$1,018	$12,509	$2,416
20%	$1,063	$12,513	$2,649
30%	$1,100	$12,518	$2,823
40%	$1,136	$12,522	$2,970
50%	$1,168	$12,526	$3,107
60%	$1,201	$12,530	$3,249
70%	$1,242	$12,534	$3,414
80%	$1,291	$12,539	$3,607
90%	$1,359	$12,547	$3,870
100%	$879	$12,565	$5,440

In this agency, programs are budgeted to the 80th percentile of the cumulative distribution function. Therefore, the "risk adjusted" estimate, including the effects of both program risks and input uncertainties, is:

		BY $M
Cost Element	**Methodology**	**20XX**
Hardware Production	AUC * Quantity	$12.95
Software Development	(Dev Hrs./Hrs. per FTE) * Cost per FTE	$3.06
Contractor Program Management Support	FTEs * Cost per FTE	$1.29
Total		$17.30

Summary

Uncertainty pertains to quantitative inputs including assumptions. There is no likelihood for an uncertainty. *Risks* are discrete events with a negative impact on program cost (or schedule). They are specified and managed by a risk management team. The likelihood of a risk is less than 1.0. The impact given occurrence of a program risk can be expressed either in percent cost increase (usual) or in dollars. Risk likelihoods and impacts and input uncertainties are random variables and should be expressed with assigned distributions and appropriate parameters. Cost impacts of both uncertainties and risks are analyzed using Monte Carlo simulation.