

A Scenario-Based Method for Cost Risk Analysis

Paul R. Garvey
The MITRE Corporation
MP 05B0000023, September 2005

Abstract

This paper presents an approach for performing an analysis of a program's cost risk. The approach is referred to as the scenario-based method (SBM). This method provides program managers and decision-makers an assessment of the amount of cost reserve needed to protect a program from cost overruns due to risk. The approach can be applied without the use of advanced statistical concepts, or Monte Carlo simulations, yet is flexible in that confidence measures for various possible program costs can be derived.

1.0 Introduction

This paper* introduces an analytical, non-Monte Carlo simulation, approach for quantifying a program's cost risks and deriving recommended levels of cost reserve. The approach is called the Scenario-Based Method (SBM). This method emphasizes the development of written scenarios as the basis for deriving and defending a program's cost and cost reserve recommendations.

The method presented in the paper grew from a question posed by a government agency. The question was *Can a valid cost risk analysis (that is traceable and defensible) be conducted with minimal (to no) reliance on Monte Carlo simulation or other statistical methods?* The question was motivated by the agency's unsatisfactory experiences in developing and implementing Monte Carlo simulations to derive "risk-adjusted" costs of future systems.

This paper presents a method that addresses the question posed by the agency. The method reflects a "minimum acceptable" approach whereby a technically valid measure of cost risk can be derived without Monte Carlo simulations or advanced statistical methods. A "statistically-light" analytical augmentation to

* *This paper was written for the United States Air Force Cost Analysis Agency.*

Version 2.0.2

this method is also presented that enables one to assess probabilities that a program's cost will (or will not) be exceeded.

2.0 Some Basic Terms and Definitions

Throughout this paper certain technical terms and distinctions between them are used. This section presents these terms and explains the subtleties between their meanings. First, we'll briefly discuss the concept of a subjective probability. This will be followed by a discussion of risk versus uncertainty and the differences between them.

Subjective Probability Assessments [1]: Probability theory is a well-established formalism for quantifying uncertainty. Its application to real-world systems engineering and cost analysis problems often involves the use of subjective probabilities. Subjective probabilities are those assigned to events on the basis of personal judgment. They are measures of a person's degree-of-belief an event will occur.

Subjective probabilities are associated with one-time, non-repeatable events, those whose probabilities cannot be objectively determined from a sample space of outcomes developed by repeated trials, or experimentation. Subjective probabilities must be consistent with the axioms of probability [1]. For instance, if an engineer assigns a probability of 0.70 to the event "*the number of gates for the new processor chip will not exceed 12000*" then it must follow the chip will exceed 12000 gates with probability 0.30. Subjective probabilities are conditional on the state of the person's knowledge, which changes with time.

To be credible, subjective probabilities should only be assigned to events by subject matter experts, persons with significant experience with events similar to the one under consideration. Instead of assigning a single subjective probability to an event, subject experts often find it easier to describe a function that depicts a distribution of probabilities. Such a distribution is sometimes called a

Version 2.0.2

subjective probability distribution. Subjective probability distributions are governed by the same mathematical properties of probability distributions associated with discrete or continuous random variables.

Subjective probability distributions are most common in cost uncertainty analysis, particularly on the input-side of the process. Because of their nature, subjective probability distributions can be thought of as “belief functions.” They describe a subject expert’s belief in the distribution of probabilities for an event under consideration. Probability theory provides the mathematical formalism with which we operate (add, subtract, multiply, and divide) on these belief functions.

Risk versus Uncertainty [1]: There is an important distinction between the terms risk and uncertainty. Risk is the chance of loss or injury. In a situation that includes favorable and unfavorable events, risk is the *probability an unfavorable event occurs*. Uncertainty is the *indefiniteness* about the outcome of a situation. We analyze uncertainty for the purpose of measuring risk.

In systems engineering the analysis might focus on measuring the risk of: failing to achieve performance objectives, overrunning the budgeted cost, or delivering the system too late to meet user needs. Conducting the analysis involves varying degrees of subjectivity. This includes defining the events of concern, as well as specifying their subjective probabilities.

Given this, it is fair to ask whether it’s meaningful to apply rigorous procedures to such analyses. In a speech before the 1955 Operations Research Society of America meeting, Charles Hitch addressed this question. He stated [2]:

Systems analyses provide a framework which permits the judgment of experts in many fields to be combined to yield results that transcend any individual judgment. The systems analyst [cost analyst] may have to be content with better rather than optimal solutions; or with devising and costing sensible methods of hedging; or merely with

Version 2.0.2

discovering critical sensitivities. We tend to be worse, in an absolute sense, in applying analysis or scientific method to broad context problems; but unaided intuition in such problems is also much worse in the absolute sense. Let's not deprive ourselves of any useful tools, however short of perfection they may fail.

Given the above, it is worth a brief review of what we mean by cost uncertainty analysis and cost risk analysis. *Cost uncertainty analysis is a process of quantifying the cost impacts of uncertainties associated with a system's technical definition and cost estimation methodologies. Cost risk analysis is a process of quantifying the cost impacts of risks associated with a system's technical definition and cost estimation methodologies. Cost risk is a measure of the chance that, due to unfavorable events, the planned or budgeted cost of a project will be exceeded.*

Why conduct the analysis? There are many answers to this question; one answer is to produce a defensible assessment of the level of cost to budget such that this cost has an acceptable probability of not being exceeded.

3.0 The Scenario-Based Method (SBM): A Non-statistical Implementation

Given the "what" and "why" of cost risk analysis, *a minimum acceptable method* is one that operates on specified scenarios that, if they occurred, would result in costs higher than the level planned or budgeted. These scenarios do not have to represent worst cases; rather, they should reflect a set of conditions a program manager or decision-maker would want to have budget to guard against, should any or all of them occur. For purposes of this discussion, we'll call this minimum acceptable method the "Scenario-Based Method" (SBM) for cost risk analysis.

The Scenario-Based Method derives from what could be called "sensitivity analysis", but with one difference. Instead of arbitrarily varying one or more variables to measure the sensitivity (or change) in cost, the Scenario-Based Method involves specifying *a well-defined set of technical and programmatic conditions that collectively affect a number of cost-related variables and associated work breakdown structure (WBS) elements in a way that increase cost beyond what was*

planned. Defining these conditions and integrating them into a coherent risk “story” for the program is what is meant by the term “scenario”.

The process of defining scenarios is a good practice. It builds the supportive rational and provides a traceable and defensible analytical basis behind a “derived” measure of cost risk; this is often lacking in traditional simulation approaches. Visibility, traceability, defensibility, and the cost impacts of specifically identified risks is a principal strength of the Scenario-Based Method.

Figure 1 illustrates the process flow behind the non-statistical SBM.

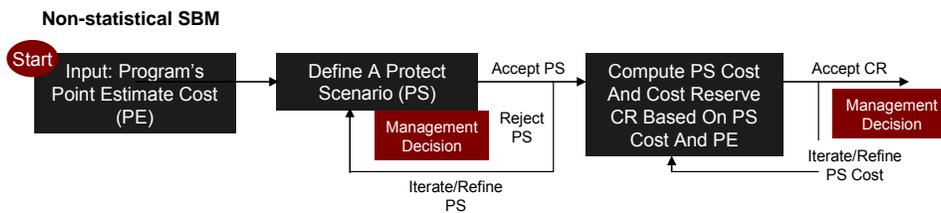


Figure 1. A Non-statistical Scenario-Based Method

The first step (see Start) is input to the process. It is the program’s point estimate cost (PE). For purposes of this paper, the point estimate cost is defined as the cost that does not include an allowance for cost reserve. It is the sum of the cost element costs summed across the program’s work breakdown structure *without adjustments for uncertainty*. Often, the point estimate is developed from the program’s cost analysis requirements description (CARD).

Next, is the effort to define a protect scenario (PS). The key to a “good PS” is one that identifies, not an extreme worst case, but a scenario that captures the impacts of the major known risks to the program – those events the program manager or decision-maker must monitor and guard the costs of the program against. Thus, the PS is not arbitrary. It should reflect the above, as well as

Version 2.0.2

provide a possible program cost that, in the opinion of the engineering and analysis team, has *an acceptable chance of not being exceeded*.

In practice, it is envisioned that management will converge on a protect scenario after a series of discussions, refinements, and iterations from the initially defined scenario. This part of the process, if executed, is to ensure all parties reach a consensus understanding of the risks the program faces and how they are best represented by the protect scenario.

Once the protect scenario has been defined and agreed to its cost is then determined. The next step is computing the amount of cost reserve dollars (CR) needed to protect the program's cost against identified risk. This step of the process defines cost reserve as the difference between the PS cost and the point estimate cost, PE. Shown in figure 1, there may be additional refinements to the cost estimated for the protect scenario, based on management reviews and considerations. This too may be an iterative process until the reasonableness of the magnitude of this figure is accepted by the management team.

A Valid Cost Risk Analysis

This approach, though simple in appearance, is a valid cost risk analysis; *why?* The process of defining scenarios is a valuable exercise in identifying technical and cost estimation risks inherent to the program. Without the need to define scenarios, cost risk analyses can be superficial with its basis not well-defined or carefully thought through. Scenario definition encourages a discourse on program risks that otherwise might not be held. It allows risks to become fully visible, traceable, and "costable" to program managers and decision-makers.

Defining, iterating, and converging on a protect scenario is valuable for understanding the "elasticity" in program costs and identifying those sets of risks (e.g., weight growth, software size increases, schedule slippages, etc.) the program must guard its costs against. Defining scenarios, in general, builds the

Version 2.0.2

supportive rational and provides a traceable and defensible analytical basis behind a “derived” measure of cost risk; this is often lacking in traditional simulation approaches. Visibility, traceability, defensibility, and the cost impacts of specifically identified risks is a principal strength of the Scenario-Based Method.

The non-statistical SBM described above does come with limits. Mentioned earlier, cost risk, by definition, is a measure of the chance that, due to unfavorable events, the planned or budgeted cost of a program will be exceeded. A non-statistical SBM does not produce confidence measures. The chance that the cost of the protect scenario, or the cost of any defined scenario, will not be exceeded is not explicitly determined. The question is *Can the design of the SBM be modified to produce confidence measures while maintaining its simplicity and analytical features?* The answer is yes. A way to do this is described in the section that follows.

4.0 The Scenario-Based Method (SBM): A Statistical Implementation

This section presents a statistical, non-Monte Carlo simulation, implementation of the SBM. It is an optional augmentation to the methodology discussed above. It can be implemented with lookup tables, a few algebraic equations, and some appropriate technical assumptions and guidance.

There are many reasons to implement a statistical SBM. These include (1) a way to develop confidence measures; specifically, confidence measures on the dollars to plan so the program’s cost has an acceptable chance of not being exceeded (2) a means where management can examine changes in confidence measures, as a function of how much reserve to “buy” to ensure program success from a cost control perspective and (3) a way to assess where costs of other scenarios of interest different than the protect scenario fall on the probability distribution of the program’s total cost.

Approach & Assumptions

Figure 2 illustrates the basic approach involved in implementing a statistical SBM. Observe that parts of the approach include the same steps required in the non-statistical SBM. So, the statistical SBM is really an augmentation to the non-statistical SBM. The following explains the approach, discusses key technical assumptions, and highlights selected steps with computational examples.

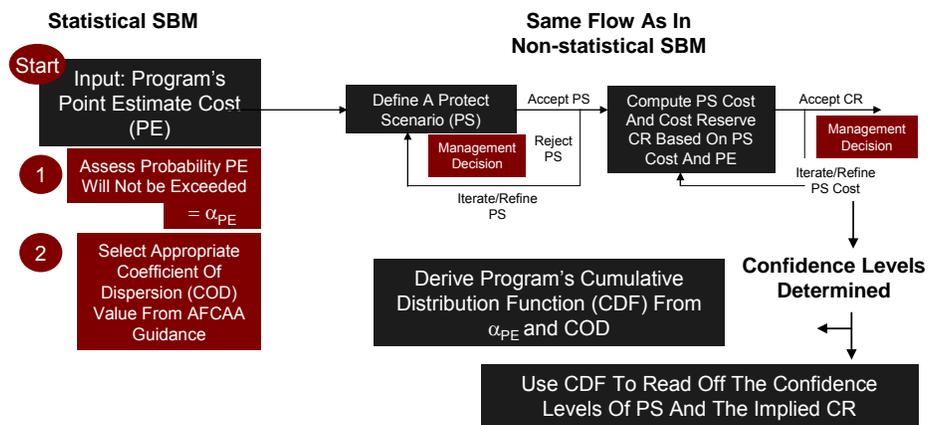


Figure 2. A Statistical Scenario-Based Method

Mentioned above, the statistical SBM follows a set of steps similar to the non-statistical SBM. In figure 2, the top three activities are essentially the same as described in the non-statistical SBM, with the following exception. Two statistical inputs are needed. They are the probability the point estimate cost (PE) will not be exceeded α_{PE} and the coefficient of dispersion (COD). We'll next discuss these a little further.

Point Estimate Probability

For the statistical SBM, we need the probability

$$P(Cost_{pgm} \leq x_{PE}) = \alpha_{PE} \tag{4-1}$$

Version 2.0.2

where $Cost_{pgm}$ is the true, but unknown, total cost of the program and x_{PE} is the program's point estimate cost (PE). Here, the probability α_{PE} is a judgmental or subjective probability. It is assessed by the engineering and analysis team. In practice, α_{PE} often falls in the interval $0.10 \leq \alpha_{PE} \leq 0.50$.

*Coefficient of Dispersion (COD)**

What is the coefficient of dispersion? The coefficient of dispersion (COD) is a statistical measure defined as the ratio of distribution's standard deviation σ to its mean μ . It is one way to look at the variability of the distribution at one standard deviation around its mean. The general form of the COD is given by equation 4-2.

$$D = \frac{\sigma}{\mu} \tag{4-2}$$

Figure 3 illustrates this statistical measure.

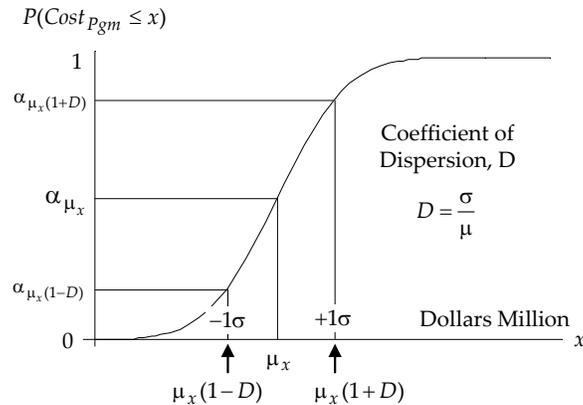


Figure 3. Coefficient of Dispersion

* The coefficient of dispersion is also known as the coefficient of variation.

Here, the COD statistic is a judgmental value but one guided by Air Force Cost Analysis Agency (AFCAA) and industry experiences with programs in various stages or phases of the acquisition process. As will be discussed later in this paper, a sensitivity analysis should be conducted on both statistical inputs, namely α_{PE} and COD, to assess where changes in assumed values affect cost risk and needed levels of reserve funds.

The next two steps along the top of the process flow, in figure 2, follow the procedures described in the non-statistical SBM. Notice these two steps do not use the statistical measures α_{PE} and COD. It is not until you reach the last step of this process that these measures come into play.

As will be shown in the forthcoming examples, the distribution function of the program's total cost can be derived from just the three values identified on the far-left side of the process flow in figure 2. Specifically, with just the point estimate cost PE, α_{PE} , and COD the underlying distribution function of the program's total cost can be determined. With this, other possible program costs, such as the protect scenario cost, can be mapped onto the function. From this, the confidence level of the protect scenario and its implied cost reserve can be seen.

This completes an overview description of the statistical SBM process. The following presents two computational examples that illustrate how the statistical SBM works.

4.1 Formulas: Statistical SBM With An Assumed Underlying Normal

Here, we assume the underlying probability distribution of $Cost_{Pgm}$ is normally distributed and the point (x_{PE}, α_{PE}) falls along this normal. If we're given just the point estimate PE, α_{PE} , COD, then the mean and standard deviation of $Cost_{Pgm}$ are given by the following equations.

$$\mu_{Cost_{Pgm}} = x_{PE} - z_{PE} \frac{Dx_{PE}}{1 + Dz_{PE}} \quad (4-3)$$

Version 2.0.2

$$\sigma_{Cost_{pgm}} = \frac{Dx_{PE}}{1 + Dz_{PE}} \quad (4-4)$$

where D is the coefficient of dispersion (COD), x_{PE} is the program's point estimate cost, z_{PE} is the value such that $P(Z \leq z_{PE}) = \alpha_{PE}$ and Z is the standard normal random variable; that is, $Z \sim N(0,1)$. The value for z_{PE} derives from the look-up table in Appendix A.

Once $\mu_{Cost_{pgm}}$ and $\sigma_{Cost_{pgm}}$ are computed, the entire distribution function of the normal can be specified, along with the probability that $Cost_{pgm}$ may take any particular outcome, such as the protect scenario cost. The following illustrates how these equations work.

Computational Example 4-1: Assumed Normal

Suppose the distribution function for $Cost_{pgm}$ is normal. Suppose the point estimate cost of the program is 100 (\$M) and this cost was assessed to fall at the 25th percentile. Suppose the type and phase of the program is such that 30 percent variability in cost around the mean has been historically seen. Suppose the protect scenario was defined and determined to cost 145 (\$M). Given this,

- a) Compute $\mu_{Cost_{pgm}}$ and $\sigma_{Cost_{pgm}}$.
- b) Plot the distribution function of $Cost_{pgm}$.
- c) Determine the confidence level of the protect scenario cost and its associated cost reserve.

Solution

- a) From the information given and from equations 4-3 and 4-4 we have

$$\mu_{Cost_{pgm}} = x_{PE} - z_{PE} \frac{Dx_{PE}}{1 + Dz_{PE}} = 100 - z_{PE} \frac{(0.3)(100)}{1 + (0.3)z_{PE}}$$

Version 2.0.2

$$\sigma_{Cost_{pgm}} = \frac{Dx_{PE}}{1 + Dz_{PE}} = \frac{(0.3)(100)}{1 + (0.3)z_{PE}}$$

We need z_{PE} to complete these computations. From the information given, we know $P(Z \leq z_{PE}) = 0.25$. Since Z is assumed to be a standard normal random variable, we can look-up the values for z_{PE} from table A-1 (refer to Appendix A). In this case, it follows that

$$P(Z \leq z_{PE} = -0.6745) = 0.25$$

therefore, with $z_{PE} = -0.6745$ we have

$$\mu_{Cost_{pgm}} = x_{PE} - z_{PE} \frac{Dx_{PE}}{1 + Dz_{PE}} = 100 - z_{PE} \frac{(0.3)(100)}{1 + (0.3)z_{PE}} = 125.4 \text{ (\$M)}$$

$$\sigma_{Cost_{pgm}} = \frac{Dx_{PE}}{1 + Dz_{PE}} = \frac{(0.3)(100)}{1 + (0.3)z_{PE}} = 37.6 \text{ (\$M)}$$

b) A plot of the distribution function of $Cost_{pgm}$ is shown in figure 4. This is a plot of a normal distribution with mean 125.4 (\$M) and standard deviation 37.6 (\$M).

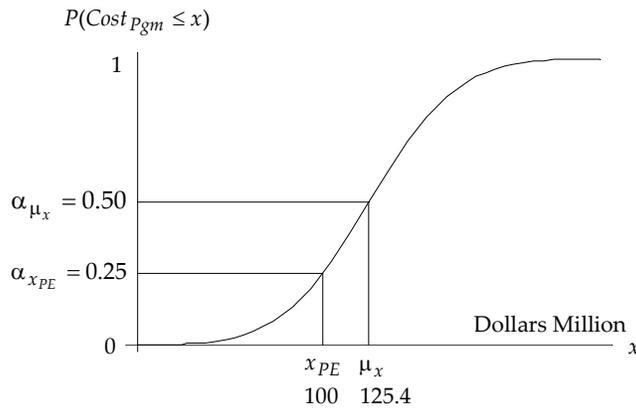


Figure 4. A Plot of the Normal Distribution: Mean 125.4, Sigma 37.6

Version 2.0.2

c) To determine the confidence level of the protect scenario we need to find $\alpha_{x_{PS}}$ such that

$$P(\text{Cost}_{pgm} \leq x_{PS} = 145) = \alpha_{x_{PS}}$$

Finding $\alpha_{x_{PS}}$ is equivalent to solving

$$\mu_{\text{Cost}_{pgm}} + z_{x_{PS}} (\sigma_{\text{Cost}_{pgm}}) = x_{PS}$$

for $z_{x_{PS}}$. From the above, we can write the expression

$$z_{x_{PS}} = \frac{x_{PS} - \mu_{\text{Cost}_{pgm}}}{\sigma_{\text{Cost}_{pgm}}} = \frac{x_{PS}}{\sigma_{\text{Cost}_{pgm}}} - \frac{1}{D} \quad (4-5)$$

Since $x_{PS} = 145$, $\mu_{\text{Cost}_{pgm}} = 125.4$, and $\sigma_{\text{Cost}_{pgm}} = 37.6$ it follows that

$$z_{x_{PS}} = \frac{x_{PS} - \mu_{\text{Cost}_{pgm}}}{\sigma_{\text{Cost}_{pgm}}} = \frac{x_{PS}}{\sigma_{\text{Cost}_{pgm}}} - \frac{1}{D} = \frac{145}{37.6} - \frac{1}{(0.3)} = 0.523$$

From the look-up table in Appendix A we see that

$$P(Z \leq z_{x_{PS}} = 0.523) \approx 0.70$$

Therefore, the protect scenario cost of 145 (\$M) falls at approximately the 70th percentile of the distribution with a cost reserve (CR) of 45 (\$M). Figure 5 shows these results graphically. This concludes example 4-1. ♣

The following provides formulas for the mean and standard deviation of Cost_{pgm} if the underlying distribution of possible program costs is represented by a lognormal. The lognormal is similar to the normal in that the $\ln(\text{Cost}_{pgm})$ is normally distributed instead of Cost_{pgm} being normally distributed. The lognormal is different than the normal distribution because it is skewed towards the positive end of the range, instead of being symmetric about the mean.

Version 2.0.2

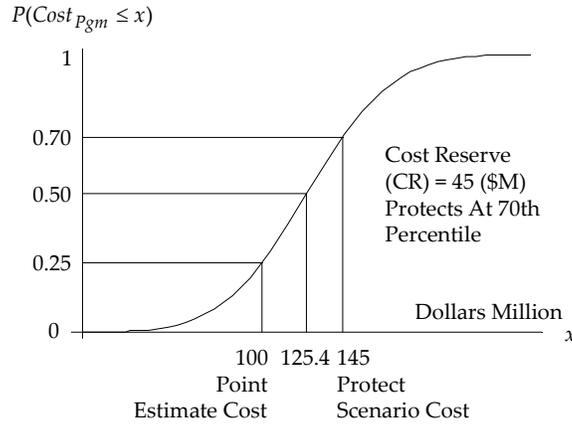


Figure 5. Example 4-1 Illustrated: Assumed *Normal* Distribution

Numerous studies [1] have empirically shown the normal or lognormal to be excellent approximations to the overall distribution function of a program’s total cost, even in the presence of correlations among cost element costs. The decision to use one over the other is really a matter of analyst judgment. In practice, it is simple enough to execute an analysis using both distributions to examine if there are significant differences between them. Then, use judgment to select the distribution that best reflects the cost and risk conditions of the program.

4.2 Formulas: Statistical SBM With An Assumed Underlying LogNormal

Here, we assume the underlying probability distribution of Cost_{pgm} is lognormally distributed and the point (x_{PE}, α_{PE}) falls along this lognormal. There are two steps involved in computing the mean and standard deviation of Cost_{pgm} . The first is to compute the mean and standard deviation of $\ln(\text{Cost}_{pgm})$. The second is to translate these values into the mean and standard deviation of Cost_{pgm} , so the units are in dollars instead of “log-dollars”.

Step 1: Formulas for the Mean and Standard Deviation of $\ln(\text{Cost}_{pgm})$

$$\mu_{\ln \text{Cost}_{pgm}} = \ln x_{PE} - z_{PE} \sqrt{\ln(1 + D^2)} \quad (4-6)$$

Version 2.0.2

$$\sigma_{\ln \text{Cost}_{pgm}} = \sqrt{\ln(1 + D^2)} \quad (4-7)$$

where D is the coefficient of dispersion (COD), x_{PE} is the program's point estimate cost, z_{PE} is the value such that $P(Z \leq z_{PE}) = \alpha_{PE}$ and Z is the standard normal random variable; that is, $Z \sim N(0,1)$. The value for z_{PE} derives from the look-up table in Appendix A.

Step 2: Once $\mu_{\ln \text{Cost}_{pgm}}$ and $\sigma_{\ln \text{Cost}_{pgm}}$ are computed, they need to be translated into "dollar-units". Equation 4-8 and equation 4-9 provide this translation [1].

$$\mu_{\text{Cost}_{pgm}} = e^{\mu_{\ln \text{Cost}_{pgm}} + \frac{1}{2}\sigma_{\ln \text{Cost}_{pgm}}^2} \quad (4-8)$$

$$\sigma_{\text{Cost}_{pgm}} = \sqrt{e^{2\mu_{\ln \text{Cost}_{pgm}} + \sigma_{\ln \text{Cost}_{pgm}}^2} (e^{\sigma_{\ln \text{Cost}_{pgm}}^2} - 1)} \quad (4-9)$$

Once $\mu_{\text{Cost}_{pgm}}$ and $\sigma_{\text{Cost}_{pgm}}$ are computed, the entire distribution function of the lognormal can be specified, along with the probability that Cost_{pgm} may take a particular outcome. The following illustrates how the last four equations work.

Computational Example 4-2: Assumed LogNormal

Suppose the distribution function for Cost_{pgm} is lognormal. Suppose the point estimate cost of the program is 100 (\$M) and this cost was assessed to fall at the 25th percentile. Suppose the type and phase of the program is such that 30 percent variability in cost around the mean has been historically seen. Suppose the protect scenario was defined and determined to cost 145 (\$M). Given this,

- a) Compute $\mu_{\text{Cost}_{pgm}}$ and $\sigma_{\text{Cost}_{pgm}}$.
- b) Plot the distribution function of Cost_{pgm} .
- c) Determine the confidence level of the protect scenario cost and its associated cost reserve.

Version 2.0.2

Solution

a) From equations 4-6 and 4-7, and example 4-1, it follows that

$$\mu_{\ln \text{Cost}_{Pgm}} = \ln x_{PE} - z_{PE} \sqrt{\ln(1 + D^2)} = \ln(100) - (-0.6745) \sqrt{\ln(1 + (0.3)^2)} = 4.80317$$

$$\sigma_{\ln \text{Cost}_{Pgm}} = \sqrt{\ln(1 + D^2)} = \sqrt{\ln(1 + (0.3)^2)} = 0.29356$$

From equations 4-8 and 4-9 we translate the above mean and standard deviation into dollar units; that is,

$$\mu_{\text{Cost}_{Pgm}} = e^{\mu_{\ln \text{Cost}_{Pgm}} + \frac{1}{2} \sigma_{\ln \text{Cost}_{Pgm}}^2} = e^{4.80317 + \frac{1}{2} (0.29356)^2} \approx 127.3 \text{ (\$M)}$$

$$\begin{aligned} \sigma_{\text{Cost}_{Pgm}} &= \sqrt{e^{2\mu_{\ln \text{Cost}_{Pgm}} + \sigma_{\ln \text{Cost}_{Pgm}}^2} (e^{\sigma_{\ln \text{Cost}_{Pgm}}^2} - 1)} \\ &= \sqrt{e^{2(4.80317) + (0.29356)^2} (e^{(0.29356)^2} - 1)} \approx 38.2 \text{ (\$M)} \end{aligned}$$

b) A plot of the distribution function of Cost_{Pgm} is shown in figure 6. This is a plot of a lognormal distribution with mean 127.3 and standard deviation 38.2

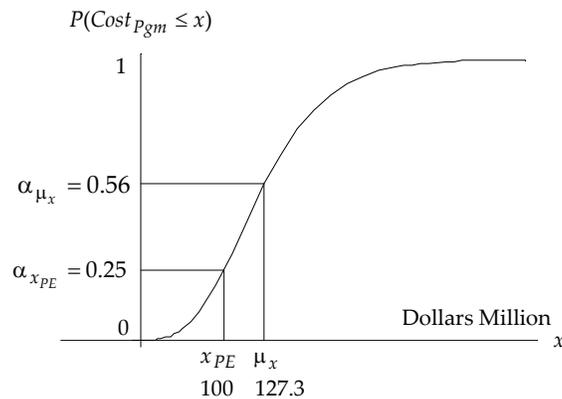


Figure 6. A Plot of the LogNormal Distribution: Mean 127.3, Sigma 38.2

c) To determine the confidence level of the protect scenario we need to find $\alpha_{x_{PS}}$ such that

Version 2.0.2

$$P(\text{Cost}_{pgm} \leq x_{PS} = 145) = \alpha_{x_{PS}}$$

Finding $\alpha_{x_{PS}}$ is equivalent to solving

$$\mu_{\ln \text{Cost}_{pgm}} + z_{x_{PS}} (\sigma_{\ln \text{Cost}_{pgm}}) = \ln x_{PS}$$

for $z_{x_{PS}}$. From the above, we can write the expression

$$z_{x_{PS}} = \frac{\ln x_{PS} - \mu_{\ln \text{Cost}_{pgm}}}{\sigma_{\ln \text{Cost}_{pgm}}}$$

Since $x_{PS} = 145$, $\mu_{\ln \text{Cost}_{pgm}} = 4.80317$, and $\sigma_{\ln \text{Cost}_{pgm}} = 0.29356$ it follows that

$$z_{x_{PS}} = \frac{\ln x_{PS} - \mu_{\ln \text{Cost}_{pgm}}}{\sigma_{\ln \text{Cost}_{pgm}}} = \frac{\ln 145 - 4.80317}{0.29356} = 0.59123$$

From the look-up table in Appendix A we see that

$$P(Z \leq z_{x_{PS}} = 0.59123) \approx 0.723$$

Therefore, the protect scenario cost of 145 (\$M) falls at approximately the 72nd percentile of the distribution with a cost reserve (CR) of 45 (\$M). Figure 7 shows these results graphically. This concludes example 4-2. ♣

4.3 A Sensitivity Analysis

There are many ways to design and perform a sensitivity analysis on the SBM, particularly the statistical SBM. In this mode, one might vary the statistical inputs, namely α_{PE} and/or the COD. From experience, we know α_{PE} will often fall in the interval $0.10 \leq \alpha_{PE} \leq 0.50$. For this paper, we set $\alpha_{PE} = 0.25$ and the COD equal to 0.30 to illustrate the statistical aspects of the SBM. In practice, these measures will vary for each program – not only as a function of the program's type (e.g., space, C4ISR) but its maturity and phase along the acquisition timeline.

Version 2.0.2

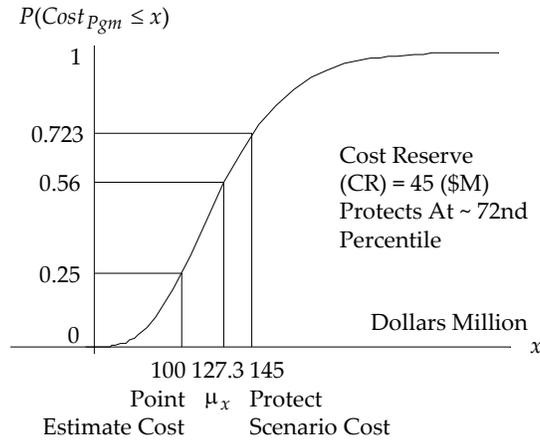


Figure 7. Example 4-2 Illustrated: Assumed *LogNormal* Distribution

The following shows a sensitivity analysis on the statistical SBM with varying levels of the coefficient of dispersion, COD. This is done in the context of example 4-2. Figure 8 illustrates how either the confidence level can vary as a function of the COD or how the dollar level can vary as a function of the COD. Here, the left-most family of lognormal distributions, in figure 8, shows for a protect scenario cost of 145 (\$M) the confidence level can range from 0.545 to 0.885 depending in the magnitude of the COD.

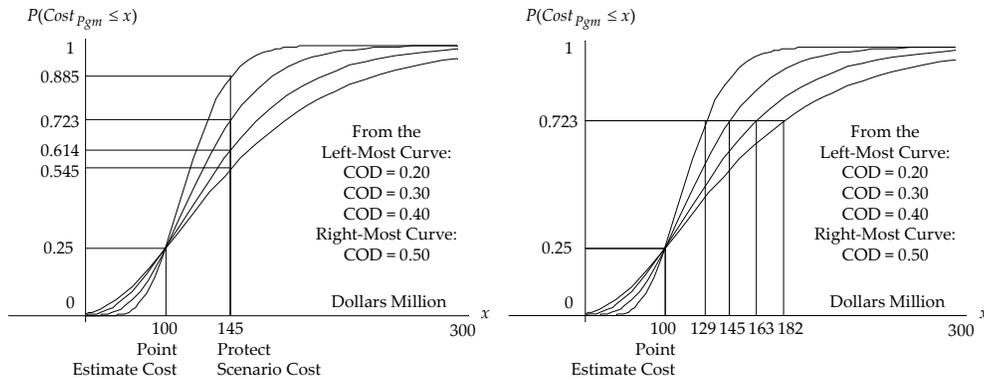


Figure 8. A Sensitivity Analysis on the Coefficient of Dispersion:
 Families of *LogNormal* Distributions

Version 2.0.2

The right-most family of lognormal distributions, in figure 8, shows for a confidence level of just over 70 percent the dollars can range from 129 (\$M) to 182 (\$M), depending on the magnitude of the COD.

The above analysis is intended to demonstrate the sensitivity of the analysis results to wide variations in the coefficient of dispersion. In practice, a program would not experience such wide swings in COD values. However, it is good practice to vary the COD by some amount around the “point” value to see what possible variations in confidence levels or dollars results*.

As a good practice point a sensitivity analysis should always be conducted, especially when implementing the statistical SBM. The analysis can signal where additional refinements to scenarios, and the underling analytical assumptions, may be needed. This is what good analysis is all about!!

5.0 Summary

This paper presented an approach for performing an analysis of a program’s cost risk. The approach is referred to as the scenario-based method (SBM). It provides program managers and decision-makers a scenario-based assessment of the amount of cost reserve needed to protect a program from cost overruns due to risk. The approach can be applied without the use of advanced statistical concepts, or Monte Carlo simulations, yet is flexible in that confidence measures for various possible program costs can be derived.

Features of this approach include the following:

** This analysis was based on the assumption that a program’s cost uncertainty could be represented by a lognormal distribution. It is important to note the lognormal is bounded by zero; hence, cost will always be non-negative. In a sensitivity analysis, such as the one presented here, it is possible the coefficient of dispersion could be so large as to drive program costs into negative values if an underlying normal is assumed, since the normal distribution is an infinite distribution at both tails. As the SBM is tested and implementation experiences with the approach are collected, it may be decided the lognormal distribution assumption is the “better” of the two, in most cases.*

Version 2.0.2

- Provides an analytic argument for deriving the amount of cost reserve needed to guard against well-defined “scenarios”;
- Brings the discussion of “scenarios” and their credibility to the decision-makers; this is a more meaningful topic to focus on, instead of statistical abstractions the classical analysis can sometimes create;
- Does not require the use of statistical methods to develop a valid measure of cost risk reserve; this is the non-statistical SBM;
- Percentiles (confidence measures) can be designed into the approach with a minimum set of statistical assumptions;
- Percentiles (as well as the mean, median (50th%), variance, etc.) can be calculated algebraically and thus can be executed in near-real time within a simple spreadsheet environment; Monte Carlo simulation is not needed;
- Does not require analysts develop probability distribution functions for all the uncertain variables in a WBS, which can be time-consuming and hard to justify;
- Correlation is indirectly captured in the analysis by the magnitude of the coefficient of dispersion applied to the analysis;
- The approach fully supports traceability and focuses attention on key risk events that have the potential to drive cost higher than expected.

In summary, the Scenario Based Method encourages and emphasizes a careful and deliberative approach to cost risk analysis. It requires the development of scenarios that represent the program’s “risk story” rather than debating what percentile to select. Time is best spent building the case arguments for how a confluence of risk events might drive the program to a particular percentile. This is where the debate and the analysis should center. This is how a program manager and decision-maker can rationalize the need for cost reserve levels that may initially exceed expectations. It is also a vehicle for identifying where risk mitigation actions should be implemented to reduce cost risk and the chances of program costs becoming out of control.

Version 2.0.2

References

- [1] Garvey, Paul R., *Probability Methods for Cost Uncertainty Analysis: A Systems Engineering Perspective*, 2000, Marcel Dekker, Inc., Taylor & Francis Group, Chapman-Hall/CRC Press, New York, N.Y., 10016.
- [2] Hitch, C. J. 1955. *An Appreciation of Systems Analysis*, P-699. Santa Monica, California: The RAND Corporation.

About the Author...

Paul R. Garvey is Chief Scientist, and a Director, for the Center for Acquisition and Systems Analysis at The MITRE Corporation. Mr. Garvey is internationally recognized and widely published in the application of decision analytic methods to problems in systems engineering risk management. His articles in this area have appeared in numerous peer-reviewed journals, technical books, and recently in John Wiley & Son's Encyclopedia of Electrical and Electronics Engineering.

Mr. Garvey authored the textbook "Probability Methods for Cost Uncertainty Analysis: A Systems Engineering Perspective", published in 2000 by Marcel Dekker, Inc., New York, NY. He is currently authoring a second textbook "Analytical Methods for Risk Management: A Systems Engineering Perspective" to be published by Taylor & Francis Group, Dekker/Chapman-Hall/CRC-Press, New York, NY.

Mr. Garvey completed his undergraduate and graduate degrees in mathematics and applied mathematics at Boston College and Northeastern University, where for ten years he was a member of the part-time faculty in the Department of Mathematics.

Version 2.0.2

Appendix A

Cumulative Distribution Function of the Standard Normal Random Variable

The tables below are values of the cumulative distribution function of the standard normal random variable “Z”. Here, $Z \sim N(0, 1)$. The columns with three-digits represent values for “z”. The columns with the eight-digits are equal to the probability given by the integral below.

$$P(Z \leq z) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy$$

Since $Z \sim N(0, 1)$ the following is true; $P(Z \leq -z) = P(Z > z) = 1 - P(Z \leq z)$.

0.00	0.5000000	0.21	0.5831661	0.42	0.6627572	0.63	0.7356528
0.01	0.5039894	0.22	0.5870644	0.43	0.6664021	0.64	0.7389138
0.02	0.5079784	0.23	0.5909541	0.44	0.6700314	0.65	0.7421540
0.03	0.5119665	0.24	0.5948348	0.45	0.6736448	0.66	0.7453732
0.04	0.5159535	0.25	0.5987063	0.46	0.6772419	0.67	0.7485712
0.05	0.5199389	0.26	0.6025681	0.47	0.6808225	0.68	0.7517478
0.06	0.5239223	0.27	0.6064198	0.48	0.6843863	0.69	0.7549030
0.07	0.5279032	0.28	0.6102612	0.49	0.6879331	0.70	0.7580364
0.08	0.5318814	0.29	0.6140918	0.50	0.6914625	0.71	0.7611480
0.09	0.5358565	0.30	0.6179114	0.51	0.6949743	0.72	0.7642376
0.10	0.5398279	0.31	0.6217195	0.52	0.6984682	0.73	0.7673050
0.11	0.5437954	0.32	0.6255158	0.53	0.7019441	0.74	0.7703501
0.12	0.5477585	0.33	0.6293000	0.54	0.7054015	0.75	0.7733727
0.13	0.5517168	0.34	0.6330717	0.55	0.7088403	0.76	0.7763728
0.14	0.5556700	0.35	0.6368306	0.56	0.7122603	0.77	0.7793501
0.15	0.5596177	0.36	0.6405764	0.57	0.7156612	0.78	0.7823046
0.16	0.5635595	0.37	0.6443087	0.58	0.7190427	0.79	0.7852362
0.17	0.5674949	0.38	0.6480272	0.59	0.7224047	0.80	0.7881447
0.18	0.5714237	0.39	0.6517317	0.60	0.7257469	0.81	0.7910300
0.19	0.5753454	0.40	0.6554217	0.61	0.7290692	0.82	0.7938920
0.20	0.5792597	0.41	0.6590970	0.62	0.7323712	0.83	0.7967307

Table A-1. Table of Standard Normal Values (continued on next page)

Example Computations

1. $P(Z \leq z = -0.525) = P(Z > z = 0.525) = 1 - P(Z \leq z = 0.525) = 1 - 0.70 = 0.30$
2. $P(Z \leq z = -0.675) = P(Z > z = 0.675) = 1 - P(Z \leq z = 0.675) = 1 - 0.75 = 0.25$
3. $P(Z \leq z = 0.525) = 0.70$

Version 2.0.2

0.84	0.7995459	1.05	0.8531409	1.26	0.8961653	1.47	0.9292191
0.85	0.8023375	1.06	0.8554277	1.27	0.8979576	1.48	0.9305633
0.86	0.8051055	1.07	0.8576903	1.28	0.8997274	1.49	0.9318879
0.87	0.8078498	1.08	0.8599289	1.29	0.9014746	1.50	0.9331928
0.88	0.8105704	1.09	0.8621434	1.30	0.9031995	1.51	0.9344783
0.89	0.8132671	1.10	0.8643339	1.31	0.9049020	1.52	0.9357445
0.90	0.8159399	1.11	0.8665004	1.32	0.9065824	1.53	0.9369916
0.91	0.8185888	1.12	0.8686431	1.33	0.9082408	1.54	0.9382198
0.92	0.8212136	1.13	0.8707618	1.34	0.9098773	1.55	0.9394292
0.93	0.8238145	1.14	0.8728568	1.35	0.9114919	1.56	0.9406200
0.94	0.8263912	1.15	0.8749280	1.36	0.9130850	1.57	0.9417924
0.95	0.8289439	1.16	0.8769755	1.37	0.9146565	1.58	0.9429466
0.96	0.8314724	1.17	0.8789995	1.38	0.9162066	1.59	0.9440826
0.97	0.8339768	1.18	0.8809998	1.39	0.9177355	1.60	0.9452007
0.98	0.8364569	1.19	0.8829767	1.40	0.9192433	1.61	0.9463011
0.99	0.8389129	1.20	0.8849303	1.41	0.9207301	1.62	0.9473839
1.00	0.8413447	1.21	0.8868605	1.42	0.9221961	1.63	0.9484493
1.01	0.8437523	1.22	0.8887675	1.43	0.9236414	1.64	0.9494974
1.02	0.8461358	1.23	0.8906514	1.44	0.9250663	1.65	0.9505285
1.03	0.8484950	1.24	0.8925122	1.45	0.9264707	1.66	0.9515428
1.04	0.8508300	1.25	0.8943502	1.46	0.9278549	1.67	0.9525403

1.68	0.9535214	1.89	0.9706211	2.10	0.9821356	2.31	0.9895559
1.69	0.9544861	1.90	0.9712835	2.11	0.9825709	2.32	0.9898296
1.70	0.9554346	1.91	0.9719335	2.12	0.9829970	2.33	0.9900969
1.71	0.9563671	1.92	0.9725711	2.13	0.9834143	2.40	0.9918025
1.72	0.9572838	1.93	0.9731967	2.14	0.9838227	2.50	0.9937903
1.73	0.9581849	1.94	0.9738102	2.15	0.9842224	2.60	0.9953388
1.74	0.9590705	1.95	0.9744120	2.16	0.9846137	2.70	0.9965330
1.75	0.9599409	1.96	0.9750022	2.17	0.9849966	2.80	0.9974448
1.76	0.9607961	1.97	0.9755809	2.18	0.9853713	2.90	0.9981341
1.77	0.9616365	1.98	0.9761483	2.19	0.9857379	3.00	0.9986500
1.78	0.9624621	1.99	0.9767046	2.20	0.9860966	3.10	0.9990323
1.79	0.9632731	2.00	0.9772499	2.21	0.9864475	3.20	0.9993128
1.80	0.9640697	2.01	0.9777845	2.22	0.9867907	3.30	0.9995165
1.81	0.9648522	2.02	0.9783084	2.23	0.9871263	3.40	0.9996630
1.82	0.9656206	2.03	0.9788218	2.24	0.9874546	3.50	0.9997673
1.83	0.9663751	2.04	0.9793249	2.25	0.9877756	3.60	0.9998409
1.84	0.9671159	2.05	0.9798179	2.26	0.9880894	3.70	0.9998922
1.85	0.9678433	2.06	0.9803008	2.27	0.9883962	3.80	0.9999276
1.86	0.9685573	2.07	0.9807739	2.28	0.9886962	3.90	0.9999519
1.87	0.9692582	2.08	0.9812373	2.29	0.9889894	4.00	0.9999683
1.88	0.9699460	2.09	0.9816912	2.30	0.9892759	5.00	0.9999997

Table A-1. Table of Standard Normal Values (concluded)