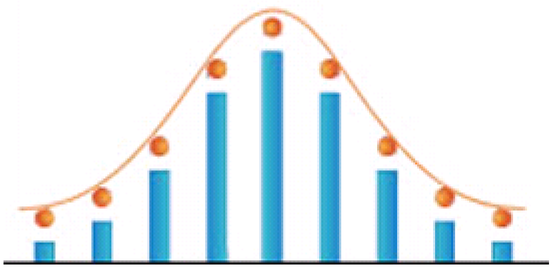
Lesson Objectives

This lesson provides an overview of various Probability Distributions and how they are used for cost elements.

- Examine how cost is treated as a probability distribution
- Explain how a Total Cost Distribution is developed
- Identify four typical types of probability distributions to represent cost elements



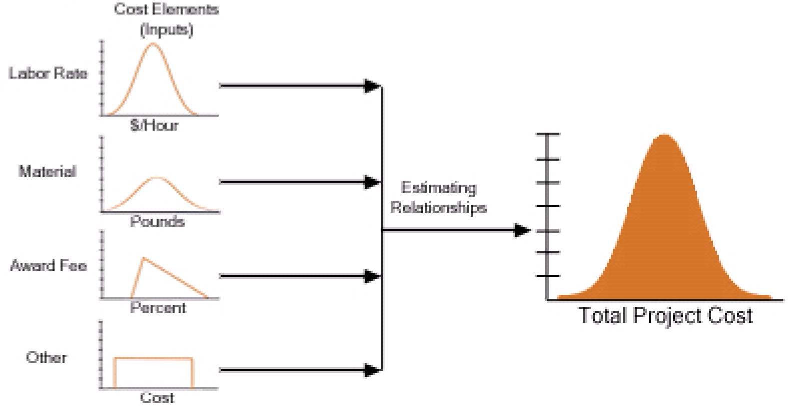
Cost as a Probability Distribution

The graphic illustrates that different random variables (with different probability distributions) and their uncertainties can be combined into a probability distribution representing the total potential project cost.

This is the finished product of a Cost Risk Analysis – a distribution representing the possible total system cost.

The cost of a system can be significantly affected by uncertainty. This uncertainty implies that costs (or any parameter) will vary over some range of values. This range of possible values allows us to think of cost as a random variable over this range.

Multiple Probability Distribution



How do we show the chance (probability) that a particular cost in this range of possible costs will be realized? **One method is with a probability distribution** - a distribution that represents a range of values and associated probabilities.

The following pages will examine the characteristics of different PD's that are typically used to represent a cost element, and methods to develop a Total Cost Distribution.

Sequential Development

We must do <u>two</u> things with respect to these PDF's in order to complete a Cost Risk Analysis.

STEP 1: Identify a Probability Density Function (PDF) for each uncertain cost element in the cost estimate.

To do this, you must:

a. Identify high, low, and most likely values (most likely values are normally the values computed in the cost estimate)

b. Choose a shape for the PDF

STEP 2: Combine the input PDFs into a Total Cost PDF

There are two ways to do this:

a. Use the method of Symmetric Approximation

b. Use the method of Monte Carlo Simulation

High, Low, and Most Likely

One of the first steps in developing a total cost distribution is to identify the PDF's for each uncertain cost element. These PDF's will be described by either two or three values consisting of a High, Low, and Most Likely.

To identify the High and Low values associated with the cost element, specialty experts are employed.

- Have experts verbalize the risks associated with each cost element

Experts need to state:

- What could go wrong
- What breakthroughs are possible
- What is certain about this element

Experts need to identify/list the inputs that affect the cost element:

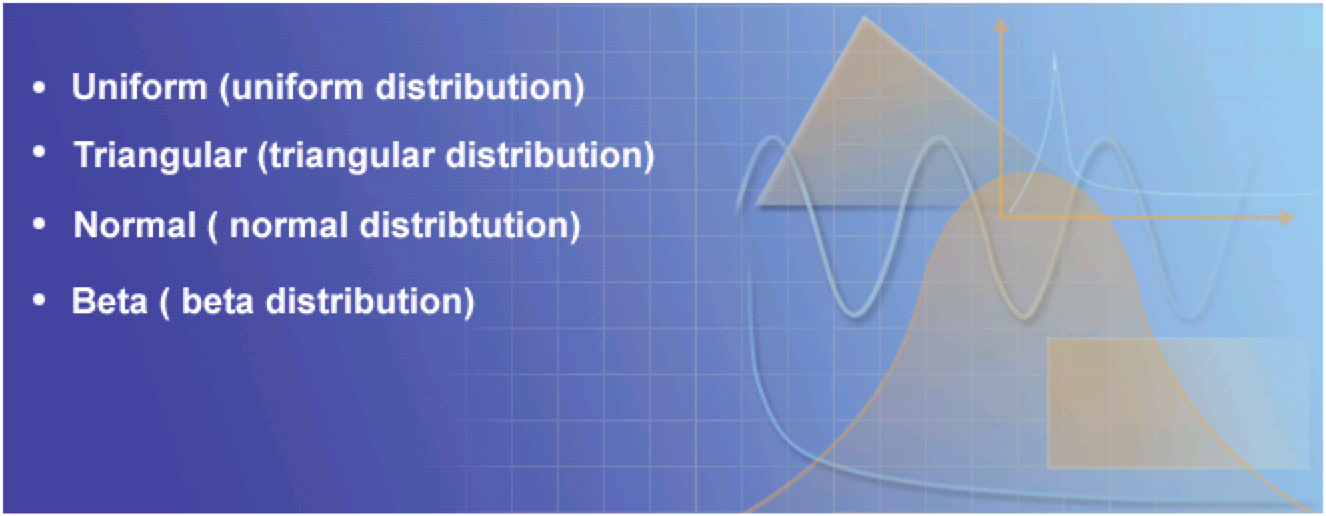
- Input examples: system weight, award fees, or composition of material
- Translate the identified risks into possible values for the inputs
    a. Translation becomes more reliable when technical experts are involved and boundaries well defined
    b. Translated risks help identify high and low values that bound the most likely value from the estimate

PDF Shapes

Next step, choose the shape of a distribution that the range of values for a specific cost element would follow.

This will be the Probability Density Function (PDF) and it represents the distribution of a cost element's potential range of values. There are four PDF shapes that are typically used to represent uncertain cost elements.

These four do not represent all possible PDFs that could be used (e.g. the lognormal distribution is sometimes seen in cost risk analysis)



Uniform Distribution

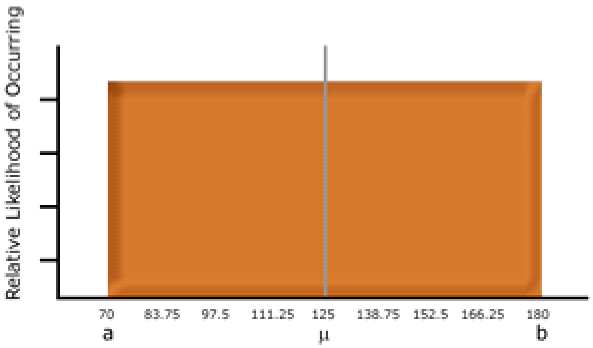
Characteristics:

- All outcomes between high and low are equally likely
- Parameters are a (the low value) and b (the high value)
- x coordinates represent costs
- y coordinates represent the likelihood of occurence

Application:

- Use when there is no information about the relative likelihood of possible outcomes across the range of possible values

Uniform Distribution

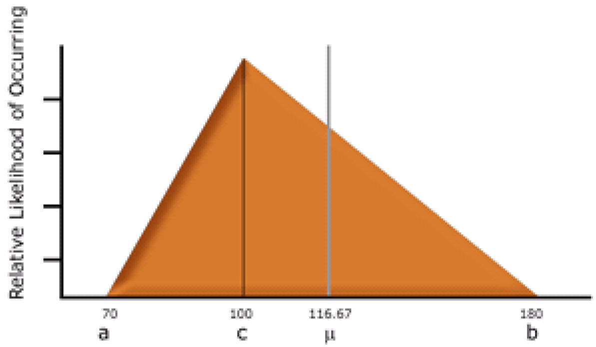


Triangular Distribution

Characteristics:

- Simple to apply
- Parameters are the high (b), low (a) and most likely (c)
- Can be of any shape between end points with varying degrees of variance and skewness (size of the tails) (PDF's with more area in the distribution tails have more probability of outcomes further from the most likely value.)
- x coordinates represent costs
- y coordinates represent the likelihood of occurrence

Triangle Distribution



Application:

- For a wide range of cost elements and variables
- Can be shaped to fit most any potential cost

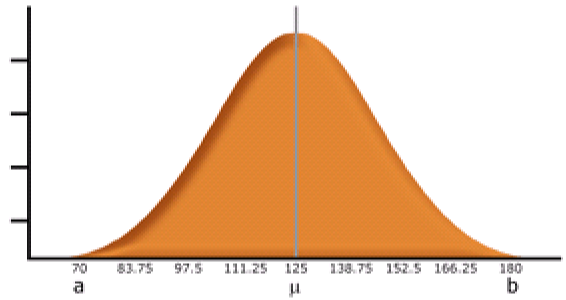
Normal Distribution

Also known as a Bell Curve.

Characteristics:

- Symmetrical – both sides of the mean are identical
- Must be used cautiously for costs because costs generally are not symmetric in nature
- Parameters required are the mean and standard deviation (these parameters will be calculated using the low, high and most likely values)
- x coordinates represent costs
- y coordinates represent the likelihood of occurrence

Normal Distribution



Application:

More accurate when measurement errors are used as such as the measurements of Mean Time Between Failures (MTBF).

Popup Text

Mean Time Between Failures (MTBF)

The mean (average) time between failures of a system. Calculations of MTBF assume that the system is fixed, after each failure, and returned to service immediately after each failure.

Beta Distribution

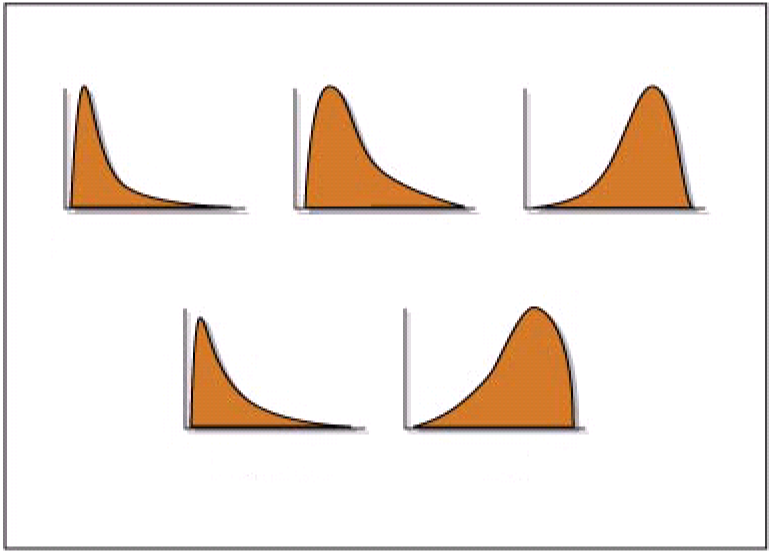
Characteristics:

- Most flexible of the distribution shapes – it can take many forms
- Difficult to specify parameters, α and β
- In practice, assume PERT Beta

PERT (Program Evaluation and Review Technique) Beta – uses low, most likely and high estimates as parameters.

Application:

- Wide range of applications – most cost functions can be described by PERT Beta
- We will use PERT Beta, which is a function of the low, high and most likely estimates (L, H, ML)



Example shapes of the Beta distribution

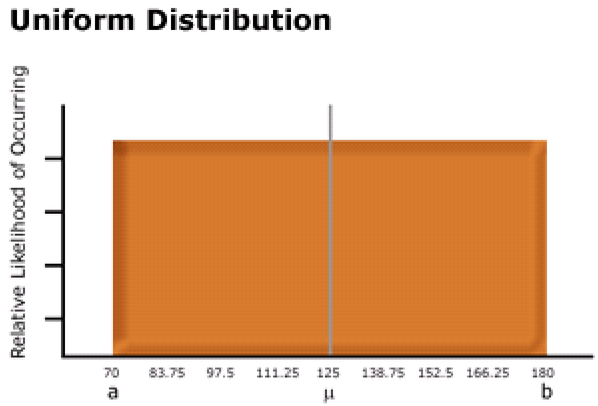
Uniform Model

To develop a Total Cost Distribution using the Symmetric Approximation technique, each element must have the mean (μ), and variance (σ^2) calculated. The high, low and most likely values are inserted into algorithms which approximate the distributions respective parameters.

Select the "next" button to view the algorithms for a **Uniform Distribution**.

Uniform Distribution:

- Note: There is no most likely value.



Uniform Algorithms

Uniform Distribution - Calculating mean and variance

where a = min value (low), b = max value (high), μ = mean (average), and σ^2 = variance

There is no mode. Reason: All y values are equivalent. Frequency does not change.

μ

Average (mean) cost = (low + high)/2

$$\mu = \frac{(a+b)}{2}$$

$$\mu = \frac{(70 + 180)}{2}$$

$$\mu = 125$$

σ^2

Variance = σ^2 = (high - low)^2/12

$$\sigma^2 = \frac{(b-a)^2}{12}$$

$$\sigma^2 = \frac{(180-70)^2}{12}$$

$$\sigma^2 = \frac{(110)^2}{12}$$

$$\sigma^2 = \frac{12,110}{12}$$

$$\sigma^2 = 1,008.33$$

D

Long Description

Uniform Algorithms

Uniform Distribution - Calculating mean and standard deviation

where a = min value (low); b = max value (high), μ = mean, and σ = s.d.
There is no mode. Reason: All y values are equivalent. Frequency does not change.

μ

Average (mean) cost = (low + high)/2

$$\mu = \frac{(a + b)}{2}$$

$$\mu = \frac{(70 + 180)}{2}$$

$$\mu = 125$$

σ^2

Variance = σ^2 = (high - low)^2/12

$$\sigma^2 = \frac{(b - a)^2}{12}$$

$$\sigma^2 = \frac{(180 - 70)^2}{12}$$

$$\sigma^2 = \frac{(110)^2}{12}$$

$$\sigma^2 = \frac{12,110}{12}$$

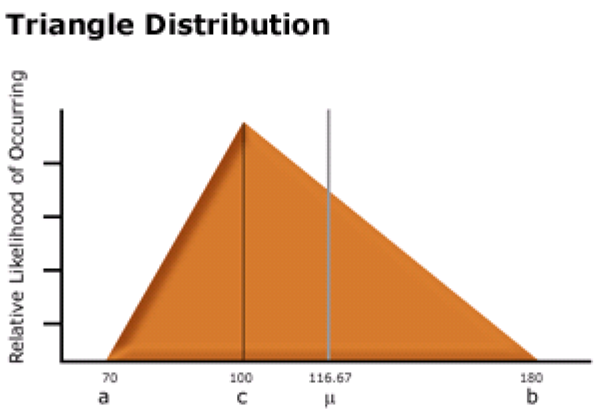
$$\sigma^2 = 1,008.33$$

Triangular Model

Triangular Distribution:

- Represents wide range of possible distribution shapes

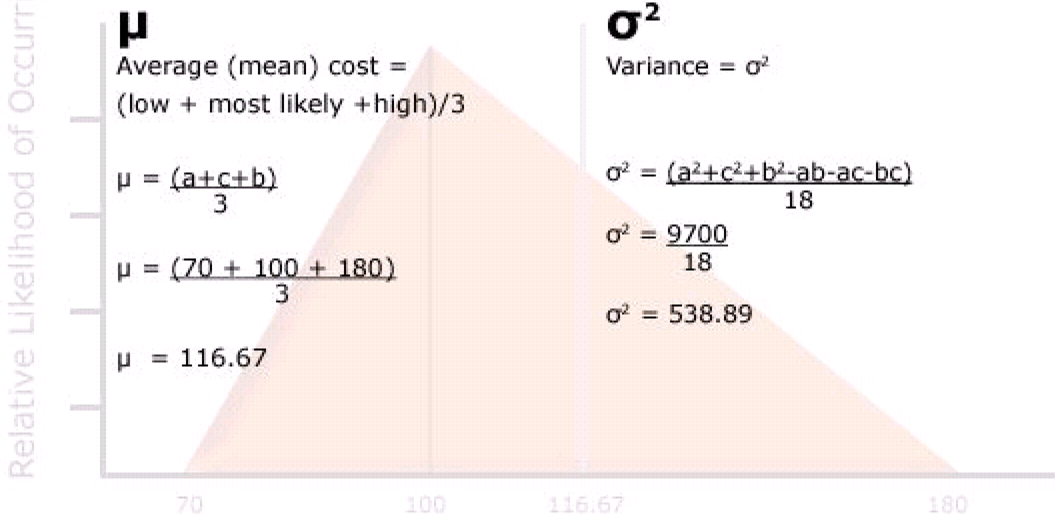
Select "next" to view the algorithms for a Triangular Distribution.



Triangular Algorithms

Triangle Distribution - Calculating mean and variance

where a = min value (cost), b = maximum value (cost), c = mode (most likely) (cost), μ = mean, and σ^2 = Variance



μ

Average (mean) cost =
(low + most likely +high)/3

$$\mu = \frac{(a+c+b)}{3}$$

$$\mu = \frac{(70 + 100 + 180)}{3}$$

$$\mu = 116.67$$

σ^2

Variance = σ^2

$$\sigma^2 = \frac{(a^2+c^2+b^2-ab-ac-bc)}{18}$$

$$\sigma^2 = \frac{9700}{18}$$

$$\sigma^2 = 538.89$$

D

Long Description

Triangle Distribution - Calculating mean and variance

where a = min value (cost), b = maximum value (cost), c = mode (most likely cost), μ = mean, and σ^2 = variance

μ | **σ^2**

Average (mean) cost = (low + most likely + high)/ 3

$$\mu = \frac{(a + c + b)}{3}$$

$$\mu = \frac{(70 + 180 + 100)}{3}$$

$$\mu = 116.67$$

Variance = σ^2

$$\sigma^2 = \frac{(a^2 + c^2 + b^2 - ab - ac - bc)}{18}$$

$$\sigma^2 = \frac{9700}{18}$$

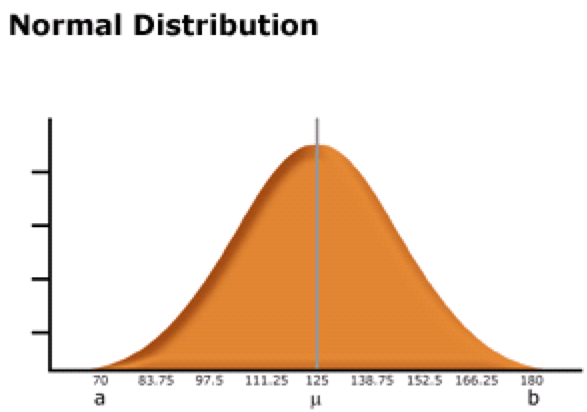
$$\sigma^2 = 538.88$$

Normal Model

Normal Distribution:

- Also known as a Bell Curve

Select the "next" button to view the algorithms for a Normal Distribution.



Normal Algorithms

Normal Distribution - Calculating mean and variance

μ

where c approximates the mean of the normal distribution

c = most likely value

Average (mean) $\mu = c$

$\mu = 125$

σ^2

where a = min value (low)
b = max value (high)
μ = mean (average)

Variance = σ^2

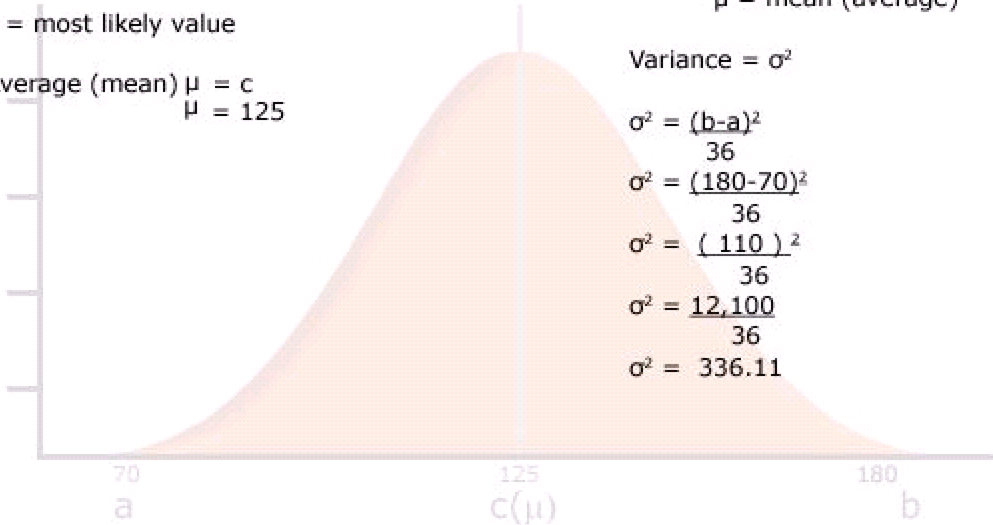
$$\sigma^2 = \frac{(b-a)^2}{36}$$

$$\sigma^2 = \frac{(180-70)^2}{36}$$

$$\sigma^2 = \frac{(110)^2}{36}$$

$$\sigma^2 = \frac{12{,}100}{36}$$

$$\sigma^2 = 336.11$$



D

Long Description

Normal Distribution - Calculating mean and variance

μ

where c approximates the mean of the normal distribution (high)

c = most likely value

Average (mean) $\mu = c$

$\mu = 125$

σ^2

where a = min value (low)

b = max value

μ = mean (average)

Variance = σ^2

$$\sigma^2 = \frac{(b - a)^2}{36}$$

$$\sigma^2 = \frac{(180 - 70)^2}{36}$$

$$\sigma^2 = \frac{(110)^2}{36}$$

$$\sigma^2 = \frac{12{,}100}{36}$$

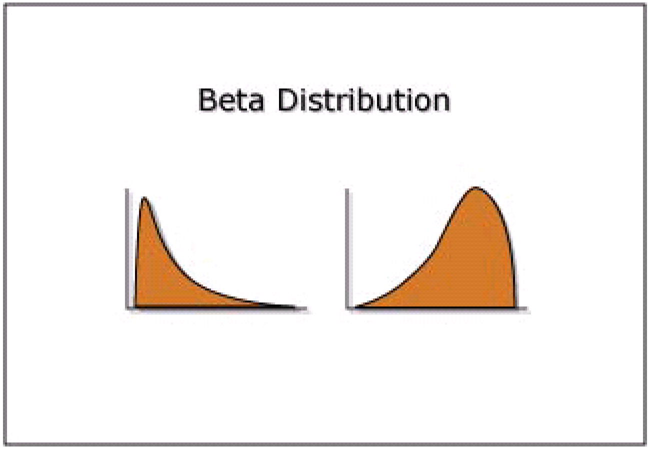
$$\sigma^2 = 336.11$$

Beta Model

Beta Distribution:

- Beta shapes are determined by α and β parameters - which are difficult to determine
- Use low, most likely and high estimates

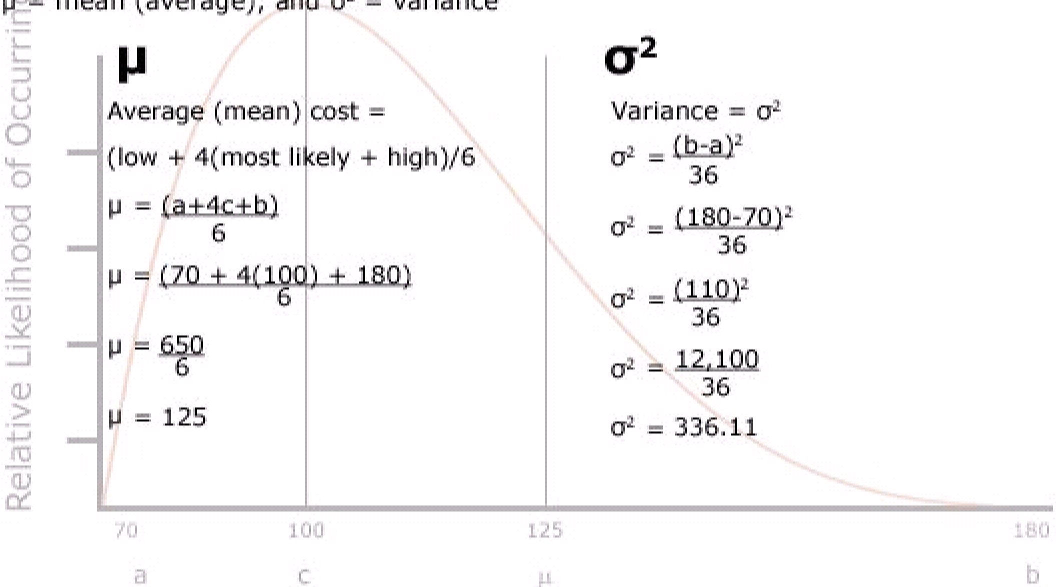
Select the "next" button to view the algorithms for a Beta Distribution.



Beta Algorithms

Beta (Pert) Distribution - Calculating mean and variance

where a = min value (low), b = max value (high), c = most likely value, μ = mean (average), and σ^2 = variance



μ

Average (mean) cost =

(low + 4(most likely + high)/6

$$\mu = \frac{(a+4c+b)}{6}$$

$$\mu = \frac{(70 + 4(100) + 180)}{6}$$

$$\mu = \frac{650}{6}$$

$$\mu = 125$$

σ^2

Variance = σ^2

$$\sigma^2 = \frac{(b-a)^2}{36}$$

$$\sigma^2 = \frac{(180-70)^2}{36}$$

$$\sigma^2 = \frac{(110)^2}{36}$$

$$\sigma^2 = \frac{12{,}100}{36}$$

$$\sigma^2 = 336.11$$

D

Long Description

Calculating mean and standard deviation

where a = min value (cost), b = maximum value (cost), c = mode (most likely cost) μ = mean, and σ ² = variance

μ

Average (mean) cost = (low + 4(most likely) + high)/6

$$\mu = \frac{(a + 4c + b)}{6}$$

$$\mu = \frac{(70 + 4(125) + 180)}{6}$$

$$\mu = \frac{750}{6}$$

$$\mu = 125$$

σ²

Variance = σ^2

$$\sigma^2 = \frac{(b - a)^2}{36}$$

$$\sigma^2 = \frac{(180 - 70)^2}{36}$$

$$\sigma^2 = \frac{(110)^2}{36}$$

$$\sigma^2 = \frac{12{,}100}{36}$$

$$\sigma^2 = 336.11$$

Symmetric Approximation

The second step in developing a total cost distribution is to combine all the identified risks and their associated ranges into a single distribution. One technique is **Symmetric Approximation**.

Symmetric Approximation is also known as the **Summation of Moments**. PDF's have four moments; 1st – Mean, 2nd – Variance, 3rd – Coefficient of Skewness (symmetry), and 4th – Coefficient of Kurtosis (height).

Data of the Symmetric Approximation method is placed in a 'linear' table of calculations. Work Breakdown Structure cost elements are listed with their distribution type, mean and variance. The means and variances are summed (the "summation of moments") which describe an approximate normal distribution. Probability statements can then be made concerning funding levels. This procedure assumes that all summed elements are independent of each other. This will not normally be the case and additional techniques must be used to determine the impact of dependence among elements. This calculation is beyond the scope of this module.

Click here to view an example.

Popup Text

Uncertainty Analysis by Symmetric Approximation Example

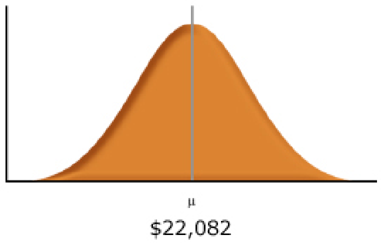
Uncertainty Analysis
by
Symmetric Approximation
Example

Component	High	Most Likely	Low	Mean μ	Standard Deviation σ	Variance σ^2	Distribution
A	$9,500	$8,000	$7,000	$8,083	$417	$173,611	Beta
B	4,000	3,250	2,500	3,250	433	187,500	Uniform
C	800	500	400	533	67	4,444	Beta
D	7,000	4,000	3,000	4,333	667	444,444	Beta
E	2,000	1,500	1,300	1,550	117	13,611	Beta
F	6,000	4,000	3,000	4,333	624	388,889	Triangular
Σ		$21,250		$22,082		$1,212,499	

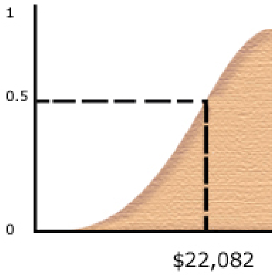
Standard Deviation = $\sigma = \sqrt{1,212,499} = 1,101.14$

Therefore, the distribution representing Total System Cost is an approximate Normal distribution with a mean (μ) of $22,082 and a Standard Deviation of 1,101.14.

The Normal probability distribution would look like...



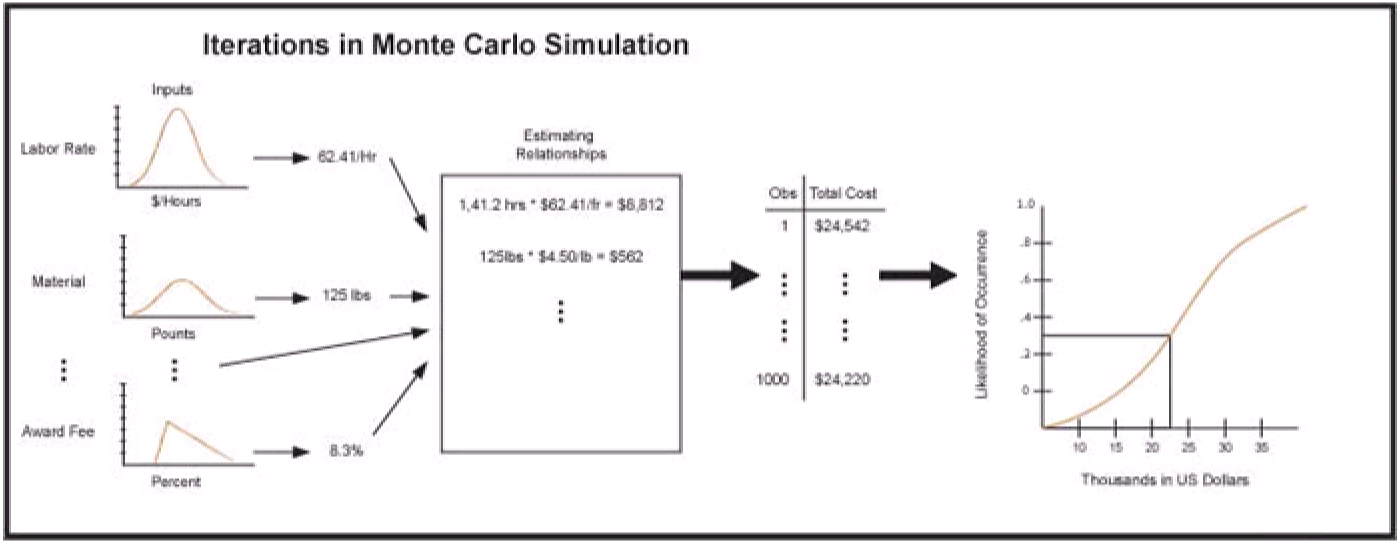
The Normal cumulative probability distribution would look like...



Probabilities of different funding levels can now be determined.

Monte Carlo Simulation

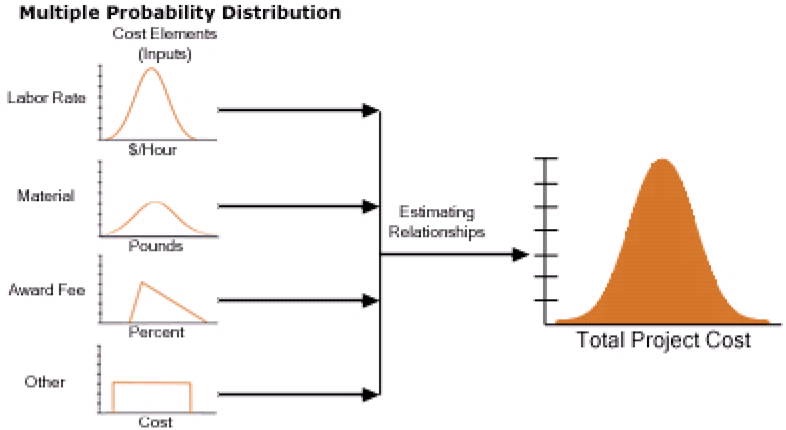
A distribution is defined for each cost element from which a random sample is drawn. The samples from each cost element's distribution are summed to a total cost. This sampling and summing process is repeated many times (e.g., 1000-10000 times). The result is a distribution representing the total cost of the system with all described uncertainties taken into account. The distribution can be displayed by a cumulative probability distribution.



Knowledge Review

Which input would **not** affect cost?

- [ ] Award fees
- [ ] Composition of materials
- [ ] System weight
- [x] None of the above



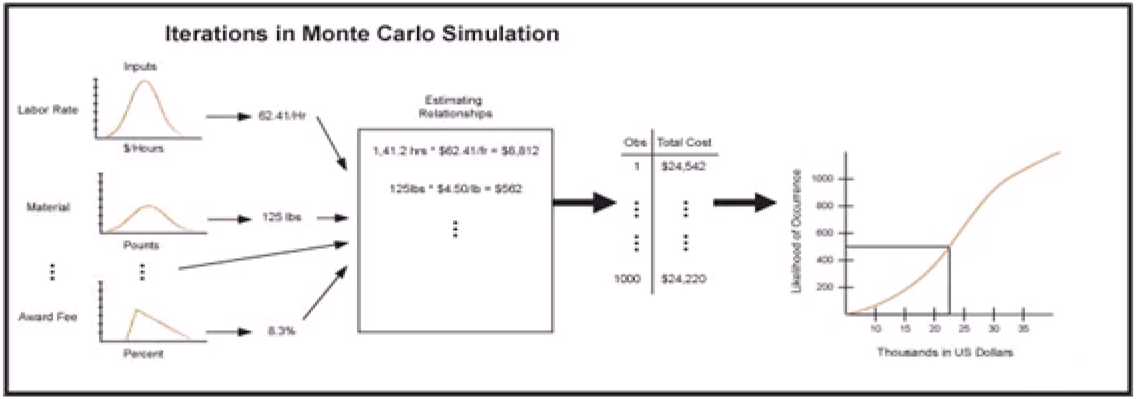
Check Answer

Any of the inputs would affect cost.

Knowledge Review

A distribution representing a system's total cost with uncertainties included is the sum of many individual cost element distributions.

- [x] True
- [ ] False



Check Answer

The answer is **True**. This is the definition of the total system cost distribution.

Knowledge Review

Each probability distribution has distinct characteristics. Choose the correct match below and submit your answer.

- [ ] High and Low values are equally likely - Normal Distribution
  More accurate when MTBF measurements are used - Uniform Distribution
  Wide range of distribution shapes with potentially large tails - Beta Distribution
  Parameters are difficult to determine, use PERT approximation - Triangular Distribution

- [x] High and Low values are equally likely - Uniform Distribution
  More accurate when MTBF measurements are used - Normal Distribution
  Wide range of distribution shapes with potentially large tails - Triangular Distribution
  Parameters are difficult to determine, use PERT approximation - Beta Distribution

Check Answer

The answer is: **High and Low values are equally likely - Uniform Distribution**; **More accurate when MTBF measurements are used - Normal Distribution**; **Wide range of distribution shapes with potentially large tails - Triangular Distribution**; and **Parameters are difficult to determine, use PERT approximation - Beta Distribution**.

Summary

To complete a Cost Risk Analysis....

- Identify a Probability Density Function for each uncertain element in the cost estimate
- Combine the element uncertainties into a Total Cost PDF

There are two mathematical approaches:

- Symmetric Approximation (Summation of Moments)
- Monte Carlo Simulation

Uniform Distribution – High and Low values are equally likely

- Used when there is no likelihood information

Triangular Distribution – Includes all three parameters; high, low and most likely value

- Tails can be 'heavy' or 'fat'

Normal Distribution – Bell Curve

- Symmetrical on both sides of the mean
- Requires mean and standard deviation

Summary, Cont.

Beta Distribution – can take on many possible shapes

- Requires α and β
- Assume PERT Beta; use low, most likely and high estimates

Table of Algorithms for calculating means and variances for four distributions

Distribution	Mean	Variance
Uniform	$\mu = \frac{(a+b)}{2}$	$\sigma^2 = \frac{(b-a)^2}{12}$
Triangular	$\mu = \frac{(a+c+b)}{3}$	$\sigma^2 = \frac{(a^2+b^2+c^2-ab-ac-bc)}{18}$
Normal	$\mu = c$	$\sigma^2 = \frac{(b-a)^2}{36}$
Beta	$\mu = \frac{(a+4c+b)}{6}$	$\sigma^2 = \frac{(b-a)^2}{36}$

D

Long Description

Table of Algorithms for each distribution shape:

Distribution: Uniform
Mean: $\mu = (a + b) / 2$
Variance: $\sigma^2 = (b-a)^2 / 12$

Distribution: Triangular
Mean: $\mu = (a + c + b) / 3$
Variance: $\sigma^2 = (a^2 + b^2 + c^2 - ab - ac - bc) / 18$

Distribution: Normal
Mean: $\mu = c$
Variance: $\sigma^2 = (b - a)^2 / 36$

Distribution: Beta
Mean: $\mu = (a + 4c + b) / 6$
Variance: $\sigma^2 = (b - a)^2 / 36$

Lesson Completion

You have completed the content for this lesson.

To continue, select another lesson from the Table of Contents on the left.

If you have closed or hidden the Table of Contents, click the Show TOC button at the top in the Atlas navigation bar.