120th ASEE Annual Conference & Exposition
FRANKLY, WE DO GIVE A D*MN
June 23-26, 2013

Paper ID #5752

Converting Point Estimates for Cost-Risk Analysis

Dr. Robert C. Creese, West Virginia University

Dr. Robert C. Creese is Professor of Industrial Engineering and Graduate Program Director in the Industrial and Management Systems Engineering Department in the Benjamin M. Statler College of Engineering and Mineral Resources at West Virginia University. He obtained his BS, MS, and PhD degrees from the Pennsylvania State University(1963), the University of California-Berkeley(1964) and the Pennsylvania State University(1972). He is a life member of ASEE, AACE-International and AFS as well as a member of ASM, AWS, SME, and ICEAA. He is a registered professional engineer in Pennsylvania and West Virginia and a certified cost engineer. He has also taught at Grove City College and The Pennsylvania State University and on sabbaticals and visits to Aalborg University in Denmark and VIT University in Vellore India. He has authored or co-authored over 100 papers and four books.

Converting Point Estimates for Cost-Risk Analysis

Abstract

The estimating process for most cost estimates follows the approach of determining the cost of each step in the production process, which is considered the "most likely" cost, and summing the cost of each of the steps, or "rolling up" the costs to estimate the total cost of the product. An amount of profit is then determined based upon the total cost to result in a quote to the perspective customer. It is generally accepted that the actual costs are typically higher than costs used in preparing the quote as the estimate is often low as some items have been omitted from the estimate, the estimating data used is not up-to-date, or the costs increase during the time between the estimate and the production of the product and delivery to the customer. A cost-risk analysis would help the supplier to better evaluate the risk to achieve the cost utilized in the preparation of the quote.

The solution to the problem would be to perform a formal cost-risk analysis. The formal cost-risk analysis would require that probability distributions be developed for each step of the process, developing correlations among these distributions, and sum the distributions statistically usually via the Monte Carlo simulation process. This would be a very expensive undertaking and may be cost effective for the aerospace and defense contractors, but it would not be cost effective to most medium and small businesses.

A simplified model was developed by Stephen Book for those who want the benefits of a cost-risk analysis, but cannot afford the cost or time to perform a formal cost risk analysis. The model has been modified to develop a Cost "S" Curve from the traditional point estimate value based upon the triangular distribution and using three parameters, H/L ratio, the percentile value for the point estimate and the percentile value for the most likely cost. This approach eliminates the need for the traditional triangular distribution parameters of the high with a specified percentile, the low with a specified percentile, and the mode. It is difficult to get estimates of the high and low values associated with percentiles, whereas the H/L ratio is easier to obtain for estimates. The results from the model include the lowest cost, the most likely cost, the median cost, the mean cost, and the highest cost estimate as well as the cost values over the entire range in five percentile increments.

The model has also been modified to develop a Cost "S" curve from the traditional point estimate value based upon the normal distribution using the H/L ratio and the percentile for the point estimate. The primary assumption made is that the distance between the high and low values is six sigma. The results from the model include the lowest cost, the mean cost, the highest cost as well as the cost values over the entire range in five percentile increments. The normal distribution is preferred for estimating the mean of a sum of components, but it is not necessarily a good estimate about the distribution.

The Triangular Distribution

The triangular distribution is a good distribution for cost estimating in that most cost estimates tend to be low and that the high cost estimate is further from the most likely or mode value than the low cost estimate. The extreme lowest cost would be zero, but the highest cost could theoretically go to infinity and thus the most likely, or mode, is closer to the low cost point than to the high cost point. The typical triangular distribution for cost estimating is illustrated in Figure 1.

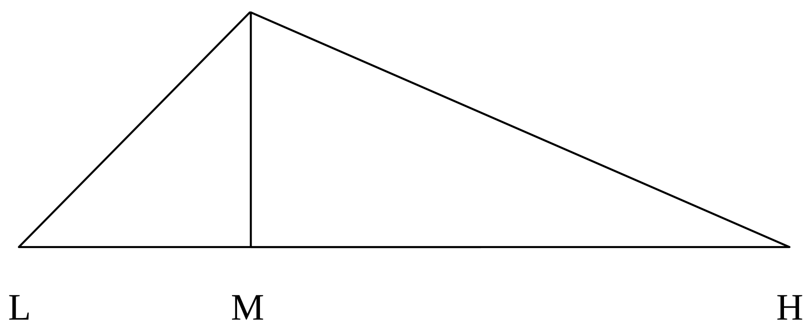

Figure 1. Triangular Distribution where L= Lowest, M=Most Likely, and H = Highest Values

To perform a risk analysis, the values for the lowest, most likely, and highest values must be known as well as the percentile for the most likely and the point estimate being evaluated for the risk analysis. The percentiles for H and L are assumed to be 100 and zero percentiles and if other values are used, they would also need to be provided. It is difficult to obtain estimates for the for the H and L values and Stephen Book[1] developed a process to calculate the distribution by using the H/L ratio, the percentile values for point estimate(P2) and the percentile for the most likely value(P1). He assumed specific values for these such as 0.25 and 0.33, but further developments have permitted these values to be varied. The modified triangle is in Figure 2 with the values for the associated percentiles included.

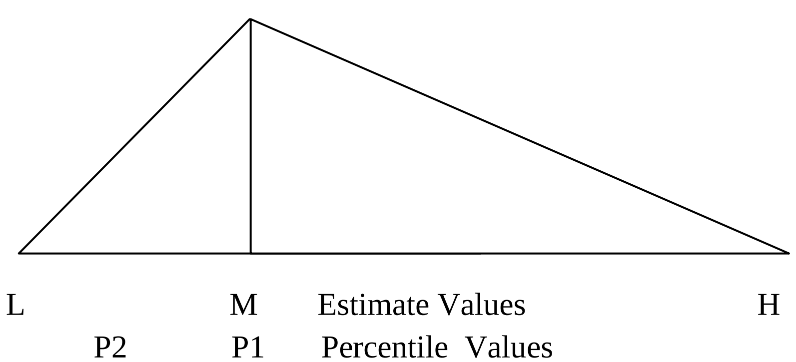

Figure 2. Triangular Distribution where L= Lowest, M=Most Likely, and H = Highest Values and P1 =Percentile for Most Likely and P2 = Percentile for Point Estimate

The two assumptions for the model are:

1 $P2 < P1 < 0.50$

2 $L < H$

The input variables for the model are described in Table 1.

Table 1. Input Variable Description and Symbols For Input Variables

Description	Symbol
1 Single Point Cost Estimate Value	PE
2 H/L Ratio (greater than 1)	$(H/L)_r$
3 The percentile value of the point estimate P2 (expressed as a decimal)	P2
4 The percentile value of the most likely value P1 (expressed as a decimal)	P1

The formulas for determination of the low, most likely or mode, high and mean values can be determined from the following expressions:

L (Low) $= PE/(1 + ((H/L)_r - 1) \times SQRT(P1 \times P2))$ (1)

H (High) $= L \times (H/L)_r$ (2)

M (Mode) $= L + P1 \times (H - L)$ (3)

Mean $= (L + H + M)/3$ (4)

To give an illustration of an application of the calculations, the input and output values obtained are presented in Table 2

Table 2. Input and Output Values for Illustrative Problem for Triangular Distribution

Input Parameter	Value	Output Parameter	Value
Single Point Estimate	1000	Lowest Value	887.1
$(H/L)_r$	1.3	Highest Value	1158.2
P1	0.45	Most Likely	1006.8
P2	0.40	Mean	1015.7
		Median	1003.7
		Standard Deviation	54.41

The risk values and the corresponding cost values are presented in Table 3 for the example. The major advantage of the $(H/L)_r$ approach is that the H and L values are calculated and one does not need to input specific values for H and L.

Table 3 Risk Probability and Cost Values for Triangular Distribution

Risk of Higher Cost	Cost	Risk of Higher Cost	Cost
0.00	887	0.50	1014
0.05	927	0.55	1021
0.10	944	0.60	1028
0.15	956	0.65	1036
0.20	967	0.70	1045
0.25	976	0.75	1055
0.30	985	0.80	1065
0.35	993	0.85	1077
0.40	1000	0.90	1091
0.45	1007	0.95	1109
0.50	1014	1.00	1153

The Normal Distribution

The triangular distribution is good for individual components, but when numerous components are summed the normal distribution is usually preferred for the estimate of the mean. It may not the best estimate for the distribution of risk, but sine nothing is know about the distributions and variances of the individual components, and since the normal distribution does not require a estimate of the percentile of the mode, one may choose to utilize the normal distribution recognizing its deficiencies. The same procedure used for the triangular distribution can be performed utilizing the H/L ratio rather than trying to obtain specific values of H and L at specific probability values. The median, mode, and mean are all at the central point of the distribution.

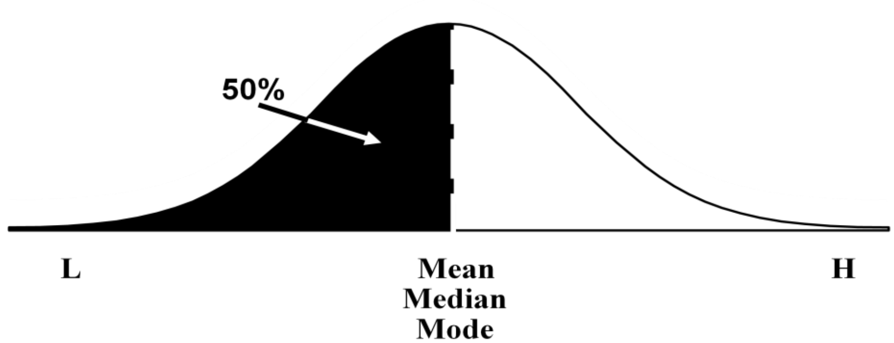

Figure 3 Normal Distribution with Low, Mean, Median, Mode, and High Values

The values for the lowest, most likely and highest values must be known as well as the percentile value for the point estimate being evaluated for the risk analysis. The percentile value for the most likely or mode is 50 percent for the normal distribution. The primary assumption made is that the high and low values are symmetric about the mean and are separated by six standard deviations. Following the approach of Dr. Book, the H/L ratio is specified rather than trying to obtain the specific high and low values. The input variables are described in Table 4.

Table 4. Variable Descriptions and Symbols for Normal Distribution Risk Analysis

Variable Description	Symbol
1. Single Point Cost Estimate	PE
2. H/L Ratio (greater than 1)	$(H/L)_r$
3. The percentile value of the point estimate P2 (expressed as a decimal)	P2
4. Lowest Estimate	L
5. Mean Estimate (mode, median)	M
6. Highest Estimate Value	H
7. Standard Normal Distribution Parameter for P2	Z(P2)
8. Standard Deviation for Distribution	σ

The formulas for determination of the low, mean, high and standard deviation values can be determined from the following expressions:

$$L = PE/((H/L)_r + 1)/2 + (Z(P2)\ x((H/L)_r - 1)/6) \tag{5}$$

$$H = L\ x\ (H/L)_r \tag{6}$$

$$M = L\ x\ (((H/L)_r /2 + 1/2) \tag{7}$$

$$\sigma = (H - L) /6 \tag{8}$$

To illustrate the calculations, the input and the output values obtained are presented in Table 5.

Table 5. Input and Output Values for Illustrative Problem with Normal Distribution

Input Parameter	Value	Output Parameter	Value
Single Point Estimate	1000	Lowest Estimate	6761
$(H/L)_r$	2.0	Highest Estimate	13522
P2	0.45	Mean Estimate	10142
		Standard Deviation	1127
		Std Normal Parameter for P2	-0.1256

The risk values and the corresponding estimate values are presented in Table.6 for the example. The major advantage of the H/L approach is that the H and L values are calculated and one does not need to input specific values.

Table 6. Risk Probability and Cost Values

Risk of Higher Cost	Cost	Risk of Higher Cost	Cost
0.00	6761	0.50	10142
0.05	8288	0.55	10283
0.10	8697	0.60	10427
0.15	8974	0.65	10576
0.20	9193	0.70	10733
0.25	9382	0.75	10902
0.30	9551	0.80	11090
0.35	9707	0.85	11310
0.40	9856	0.90	11586
0.45	10000	0.95	11995
0.50	10142	1.00	13522

Conclusions

The traditional process for doing risk analysis involves giving estimates for the high and low values of the variable and the percentile values for those estimates and it is difficult to obtain the specific values required. In addition, the point estimate is usually lower than the mean or mode and thus a percentile must be specified for that as well. The process demonstrated uses the H/L ratio instead of the specific H and L values and uses this value to calculate the H and L values. The process has been demonstrated for both the triangular and normal distribution, which are most commonly used in risk analysis. Caution must be utilized in using the normal distribution as although it is the best estimate for the mean, it is not a good estimate for a group of asymmetric estimates. This permits an easy method to do risk analysis and calculates the high and low values instead of trying to estimate the high and low values and their corresponding percentile values.

Bibliography

1. Stephen A. Book, "How to Make Your Point Estimate Look Like a Cost-Risk Analysis" , SCEA 2004 National Conference, Society of Cost Estimating and Analysis, Manhattan Beach, CA 15-18 June 2004. { How to **Make Your Point Estimate Look Like** - Society of **Cost ...** }