



How to Make Your Point Estimate Look Like a Cost-Risk Analysis (so It Can be Used for Decisionmaking)

Stephen A. Book

MCR, LLC

390 No. Sepulveda Blvd.

El Segundo, CA 90245

(310) 640-0005 x244

Fax (310) 640-0003

sbook@mcri.com

**SCEA 2004 National Conference
Society of Cost Estimating and Analysis
Manhattan Beach, CA
15-18 June 2004**



Abstract

Until recently, sponsors and managers of space-system and other high-technology systems under consideration expected cost estimators to provide “point” estimates of costs of various architectures and/or designs at each program decision point, from the initial trade-study stage to source selection and right through to the completion of the project. Unfortunately, the “point estimate” was never precisely defined – the term never meant the same thing to everybody. For example, was it the “most likely” cost, the 50% confidence cost, the “average” cost, or what? Uncertainty and outright confusion regarding what useful information about system cost this “point” estimate was conveying made it virtually useless as a figure of merit for making decisions regarding comparing competing options, as well as for planning system budgets.

Until well into the 1990s, estimating was done using the “accounting” approach, whereby each WBS element’s estimated cost is considered to be its “most likely” cost, and then the most likely costs of all WBS elements are summed (“rolled up”) to yield an estimate of total program cost. While this particular estimate was often advertised as the “most likely” program cost, it was in fact not and, in addition, the confidence that the program could be delivered at that cost was typically somewhere between 20% and 30%. While generally thought to be a new phenomenon in the 1990s, the difficulties associated with this particular method were recognized in France as early as 1952 (R. Giguet and G. Morlat, “The Causes of Systematic Error in the Cost Estimates of Public Works,” *Annals of Bridges and Roads*, No. 5, September-October 1952, Paris, France; Translated into English from the French by W.W. Taylor, U.S. Air Force Project RAND, Santa Monica CA, March 1958).

The solution to the problem lies in treating the entire cost-estimating process statistically, a course of action that has come to be known as “cost-risk analysis.” (In fact, even the use of the term “most likely” *a fortiori* implies that other estimates are “less likely” and is therefore an admission that we are facing a statistical situation.) Probability distributions must be established to model the cost of each WBS element, correlations among these distributions estimated, and the distributions summed *statistically*, typically by Monte Carlo sampling. The result will be a probability distribution of total system cost, from which meaningful estimates of the median (50% confidence), 70th percentile (70% confidence), and other relevant quantities can be obtained. Without these confidence levels associated with each possible dollar value of cost, it is difficult to use cost estimates in the decisionmaking process.

Nevertheless, it is still quite common for cost estimates in the defense and aerospace sector to be based on the procedure of summing most likely costs. For those who don’t want to do a cost-risk analysis, but who want the benefits of having done so, such as levels of confidence associated with a range of estimates, this report offers a method of doing so.



Contents

- **Cost-Risk Analysis**
 - Why Do It?
 - What is a “Point” Estimate and Where Does It Fit?
- **What Information is Needed to Do a Cost-Risk Analysis?**
- **What Assumptions Can Circumvent the Need for Real Information?**
- **The Recommended Technique**
 - The Model
 - The Mathematics
 - The Results
- **An Example**
- **Summary and Conclusion**



Contents

- **Cost-Risk Analysis**
 - Why Do It?
 - What is a “Point” Estimate and Where Does It Fit?
- **What Information is Needed to Do a Cost-Risk Analysis?**
- **What Assumptions Can Circumvent the Need for Real Information?**
- **The Recommended Technique**
 - The Model
 - The Mathematics
 - The Results
- **An Example**
- **Summary and Conclusion**



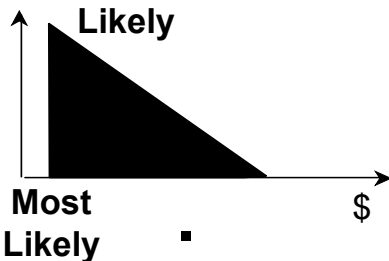
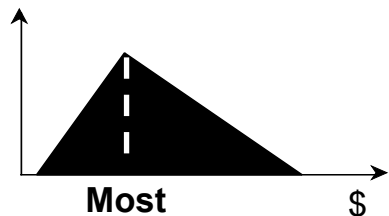
Why Do Cost-Risk Analysis?

- **The Theoretical Reason**
 - Sum of WBS-Element Most Likely Costs is NOT Most Likely Total Cost
 - For Space-Related and Other High-Technology Projects, Sum of WBS-Element Most Likely Costs is Almost Certainly an Underestimate of Actual Project Cost
 - Therefore Every Cost-Analysis Job Requires a Risk Analysis
- **The Practical Reason**
 - “Point” Estimate Not Useful for Decisionmaking, Because No Success Probability Can be Associated with It
 - Decisionmaker Must Understand Implications of Choosing One Particular Funding Level Over Several Others
 - The Only Way to Provide this Kind of Information to the Decisionmaker is to Associate a Success Probability with Each Possible Funding Level
- **Why Do Decisionmakers Need Cost-Risk Analysis?**
 - Only Percentiles Are Meaningful to Decision Makers for Budgeting and Program-Control Purposes
 - If Budget is Set at 50th-Percentile Cost, Probably of Overrun is 50%
 - If Budget is Set at 80th-Percentile Cost, Probably of Overrun is 20%

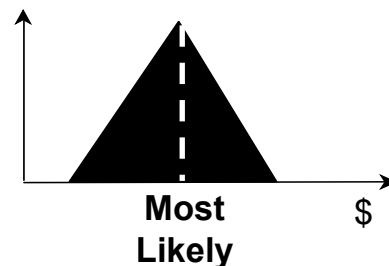


Where Does the “Point” Estimate Fit?

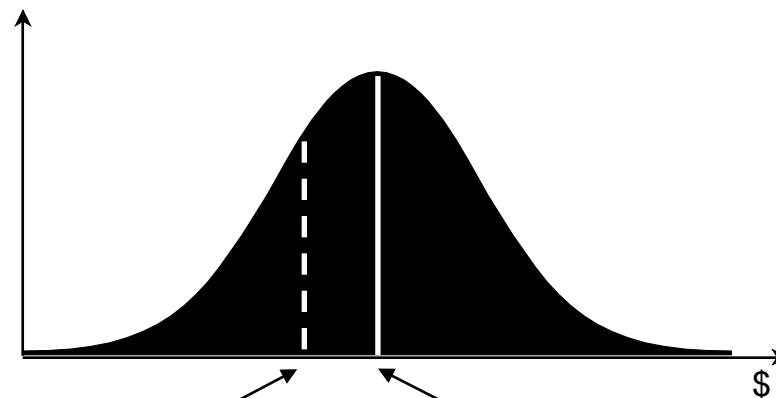
WBS-ELEMENT TRIANGULAR COST DISTRIBUTIONS



▪
▪
▪



MERGE WBS-ELEMENT COST DISTRIBUTIONS INTO TOTAL-COST NORMAL DISTRIBUTION



POINT ESTIMATE IS TYPICALLY
ROLL-UP OF MOST LIKELY
WBS-ELEMENT COSTS
(Historically and by Simulation
between the
20th and 30th Percentiles)

MOST LIKELY
TOTAL COST



Contents

- **Cost-Risk Analysis**
 - Why Do It?
 - What is a “Point” Estimate and Where Does It Fit?
- **What Information is Needed to Do a Cost-Risk Analysis?**
- **What Assumptions Can Circumvent the Need for Real Information?**
- **The Recommended Technique**
 - The Model
 - The Mathematics
 - The Results
- **An Example**
- **Summary and Conclusion**



A Project's Technical Description is Not Sufficient for Cost Estimating

- **A Technical Description (as provided in a CARD, for example) Does not Contain All Information Needed for a Realistic Cost Estimate**
- **The Technical Description Does not Describe How Difficult It is to Build the System, vis-à-vis ...**
 - ... Beyond State-of-the-Art Technology
 - ... Software Development, Integration, and Test
 - ... Other Risk Issues
- **Yet System Cost Depends Heavily on How Difficult it is to Overcome the Risk Issues**
 - Difficulty Can be Translated into Additional Money and/or Additional Time
 - Ignoring Such Difficulty Can (and Does) Lead to Cost Overruns and Schedule Slips



A Risk-Management Plan is Needed

- **A Project's Risk-Management Plan Provides Additional Information to Project Managers**
 - A “Watch List” of Risk Issues that May Cause Problems in Bringing the Project to a Successful Conclusion within Budget and on Time
 - An Assessment of How Each Listed Risk Issue Can be Circumvented or Satisfactorily Resolved
 - An Estimate of Additional Time and Resources, Including Personnel, that May Have to be Applied to Each Risk Issue
- **Information from the Risk-Management Plan also Supports the Cost-Estimating Process**
 - $(\text{Additional Time}) \times (\text{Additional Personnel}) = \text{Additional Cost}$
 - But (Statistically) Not All Risks Will Come to Pass – That's Why They are Discussed in the Risk-Management Plan, Rather than the Technical Description



Typical Technical Risk Drivers

- **Beyond-State-of-the-Art Technology Needs**
 - Data-Processing Capability
 - Communication Links
 - Temperature and other Environmental Conditions
 - Power Requirements
 - Software Solutions
- **Special Circumstances**
 - Geographic Distribution of Production Sites
 - Security Concerns
 - Procurement Quantities
- **Tight Schedules**
 - Requirements for Technology Development
 - Software Development and Testing
 - Integration and Testing of COTS Software



Typical Programmatic Risk Drivers

- **System Integration**
 - Multicontractor Teams and Organizational Interfaces
 - Conflicting Schedules and Workload
 - System Testing and Retesting
- **Limited Resources**
 - Funding
 - Trained Personnel
 - Program Funding Stretch-Out
- **Security Arrangements**
- **Supplier Viability**
- **Unforeseen Events**



Typical Cost-Estimating Risk Drivers

- **Statistical Character of Cost-Estimating Relationships (CERs)**
 - Standard Error of the Estimate
 - Bias
- **Analogies Pushed Beyond Realistic Limits**
- **Estimates Inconsistent with Historical Records**
 - Hardware: Dollars per Pound
 - Software: Lines of Code per Developer-Month
 - Communication Speed
 - Data Throughput
 - Staffing, Labor Needs and Costs
 - Maintenance Costs



Software Risk Concerns

- **Standard Cost-Estimating Paradigm for Hardware is not Applicable to Software**
 - Software Development Is Uniquely Personnel-intensive – Even Within Same Company or Workgroup, Productivity May Vary As Much As 100 to 1 Among Programmers
 - There Are No “Technical” Characteristics Such As Weight, Power, etc., that Play the Role of Cost Driver
- **Primary “Measurable” Cost Driver, Number of Lines of Code, is Notoriously Difficult to Estimate**
 - Software Requirements Cannot Be Fully Captured in Any Finite List – True List of Requirements Is Virtually Infinite
 - Software Engineers’ High Self-Esteem or Traditional Optimism Underestimates How Much Code Is Needed
 - Software Tasks Are All Nonrecurring Development – All Research and Testing, No Production
 - Initial Delivered Code Often Performs Inadequately and Fundamental Modification Costs Are Prohibitive
 - Solution is to Write More Lines of Code



Risks are Correlated

- **Resolving One WBS Element's Risk Issues by Spending More Money Often Involves Increasing Cost of Several Other Elements**
- **For Example, Technical Risks in Radar Subsystem Will Induce Weight (and Therefore) Cost Growth in Power, Platform, Software, and Other Subsystems**
- **Schedule Slippage Due to Problems in One WBS Element Lead to Cost Growth in Other Elements ("Standing Army Effect")**
- **Hardware Problems Discovered Late in Program Often Have to Be Circumvented by Making Expensive Last-minute Fixes to the Software (Which Then Show up As Software Cost Overruns)**

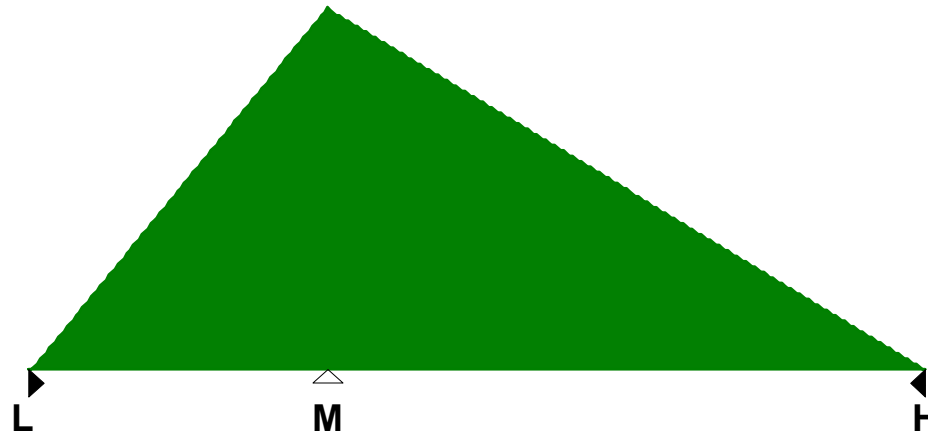


Contents

- **Cost-Risk Analysis**
 - Why Do It?
 - What is a “Point” Estimate and Where Does It Fit?
- **What Information is Needed to Do a Cost-Risk Analysis?**
- **What Assumptions Can Circumvent the Need for Real Information?**
- **The Recommended Technique**
 - The Model
 - The Mathematics
 - The Results
- **An Example**
- **Summary and Conclusion**

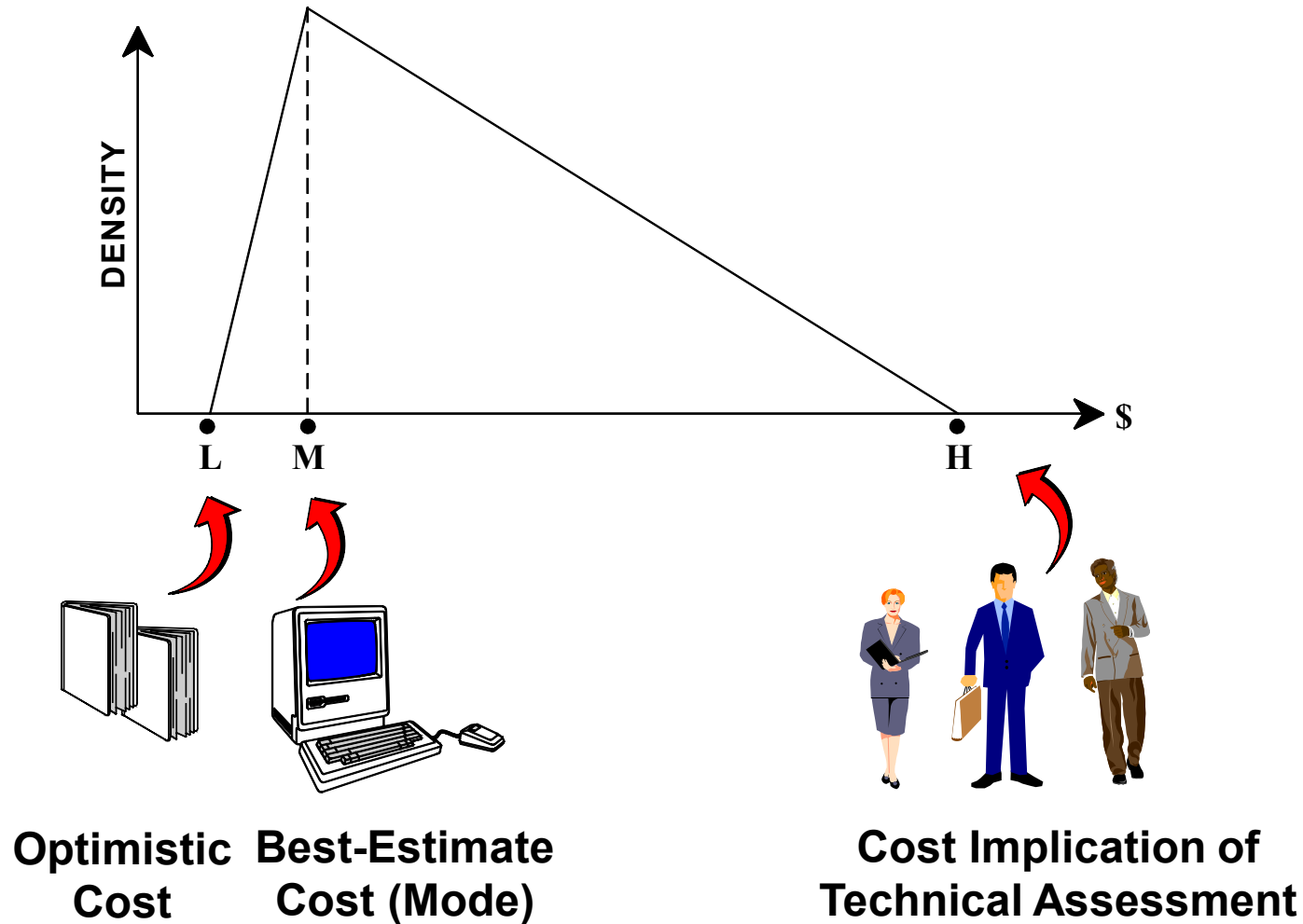
To Make Your Estimate Look Like the Result of a Cost-Risk Analysis ...

- Assume that the Total Cost Associated with Your Point Estimate has a Triangular Distribution
- Assume that Most Likely Cost is 1/3 of the Way from the Low to the High: Algebraically, $M = (2/3)L + (1/3)H$.



- Estimate, from Historical Experience or Otherwise, the Ratio of Worst-Case Cost (H) to Most Optimistic Cost (L)
- Define the Point Estimate to be at the 25th Percentile

Triangular Distribution of Element Cost

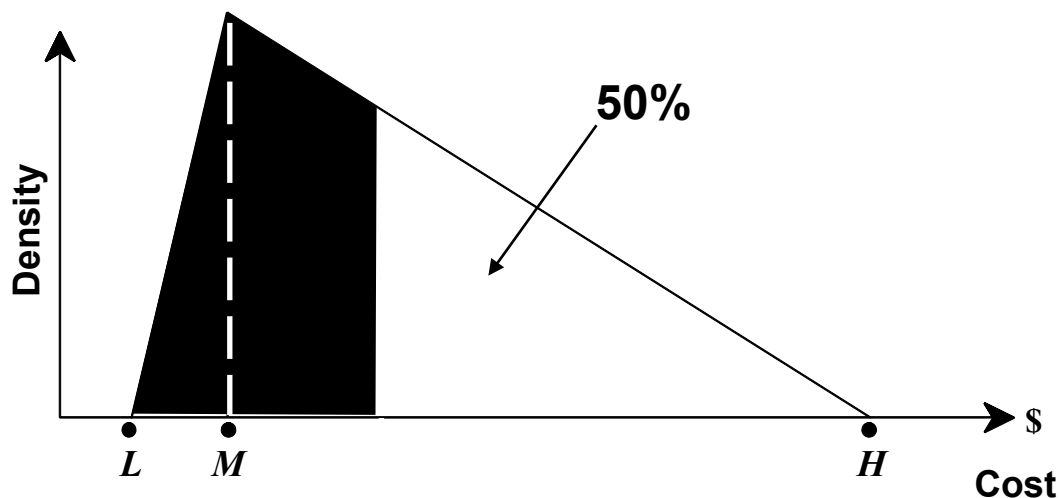




Contents

- **Cost-Risk Analysis**
 - Why Do It?
 - What is a “Point” Estimate and Where Does It Fit?
- **What Information is Needed to Do a Cost-Risk Analysis?**
- **What Assumptions Can Circumvent the Need for Real Information?**
- **The Recommended Technique**
 - **The Model**
 - **The Mathematics**
 - **The Results**
- **An Example**
- **Summary and Conclusion**

Triangular Distribution of WBS-Element Risk Impact on Cost



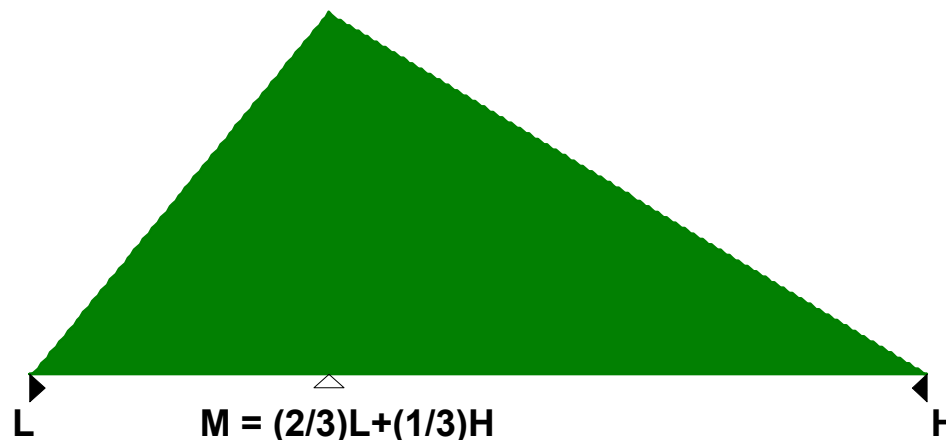
- **Graph of Probability Density Function**
- **Total Area of Triangle = 1.00**
- **Three Parameters L , M , H Completely Specify Distribution**
- **Mean, Median, Sigma, All Percentiles Can be Expressed in Terms of L , M , and H**



Recall the Basic Assumptions of the Model

- **Assume that the Total Cost Associated with Your Point Estimate has a Triangular Distribution**
- **Assume that Most Likely Cost is 1/3 of the Way from the Low to the High: Algebraically, $M = (2/3)L + (1/3)H$.**
- **Estimate, from Historical Experience or Otherwise, the Ratio of Worst-Case Cost (H) to Most Optimistic Cost (L)**
- **Define the Point Estimate to be at the 25th Percentile**

The Basic Assumptions Expressed in Mathematical Terms



- $H-L = 3(M-L)$
- $H = kL$, where k is to be Estimated
- Set Point Estimate to be the Unit of Measure, so that All Percentiles and Other Statistics are Expressed as Multiples of the Point Estimate



Statistical Metrics of the Triangular Distribution

- **Mode = M** (most likely value of cost)
- **Median = $T_{.50} = L + \sqrt{0.50(M - L)(H - L)}$** if $M - L \geq 0.50(H - L)$
 $= H - \sqrt{0.50(H - L)(H - M)}$ if $M - L \leq 0.50(H - L)$
- **$T_p =$ Dollar Value at Which $P\{\text{Cost} \leq T_p\} = p$**

$$T_p = H - \sqrt{(1 - p)(H - L)(H - M)} \quad \text{if } p \geq \frac{M - L}{H - L}$$

$$T_p = L + \sqrt{p(M - L)(H - L)} \quad \text{if } p \leq \frac{M - L}{H - L}$$

- **Mean = $\frac{L + M + H}{3}$**

$$\text{Standard Deviation} = \sigma = \sqrt{\frac{L^2 + M^2 + H^2 - LM - LH - MH}{18}}$$



Immediate Consequences of the Basic Assumptions

- **Relationships among the Triangular Parameters**

- $H-L = kL-L = (k-1)L$

- $M-L = (H-L)/3 = \left(\frac{k-1}{3}\right)L$

- $H-M = (H-L)-(M-L) = (k-1)L - [(k-1)/3]L = (2/3)(k-1)L$

- **Useful Statistical Expressions**

- For $p < (M-L)/(H-L) = 0.333333\dots (= 1/3)$,

$$T_p = L + \sqrt{p(M-L)(H-L)}$$

- For $p > (M-L)/(H-L) = 0.333333 (= 1/3)$,

$$T_p = H - \sqrt{(1-p)(H-L)(H-M)}$$



Expressions for the Percentiles

- If $p < .33333333\dots$, then

$$T_p = L + \sqrt{p(M-L)(H-L)} = L + \sqrt{\frac{p(k-1)^2}{3}} L^2$$

$$= L \left\{ 1 + (k-1) \sqrt{\frac{p}{3}} \right\}$$

- If $p > .33333333\dots$, then

$$T_p = H - \sqrt{(1-p)(H-L)(H-M)} = H - \sqrt{\frac{2(1-p)(k-1)^2}{3}} L^2$$

$$= kL - (k-1)L \sqrt{\frac{2(1-p)}{3}} = L \left\{ k - (k-1) \sqrt{\frac{2(1-p)}{3}} \right\}$$



Now to Fit the Point Estimate into This Framework ...

- **The Point Estimate is Postulated to be the 25th Percentile Estimate, so that (since $.25 < .333333\dots$)**

$$T_{0.25} = L + \sqrt{0.25(M-L)(H-L)} = L + \sqrt{\frac{0.25(k-1)^2}{3}} L^2$$
$$= L \left\{ 1 + (k-1) \sqrt{\frac{1}{12}} \right\}$$

- **From which it Follows that ...**

$$L = \frac{T_{0.25}}{\left\{ 1 + (k-1) \sqrt{\frac{1}{12}} \right\}}$$



L, M, and H Can Then be Calculated

- Start with $L = \frac{T_{0.25}}{\left\{1 + (k-1)\sqrt{\frac{1}{12}}\right\}}$

- Then $M = L \left\{1 + \frac{k-1}{3}\right\} = \frac{T_{0.25} \left\{1 + \frac{k-1}{3}\right\}}{1 + (k-1)\sqrt{\frac{1}{12}}}$

and $H = kL = \frac{kT_{0.25}}{1 + (k-1)\sqrt{\frac{1}{12}}}$



... as Well as All the Percentiles

- For $p = 0.05, 0.10, 0.15, 0.20, 0.25,$ and 0.30 (all of which are less than $0.3333333... = 1/3$),

$$T_p = L \left\{ 1 + (k - 1) \sqrt{\frac{p}{3}} \right\}$$

- For All Values of $p > 0.3333333$, Namely $p = 0.35$ and Higher,

$$T_p = L \left\{ k - (k - 1) \sqrt{\frac{2(1 - p)}{3}} \right\}$$



Applying the Assumptions and Formulas, the Triangular Descriptor Multiples are ...

If High Cost =	2*L	3*L	4*L	5*L	6*L	7*L	8*L	9*L	10*L
Then Multiples for Triangular-Distribution Parameters are as follows:									
Given: 25th	1	1	1	1	1	1	1	1	1
Low	0.775991	0.633975	0.535898	0.464102	0.40927	0.366025	0.331046	0.302169	0.277926
Most Likely	1.034654	1.056624	1.071797	1.082904	1.091386	1.098076	1.103488	1.107955	1.111705
High	1.551982	1.901924	2.143594	2.320508	2.455619	2.562178	2.64837	2.719525	2.779263
(M-L)/(H-L)	0.333333	0.333333	0.333333	0.333333	0.333333	0.333333	0.333333	0.333333	0.333333

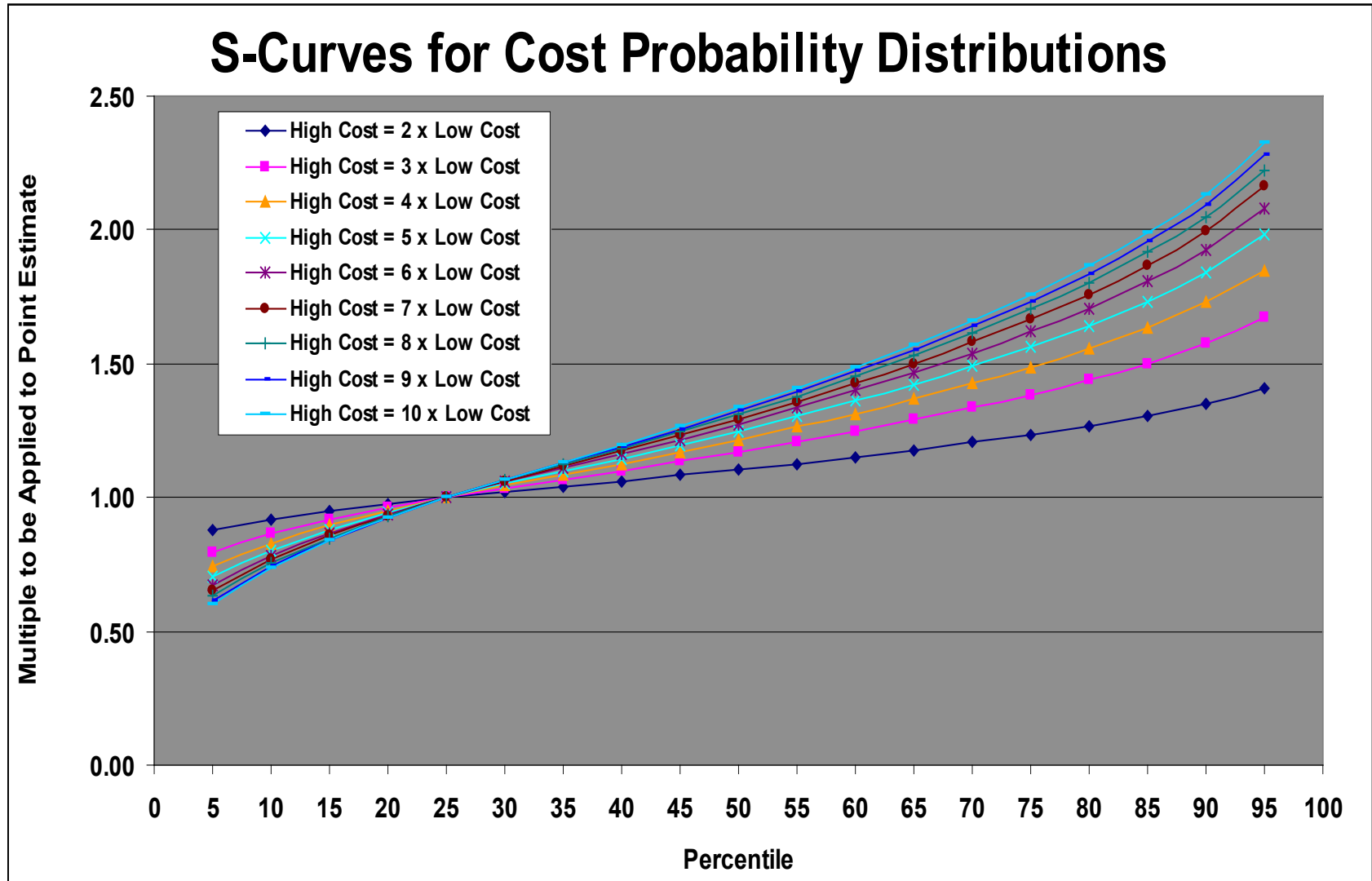


... and the Percentiles are ...

If High Cost =	2*L	3*L	4*L	5*L	6*L	7*L	8*L	9*L	10*L
Then Multiples for Percentiles and other Statistics are as follows:									
Low	0.775991	0.633975	0.535898	0.464102	0.40927	0.366025	0.331046	0.302169	0.277926
5th	0.876171	0.797666	0.743451	0.703763	0.673452	0.649547	0.630211	0.614249	0.600847
10th	0.917667	0.865469	0.829422	0.803034	0.78288	0.766986	0.75413	0.743516	0.734606
15th	0.949508	0.917497	0.89539	0.879207	0.866847	0.8571	0.849216	0.842707	0.837242
20th	0.976351	0.961358	0.951003	0.943424	0.937635	0.93307	0.929377	0.926328	0.923769
Given: 25th	1	1	1	1	1	1	1	1	1
30th	1.021381	1.034935	1.044296	1.051149	1.056382	1.06051	1.063848	1.066605	1.068918
Most Likely	1.034654	1.056624	1.071797	1.082904	1.091386	1.098076	1.103488	1.107955	1.111705
35th	1.041162	1.067257	1.085279	1.098472	1.108547	1.116493	1.122921	1.128227	1.132682
40th	1.061202	1.100002	1.126798	1.146414	1.161394	1.173209	1.182766	1.190655	1.197279
45th	1.082096	1.134142	1.170086	1.196398	1.216493	1.232341	1.245161	1.255744	1.264628
50th (Median)	1.103963	1.169873	1.21539	1.248711	1.274159	1.294229	1.310463	1.323864	1.335116
Mean	1.120876	1.197508	1.25043	1.289171	1.318758	1.342093	1.360968	1.37655	1.389631
55th	1.126954	1.207439	1.263023	1.303712	1.334787	1.359296	1.37912	1.395485	1.409224
60th	1.151262	1.247158	1.313383	1.361864	1.398889	1.42809	1.451709	1.471208	1.487578
65th	1.177142	1.289446	1.367003	1.423779	1.467139	1.501335	1.528996	1.551832	1.571003
70th	1.204948	1.33488	1.42461	1.490298	1.540464	1.580029	1.612031	1.638451	1.660631
75th	1.235185	1.384286	1.487255	1.562633	1.620201	1.665602	1.702327	1.732644	1.758097
80th	1.26863	1.438934	1.556546	1.642644	1.708398	1.760256	1.802203	1.836832	1.865904
85th	1.306592	1.500963	1.635196	1.733461	1.808507	1.867693	1.915568	1.95509	1.988271
90th	1.351622	1.574541	1.728488	1.841186	1.927254	1.995134	2.05004	2.095367	2.133421
95th	1.410306	1.670429	1.85007	1.981576	2.082009	2.161217	2.225287	2.278179	2.322583
High	1.551982	1.901924	2.143594	2.320508	2.455619	2.562178	2.64837	2.719525	2.779263



The Family of Cost-Risk “S-Curves”





Contents

- **Cost-Risk Analysis**
 - Why Do It?
 - What is a “Point” Estimate and Where Does It Fit?
- **What Information is Needed to Do a Cost-Risk Analysis?**
- **What Assumptions Can Circumvent the Need for Real Information?**
- **The Recommended Technique**
 - The Model
 - The Mathematics
 - The Results
- **An Example**
- **Summary and Conclusion**



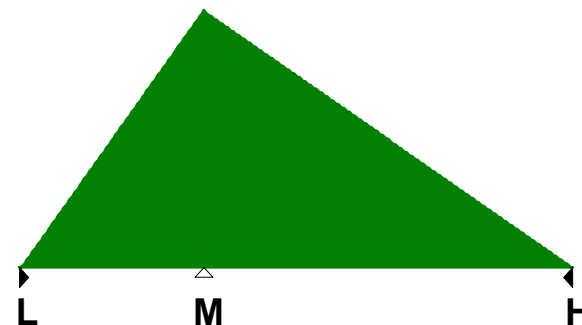
When to Apply the Technique

- **You've Computed a Point Estimate for the "Pliny the Elder" Launch Integration System, Delivering It (the Estimate) the Day It Was Due**
 - Phase I: \$23,045,658 for Hardware
 - Phase II: \$16,444,867 for Software
 - Phase III: \$8,506,213 for Testing
- **Your Estimating Team Members, Now Dispersed, are Working on Other Projects**
- **Unfortunately, the Funding Agency to Whom the Estimate is Briefed Asks ...**
 - Did You Do a Risk Analysis Focusing on Cost?
 - Do You Have Sufficient Money in the Estimate to Cover Risk?
 - How Much Money is Required, for Each Phase, to Have 50% Confidence in the Estimate for that Phase?
 - How Much Money is Required, for Each Phase, to Have 80% Confidence in the Estimate for that Phase?
- **Apply the Technique to Avoid the Embarrassment of Having to Answer "No", "I Don't Think So", "I Don't Know", "I Don't Know", Respectively**

How to Apply the Technique

- **Recall Your Point Estimates**

- Phase I: \$23,045,658 for Hardware
- Phase II: \$16,444,867 for Software
- Phase III: \$8,506,213 for Testing



- **Assess, Based on the Historical Record of Cost Experience, the Ratio H/L**

- Hardware: $H = 3L$
- Software: $H = 8L$
- Testing: $H = 2L$

- **Find the Relevant Columns You Need in the Reference Table of Percentile Multipliers**



Reference Table of Multipliers

If High Cost =	2*L	3*L	4*L	5*L	6*L	7*L	8*L	9*L	10*L
Then Multiples for Percentiles and other Statistics are as follows:									
Low	0.775991	0.633975	0.535898	0.464102	0.40927	0.366025	0.331046	0.302169	0.277926
5th	0.876171	0.797666	0.743451	0.703763	0.673452	0.649547	0.630211	0.614249	0.600847
10th	0.917667	0.865469	0.829422	0.803034	0.78288	0.766986	0.75413	0.743516	0.734606
15th	0.949508	0.917497	0.89539	0.879207	0.866847	0.8571	0.849216	0.842707	0.837242
20th	0.976351	0.961358	0.951003	0.943424	0.937635	0.93307	0.929377	0.926328	0.923769
Given: 25th	1	1	1	1	1	1	1	1	1
30th	1.021381	1.034935	1.044296	1.051149	1.056382	1.06051	1.063848	1.066605	1.068918
Most Likely	1.034654	1.056624	1.071797	1.082904	1.091386	1.098076	1.103488	1.107955	1.111705
35th	1.041162	1.067257	1.085279	1.098472	1.108547	1.116493	1.122921	1.128227	1.132682
40th	1.061202	1.100002	1.126798	1.146414	1.161394	1.173209	1.182766	1.190655	1.197279
45th	1.082096	1.134142	1.170086	1.196398	1.216493	1.232341	1.245161	1.255744	1.264628
50th (Median)	1.103963	1.169873	1.21539	1.248711	1.274159	1.294229	1.310463	1.323864	1.335116
Mean	1.120876	1.197508	1.25043	1.289171	1.318758	1.342093	1.360968	1.37655	1.389631
55th	1.126954	1.207439	1.263023	1.303712	1.334787	1.359296	1.37912	1.395485	1.409224
60th	1.151262	1.247158	1.313383	1.361864	1.398889	1.42809	1.451709	1.471208	1.487578
65th	1.177142	1.289446	1.367003	1.423779	1.467139	1.501335	1.528996	1.551832	1.571003
70th	1.204948	1.33488	1.42461	1.490298	1.540464	1.580029	1.612031	1.638451	1.660631
75th	1.235185	1.384286	1.487255	1.562633	1.620201	1.665602	1.702327	1.732644	1.758097
80th	1.26863	1.438934	1.556546	1.642644	1.708398	1.760256	1.802203	1.836832	1.865904
85th	1.306592	1.500963	1.635196	1.733461	1.808507	1.867693	1.915568	1.95509	1.988271
90th	1.351622	1.574541	1.728488	1.841186	1.927254	1.995134	2.05004	2.095367	2.133421
95th	1.410306	1.670429	1.85007	1.981576	2.082009	2.161217	2.225287	2.278179	2.322583
High	1.551982	1.901924	2.143594	2.320508	2.455619	2.562178	2.64837	2.719525	2.779263



Now Onto the Cost-Risk Estimates ...

- **Hardware (H = 3L)**
 - Point Estimate: \$23,045,658
 - 50th Percentile Estimate = $\$23,045,658 \times 1.169873$
= \$26,960,493
 - 80th Percentile Estimate = $\$23,045,658 \times 1.438934$
= \$33,161,181
- **Software (H = 8L)**
 - Point Estimate: \$16,444,867
 - 50th Percentile Estimate = $\$16,444,867 \times 1.310463$
= \$21,550,390
 - 80th Percentile Estimate = $\$16,444,867 \times 1.802203$
= \$29,636,989
- **Testing (H = 2L)**
 - Point Estimate: \$8,506,213
 - 50th Percentile Estimate = $\$8,506,213 \times 1.103963$
= \$9,390,544
 - 80th Percentile Estimate = $\$8,506,213 \times 1.268630$
= \$10,791,237



Your Response (15 minutes later) to the Funding Agency

Estimate\Phase	Hardware	Software	Testing
Point	\$23,045,658	\$16,444,867	\$8,506,213
50% Confidence	\$26,960,493	\$21,550,390	\$9,390,544
80% Confidence	\$33,161,181	\$29,636,989	\$10,791,237

- **What Other Information Do You Need, Ladies and Gentlemen?**



The Funding Agency Has Another Problem with Your Estimate

- **The Funding Agency: “Your Cost Estimate is Not Useful to Us, because We Don’t Have that Much Money in the Budget. In fact, All We Have Available for this Program is the Following:**
 - Phase I: \$25,000,000 for Hardware
 - Phase II: \$20,000,000 for Software
 - Phase III: \$10,000,000 for Testing

Can You Help Us with This Problem?”

- **You: “Now, if You Will Allow Me Another Few Minutes, I’ll Give You an Estimate of the Probability that the Available Amount for Each Phase Will be Sufficient to Fund that Phase.”**



Estimating the Level of Confidence Associated with Available Funding

- **First We Calculate the Point-Estimate Multiplier for the Available Funding**
 - Hardware: $\$25,000,000/\$23,045,658 = 1.084803$
 - Software: $\$20,000,000/\$16,444,867 = 1.216185$
 - Testing: $\$10,000,000/\$8,506,213 = 1.175611$
- **These Multiplier Values, when Compared with the Complete Table of Multipliers, Shows that**
 - \$25,000,000 for Hardware: Between 35th and 40th Percentiles
 - \$20,000,000 for Software: Between 40th and 45th Percentiles
 - \$10,000,000 for Testing: Between 60th and 65th Percentiles

If High Cost =	2*L	3*L	8*L
Then Multiples for Percentiles are as follows:			
Low	0.775991	0.633975	0.331046
5th	0.876171	0.797666	0.630211
10th	0.917667	0.865469	0.75413
15th	0.949508	0.917497	0.849216
20th	0.976351	0.961358	0.929377
Given: 25th	1	1	1
30th	1.021381	1.034935	1.063848
Most Likely	1.034654	1.056624	1.103488
35th	1.041162	1.067257	1.122921
40th	1.061202	1.10000	1.182766
45th	1.082096	1.134142	1.245161
50th (Median)	1.103963	1.169873	1.310463
Mean	1.120876	1.197508	1.360968
55th	1.126954	1.207439	1.37912
60th	1.151262	1.247158	1.451709
65th	1.177142	1.289446	1.528996
70th	1.204948	1.33488	1.612031
75th	1.235185	1.384286	1.702327
80th	1.26863	1.438934	1.802203
85th	1.306592	1.500963	1.915568
90th	1.351622	1.574541	2.05004
95th	1.410306	1.670429	2.225287
High	1.551982	1.901924	2.64837



Formula for the Level of Confidence

- The Multipliers Associated with the Available Funding Show that $p > 0.333333$ for All Phases
- Therefore We Use the Formula

$$T_p = L \left\{ k - (k-1) \sqrt{\frac{2(1-p)}{3}} \right\}$$

- We Use this Formula Differently, However, this Time Calculating p from the Known Value of T_p :

$$p = 1 - \frac{3}{2L^2} \left(\frac{kL - T_p}{k-1} \right)^2$$



Calculating the Level of Confidence

Phase\Item	L	$k = H/L$	T_p	p
Hardware	0.633975	3	1.084803	0.377
Software	0.331046	8	1.216185	0.427
Testing	0.775991	2	1.175611	0.647

- Therefore the Confidence Levels Associated with Each Phase of Available Funding are as Follows:
 - We are 37.7% Confident that the Hardware Phase of the Program Can be Completed for \$25,000,000
 - We are 42.7% Confident that the Software Phase of the Program Can be Completed for \$20,000,000
 - We are 64.7% Confident that the Testing Phase of the Program Can be Completed for \$10,000,000



Probability that Phase Costs Will Exceed Available Funding

- **In Other Words ...**
 - We Estimate that the Probability is $100\% - 37.7\% = 62.3\%$ that the Hardware Phase of the Program Will Overrun the Available Funding of \$25,000,000
 - We Estimate that the Probability is $100\% - 42.7\% = 57.3\%$ that the Software Phase of the Program Will Overrun the Available Funding of \$20,000,000
 - We Estimate that the Probability is $100\% - 64.7\% = 35.3\%$ that the Testing Phase of the Program Will Overrun the Available Funding of \$10,000,000
- **It Really Looks Like We Did a Risk Analysis, Doesn't It?**



Contents

- **Cost-Risk Analysis**
 - Why Do It?
 - What is a “Point” Estimate and Where Does It Fit?
- **What Information is Needed to Do a Cost-Risk Analysis?**
- **What Assumptions Can Circumvent the Need for Real Information?**
- **The Recommended Technique**
 - The Model
 - The Mathematics
 - The Results
- **An Example**
- **Summary and Conclusion**



Cost-Risk Analysis

- **“Cost Risk”: A Working Definition**
 - Inadequacy of Forecasted Funding Requirements to Assure That Program Can Be Completed and Meet Its Stated Objectives
- **“Cost-Risk Analysis”: A Procedure**
 - Model WBS-element Costs As Uncertain Quantities (i.e., Random Variables) That Have Probability Distributions
 - Combine WBS-element Cost Distributions Statistically (e.g., by Monte Carlo Sampling) to Generate Cumulative Distribution of Total Program Cost
 - Read off 70th Percentile Cost, 90th Percentile Cost, etc., from Cumulative Distribution to Estimate Additional Amount of Dollars Needed to Cover Risk
 - Quantify Confidence in Anybody’s “Point” Estimate of Program Cost or in Budgeted Funding



Summary

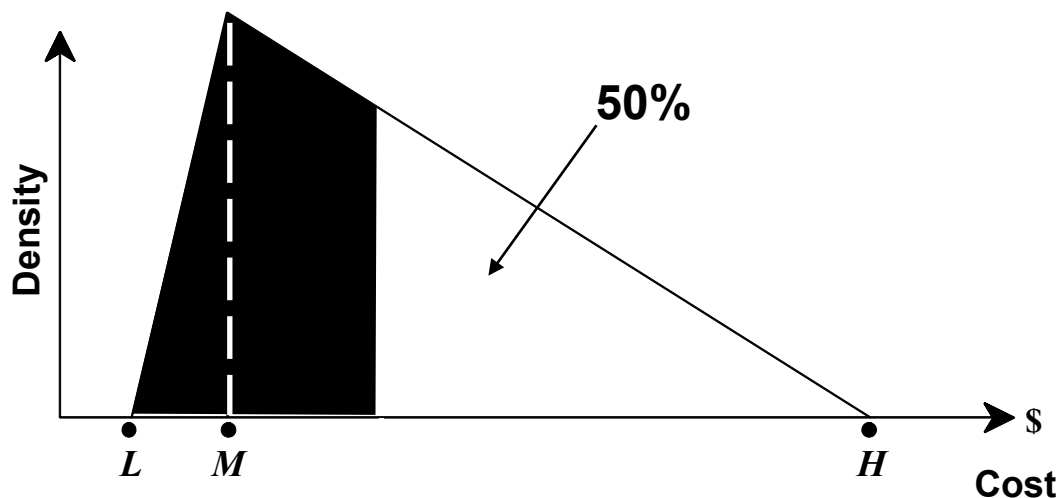
- **Do Not Sum “Most Likely” Costs**
 - Point Estimate Derived that Way is Usually Near the 25th Percentile
 - For Space-Related and Other High-Technology Programs, Point Estimate is Almost Certainly an Underestimate of Actual Project Cost
 - Therefore Every Cost-Analysis Job Requires a Risk Analysis
- **Costs are Random Variables, Not Deterministic Numbers**
 - Program Cost not Well Represented by Any Single Number
 - Triangular Probability Distribution Available to Serve as Simple Model for Cost Impacts of Technical, Programmatic Risks
- **Simple Technique Offered Here Allows Quick Calculation of Cost Percentiles, All Other Statistics, as Well as Level of Confidence Associated with Available Funding**
 - Based on Theoretical Relationships and Historical Experience
 - No Formal Detailed Technical and Programmatic Risk Analysis of the Program Itself
 - No Monte Carlo Sampling from a Variety of Distributions
 - No Credibility with People who Really Understand Cost Analysis
 - **But to Everyone Else It Looks Like You Did a Risk Analysis!**



Appendix

Basic Formulas Associated with the Triangular Probability Distribution

Triangular Distribution of WBS-Element Risk Impact on Cost



- **Graph of Probability Density Function**
- **Total Area of Triangle = 1.00**
- **Three Parameters L , M , H Completely Specify Distribution**
- **Mean, Median, Sigma, All Percentiles Can be Expressed in Terms of L , M , and H**



Statistical Metrics of the Triangular Distribution

- **Mode = M** (most likely value of cost)
- **Median = $T_{.50} = L + \sqrt{0.50(M - L)(H - L)}$** if $M - L \geq 0.50(H - L)$
 $= H - \sqrt{0.50(H - L)(H - M)}$ if $M - L \leq 0.50(H - L)$
- **$T_p =$ Dollar Value at Which $P\{\text{Cost} \leq T_p\} = p$**

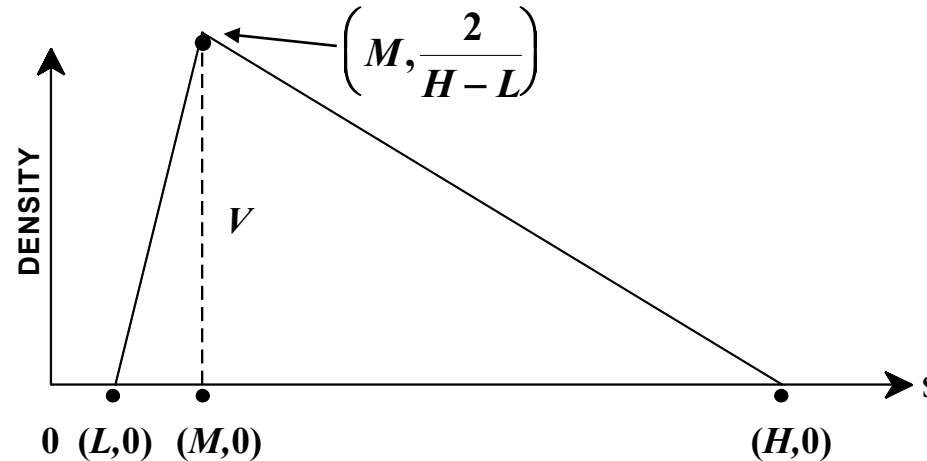
$$T_p = H - \sqrt{(1 - p)(H - L)(H - M)} \quad \text{if } p \geq \frac{M - L}{H - L}$$

$$T_p = L + \sqrt{p(M - L)(H - L)} \quad \text{if } p \leq \frac{M - L}{H - L}$$

- **Mean = $\frac{L + M + H}{3}$**

$$\text{Standard Deviation} = \sigma = \sqrt{\frac{L^2 + M^2 + H^2 - LM - LH - MH}{18}}$$

Analytic Geometry of the Triangle



- **Area of Triangle** = $\frac{1}{2}V(H-L) = 1$, so that $V = \frac{2}{H-L}$
- **Straight Line Joining** $(L, 0)$ and $\left(M, \frac{2}{H-L}\right)$ has Equation

$$\frac{y-0}{x-L} = \frac{\frac{2}{H-L} - 0}{M-L}, \quad \text{i.e., } y = \frac{2(x-L)}{(M-L)(H-L)}$$

- **Straight Line Joining** $\left(M, \frac{2}{H-L}\right)$ and $(H, 0)$ has Equation

$$\frac{y-0}{x-H} = \frac{\frac{2}{H-L} - 0}{M-H}, \quad \text{i.e., } y = \frac{2(H-x)}{(H-L)(H-M)}$$



Triangular Density Function

$$\bullet \quad f(x) = \frac{2(x-L)}{(M-L)(H-L)} \quad \text{for } L \leq x \leq M$$

$$= \frac{2(H-x)}{(H-L)(H-M)} \quad \text{for } M \leq x \leq H$$

$$= 0 \quad \text{for other values of } x$$

$$\bullet \quad P\{\text{Cost} \leq M\} = \int_L^M f(x) dx = \frac{2}{(M-L)(H-L)} \int_L^M (x-L) dx$$

$$= \frac{2}{(M-L)(H-L)} \left[\frac{x^2}{2} - Lx \right]_L^M = \frac{2 \left(\frac{M^2}{2} - LM - \frac{L^2}{2} + L^2 \right)}{(M-L)(H-L)}$$

$$= \frac{(M-L)^2}{(M-L)(H-L)} = \frac{M-L}{H-L}$$

Derivation of Formulas for Percentiles of the Triangular Distribution

- T_p = Dollar Value at Which $P\{\text{Cost} \leq T_p\} = p$

- $P\{\text{Cost} \leq T_p\} = \int_L^{T_p} f(x)dx$ if $p \leq \frac{M-L}{H-L}$

$$\text{i.e., } p = \frac{(T_p - L)^2}{(M - L)(H - L)} \quad \therefore T_p = L + \sqrt{p(M - L)(H - L)}$$

- $P\{\text{Cost} \leq T_p\} = \int_L^{T_p} f(x)dx = \frac{M-L}{H-L} + \int_M^{T_p} f(x)dx$ if $p \geq \frac{M-L}{H-L}$

$$= 1 - \int_{T_p}^H f(x)dx = 1 - \frac{2}{(H-L)(H-M)} \int_{T_p}^H (H-x)dx$$

$$\text{i.e., } p = 1 - \frac{2}{(H-L)(H-M)} \left[Hx - \frac{x^2}{2} \right]_{T_p}^H = 1 - \frac{(H - T_p)^2}{(H-L)(H-M)}$$

$$\therefore T_p = H - \sqrt{(1-p)(H-L)(H-M)}$$

Derivation of Formula for the Mean of the Triangular Distribution

$$\begin{aligned}
 \bullet \text{ Mean} &= \int_L^H xf(x)dx = \int_L^M \frac{2x(x-L)}{(M-L)(H-L)} dx + \int_M^H \frac{2x(H-x)}{(H-L)(H-M)} dx \\
 &= \frac{2}{(M-L)(H-L)} \left[\frac{x^3}{3} - \frac{Lx^2}{2} \right]_L^M + \frac{2}{(H-L)(H-M)} \left[\frac{Hx^2}{2} - \frac{x^3}{3} \right]_M^H \\
 &= \frac{2}{(M-L)(H-L)} \left(\frac{M^3}{3} - \frac{LM^2}{2} - \frac{L^3}{3} + \frac{L^3}{2} \right) + \frac{2}{(H-L)(H-M)} \left(\frac{H^3}{2} - \frac{H^3}{3} - \frac{HM^2}{2} + \frac{M^3}{3} \right) \\
 &= \frac{2}{(M-L)(H-L)} \left(\frac{2M^3 - 3LM^2 + L^3}{6} \right) + \frac{2}{(H-L)(H-M)} \left(\frac{H^3 - 3HM^2 + 2M^3}{6} \right) \\
 &= \frac{2}{H-L} \left(\frac{2M^2 - ML - L^2}{6} \right) + \frac{2}{H-L} \left(\frac{H^2 + MH - 2M^2}{6} \right) = \frac{2}{H-L} \left(\frac{H^2 - L^2 + M(H-L)}{6} \right) \\
 &= \frac{L + M + H}{3}
 \end{aligned}$$

Formula for the Second Moment of the Triangular Distribution

$$\begin{aligned}
 E(X^2) &= \int_L^H x^2 f(x) dx = \int_L^M \frac{2x^2(x-L)}{(M-L)(H-L)} dx + \int_M^H \frac{2x^2(H-x)}{(H-L)(H-M)} dx \\
 &= \frac{2}{(M-L)(H-L)} \int_L^M (x^3 - Lx^2) dx + \frac{2}{(H-L)(H-M)} \int_M^H (Hx^2 - x^3) dx \\
 &= \frac{2}{(M-L)(H-L)} \left[\frac{x^4}{4} - \frac{Lx^3}{3} \right]_L^M + \frac{2}{(H-L)(H-M)} \left[\frac{Hx^3}{3} - \frac{x^4}{4} \right]_M^H \\
 &= \frac{2}{(M-L)(H-L)} \left(\frac{M^4}{4} - \frac{LM^3}{3} - \frac{L^4}{4} + \frac{L^4}{3} \right) + \frac{2}{(H-L)(H-M)} \left[\frac{H^4}{3} - \frac{H^4}{4} - \frac{HM^3}{3} + \frac{M^4}{4} \right] \\
 &= \frac{2}{(M-L)(H-L)} \left[\frac{3M^4 - 4LM^3 + L^4}{12} \right] + \frac{2}{(H-L)(H-M)} \left[\frac{H^4 - 4HM^3 + 3M^4}{12} \right] \\
 &= \frac{3M^3 - LM^2 - L^2M - L^3}{6(H-L)} + \frac{H^3 + H^2M + HM^2 - 3M^3}{6(H-L)} \\
 &= \frac{(H^3 - L^3) + M(H^2 - L^2) + M^2(H-L)}{6(H-L)} = \frac{H^2 + HL + L^2 + M(H+L) + M^2}{6}
 \end{aligned}$$



Derivation of Formula for the Sigma Value of the Triangular Distribution

- $Var(X) = E(X^2) - \{E(X)\}^2 = E(X^2) - \left(\frac{L + M + H}{3}\right)^2$
- $\sigma^2 = E(X^2) - [E(X)]^2$

$$= \frac{H^2 + HL + L^2 + MH + ML + M^2}{6} - \left(\frac{L + M + H}{3}\right)^2$$

$$= \frac{3H^2 + 3HL + 3L^2 + 3MH + 3ML + 3M^2}{18}$$

$$- \frac{2L^2 + 2M^2 + 2H^2 + 4LM + 4LH + 4MH}{18}$$

$$= \frac{L^2 + M^2 + H^2 - LM - LH - MH}{18}$$
- $\sigma = \sqrt{\frac{L^2 + M^2 + H^2 - LM - LH - MH}{18}}$



Speaker's Bio

Dr. Stephen A. Book is Chief Technical Officer of MCR, LLC. In that capacity, he is responsible for ensuring technical excellence of MCR products, services, and processes by encouraging process improvement, maintaining quality control, and training employees and customers in cost and schedule analysis and associated program-control disciplines. Dr. Book joined MCR in January 2001 after 21 years with The Aerospace Corporation, holding the title “Distinguished Engineer” during 1996-2000 and having served as Director, Resource and Requirements Analysis Department, during 1989-1995. He has given numerous technical and tutorial presentations on cost-risk analysis and other statistical aspects of cost and economics to DoD, NASA, and EACE (European Aerospace Working Group on Cost Engineering) Cost Symposia, the AF/NASA/ESA Space Systems Cost Analysis Group (SSCAG), the U.S. Army Conference on Applied Statistics (ACAS), and professional societies such as the International Society of Parametric Analysts (ISPA), Society for Cost Estimating and Analysis (SCEA), Military Operations Research Society (MORS), U.K. Association of Cost Engineers (ACostE), and the American Institute of Aeronautics and Astronautics (AIAA). He has served on national panels as an independent reviewer of NASA programs such as the 1997-98 Cost Assessment and Validation Task Force on the International Space Station (“Chabrow Committee”) and the 1998-99 National Research Council Committee on Space Shuttle Upgrades. He is the immediate past chairman of the Risk Subgroup of SSCAG and is a member of the Economics Technical Committee of the AIAA. Dr. Book earned his Ph.D. in mathematics, with concentration in probability and statistics, at the University of Oregon.