"Creating Customer-Focused Success"

Cost Risk as a Discriminator in Trade Studies

Stephen A. Book
Chief Technical Officer
MCR, LLC
sbook@mcri.com
(310) 640-0005 x244

Joint ISPA-SCEA International Conference & Workshop
International Society of Parametric Analysts
Society of Cost Estimating and Analysis
New Orleans LA
12-16 June 1007

Abstract

Prior to formal program initiation, analysts typically undertake trade studies to investigate which of several candidate architectures or designs can best provide a desired capability at minimum cost. However, the various candidates typically differ significantly in risk as well as in cost, but members of the Government or industry trade-study team do not have the time, and the candidate solutions usually aren't sufficiently detailed at this stage, to conduct a thorough risk analyses. Yet, those differences in risk, as well as in cost, should be taken into account to the extent possible during the trade-study decision process. Because timeliness and simplicity are key requirements of analyses undertaken in support of trade studies, what usually happens is that a "point" cost estimate, or perhaps a 50%-confidence estimate, is established for each candidate, and the go-ahead decision is made on the basis of that estimate. But a nagging question remains: "What if Candidate A, the lower-cost option based on those estimates, faces risk issues that make its 70th-percentile cost higher than that of Candidate B?" In other words, Candidate B would be the lower-cost option if the cost comparison were made at the 70% confidence level. This is the classic situation in which the decision maker must choose between a low-cost, high-risk option and a high-cost, low-risk option. This report describes a methodology that allows the program manager take account of all risk scenarios by making use of all cost percentiles simultaneously, namely the entire cost probability distribution of each candidate, not simply the point estimate or the 70% confidence cost. As it turns out, the expression of system cost in terms of a lognormal or simulation-generated probability distribution makes it possible to estimate the <u>probability</u> that each candidate will turn out to be the least costly of all the options, and probabilities of that kind are the basis on which an informed decision can be made.

Contents

- **Cost as a Discriminator**
- **Special Nature of Trade Studies**
- **The Triangular Distribution**
- **Peculiarities of the Lognormal Distribution**
- **Cost Risk as a Discriminator**
- **Summary**
- **Back-up Charts**

Contents

- **Cost as a Discriminator**
- **Special Nature of Trade Studies**
- **The Triangular Distribution**
- **Peculiarities of the Lognormal Distribution**
- **Cost Risk as a Discriminator**
- **Summary**
- **Back-up Charts**

"Point" Cost Estimates

- **Planning and Acquisition Teams Require Cost Estimates to Recommend Decisions in Trade Studies, Because Cost Is a Significant Criterion that Discriminates Among the Candidates**
- **But, at the Trade-Study Stage, Confidence in any "Point" Estimate of Project Cost is Necessarily Low, due to**
  - **Technological (Im)maturity of Proposed Candidates**
  - **Design Uncertainties**
  - **Programmatic Considerations**
  - **Probable Schedule Slips of Unknown Duration**
  - **Test Failures and Other Unforeseen Events**
- **"Actual" Project Cost Will Fall Within Some Range Surrounding Any Given "Point" Estimate (with Some Degree of Confidence)**
  - **The Best We Can Hope to Do is to Understand the Uncertainty**
  - **Decisionmakers Should Take this Uncertainty Into Account**

5

Another Problem: What Does the Term "Point" Estimate Mean?

- **The "Best" Estimate! (What Does *That* Mean?)**

- **The "Most Likely" Cost? ("Mode")**

- **The 50th-Percentile Cost? ("Median")**

- **The "Expected" Cost? ("Mean")**

- **The 4th-Percentile Cost?**

- **Something Else?**

- **These Numbers are Almost Always Different**

Fact: Costs Have Probability Distributions

- **"Actual" Project Cost is an Uncertain Quantity (Technically, a "Random Variable")**

- **The "Point" Estimate is not the Only Possible Estimate – There are Others**

- **The "Best" Estimate is not the Only Possible Estimate Either – Other Estimates are Presumably "Worse"**

- **Common Use of Phrase "Most Likely" or "Most Probable" Cost Implicitly Assumes that Other Cost Levels are "Less Likely" or "Less Probable"**

- **This Whole Discussion Implies that Costs Really are Probabilistic in Nature**

Is "Cost" a Good Discriminator?

- **"Cost" Is an Important Decision Criterion in Trade Studies**
  - **Other Things (e.g., Performance, Schedule) Being Equal, the Candidate Architecture or Design that "Costs Less" is Often Selected**
  - **When "Other Things" Are Not Equal, Differences Have to Be Balanced off Against Cost to Find Out if They Are "Worth the Difference" in Cost (to Give the Government "Best Value")**

- **But, for Each Candidate, "Cost" Is Often Represented by Only One Number – Is That Number ...**
  - **The "Best" Estimate?**
  - **The 50th-percentile Estimate?**
  - **The 4th-percentile Estimate?**
  - **... Some Other Estimate?**

- **Is It Wise to Base a Decision on One Number That Represents "Cost," if Many Other Numbers Would Be Equally Good or Even Better Representatives?**
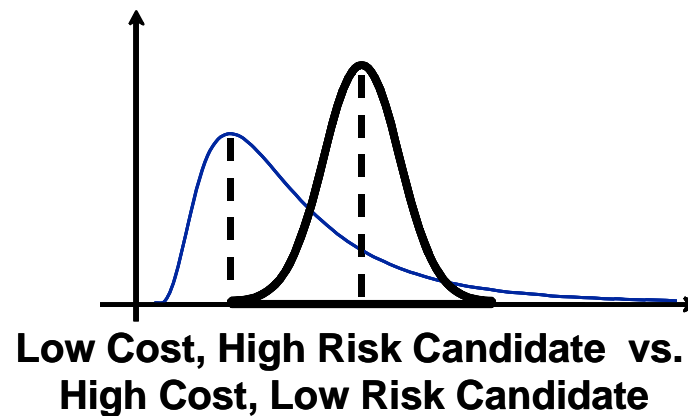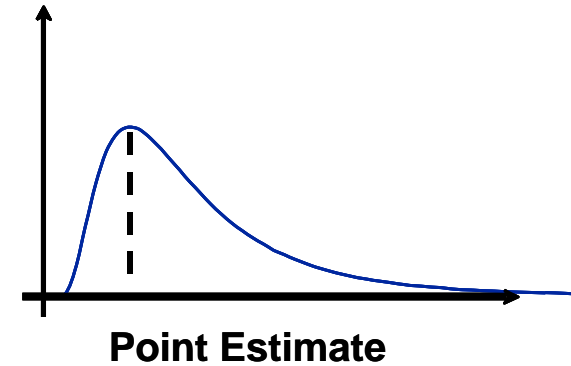
Cost Distributions of Candidates in Trade Study or Source Selection

Low Risk Candidate

Point Estimate

High Risk Candidate

Point Estimate
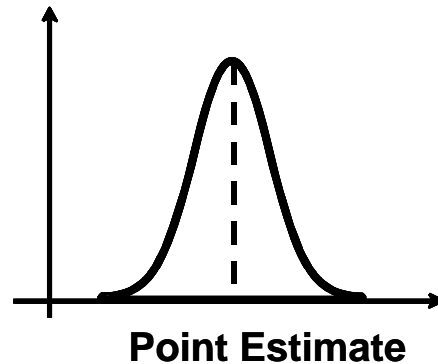
Low Cost, High Risk Candidate vs. High Cost, Low Risk Candidate

Which Candidate Has Lower Cost?

- **Sorry, but No One Can Answer *That* Question at the Trade-Study Stage**
- **Current Standard Practice is to Base the Comparison on the "Point Estimate," Mean Cost, or Some Percentile such as the 50th**
- **If We Know (or Believe We Know) the Total-Cost Probability Distribution of Each Candidate, then,**
  - **For each Pair of Candidates A and B, We Can Actually Calculate the Probability that Candidate A will Cost More than Candidate B**
  - **In Fact, for each Candidate, We Can Determine by Computer Simulation the Probability that that Candidate is the Lowest-Cost Candidate**

Which Candidate *Probably* Has Lower Cost?

- **Now *That's* a Question We *Can* Answer**
- **First, Establish the Total-Cost Probability Distribution for Each Architecture or Design Candidate**
- **Second, For each Pair of Candidates A and B, Calculate Probability that A will Cost More than B**
- **Third, Define a "Cost-Risk Figure of Merit": For Each Candidate, "Combine" the Probabilities that *it* will Cost More than Each of the Others**
- **Finally, the Candidate Whose Figure of Merit is Optimal (in the sense of signifying probable lower cost) has the Greatest Likelihood of Being the Least-Cost Candidate**

11

Contents

- **Cost as a Discriminator**
- **Special Nature of Trade Studies**
- **The Triangular Distribution**
- **Peculiarities of the Lognormal Distribution**
- **Cost Risk as a Discriminator**
- **Summary**
- **Back-up Charts**

Special Efficiency is Required for Trade Studies

- **Many Architecture or Design Candidates Must Be Investigated: For Trade Studies, 10 to 15 Is Very Common and 40 Is Not Unheard of**
- **Each Candidate Is Only Roughly Defined**
  - **Cost Analysis Requirements Description (CARD), Risk Management Plan (RMP) Not Always Available – System Not Fully Defined**
  - **Costs of High-Level WBS Elements Only Can be Estimated**
- **Monte-Carlo Analysis Not Appropriate**
  - **Turnaround Time Often Too Long If Many Estimates Have to Be Completed Within Very Short Time Span**
  - **High-Level, No-CARD, No-RMP Estimate Neither Deserves nor Benefits From High-Precision Monte-Carlo (However, if Monte-Carlo simulation is done, that information can be used)**
- **Analytic Approximation Combining High-Level WBS-Element Costs Into Total-System Cost is Quick and Adequate**

13

Cost-Risk Set-up for Trade Studies

- ## Use Triangular Distributions to Model Subsystem Costs
  - **This Works Well for Any System Architecture or Design**
  - **More Detailed Representation Difficult to Justify at This Stage of Project Definition**

- ## Model Total-System Cost Distribution as Lognormal
  - **Central Limit Theorem Requires Large Number of WBS Elements to Force Total-Cost Distribution to be Normally Distributed – Trade Studies Generally Involve only 10 or so High-Level WBS Elements**
  - **For Small Number of WBS Elements, Total-Cost Distribution Retains Skewness, Making Lognormal Distribution a Better Fit**
  - **Studies at MITRE and Aerospace Indicate that Lognormal Reasonably Approximates Statistical Sum of Triangles ("FRISK" Methodology)**
  - **"FRISK" Methodology is Built into NAFCOM, NASA's Primary Cost Model, Developed for Marshall Space Flight Center by SAIC Huntsville**

- ## Mathematics of Lognormal Distribution Allows Quick Computation of Figure of Merit
  - **Easy to Estimate Probabilities That One Candidate Will Cost More Than Each of the Other Candidates**
  - **Simple Mathematics of Lognormal Distribution Facilitates Rank-ordering of Candidates According to Figure of Merit**
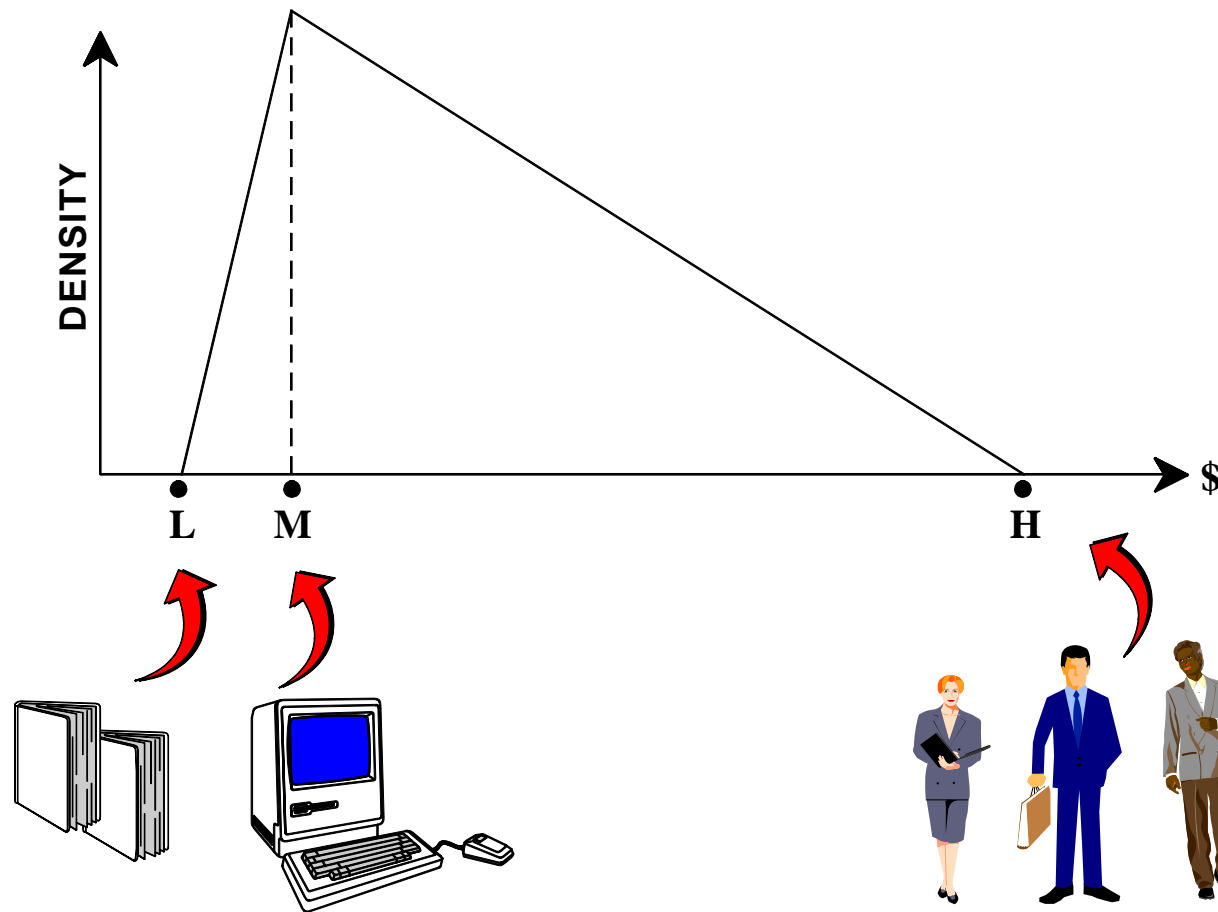
Contents

- **Cost as a Discriminator**
- **Special Nature of Trade Studies**
- **The Triangular Distribution**
- **Peculiarities of the Lognormal Distribution**
- **Cost Risk as a Discriminator**
- **Summary**
- **Back-up Charts**

Triangular Cost Distribution for Subsystem Elements

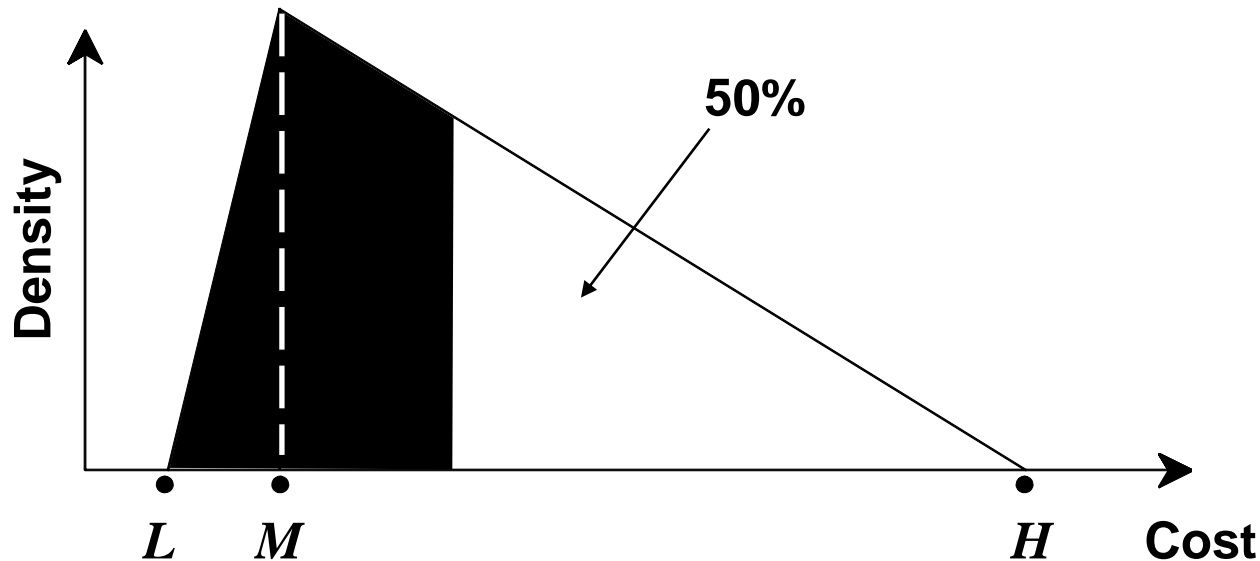**Optimistic Cost**　**Best-Estimate Cost (Mode)**　**Cost Implication of Technical, Programmatic, Estimating Assessment**

Triangular Distribution Description

- **Probability Density Function is Triangular with Total Area under "Curve" = 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**

17

Triangular Distribution Statistics

- **Mode = M (most likely value of cost)**

- **Median = $T_{.50} = L + \sqrt{0.50(M-L)(H-L)}$ if $M - L \geq 0.50(H-L)$**

$$= H - \sqrt{0.50(H-L)(H-M)}   \text{ if } M - L \leq 0.50(H-L)$$

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

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

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

- **Mean = $\dfrac{L+M+H}{3}$, $\sigma = \sqrt{\dfrac{L^2 + M^2 + H^2 - LM - LH - MH}{18}}$**

Model the Total Cost Analytically

- **Approximate the Statistical Sum of Triangular Distributions by a Lognormal Distribution**

WBS-ELEMENT TRIANGULAR COST DISTRIBUTIONS

Most Likely

$

Most Likely

$

Most Likely

$
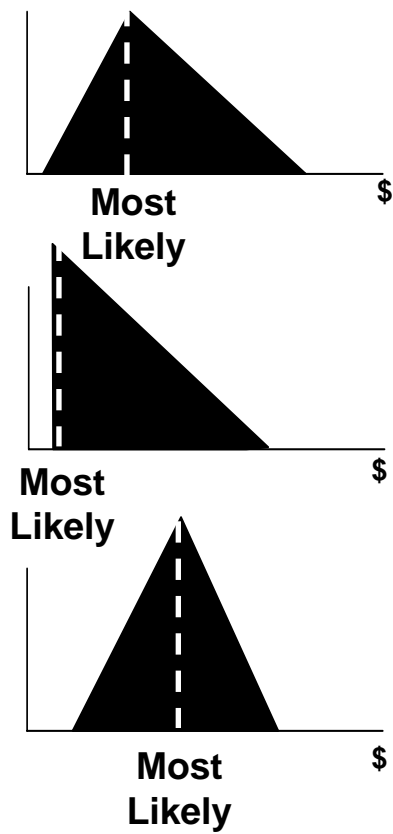
MERGE WBS-ELEMENT COST DISTRIBUTIONS INTO TOTAL-COST LOGNORMAL DISTRIBUTION
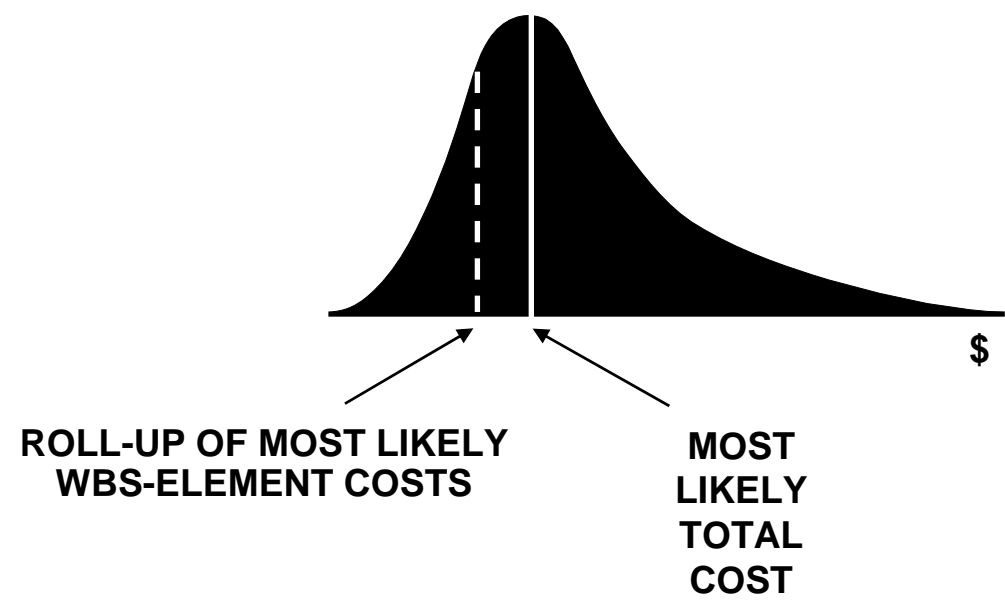
$

ROLL-UP OF MOST LIKELY WBS-ELEMENT COSTS

MOST LIKELY TOTAL COST

19

Contents

- **Cost as a Discriminator**
- **Special Nature of Trade Studies**
- **The Triangular Distribution**
- **Peculiarities of the Lognormal Distribution**
- **Cost Risk as a Discriminator**
- **Summary**
- **Back-up Charts**

A Picture of the Transition From Normal to Lognormal

50%

P
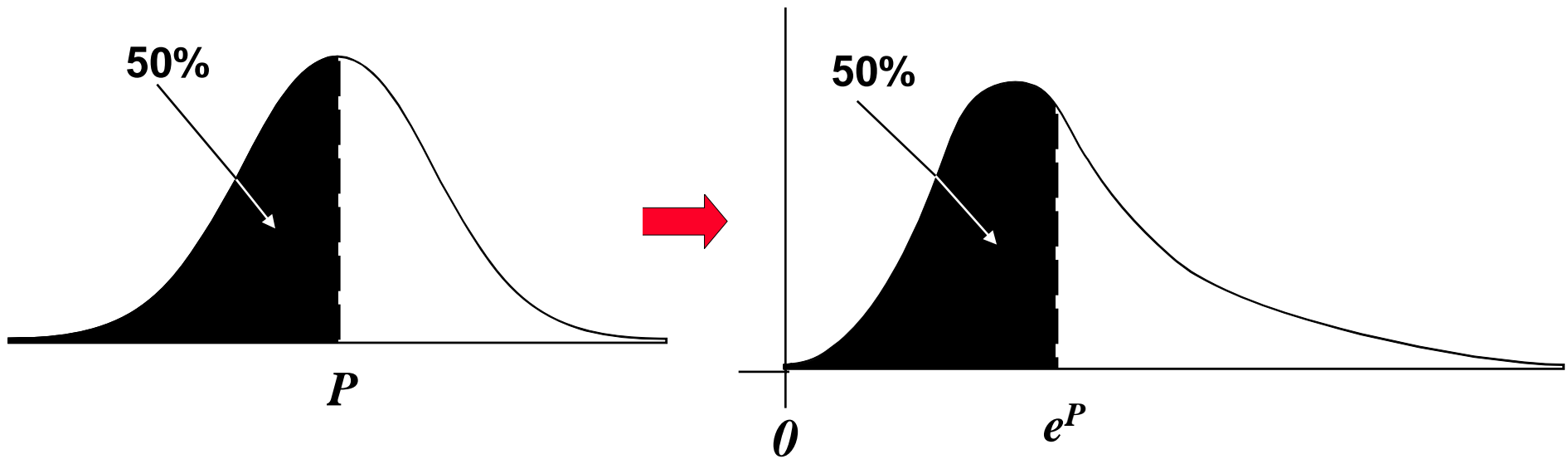
X

Normal Distribution

50%

0 e^P

$Y = e^X$

Lognormal Distribution

Lognormal Distribution Description
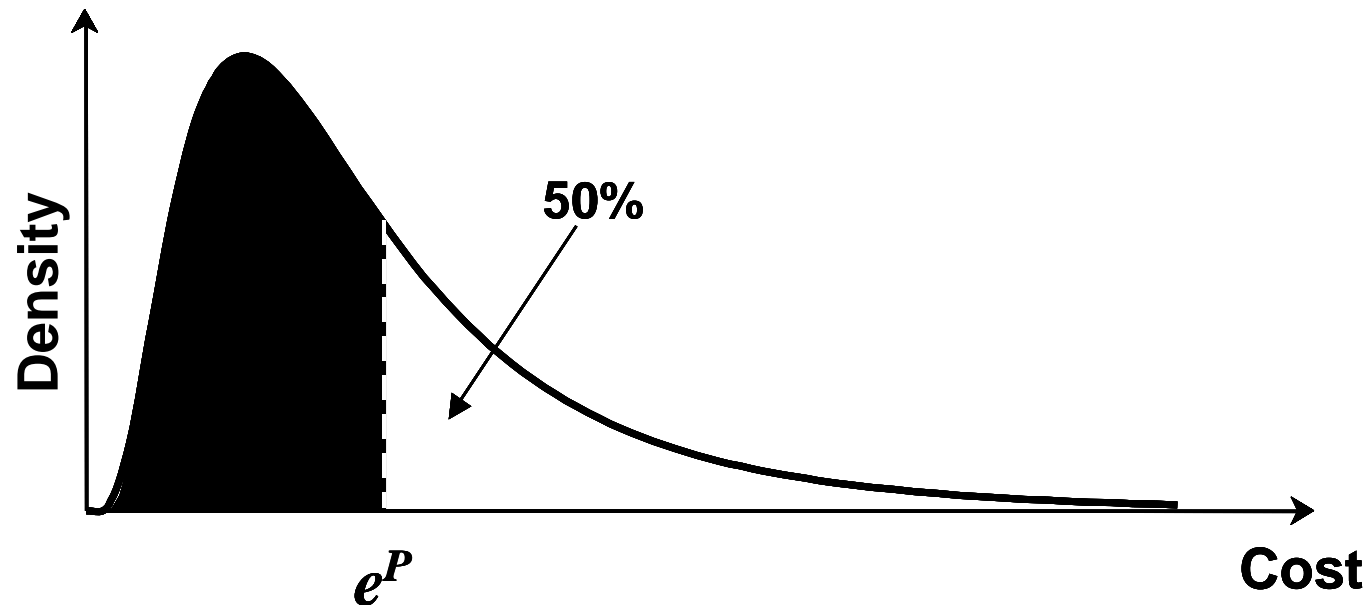
- **Probability Density Function with Total Area under Curve = 1.00**
- **Two Parameters (P,Q or μ,σ or other pairs) Completely Specify Distribution**
  - P,Q are Mean, Standard Deviation of "Underlying" Normal Distribution
  - μ,σ are Mean, Standard Deviation of Lognormal Distribution Itself
- **Mean, Median, Sigma, All Percentiles can be Expressed in Terms of P and Q (or μ and σ)**

Lognormal Distribution Analytics

- **If Y is to be a Lognormal Random Variable with Mean μ and Standard Deviation σ, then**

$$Y = e^{X}$$

  where X is Normal with Mean P and Standard Deviation Q

- **Therefore $\log(Y) = X$ has Normal (Gaussian) Distribution**

- **Lognormal Density Function is, in Terms of P and Q,**

$$h(y) = \frac{1}{yQ\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\log y - P}{Q}\right)^{2}} \quad \text{for} \quad 0 < y < \infty$$

$$= 0 \qquad\qquad\qquad\qquad \text{for} \quad -\infty < y < 0$$

- $P\left\{\text{Cost} \leq e^{P}\right\} = P\left\{e^{X} \leq e^{P}\right\} = P\left\{X \leq P\right\} = 0.50$ **so that Median of Lognormal is e^{P}**

Percentiles of Lognormal Distributions

PERCENTILE	$1-\alpha$	z_α
95	.95	1.64485
90	.90	1.28155
80	.80	0.84162
70	.70	0.52440
60	.60	0.25335
50	.50	0.00000
40	.40	-0.25335
30	.30	-0.52440
20	.20	-0.84162
10	.10	-1.28155
5	.05	-1.64485

- **In the Table to the Right are the Percentiles z_α of Standard Normal Distribution (like those in the back of your statistics textbook)**

- **$L_{1-\alpha}$ is the Analogous Percentile of the Lognormal Distribution**

- **$L_{1-\alpha}$ = Dollar Value at Which**

  $$P\{\textbf{Cost} \le L_{1-\alpha}\} \text{ is } 1-\alpha$$

- $$P\{\textbf{Cost} \le L_{1-\alpha}\} = P\{e^X \le L_{1-\alpha}\} = P\{X \le \lambda n\ L_{1-\alpha}\}$$

  $$= P\left\{\frac{X-P}{Q} \le \frac{\lambda n\ L_{1-\alpha} - P}{Q}\right\} = 1-\alpha \ \ \textbf{if} \ \ \frac{\lambda n\ L_{1-\alpha} - P}{Q} = z_\alpha$$

- $L_{1-\alpha} = e^{P+z_\alpha Q}$ **where z_α is the $(1-\alpha)$th Percentile of the Standard Gaussian Distribution**

24

Lognormal Distribution Statistics

- **Mode** $= e^{P-Q^2}$

- **Median** = 50th Percentile = e^P

- $L_{1-\alpha} =$ **Dollar Value at Which** $P\{\text{Cost} \leq L_{1-\alpha}\}$ **is** $1-\alpha$

$$= e^{P+z_\alpha Q} \quad \text{where } z_\alpha \text{ is the (1-α)th Percentile of the}$$

Standard Gaussian Distribution

- **Mean** $= \mu = e^{P+\frac{1}{2}Q^2}, \quad \sigma = e^{P+\frac{1}{2}Q^2}\sqrt{e^{Q^2}-1}$

25

System Inputs to Trade Study

- D = **Number of Independent Candidates**

- N_j = **Number of Cost Elements in Candidate j's WBS**

- **NOTE: The Candidates Need Not Have the Same WBS**

Cost Element Inputs

- L_{ji} = **Minimum Cost for Candidate** j, **Cost Element** i

- M_{ji} = **Most Likely (Mode) Cost for Candidate** j, **Cost Element** i

- H_{ji} = **Maximum Cost for Candidate** j, **Cost Element** i

- $\rho_{ji_1i_2}$ = **Correlation between Pair of Cost Elements for Candidate** j, **Elements** $i_1 = 1, ..., N_j$; $i_2 = 1, ..., N_j$

  Note: $\rho_{jii} = 1$, **(Diagonal Elements** $= 1$**),** $i = 1, \ \mathrm{K} \ , \ N_j$

- **Correlation "Matrix" Must Be Nonnegative Definite and Have** $-1 \leq \rho_{ji_1i_2} \leq 1$, **(All Correlations are Between** -1 **and** 1**)** $i_1 \neq i_2$

First Step of
Analytic Approximation

- **Use L_{ji}, M_{ji}, H_{ji} to Calculate Mean μ_{ji}, Variance σ^2_{ji} of Cost Distribution of Candidate j, Cost Element i, where $j = 1,...,D; i = 1,...,N_j$**

- $\mu_{ji} = \dfrac{L_{ji} + M_{ji} + H_{ji}}{3}$

$$\sigma^2_{ji} = \frac{L^2_{ji} + M^2_{ji} + H^2_{ji} - L_{ji}M_{ji} - L_{ji}H_{ji} - M_{ji}H_{ji}}{18}$$

28

Total Cost Statistics for Candidate X

- **Candidate X is one of the D Candidates, $j = 1, \ldots, D$**

- $C_{tX} = \sum_{i=1}^{N_X} C_{Xi}$ = **Total Cost for Candidate X**

- C_{tX} **is a Random Variable**

- $\mu_{tX} = E(C_{tX}) = \sum_{i=1}^{N_X} \mu_{Xi}$ = **Mean of Candidate X'sTotal Cost**

- $\sigma_{tX}^{2} = Var(C_{tX}) = \sum_{i=1}^{N_X} \sigma_{Xi}^{2} + \sum \sum_{i_1 > i_2} \rho_{Xi_1 i_2} \sigma_{Xi_1} \sigma_{Xi_2}$

  = **Variance of Total Cost for Candidate X**

Second Step of Analytic Approximation

- **Use μ_{tX}, σ_{tX}^2 as Lognormal Parameters to Solve for Underlying Normal-Distribution Parameters P_{tX}, Q_{tX}**

- **Apply "Method of Moments"**

- $$P_{tX} = \frac{1}{2}\lambda n \left\{ \frac{\mu_{tX}^4}{(\mu_{tX}^2 + \sigma_{tX}^2)} \right\}$$

$$Q_{tX} = \left\{ \lambda n \left(1 + \frac{\sigma_{tX}^2}{\mu_{tX}^2} \right) \right\}^{\frac{1}{2}}$$

Where We Are at This Point

- **Total Cost of Each Candidate Has Been Modeled as a Lognormal Random Variable with "Correct" Mean and Variance**
- **"Method of Moments" Makes Use of Only Means, Variances, and Inter-Element Correlations to Calculate Statistical Descriptors of Total-Cost Lognormal Distribution**
  - Therefore Modeling Individual WBS-Element Costs as Triangular Distributions is Not Really Required
  - …But It's Still The Easiest Way To Do It
- **Payoff for Using a Lognormal Model for Total Cost will soon be Evident, because its Unique Mathematical Properties Allow Simple Calculation of Cost-Risk Figure of Merit**

Contents

- **Cost as a Discriminator**
- **Special Nature of Trade Studies**
- **The Triangular Distribution**
- **Peculiarities of the Lognormal Distribution**
- **Cost Risk as a Discriminator**
- **Summary**
- **Back-up Charts**

Comparing Costs of Two Candidates

- **"Cost Ratio" for Candidate Pair (A,B)**

$$R_{A,B} = \frac{C_{tA}}{C_{tB}} \; for \; A = 1, \; K, \; D; \; B = 1, \; K, \; D; \; A \neq B$$

  - **"Cost Ratio" is Itself a Random Variable**

- $P\left\{R_{A,B} > 1\right\} = P\left\{\ln\left(R_{A,B}\right) > 0\right\} = r_{AB}$ **is the Probability that the Cost of Candidate A Exceeds the Cost of Candidate B**

- $r_{BA} = 1 - r_{AB}$ **if Candidates A and B are not the same**

Probabilistic Relationships Between Candidates

- **For Two Different Candidate Architectures or Designs A and B, We Can Have**
  - A **Probably Costs More than** B
  - A **Probably Costs the Same as** B **(statistically impossible)**
  - A **Probably Costs Less than** B

- **If A Probably Costs More than B, then $r_{AB} > \frac{1}{2}$**

- **If A Probably Costs Less than B, then $r_{AB} < \frac{1}{2}$**

- **But A Costs Exactly the Same as Itself – Therefore We Should Define $r_{AA} = \frac{1}{2}$ in Order to Ensure the Correct Ranking of Candidates According to Probable Cost**

Probability Distribution of the "Cost Ratio"

- **Logarithm of Cost Ratio is**

$$\lambda n \left(R_{A,B}\right) = \lambda n \left(\frac{C_{tA}}{C_{tB}}\right) = \lambda n \left(C_{tA}\right) - \lambda n \left(C_{tB}\right)$$

- **Difference of Two Independent Normal Random Variables has Normal Distribution**

- **Therefore Ratio of Two Independent Lognormal Random Variables has Lognormal Distribution**

- $R_{A,B}$ **is Lognormally Distributed, i.e.,** $\lambda n \left(R_{A,B}\right)$ **is Normally Distributed, and has Parameters**

$$P_{AB} = P_{tA} - P_{tB}$$

$$Q^2_{AB} = Q^2_{tA} + Q^2_{tB}$$

35

Calculating the Cost Ratio

- $R_{A,B}$ **has a Lognormal Distribution with Underlying Gaussian Parameters** P_{AB} **and** Q_{AB}

- $r_{AB} = P\{R_{A,B} > 1\} = P\{\ln(R_{A,B}) > 0\}$

$$= \frac{1}{Q_{AB}\sqrt{2\pi}} \int_0^\infty e^{-\frac{1}{2}\left(\frac{x - P_{AB}}{Q_{AB}}\right)^2} dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\frac{P_{AB}}{Q_{AB}}}^\infty e^{-\frac{u^2}{2}} du = \Phi\left(\frac{P_{AB}}{Q_{AB}}\right),$$

$$u = \frac{x - P_{AB}}{Q_{AB}}$$

$$du = \frac{dx}{Q_{AB}}$$

which Can be Looked up in a Table of the Standard Normal Distribution Function $\Phi(x)$ **or Calculated in Excel Using the NORMSDIST Function**

36

Comparing Cost Ratios

- $r_{AB} = \Phi\left(\dfrac{P_{AB}}{Q_{AB}}\right)$
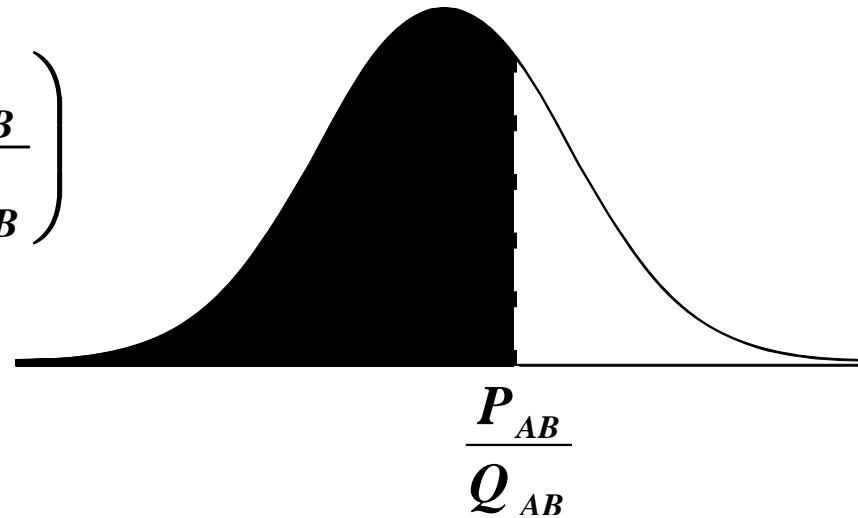
$$\frac{P_{AB}}{Q_{AB}}$$

- **Lower Cost Ratio Corresponds to Lower $\dfrac{P_{AB}}{Q_{AB}}$**

- **Lower $\dfrac{P_{AB}}{Q_{AB}}$ Therefore Corresponds to a Lower Probability that Cost of Candidate A Exceeds Cost of Candidate B**

37

A Simple Figure of Merit

- **Sum All Cost Probabilities for Candidate A as Follows:**

$$S_A = \sum_j P\{Cost(A) > Cost(j)\} = \sum_j \varPhi\left(\frac{P_{Aj}}{Q_{Aj}}\right)$$

- **Calculate Sum like This for all Other Candidates**

- **Rank-Order the Sequence $\{S_j\}$ to get $S_1 < S_2 < ... < S_D$, the Figures of Merit Ranked Best (lowest probable cost) to Worst (highest probable cost)**

- **S_1 Corresponds to Candidate with Lowest Figure of Merit (lowest probability of exceeding other individual candidates in cost)**

- **S_D Corresponds to Candidate with Highest Figure of Merit**

An Example

- **Consider a Trade Study Comparing Four Candidates A, B, C, and D, whose Distributions are Lognormal with Means and Standard Deviations as in the Table Below:**

Candidate	A	B	C	D
Distribution	LN	LN	LN	LN
Mean	95	85	100	90
Std Dev	50	5	30	70
P	4.431617	4.440924	4.562081	4.263267
Q	0.49449	0.058773	0.29356	0.687812

- **The Cost Ratios are then Calculated as Follows:**

$$r_{AB} = \Phi\left(\frac{P_A - P_B}{\sqrt{Q_A^2 + Q_B^2}}\right) = 0.492544 \qquad r_{BC} = \Phi\left(\frac{P_B - P_C}{\sqrt{Q_B^2 + Q_C^2}}\right) = 0.342854$$

$$r_{AC} = \Phi\left(\frac{P_A - P_C}{\sqrt{Q_A^2 + Q_C^2}}\right) = 0.410262 \qquad r_{BD} = \Phi\left(\frac{P_B - P_D}{\sqrt{Q_B^2 + Q_D^2}}\right) = 0.511511$$

$$r_{AD} = \Phi\left(\frac{P_A - P_D}{\sqrt{Q_A^2 + Q_D^2}}\right) = 0.578764 \qquad r_{CD} = \Phi\left(\frac{P_C - P_D}{\sqrt{Q_C^2 + Q_D^2}}\right) = 0.655264$$

39

Computing the "Simple Figure of Merit"

- **The Cost Ratios Lead to the Following Sums, from which it**

Probability that:	... Will Cost More than ...				Sums
	A	B	C	D	
A ...	-	0.492544	0.410262	0.578764	1.481571
B ...	0.507456	-	0.342854	0.511511	1.361822
C ...	0.589738	0.657146	-	0.655264	1.902147
D ...	0.421236	0.488489	0.344736	-	1.254461

Follows that $S_A = 1.481571$, $S_B = 1.361822$, $S_C = 1.902147$, and $S_D = 1.254461$.

- **We Rank-Order the Sequence of Simple Figures of Merit to get $S_1 < S_2 < ... < S_n$, the Risk-Sensitive Figures of Merit Ranked Best (lowest) to Worst (highest)**

Rank Order ("Best" to "Worst")	Figure of Merit
S_D	1.254461
S_B	1.361822
S_A	1.481571
S_C	1.902147

40

More "Mathematically Correct" Figures of Merit Can be Used

- **Exact Probabilities**: For Candidate A (and then all other Candidates), Calculate the Probability that it is the Least-Cost System:

$$L_A = P\left\{ \underset{j}{\mathrm{I}}\ \left[\mathrm{Cost}\ (A) > \mathrm{Cost}\ (j)\right]\right\}$$

  – Unfortunately, This Number is Difficult to Calculate because the Events in the Intersection are not Independent, so We Cannot Simply Multiply the Individual Probabilities
  – However, We **Can** Simulate It (see below)

- **Rank All Candidates by Cost**: Calculate All Pairwise Probabilities of the form

$$r_{AB} = P\left\{Cost\ (A) > Cost\ (B)\right\}$$

   and Apply a Standard Mathematical Method that Converts Pairwise Comparisons into a Ranked List

Paired-Comparison Matrices

- **T.L. Saaty (References 5 and 6) Introduced the Concept of a Paired-Comparison Matrix in Connection with a Proposed Method of Analysis that he Called the "Analytic Hierarchy Process"**
- **If n Items are to be Compared, a Paired-Comparison Matrix is a Square $n \times n$ Matrix whose Entries Meet the Following Requirements:**
  - **Each Entry $a_{ij} > 0$ for $1 \leq i \leq n$ and $1 \leq j \leq n$**
  - **All Diagonal Elements $a_{ii} = 1$**
  - **The Transposed Elements $a_{ji} = 1/a_{ij}$ for all Comparison Pairs i and j**
- **The Following Matrix is an Example of a Paired-Comparison Matrix:**

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} = \begin{bmatrix} 1 & 0.5 & 0.2 & 0.4 \\ 2 & 1 & 0.8 & 0.5 \\ 5 & 1.25 & 1 & 0.4 \\ 2.5 & 2 & 2.5 & 1 \end{bmatrix}$$

42

Bad News: The r_{AB} Values Do Not Comprise a Paired-Comparison Matrix

- **For the Problem We are Working on, the Paired Comparisons are the Probabilities**

$$r_{AB} = P\{Cost(A) > Cost(B)\}$$

- **Suppose We Have Four Candidate Architectures or Designs to Compare – Then Our Paired Comparisons are of the Form**

$$r_{12} = P\{Cost(1) > Cost(2)\}, \quad r_{13} = P\{Cost(1) > Cost(3)\}, \quad r_{23} =$$

$$P\{Cost(2) > Cost(3)\},... \quad \text{etc.}$$

- **Note that the r_{ij} Values do not Meet the Second and Third of the Three Requirements for a Paired-Comparison Matrix**
  - $a_{ii} = P\{Cost(i) > Cost(i)\} = \frac{1}{2}$ **(by definition), not** *1*
  - $a_{ji} = 1 - a_{ij}$, **not** $1/a_{ij}$
- **We are Therefore Going to Have to Make Some Adjustments**

Making the r_{AB} Values Fit a Paired-Comparison Matrix

- **Start with $r_{ij} = P\{Cost(i) > Cost(j)\}$ and $r_{ii} = \frac{1}{2}$**

- **Then Define $b_{ij} = r_{ij} - \frac{1}{2}$**

- **Note that, if i and j are not the Same,**

$$b_{ji} = r_{ji} - \frac{1}{2} = (1 - r_{ij}) - \frac{1}{2} = (1 - \frac{1}{2}) - r_{ij}$$

$$= \frac{1}{2} - r_{ij} = -(r_{ij} - \frac{1}{2}) = -b_{ij}$$

- **Now Define $a_{ij} = exp(b_{ij})$, i.e.** $\quad a_{ij} = e^{b_{ij}}$

- **Now Note that $a_{ji} = exp(b_{ji}) = exp(-b_{ij}) = 1/exp(b_{ij}) = 1/a_{ij}$ if i and j are not the Same**

- **Note that $a_{ii} = exp(b_{ii}) = exp(r_{ii} - \frac{1}{2}) = exp(\frac{1}{2} - \frac{1}{2}) = exp(0) = 1$ if the Two Candidates Being Compared are the Same**

44

The Same Example

- **Our Trade Study Compares Four Candidates A, B, C, and D, whose Distributions are Lognormal with Means and Standard Deviations as in the Table Below:**

Candidate	A	B	C	D
Distribution	LN	LN	LN	LN
Mean	95	85	100	90
Std Dev	50	5	30	70
P	4.431617	4.440924	4.562081	4.263267
Q	0.49449	0.058773	0.29356	0.687812

- **The Cost Ratios are then Calculated as Follows:**

$$r_{AB} = \Phi\left(\frac{P_A - P_B}{\sqrt{Q_A^2 + Q_B^2}}\right) = 0.492544 \qquad r_{BC} = \Phi\left(\frac{P_B - P_C}{\sqrt{Q_B^2 + Q_C^2}}\right) = 0.342854$$

$$r_{AC} = \Phi\left(\frac{P_A - P_C}{\sqrt{Q_A^2 + Q_C^2}}\right) = 0.410262 \qquad r_{BD} = \Phi\left(\frac{P_B - P_D}{\sqrt{Q_B^2 + Q_D^2}}\right) = 0.511511$$

$$r_{AD} = \Phi\left(\frac{P_A - P_D}{\sqrt{Q_A^2 + Q_D^2}}\right) = 0.578764 \qquad r_{CD} = \Phi\left(\frac{P_C - P_D}{\sqrt{Q_C^2 + Q_D^2}}\right) = 0.655264$$

Setting Up the Paired Comparisons

- **Consider a Trade Study Comparing Four Candidates A, B, C, and D**
- **The Matrix of r_{ij} Values, where (except for the separately defined diagonal entries) $r_{ij} = P\{Cost(i) > Cost(j)\}$, is the Following:**

$$
\begin{bmatrix}
r_{AA} & r_{AB} & r_{AC} & r_{AD} \\
r_{BA} & r_{BB} & r_{BC} & r_{BD} \\
r_{CA} & r_{CB} & r_{CC} & r_{CD} \\
r_{DA} & r_{DB} & r_{DC} & r_{DD}
\end{bmatrix}
=
$$

0.5000	0.4925	0.4103	0.5788
0.5075	0.5000	0.3429	0.5115
0.5897	0.6571	0.5000	0.6553
0.4212	0.4885	0.3447	0.5000

- **The Matrix of b_{ij} Values, where $b_{ij} = r_{ij} - \frac{1}{2}$, is the Following:**

$$
\begin{bmatrix}
b_{AA} & b_{AB} & b_{AC} & b_{AD} \\
b_{BA} & b_{BB} & b_{BC} & b_{BD} \\
b_{CA} & b_{CB} & b_{CC} & b_{CD} \\
b_{DA} & b_{DB} & b_{DC} & b_{DD}
\end{bmatrix}
=
$$

0.0000	-0.0075	-0.0897	0.0788
0.0075	0.0000	-0.1571	0.0115
0.0897	0.1571	0.0000	0.1553
-0.0788	-0.0115	-0.1553	0.0000

46

The Paired-Comparison Matrix

- **The Paired-Comparison Matrix of a_{ij} Values, where $a_{ij} = e^{b_{ij}}$, is the Following:**

$$
\begin{bmatrix}
a_{AA} & a_{AB} & a_{AC} & a_{AD} \\
a_{BA} & a_{BB} & a_{BC} & a_{BD} \\
a_{CA} & a_{CB} & a_{CC} & a_{CD} \\
a_{DA} & a_{DB} & a_{DC} & a_{DD}
\end{bmatrix}
=
$$

1.0000	0.9926	0.9142	1.0819
1.0075	1.0000	0.8546	1.0116
1.0939	1.1702	1.0000	1.1680
0.9243	0.9886	0.8562	1.0000

47

The Decision Criterion

- **T.L. Saaty (References 5 and 6) Proposed the "Analytic Hierarchy Process" (AHP) as a Technique for Converting the Information in the Paired-Comparison Matrix into a Ranked List of Candidates**
- **Later, However, G. Crawford and C. Williams (Reference 1) and G. Crawford (Reference 2) Pointed Out Some Technical Problems with the AHP Approach**
  - They Suggested a Method Based on the Geometric Mean as a More Accurate Way of Deriving the Ranked List of Candidates
  - Additional Analysis Supporting the Use of the Geometric-Mean Technique was Later Published by J.M. Hihn and C. Johnson (Reference 3) and E. Miranda (Reference 4)
- **The Geometric-Mean Technique is Also Easier to Apply**
- **We Shall Use that Method on the Following Charts to Establish the Ranked List of Candidates**

The Geometric Mean

- **The Geometric Mean of Numbers $x_1, x_2, …, x_n$ is Defined as the nth Root of the Product of the n Numbers – Its Algebraic Expression is**

$$G = \left(\prod_{i=1}^{n} x_i \right)^{\frac{1}{n}}$$

- **To Obtain a Ranking of the Candidates in Probable Cost Order (highest to lowest), We Calculate the Geometric Mean of Each Row of the Paired- Comparison Matrix:**

  - $G_A = (1.0000 \times 0.9926 \times 0.9142 \times 1.0819)^{1/4} = 0.9954$
  - $G_B = (1.0075 \times 1.0000 \times 0.8546 \times 1.0116)^{1/4} = 0.9660$
  - $G_C = (1.0939 \times 1.1702 \times 1.0000 \times 1.1680)^{1/4} = 1.1058$
  - $G_D = (0.9243 \times 0.9886 \times 0.8562 \times 1.0000)^{1/4} = 0.9405$

Solution – The Candidates Ranked in Probable Cost Order

- **The Geometric-Mean Vector is Therefore**

$$\begin{bmatrix} A \\ B \\ C \\ D \end{bmatrix} = \begin{bmatrix} 0.9954 \\ 0.9660 \\ 1.1058 \\ 0.9405 \end{bmatrix}$$

- **The Analysis Indicates that …**
  - **Candidate C is the Probable Highest Cost Option (1.1058)**
  - **Candidate A is the Probable Second-Highest Cost Option (0.9954)**
  - **Candidate B is the Probable Third-Highest Cost Option (0.9660)**
  - **Candidate D is the Probable Lowest Cost Option (0.9405)**

- **I Hope No One Thinks of This, but Remember the Earlier Result that $r_{AB} = 0.492544$? – What Does That Mean?**

The Same Example Worked Using Computer Simulation

- **Recall that the Four Candidates A, B, C, and D Have Lognormal Cost Distributions with Known Means and Standard Deviations**
- **We Can Draw 10,000 (or any number of) Independent Random (Monte Carlo or Latin Hypercube) Samples from Each of These Distributions to Simulate the Costs of Each of the Candidates:**

Trial	Cost(A)	Cost(B)	Cost(C)	Cost(D)	Minimum	Least Expensive	Numerical Value
1	214.16	83.15	86.20	114.68	83.15	B	0.001000
2	135.94	87.67	105.34	114.41	87.67	B	0.001000
3	101.53	83.36	85.13	58.04	58.04	D	1,000.000000
4	114.80	79.87	113.79	104.86	79.87	B	0.001000
5	52.68	84.32	50.09	61.10	50.09	C	1.000000
6	122.70	78.71	128.86	220.20	78.71	B	0.001000
7	129.13	78.10	49.62	74.52	49.62	C	1.000000
8	105.91	84.91	82.64	57.66	57.66	D	1,000.000000
9	89.89	81.96	66.38	56.16	56.16	D	1,000.000000
10	54.40	81.09	83.76	20.76	20.76	D	1,000.000000
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
9,990	78.44	76.21	147.71	17.07	17.07	D	1,000.000000
9,991	68.31	91.44	132.34	70.03	68.31	A	0.000001
9,992	98.67	90.27	55.72	122.22	55.72	C	1.000000
9,993	49.94	86.12	62.48	35.91	35.91	D	1,000.000000
9,994	33.52	92.24	92.91	32.62	32.62	D	1,000.000000
9,995	72.86	87.39	103.97	41.78	41.78	D	1,000.000000
9,996	132.37	84.02	163.71	43.62	43.62	D	1,000.000000
9,997	72.87	83.28	97.51	41.03	41.03	D	1,000.000000
9,998	68.87	87.07	85.23	264.63	68.87	A	0.000001
9,999	80.72	82.47	94.94	18.94	18.94	D	1,000.000000
10,000	91.86	89.36	90.98	107.13	89.36	B	0.001000
						Sum =	11,003.005002

How to Determine the Most Probable Lowest-Cost Candidate

- **Look at the "Sum" at the Bottom of the Far-Right Column of the Table on the Previous Chart – That is the Sum of the 21 Numerical Values Displayed**
- **The Sum that in that Position for all 10,000 Samples Indicates that**
  - **A was the Lowest-Cost Candidate 2,894 Times**
  - **B was the Lowest-Cost Candidate 1,299 Times**
  - **C was the Lowest-Cost Candidate 1,118 Times**
  - **D was the Lowest-Cost Candidate 4,689 Times**

What Can We Conclude from the Results of the Simulation?

- **The Simulation Results Indicate that**
  - D Has Probability 0.4689 of Being the Lowest-Cost Candidate
  - A Has Probability 0.2894 of Being the Lowest-Cost Candidate
  - B Has Probability 0.1299 of Being the Lowest-Cost Candidate
  - C Has Probability 0.1118 of Being the Lowest-Cost Candidate
- **Compare these Results with Those of the Paired-Comparison Analysis Done Earlier**
  - D is the Most Probable Lowest Cost Option (0.9405)
  - B is the Second Most Probable Lowest Cost Option (0.9660)
  - A is the Third Most Probable Lowest Cost Option (0.9954)
  - C is the Fourth Most Probable Lowest Cost Option (1.1058)
- **… and with the Results of the Simple Summation Figure of Merit**
  - D is the Most Probable Lowest Cost Option (1.2545)
  - B is the Second Most Probable Lowest Cost Option (1.3618)
  - A is the Third Most Probable Lowest Cost Option (1.4816)
  - C is the Fourth Most Probable Lowest Cost Option (19021)

Discussion of the Results

- **Notice that the Multiple-Comparison Simulation Analysis Led to Ranking the Trade-Study Candidates in the Order D, A, B, C in Probability of Being the Lowest-Cost Solution**

- **Both the Simple Figure of Merit and the Paired-Comparison Analysis, However, Ranked the Candidates in the Order D, B, A, C**

- **Which is More Likely the Correct Ranking?**
  - **Both the Simple Figure of Merit and Paired-Comparison Methods are Based on the Same Information, Namely Paired Comparisons of the Form** $P\{Cost(A) > Cost(B)\}$
  - **But the Simulation Method is Based on Multiple Comparisons of the Form**

  $$P\{Cost(A) > Cost(B) \text{ and } Cost(A) > Cost(C) \text{ and } Cost(A) > Cost(D)\}$$

Recommendation

- **The Simulation Method Makes Use of <u>All</u> the Multiple-Comparison Information that is Available about the Relative Costs of the Candidates**
- **The Other Two Methods Make Use of Only <u>Some</u> of the Information, Namely the Pairwise Information**
- **Consider the Following Analogy**
  - **Using Only the Pairwise Information is Somewhat Like Looking at the Two-Dimensional Shadows that a Three-Dimensional Object Casts on Several Walls of a Room and Trying to Imagine How the Object Looks in Three Dimensions**
  - **However, Using the Multiple-Comparison Simulation Information is Like Looking Directly at the Three-Dimensional Object Itself**
- **Therefore, the Recommendation of this Study is that Paired-Comparison Methods be Eschewed, in Favor of the Simulation Method**

Solution – The Candidates Ranked in Probable Cost Order

Candidate	A	B	C	D
Distribution	LN	LN	LN	LN
Mean	95	85	100	90
Std Dev	50	5	30	70

Rank According to Probable Lowest Cost	Candidate Architecture or Design
Most Probable Lowest Cost	D
Second Most Probable Lowest Cost	A
Third Most Probable Lowest Cost	B
Fourth Most Probable Lowest Cost	C

Another Advantage of the Simulation Method

- **Remember All that Discussion about How the Lognormal Distribution is a Good Approximation for the Statistical Sum of Triangular Distributions?**

- **Well, if You Don't Go for Analytic Approximations, You Can Sum the Triangular Distributions by Computer Simulation (Monte Carlo or Latin Hypercube)**

- **Then, Instead of the 10,000 Samples from Each of the Lognormal Distributions, Use the 10,000 Samples of Each of the Simulations of the Actual Sums of the Triangular Distributions**

  – **They'll Probably Fit a Lognormal Distribution Pretty Well Anyway, but, if They Don't, You Won't Have to Worry About That Issue**

  – **And You Will be Able to Carry Out the Full Simulation Method of Multiple Comparisons**

A Final Comment

- **In the References, Numbers Derived by Calculating Geometric Means of the Rows of the Paired-Comparison Matrix are Interpreted as "Relative" Measures of Magnitude of the Characteristic Being Compared**

- **In Our Situation, However, the Numbers Themselves <u>Do Not Have a Specific Meaning</u> Associated with the Cost of Each Candidate and Can be Used Only for Ranking**

- **The Reason for This is that We Had to Transform Our Numbers Several Times so that We Could Establish a Matrix that Met the Criteria for Being a Paired-Comparison Matrix**
  - **We Did Not Obtain Our Initial Numbers via Paired Comparisons, but Rather via Calculation of Probabilities, the Resulting Matrix of which is Not a Paired-Comparison Matrix**
  - **In the References, on the Other Hand, the Initial Numbers were Derived from Paired Comparisons and so Met the Criteria Immediately**

Contents

- **Cost as a Discriminator**
- **Special Nature of Trade Studies**
- **The Triangular Distribution**
- **Peculiarities of the Lognormal Distribution**
- **Cost Risk as a Discriminator**
- **Summary**
- **Back-up Charts**

Summary

- **Costs are Random Variables, Not Deterministic Numbers, so System Cost is not Well Represented by Any Single Possible Value of the Random Variable**
- **In a Trade Study Involving Several Candidates, a Cost-Risk Figure of Merit Addresses not only Different "Most Likely" Costs, but also Different Risk Characteristics, of the Various Candidates**
- **Easy-to-Carry-Out Risk-Impact Simulation Method Leads to Probabilities that One Candidate Will Cost More than Each of Several Others**
  - **Method of Moments Makes for Easy Transition From Costs of Individual WBS Elements to Total System Cost Modeled As Lognormal Distribution**
  - **Lognormal Probabilities Easy to Work With Due to Their Unique Relationship With Normal Probabilities**
  - **Multiple-Comparison Simulation Method's Ranking Criterion Supports Decision Making**

Contents

- **Cost as a Discriminator**
- **Special Nature of Trade Studies**
- **The Triangular Distribution**
- **Peculiarities of the Lognormal Distribution**
- **Cost Risk as a Discriminator**
- **Summary**
- **Back-up Charts**

References

1. G. Crawford and C. Williams, "The Analysis of Subjective Judgment Matrices," Project Air Force Report R-2572-1-AF, The RAND Corporation, May 1985, 34+xii pages.

2. G. Crawford, "The Geometric Mean Procedure for Estimating the Scale of a Judgment Matrix," *Mathematical Modelling*, Vol. 9 (1987), pages 327-334

3. J.M. Hihn and C.R. Johnson, "Evaluation Techniques for Paired Ratio-Comparison Matrices in a Hierarchical Decision Model," *Measurement in Economics*, Physica-Verlag Heidelberg, 1988, pages 269-288.

4. E. Miranda, "Improving Subjective Estimates Using Paired Comparisons," *IEEE Software*, January/February 2001, pages 87-91.

5. T.L. Saaty, "A Scaling Method for Priorities in a Hierarchical Structure," *Journal of Mathematical Psychology*, Vol. 15 (1977), pages 234-281.

6. T.L. Saaty, *The Analytic Hierarchy Process*, New York, NY: McGraw-Hill, 1980.

Backup Charts –
Details of the Mathematics

Analytic Geometry of the Triangle

"Creating Customer-Focused Success"

MCR
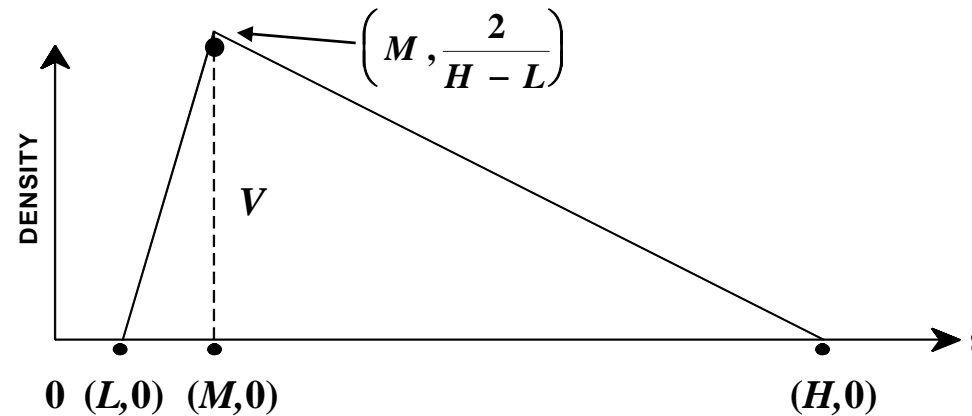
- **Area of Triangle** $= \dfrac{1}{2} V(H - L) = 1$, **so that** $V = \dfrac{2}{H - L}$

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

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

64

Triangular Density Function

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

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

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

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

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

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

Percentiles of Triangular Distribution

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

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

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

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

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

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

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

66

Mode, Median of Triangular Distribution

- **Mode =** M **(most likely value of cost)**

- **Median =** $T_{.50}$ **, where**

$$T_{.50} = L + \sqrt{0.50(M-L)(H-L)} \quad \text{if } M - L \geq 0.50(H-L)$$

$$= H - \sqrt{0.50(H-L)(H-M)} \quad \text{if } M - L \leq 0.50(H-L)$$

67

Mean of Triangular Distribution

- **Mean** $= \int_L^H xf(x)dx = \int_L^M \frac{2x(x-L)}{(M-L)(H-L)}dx + \int_M^H \frac{2x(H-x)}{(H-L)(H-M)}dx$

$$= \frac{2}{(M-L)(H-L)}\left[\frac{x^3}{3} - \frac{Lx^2}{2}\right]_L^M + \frac{2}{(H-L)(H-M)}\left[\frac{Hx^2}{2} - \frac{x^3}{3}\right]_M^H$$

$$= \frac{2}{(M-L)(H-L)}\left(\frac{M^3}{3} - \frac{LM^2}{2} - \frac{L^3}{3} + \frac{L^3}{2}\right) + \frac{2}{(H-L)(H-M)}\left(\frac{H^3}{2} - \frac{H^3}{3} - \frac{HM^2}{2} + \frac{M^3}{3}\right)$$

$$= \frac{2}{(M-L)(H-L)}\left(\frac{2M^3 - 3LM^2 + L^3}{6}\right) + \frac{2}{(H-L)(H-M)}\left(\frac{H^3 - 3HM^2 + 2M^3}{6}\right)$$

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

$$= \frac{L + M + H}{3}$$

Second Moment

$$E\left(X^2\right) = \int_L^H x^2 f(x) dx = \int_L^M \frac{2x^2(x-L)}{(M-L)(H-L)} dx + \int_M^H \frac{2x^2(H-x)}{(H-L)(H-M)} dx$$

$$= \frac{2}{(M-L)(H-L)} \int_L^M \left(x^3 - Lx^2\right) dx + \frac{2}{(H-L)(H-M)} \int_M^H \left(Hx^2 - x^3\right) dx$$

$$= \frac{2}{(M-L)(H-L)} \left[\frac{x^4}{4} - \frac{Lx^3}{3}\right]_L^M + \frac{2}{(H-L)(H-M)} \left[\frac{Hx^3}{3} - \frac{x^4}{4}\right]_M^H$$

$$= \frac{2}{(M-L)(H-L)} \left(\frac{M^4}{4} - \frac{LM^3}{3} - \frac{L^4}{4} + \frac{L^4}{3}\right) + \frac{2}{(H-L)(H-M)} \left[\frac{H^4}{3} - \frac{H^4}{4} - \frac{HM^3}{3} + \frac{M^4}{4}\right]$$

$$= \frac{2}{(M-L)(H-L)} \left[\frac{3M^4 - 4LM^3 + L^4}{12}\right] + \frac{2}{(H-L)(H-M)} \left[\frac{H^4 - 4HM^3 + 3M^4}{12}\right]$$

$$= \frac{3M^3 - LM^2 - L^2M - L^3}{6(H-L)} + \frac{H^3 + H^2M + HM^2 - 3M^3}{6(H-L)}$$

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

MCR

Sigma Value
of Triangular Distribution

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

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

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

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

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

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

Mode, Median of Lognormal Distribution

- **Mode = y Value for Which Density Function $h(y)$ Attains its Maximum**

$$h(y) = \frac{1}{yQ\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln\ y-P}{Q}\right)^2}$$

$$\ln\ h(y) = -\frac{1}{2}\left(\frac{\ln\ y-P}{Q}\right)^2 - \ln\left(yQ\sqrt{2\pi}\right)$$

$$\frac{h'(y)}{h(y)} = -\left(\frac{\ln\ y-P}{Q}\right)\left(\frac{1}{Qy}\right) - \frac{1}{y} = 0\ \text{ when }\ y = e^{P-Q^2}$$

- **Therefore, Mode $= e^{P-Q^2}$**

- **Median = 50th Percentile $= e^P$ because**

$$L_{0.50} = e^{P+z_{0.50}Q} = e^P\ \left(setting\ z_{0.50} = 0\right)$$

71

Mean of Lognormal Distribution

- **Mean**
$$= \int_0^\infty yh(y)dy = \int_0^\infty \frac{1}{Q\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln y - P}{Q}\right)^2} dy$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty e^{-\frac{1}{2}u^2} e^{P+Qu} du$$

$$= \frac{e^P}{\sqrt{2\pi}} \int_{-\infty}^\infty e^{-\frac{1}{2}\left(u^2 - 2Qu\right)} du = \frac{e^P}{\sqrt{2\pi}} \int_{-\infty}^\infty e^{-\frac{1}{2}(u-Q)^2 + \frac{1}{2}Q^2} du$$

$$= e^{P+\frac{1}{2}Q^2} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty e^{-\frac{z^2}{2}} dz$$

$$= e^{P+\frac{1}{2}Q^2}$$

$$u = \frac{\ln y - P}{Q}$$

$$du = \frac{dy}{Qy}$$

$$dy = Qy\,du$$

$$= Qe^{P+Qu} du$$

$$z = u - Q$$

"Creating Customer-Focused Success"

MCR

Second Moment, Sigma Value of Lognormal Distribution

- $E\left(X^2\right) = \int_0^\infty y^2 h(y)\,dy = \int_0^\infty \dfrac{y}{Q\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln\ y-P}{Q}\right)^2}\,dy$

$ = \dfrac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty \left(e^{P+Qu}\right)^2 e^{-\frac{1}{2}u^2}\,du = \dfrac{e^{2P}}{\sqrt{2\pi}} \int_{-\infty}^\infty e^{-\frac{1}{2}\left(u^2-4Qu\right)}\,du$

$$u = \frac{\ln\ y - P}{Q}$$

$$y = e^{P+Qu}$$

$ = \dfrac{e^{P+2Q^2}}{\sqrt{2\pi}} \int_{-\infty}^\infty e^{-\frac{1}{2}(u-2Q)^2}\,du$

$ = e^{2P+2Q^2} \dfrac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty e^{-\frac{z^2}{2}}\,dz = e^{2P+2Q^2}$

$$z = u - 2Q$$

$$\sigma^2 = E\left(X^2\right) - \left(E(X)\right)^2 = e^{2P+2Q^2} - \left(e^{P+\frac{1}{2}Q^2}\right)^2 = e^{2P+2Q^2} - e^{2P+Q^2} = e^{2P+Q^2}\left(e^{Q^2} - 1\right)$$

$$\sigma = e^{P+\frac{1}{2}Q^2}\sqrt{e^{Q^2} - 1}$$

Speaker Biography

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. He was a principal contributor to several Air Force cost studies of national significance, including the DSP/FEWS/BSTS/AWS/Brilliant Eyes Sensor Integration Study (1992) and the ALS/Spacelifter/EELV Launch Options Study (1993). He has served on national panels as an independent reviewer of NASA programs, for example the 2005 Senior External Review Team on cost-estimating methods for the Exploration Systems Mission Directorate, 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. Dr. Book joined MCR in January 2001 after 21 years with The Aerospace Corporation, where he held the title "Distinguished Engineer" during 1996-2000 and served as Director, Resource and Requirements Analysis Department, during 1989-1995. Dr Book is the current editor of ISPA's *Journal of Parametrics*, and the 2005 recipient of ISPA's Freiman Award for Lifetime Achievement. He earned his Ph.D. in mathematics, with concentration in probability and statistics, at the University of Oregon.