Automated Cost Estimating Integrated Tools

Basic Risk Overview

Sabrina Beane and Gina Fennell
ACEIT Users Conference, Santa Barbara, CA
January 2009

Tecolote Research, Inc.

Abstract

Presenters: Gina Fennell & Sabrina Beane *(Tecolote – Software Products and Services Group)*

Providing risk-adjusted estimates is becoming more important as cost estimating evolves. ACE contains a very comprehensive risk package that allows you to model many different distributions in a multitude of ways. New users are often intimidated by the many risk features of ACE.

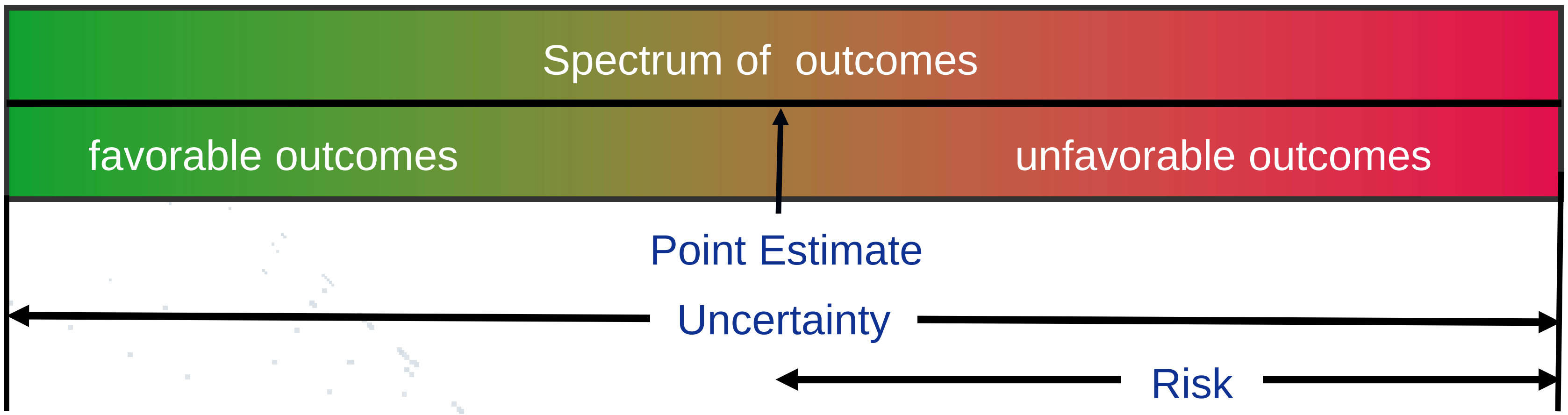
This presentation breaks down risk into the basics. It explains what risk is and how to model it in ACE. In addition, it will provide step-by-step instructions on how to conduct a risk analysis on a session and get risk-adjusted results.

Outline

- What is Risk?
- Why perform an uncertainty analysis?
- When do I perform uncertainty analysis?
- Where do I enter uncertainty?
  - Identify rows that have uncertainty
- How do I perform uncertainty analysis?
  - What do I know?
  - Model uncertainty on identified rows
  - Check for consistency
  - Measure correlation
  - Apply correlation
  - Report

What is Risk?

Situation with multiple outcome possibilities



Risk is the likelihood that an unfavorable outcome occurs.

Uncertainty is the range of both favorable and unfavorable outcomes.

Sources of Uncertainty

"In this world nothing can be said to be certain, except death and taxes."

- Benjamin Franklin

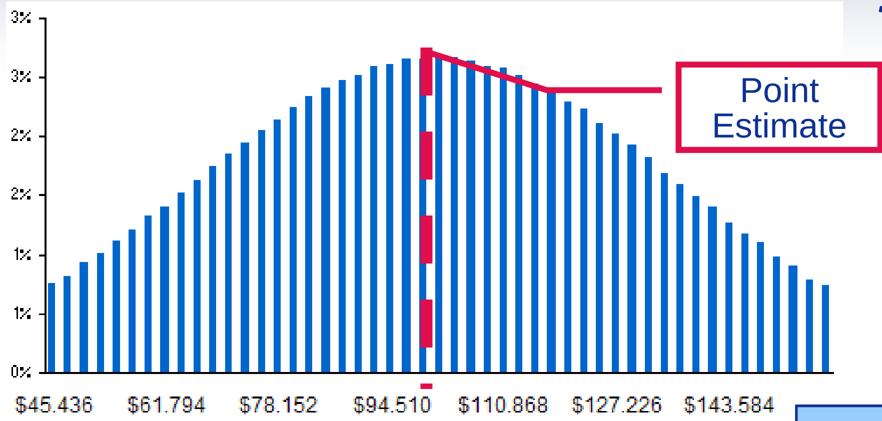
- **What elements are <u>certain</u>?**
  - Pure math
    - Cost * Qty
  - Elements that are firm fixed price
  - Referenced elements that already have uncertainty
- **What elements are <u>uncertain</u>?**
  - Estimating Methods
    - CERs, Factors
  - Inputs
    - Weight, Range, Durations

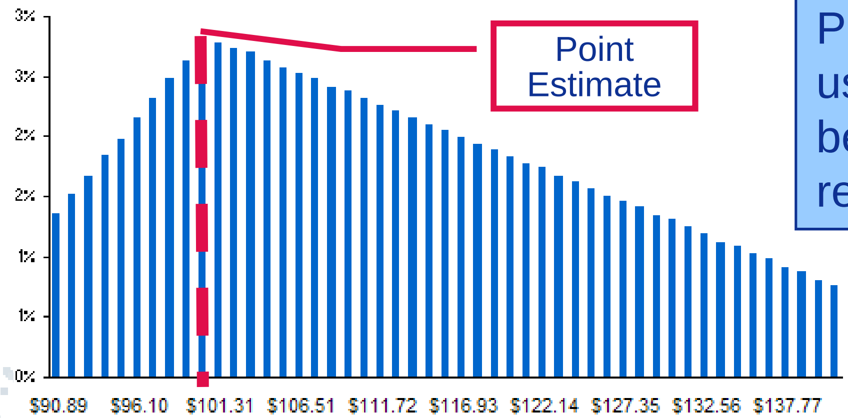
Why perform a Cost Uncertainty Analysis?

- **To put context on the point estimate**
  - Point estimate is one possible outcome in a range of results
- **To identify those items that contribute most to uncertainty of the project**
  - Provides data on where to focus efforts for risk mitigation
- **Budget obligations entail commitments which must be made in discrete dollars, not ranges of dollars**
  - Costs can be presented at a specific confidence level
- **The sum of the individual elements is not the most likely total result**

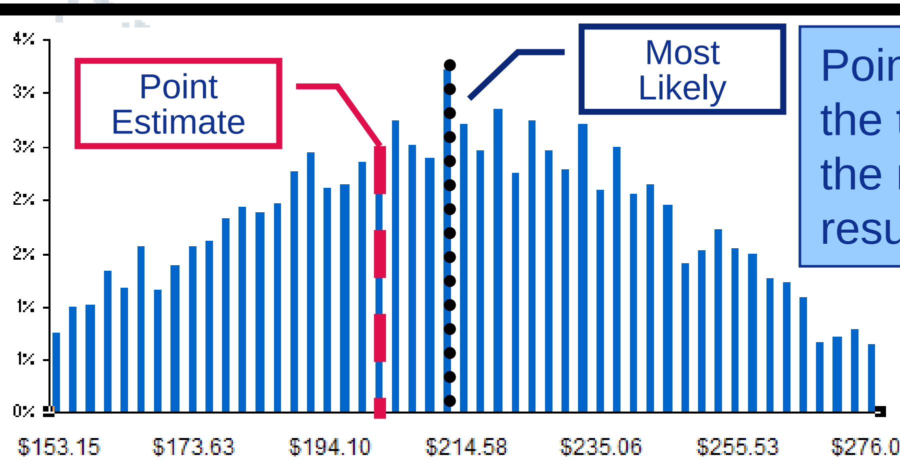
The sum of the elements is not the most likely result

	Point Estimate
Item 1	100
Item 2	100
+	
Total	200



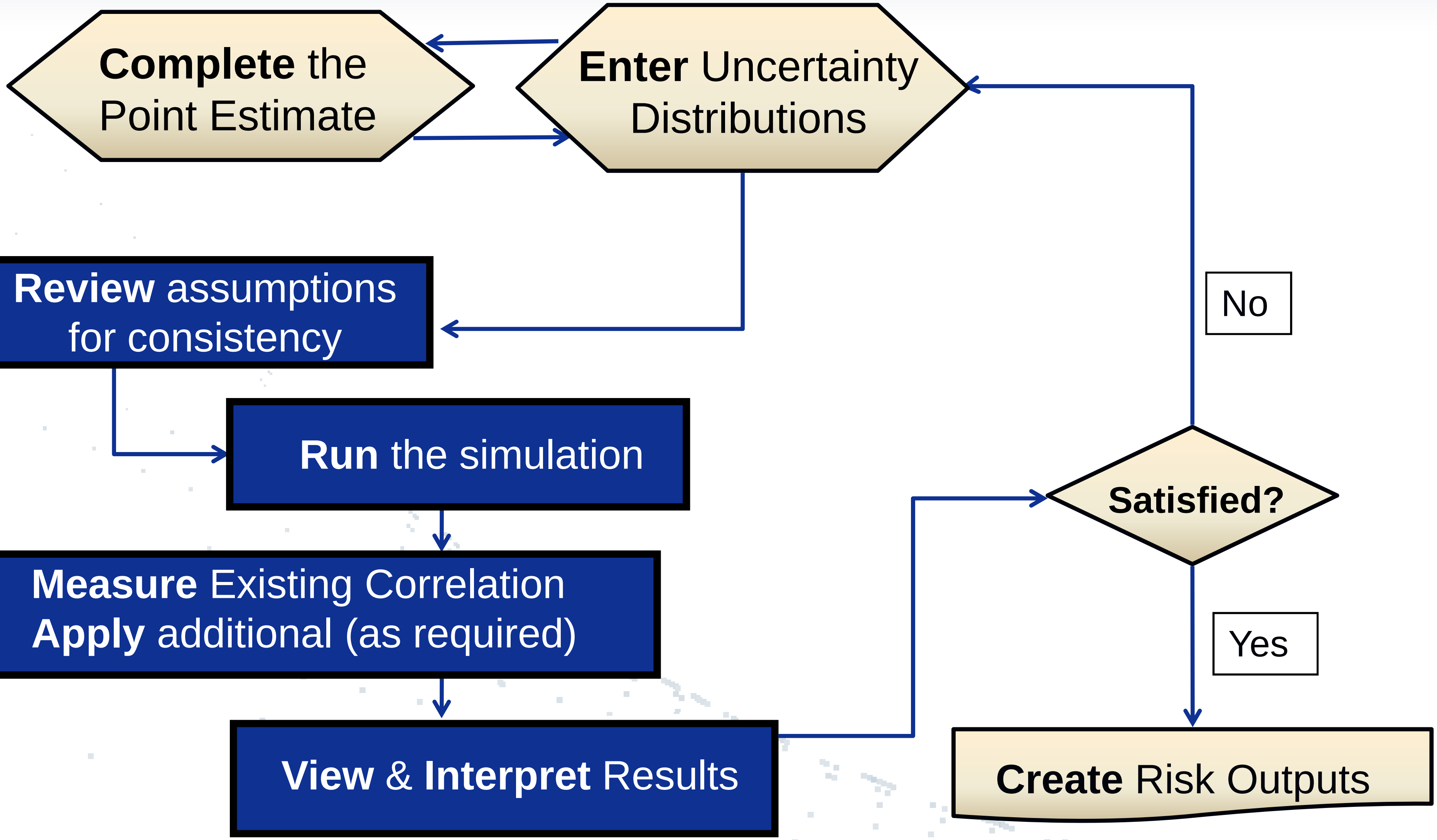


Point Estimates are usually assumed to be the most likely results



Point Estimates of the total are **not** the most likely result

When to Enter Uncertainty



Where to Enter Uncertainty

- **Identify rows that contain uncertainty**

Estimating Method	Example	Enter Uncertainty on the Equation?	Enter Uncertainty on the Inputs?
Parametric CER	30.15 + 1.049 * WarheadWt	Yes	Yes
Analogies	SystemX * ComplexFactor	No	Yes
Factor	0.15*PMP	Yes	Maybe
Build-up	EffortHours * LaborRate	No	Yes
Throughputs	$10,500	Yes	NA
Third Party Tools	Effort Hours Estimate Derived from another tool	Yes	NA

- **No Uncertainty on Quantity rows; treat as discrete what-if cases.**

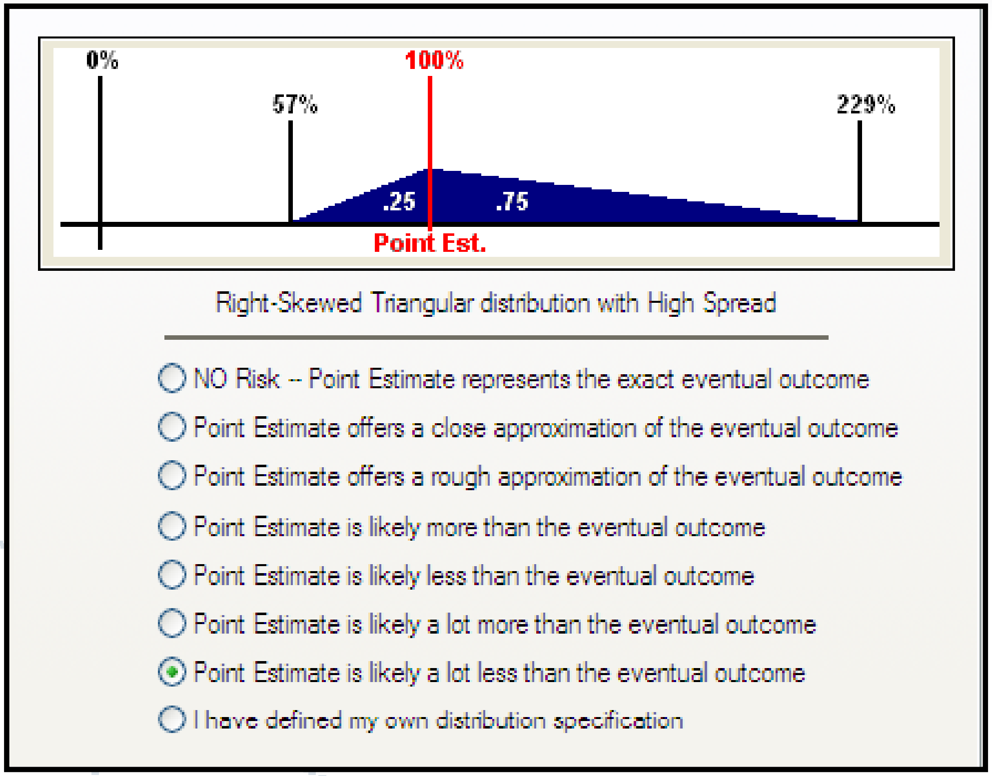
How to Enter Uncertainty

- **Identify rows that have uncertainty**
- **What do I know about the uncertainty?**
  - Not much
  - Some
  - I have a CER with Statistics
- **Model uncertainty on identified rows**
  - Think about the shape of possible outcomes
    - Is there a greater chance of overrun or under-run
  - Identify data range/bounds

Defining Distributions with Little or No Information

- **When you have little or no information about a row's uncertainty, consider the following questions:**
  - How wide is the range of possible values?
    - ACE uses the term ***Spread*** to quantify the range of the data.
      - Low – models a small range
      - Medium – models a moderate range
      - High – models a wide range
  - How good is your point estimate?
    - ACE uses the term ***Skew*** to show where more of the uncertainty lies.
      - Is the point estimate in the **center** of the range?
      - Is the uncertainty to the **right** of the point estimate?
      - Is the uncertainty to the **left** of the point estimate?

Input All Form – Basic Mode



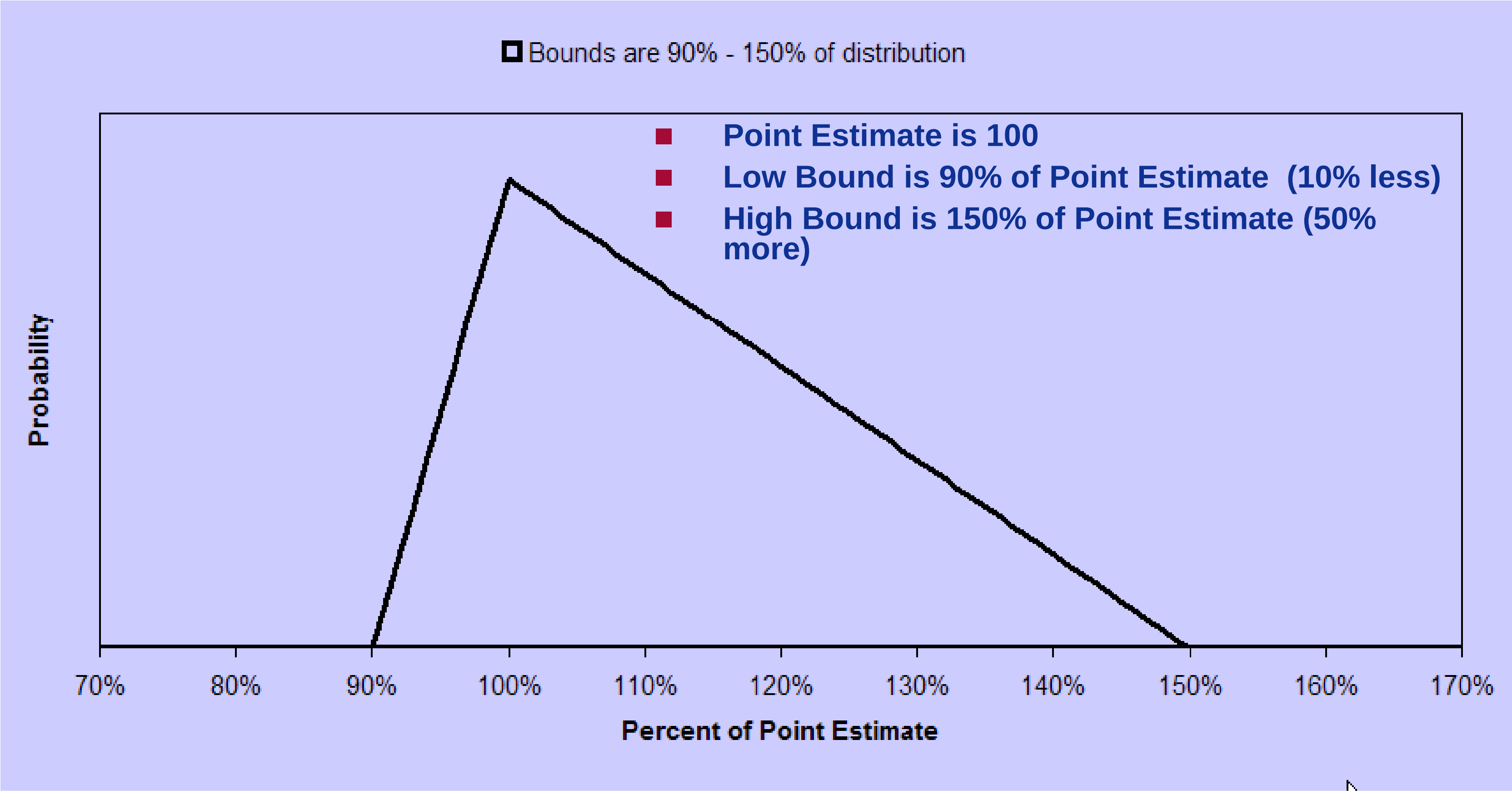
- The above picture shows a wide range of values where the point estimate is on the lower end of the range of values
- ACE models this as **High Spread** and **Right Skew**

Defining Distributions with Limited Information

- **If you have information on the range of the data, consider the following questions:**
  - What are the high and low bounds?
    - Are they a percentage of the point estimate or actual values?
    - What do the values actually represent?
  - What is the distribution shape?
    - Triangular
    - LogNormal
    - Uniform
    - Normal

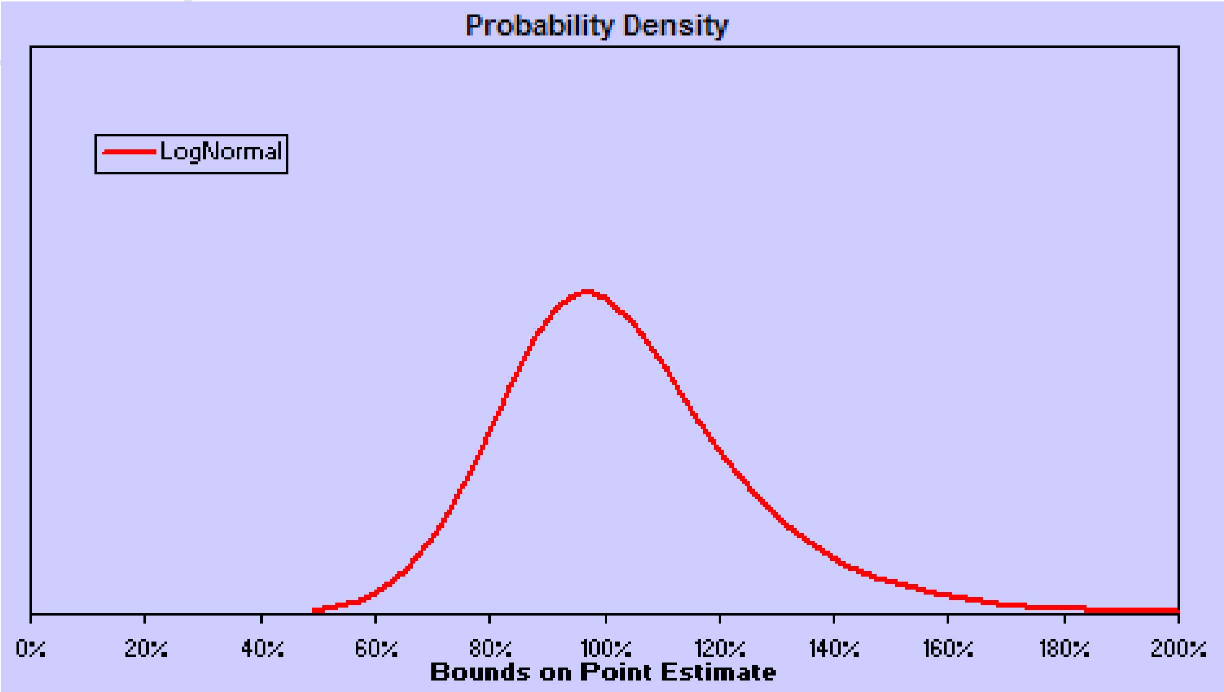
Triangular Distribution

- **Requires 3 pieces of information**
  - Most Likely Result
  - Low Bound
  - High Bound
- **The most likely is the point estimate**



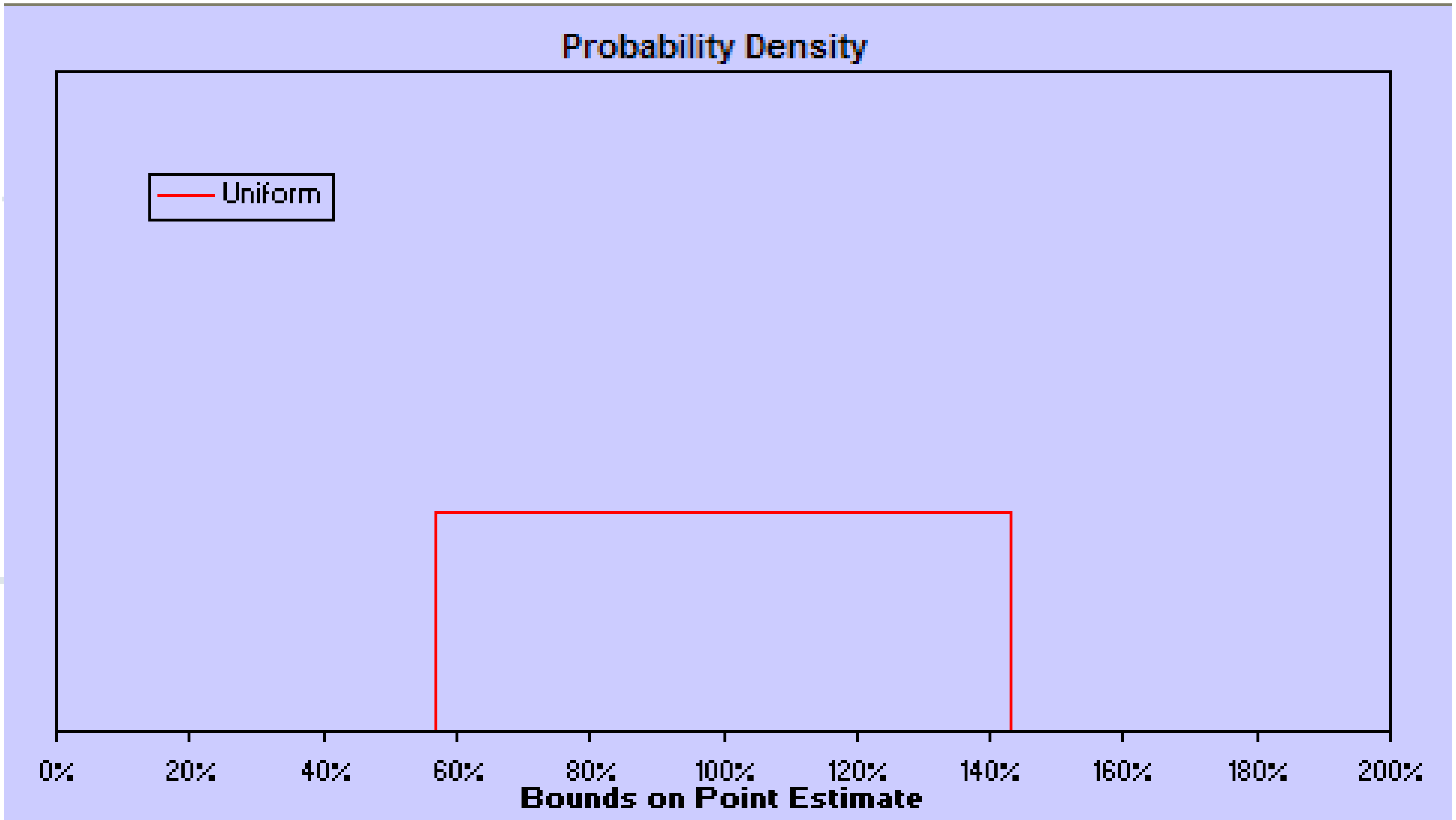
Lognormal Distribution

- **Bell Shaped curve with one finite tail and one infinite tail**
- **Requires 2 pieces of information**
  - Point Estimate (Most likely value)
  - Upper **or** lower bound (the log space mirror image is calculated for the other bound)
- **Why choose LogNormal?**
  1. The value can increase without limits, but cannot fall below zero.
  2. The value is positively skewed, with most of the values near the lower limit.



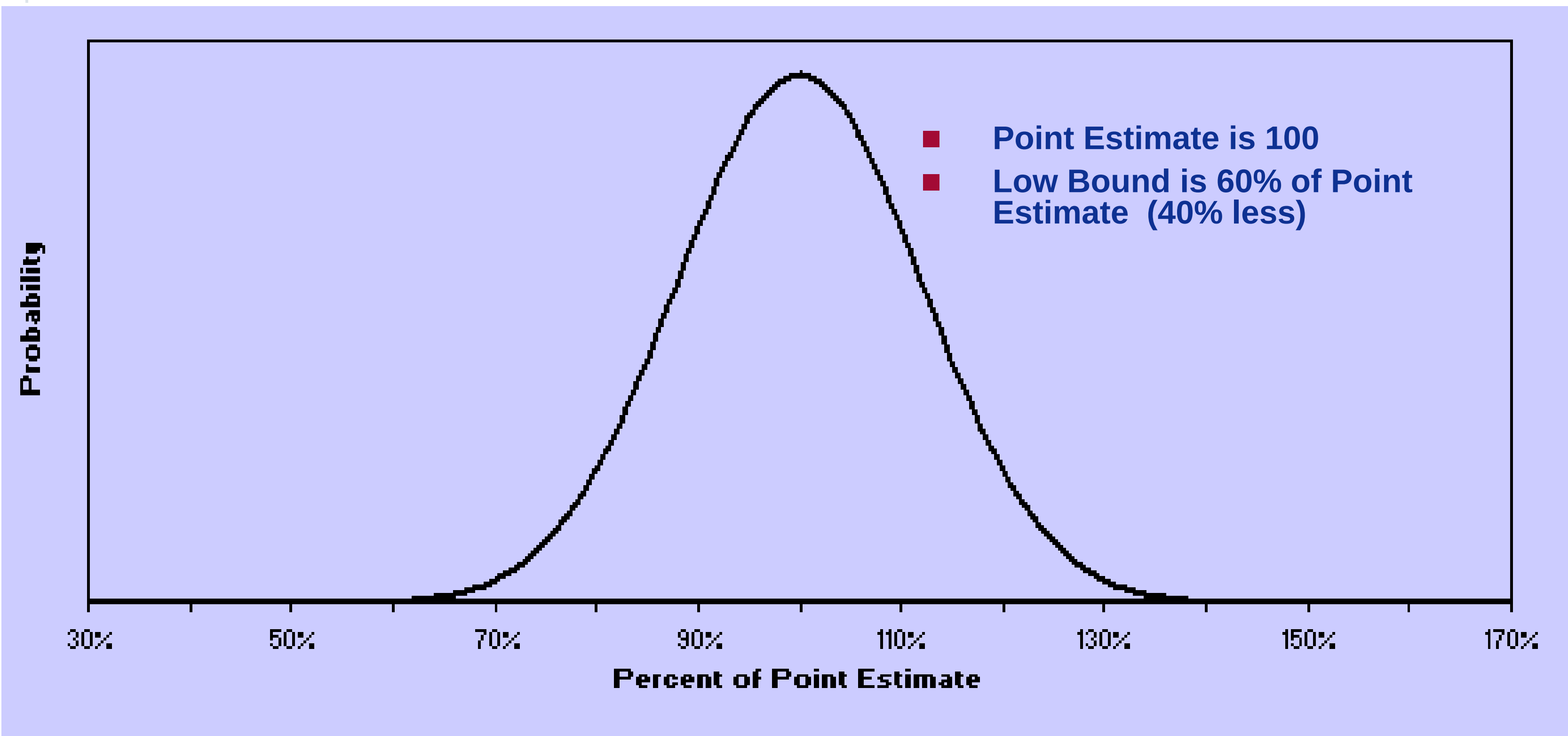
Uniform Distribution

- All outcomes have the same probability of occurrence
- Point Estimate is mode by default



Normal Distribution

- **Bell Shaped symmetric curve with tails that are infinite**
- **The point estimate is the mean/median/mode.**
- **Requires 2 pieces of information**
  - Point Estimate (most likely value)
  - Upper **or** lower bound (the mirror image is calculated for the other bound)



How do I pick a Distribution Shape?

- **Triangular**
  - Use with Engineering data or analogy estimate
  - Easy to use and understand
  - Use when likelihood decreases with distance from the point estimate
- **LogNormal**
  - Use with Log-linear CERs or in absence of better information
  - Use when value can increase without limits but cannot fall below zero and when most values are near the mode
- **Uniform**
  - Use with Engineering data or analogy estimate
  - Use when every value has an equal likelihood of occurrence
- **Normal**
  - Use with Linear CERs or a CER with an additive error
  - Symmetric distribution means point estimate has equal likelihood of overrun or under-run

Defining Distributions when you have a CER and Supporting Statistics

- **Determine the form of the CER**
  - Is it a linear, a log linear, learning curve or other equation form?
- **Identify the distribution shape**
  - For linear CER, use **Normal**
  - Otherwise, use **LogNormal**
- **Use the CER statistical data to define the spread of the distribution**
  - LogNormal – Use the **Adjusted Standard Error**, CV, or the low/high bounds.
  - Normal – Use the **CV**, Standard deviation, or the low/high bounds.

Coefficient of Variation (CV) = StdDev/Mean

Review the Assumptions for Consistency

- **Review the model for standard distribution selection**
  - Use the same distribution in similar situations throughout the estimate (e.g., for all factor method rows, use Triangular)
- **Review the model for a consistent approach when applying bounds**
  - In general, be consistent when specifying the range for a distribution type (e.g., always use high/low with Triangular or always use statistics with Normal)
- **Are the inputs easy to understand?**
  - Can you explain why you selected the shape and bounds?
  - Is it documented?

Run the Simulation

- ✓ Point Estimate Complete
- ✓ Distributions specified
- ✓ Consistent specification of uncertainty

Now what?

1. **Calculate with Risk (Calc > Calc with RI$K)**
2. **Check out Confidence Level at each row that contains uncertainty**
   - e.g., $ 130,123.192 (38%)
3. **View RI$K Statistics**
   - Open Inputs/Results Viewer
   - Select BY RI$K Statistics from View menu

So What does it Mean?

Confidence Level: Probability of the point estimate value. Typical values for the Total row are 15%-45%.

Mean: Average value for the row after the simulation

Coefficient of Variation (CV): a normalized measure of dispersion . Typical values at the total are .15 -.35. Less than .15 indicates very optimistic ranges or a lack of correlation

WBS/CES Description	Point Estimate	Mean	Std Dev	CV	5%	10%	15%
*Estimate							
Total	$ 130,123,192 (38%)	$ 144,595,644	$ 39,880,575	0.2758	$ 89,012,763	$ 98,293,007	$ 105,043,194
Manufacturing	$ 94,600,493 (49%)	$ 97,667,970	$ 25,459,008	0.2607	$ 61,486,116	$ 68,257,258	$ 73,037,208
Air Vehicle	$ 82,261,299 (47%)	$ 86,262,689	$ 22,204,794	0.2574	$ 54,968,932	$ 60,455,059	$ 64,387,796
Integration	$ 12,339,195 (63%)	$ 11,405,281	$ 3,784,880	0.3319	$ 6,138,618	$ 7,017,547	$ 7,599,571
SEPM	$ 35,002,183 (30%)	$ 46,406,114	$ 18,249,922	0.3933	$ 23,192,549	$ 26,020,129	$ 29,025,602
Other	$ 520,517 (50%)	$ 521,560	$ 180,223	0.3455	$ 224,985	$ 288,800	$ 332,852
*INPUT VARIABLES Air Vehicle Unit Cost	$ 9,140,144 (47%)	$ 9,584,743	$ 2,467,199	0.2574	$ 6,107,659	$ 6,717,229	$ 7,154,200
Air Vehicle Buy Quantity	9	9			9	9	9
*Technical/Performance Characterist							
Air Vehicle Takeoff Weight (lbs)	12000.00 (25%)	12800.05	1020.80	0.0797	11340.23	11560.40	11728.54
Air Vehicle Range (nmi)	250.00 (42%)	256.67	24.63	0.0960	217.21	224.52	230.00

Std Dev: a measure of the dispersion of the simulation results

Confidence Levels: each row's value at the selected confidence . NOTE: Child rows do not sum to parent.

Add Correlation (almost always)

- **Correlation measures the strength of the relationship between elements**
  - ***Functional correlation*** is correlation that is built-in based on the estimating methodology
    - E.g., SEPM = 0.15 * Air Vehicle
  - ***Defined correlation*** is correlation manually entered
- **Correlation ranges from -1 to 1**
  - +/-1.0 is a perfect relationship
  - +/- 0.7 is a moderate relationship
  - +/- 0.5 is a weak relationship
- **Correlation can only be defined for elements that have risk distributions**
- Use **Tools>RI$K Grouping and Correlation** to add defined correlation

How do we Model Correlation?

- **Steps to Model Correlation**
  1. Measure existing correlation in the session
     - Run the Correlation Report
     - Identify rows where the correlation measure is minimal (e.g., values are close to 0)
     - Identify rows that have correlation measures that are strong but shouldn't be (e.g., values are closer to 1 or -1)
  2. Remove unintentional correlation and add correlation where needed
     - Use the Risk Grouping and Correlation tool
     - Create groups of related elements and assign correlation
     - To remove unwanted correlation, find source and adjust estimate as necessary to remove correlation
  3. Rerun the Correlation report to verify correct correlation

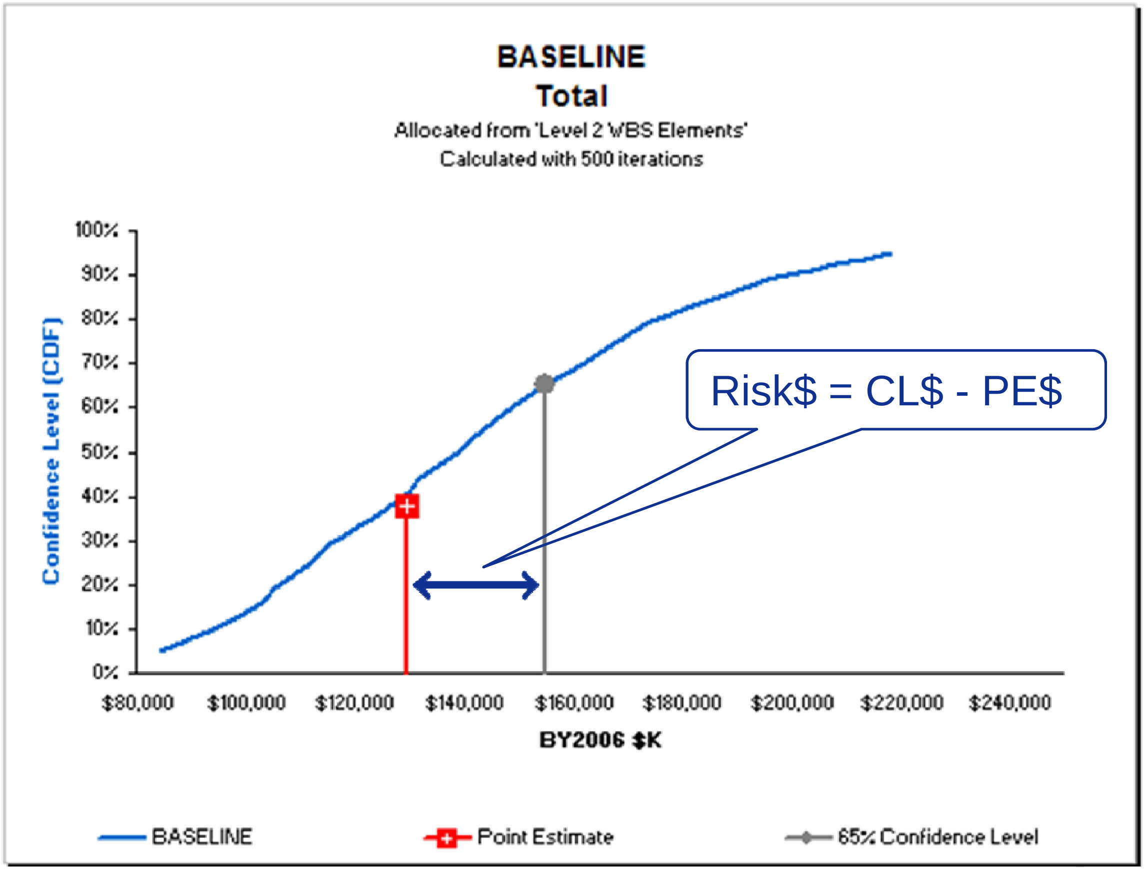
How do we Model Correlation the Easy Way?

- **In the absence of time and better information**
  1. Measure existing correlation in the session
     - If numbers are significantly different from 0, correlation already exists (e.g., > .2)
  2. Add correlation to elements that you KNOW are related
     - E.g., Air Vehicle Weight and Air Vehicle Range are inversely correlated – as weight increases, range decreases
     - Create group of these elements and assign correlation
  3. Add correlation to elements not already associated with a group
     - Create group for leftover input variables
       - Assign minimal correlation (e.g., .25)
     - Create group for leftover elements in the WBS
       - Assign minimal correlation (e.g., .25)
  4. Rerun the Correlation report to verify reasonable correlation

- **Clearly communicate the simulation results**
  - Show an S-Curve - a graphical representation of where the point estimate falls within the risk range
  - Show Phased results at a particular confidence level in Then Year dollars (for budgeting purposes)

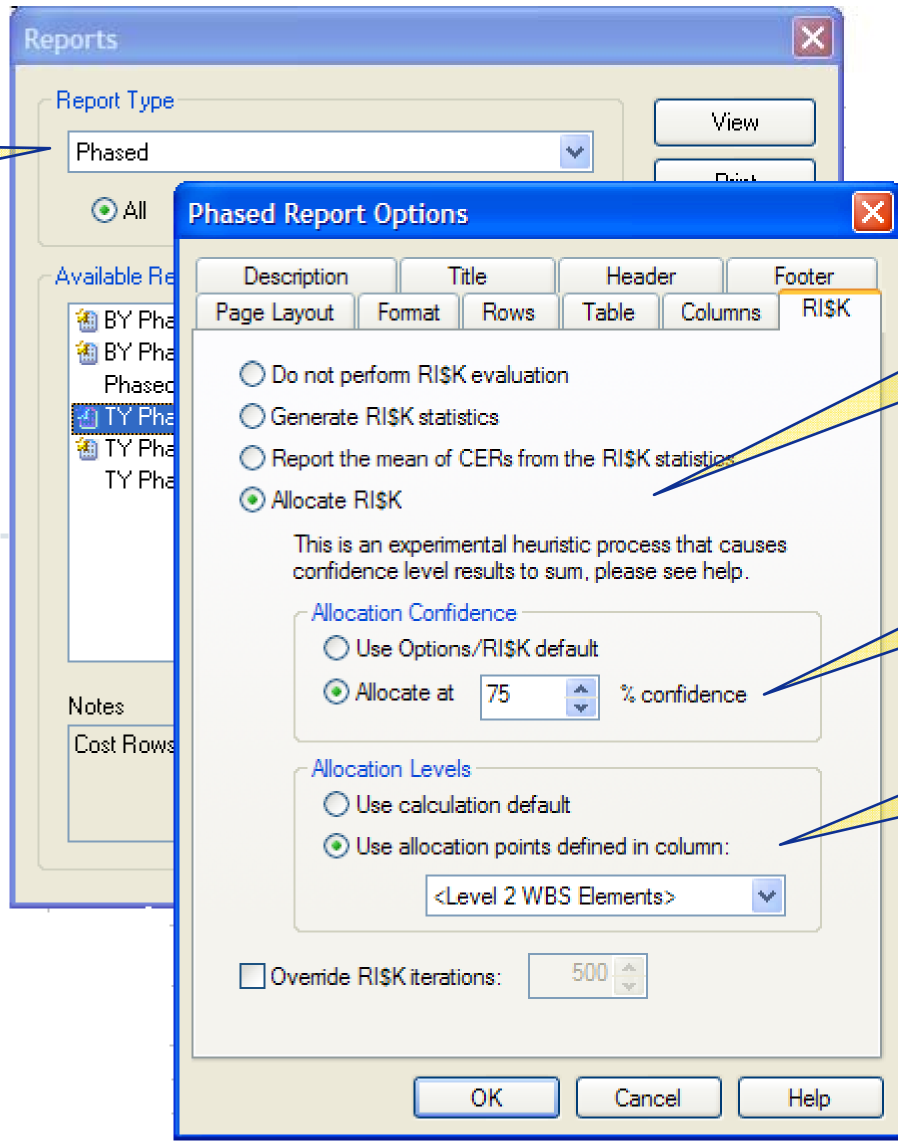
What does the S-curve tell you?

- POST RI$K Chart plots the confidence level results
- Gives you the risk dollars at a specified confidence level



How do I create a Phased Allocated RI$K report?

Edit/Create a Phased Report



Select Allocate Risk

Select the Confidence Level

Select Allocation Level

Time-phased Allocated RI$K Report

- This report show TY phased costs at the 75% confidence level
- Allocated at level 2 means these elements will be at exactly 75% confidence

02a - Basic Risk.aceit - TY Phased Costs (TY $K, Time Phased, Case: Point Estimate, 75% CL allocated at Level 2)

	Cost Element	Approp	Total	FY 2006	FY 2007	FY 2008	FY 2009
19	*Estimate						
20	Total		$ 184,410.056 (~76%)	$ 19,414.349	$ 19,873.497	$ 40,487.831	$ 41,371.320
21	Manufacturing		$ 122,602.915 (75%)	$ 12,856.168	$ 13,162.558	$ 26,926.745	$ 27,516.297
22	Air Vehicle	3010	$ 107,702.772 (74%)	$ 11,293.736	$ 11,562.889	$ 23,654.291	$ 24,172.194
23	Integration	3010	$ 14,900.143 (74%)	$ 1,562.432	$ 1,599.668	$ 3,272.454	$ 3,344.103
24	SEPM	3010	$ 61,127.040 (75%)	$ 6,409.795	$ 6,562.553	$ 13,425.066	$ 13,719.003
25	Other	3080	$ 680.101 (75%)	$ 148.386	$ 148.386	$ 136.020	$ 136.020
26							
27	*INPUT VARIABLES						
28	Air Vehicle Unit Cost	3010	$ 11,027.426 (75%)				
29	Air Vehicle Buy Quantity		9	1	1	2	2
30							
31	*Technical/Performance Characte						
32	Air Vehicle Takeoff Weight (lbs)		13520.43 (75%)				
33	Air Vehicle Range (nmi)		274.16 (75%)				

Summary

- **You can easily perform a risk analysis using ACE**
  - Identify rows that have uncertainty
  - Model uncertainty on identified rows
    - Distribution types: Triangular, LogNormal, Normal, and Uniform
    - Bounds: specify using Spread and Skew, Low and High values, or statistics from a regression analysis
  - Check your risk specifications for consistency
  - Measure and apply correlation if needed
  - Report results

THANK YOU