Expert Elicitation of a Maximum Duration using Risk Scenarios

Presented by:

Marc Greenberg
Cost Analysis Division (CAD)
National Aeronautics and Space Administration
How Does A Cost Analyst REALLY Quantify the Unquantifiable?

A. Yell out a cool sounding number with conviction!
B. Divide what’s available in your budget by 1, then multiply it by 0.78
C. Apply common estimating methods (e.g., analogy & parametric)
D. Use subject matter expert opinion
E. Incorporate cost risk & uncertainty analysis techniques
F. C, D or E (or any combination of C, D and E)
Outline

• Purpose of Presentation
• Background
  – The Uncertainty Spectrum
  – Five Expert Elicitation (EE) Phases
• Case Study: Estimate Morning Commute Time
  – Establish Framework of Interview Session
  – 1: Direct Input (DI) Method
    • The Risk Reference Table (note: Also used for SB-RRW Method)
  – 2: Scenario Based Relative Risk Ratio (SB-RRW) Method
• Suggested use of DI and SB-RRW Methods in Practice
• Conclusion
Demonstrate two expert elicitation methods that ...

1. Model expert’s inputs as a triangular distribution
   - Direct Input (DI) Method
     • Q&A to elicit Min, Most-Likely & Max from expert, and then adjust for expert bias.
   - Scenario Based Relative Risk Weighting (SB-RRW) Method
     • Expert-derived scenario-based factors applied to Most-Likely to estimate Min & Max.

2. Incorporate techniques to account for expert bias
   - DI: Q&A elicits likelihood to be below Min & above Max
   - SB-RRW: Use of pairwise comparison helps prevent ‘gaming’ the outcome
   - For both methods, use of visual aids helps expert calibrate original inputs

3. Are structured in a way to justify expert inputs
   - DI: Each response to each question requires a rationale from the expert
   - SB-RRW: Output provides each risk factor’s contribution to uncertainty

These two methods are set up so that they are not too complex to be impractical & not too simple to be too subjective.
Expert judgment should only be used when there is (i) lack of time for collection & analysis of historical data, (ii) lack of available historical data or (iii) the design is incomplete.
Contrary to popular belief, this Dilbert Cartoon does NOT give the best definition of Expert “Judgment” 😊

Try this one instead …

Expert Judgment (for estimating) are value estimates developed solely on the basis of a person’s experience & knowledge of the process or product being estimated.
**Expert Elicitation (EE) Phases**

Expert Elicitation consists of five phases:

*(note that Phases 4 & 5 are iterative)*

1. Motivating the expert

2. Structuring objective, assumptions & process

3. Training (conditioning) the expert

4. **Assessing (encoding) expert’s responses**
   - Q&A – Expert’s experienced-based opinion is elicited
   - Quantitative results w/ documented rationale

5. Verifying encoded values & documentation

---

This majority of this presentation covers only Phase 4
Example: Estimate Commute Time

• Why this example?
  – Fairly easy to find a subject matter expert (SME)
  – It is a parameter that is measurable
  – Most experts can estimate a most likely time
  – Factors that drive uncertainty can be readily identified
  – People generally care about their morning commute time!

Assume only Given a Most-Likely Commute = 55 minutes
EE Phases 1 and 2: *Framework of Interview*

**EE Phase 1: Motivating the expert**
- Explain the importance & reasons for collecting the data
- Explore stake in decision & potential for motivational bias

**EE Phase 2: Structuring objective, assumptions & process**
- Must be explicit about what you want to know & why you need to know it
  - Clearly define variable & avoid ambiguity and explain data values that are required (e.g. hours, dollars, %, etc)

You should have worked with SME to develop the Objective and up to 6 Major Assumptions in the table below

<table>
<thead>
<tr>
<th>Objective:</th>
<th>Develop an uncertainty distribution associated with time (minutes) it will take for your morning commute starting 1 October 2015.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assumption 1:</td>
<td>Your commute estimate includes only <strong>morning</strong> driving time</td>
</tr>
<tr>
<td>Assumption 2:</td>
<td>Period of commutes occur in <strong>FY15</strong> (from 1 Oct 2015 thru 30 Sep 2016)</td>
</tr>
<tr>
<td>Assumption 3:</td>
<td>Commute time will be measured in <strong>minutes</strong></td>
</tr>
<tr>
<td>Assumption 4:</td>
<td>‘<strong>Most Likely</strong>’ commute time reflects the time expected to occur most often</td>
</tr>
<tr>
<td>Assumption 5:</td>
<td>The commute ‘process’ will be <strong>analogous</strong> to the one you've been doing</td>
</tr>
<tr>
<td>Assumption 6:</td>
<td>Unless prompted by interviewer, do not try to account for extremely rare &amp; unusual scenarios</td>
</tr>
</tbody>
</table>
EE Phase 3: Overarching Interview Process

3. Training (conditioning) the expert
   • Go over instructions for Q&A process
   • Emphasize benefits of time constraints & iterations

Instructions: This interview is intended to be conducted in up to 3 iterations. Each iteration should take no longer than 20 minutes.

A. Based on your experience, please answer all interview questions.
B. Once you've completed the questions, review them & take a 15 minute break.
C. If required, use the graphics to assist you to answer select questions again.
D. Your interviewer is also here to assist you at any point during the interview.

Notes on 2nd and 3rd iterations (if needed):

A. The 2nd iteration is intended to be a refinement of your 1st round answers.
   - Use lessons-learned from the 1st iteration to assist you in the 2nd iteration.
B. The 3rd iteration is intended to be a refinement of your 2nd round answers.
   - Use lessons-learned from the 2nd iteration to assist you in the 3rd iteration.
**Direct Input (DI) Method.**

The DI Method elicits the Most-Likely, Lowest & Highest values from a subject matter expert (SME) in ‘round 1’ then revisits these questions with the assistance of graphics and a “risk reference” table.

**Pros:**
- a) Relatively fast/efficient way to use SME opinion to get min, most likely & max
- b) Easy to explain to stakeholders and decision-makers
- b) Enables SME to iterate using graphics, risk factors and risk scenarios

**Cons:**
- a) SME is required to provide initial estimates of low, most likely and high values
- b) DI Method typically must counter SME anchoring to her most likely estimate
  - i.e., DI Method nearly always requires adjustment to account for expert bias
- c) Expert must recall (& later explain) duration or cost extremes
- d) Risk factors affecting dispersion are described after 1st iteration
- e) Takes time to set up “risk reference” table
Question 1a and 1b: Expert creates “value-scale” tailored his/her bias …

In the context of your morning commute time …
What probability would you assign to a commute time that's Very Unlikely?
What probability would you assign to a commute time that's Extremely Unlikely?

Available Selection of Values to the Expert (shaded cells were selected by expert):

<table>
<thead>
<tr>
<th>Very Likely</th>
<th>Very Unlikely</th>
<th>Extremely Likely</th>
<th>Extremely Unlikely</th>
</tr>
</thead>
<tbody>
<tr>
<td>80.0%</td>
<td>20.0%</td>
<td>96.0%</td>
<td>4.0%</td>
</tr>
<tr>
<td>82.5%</td>
<td>17.5%</td>
<td>97.0%</td>
<td>3.0%</td>
</tr>
<tr>
<td>85.0%</td>
<td>15.0%</td>
<td>98.0%</td>
<td>2.0%</td>
</tr>
<tr>
<td>87.5%</td>
<td>12.5%</td>
<td>98.5%</td>
<td>1.5%</td>
</tr>
<tr>
<td>90.0%</td>
<td>10.0%</td>
<td>99.0%</td>
<td>1.0%</td>
</tr>
<tr>
<td>92.5%</td>
<td>7.5%</td>
<td>99.5%</td>
<td>0.5%</td>
</tr>
<tr>
<td>95.0%</td>
<td>5.0%</td>
<td>99.9%</td>
<td>0.1%</td>
</tr>
</tbody>
</table>
EE Phase 4: DI Method (iteration 1)

Question 1a and 1b: Expert creates a “value-scale” tailored his/her bias …

What probability would you assign to a commute time that's **Very Unlikely** = 10.0%

What probability would you assign to a commute time that's **Extremely Unlikely** = 1.0%

<table>
<thead>
<tr>
<th>Descriptor</th>
<th>Explanation</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absolutely Impossible</td>
<td>No possibility of occurrence</td>
<td>0.0%</td>
</tr>
<tr>
<td>Extremely Unlikely</td>
<td>Nearly impossible to occur; very rare</td>
<td>1.0%</td>
</tr>
<tr>
<td>Very Unlikely</td>
<td>Highly unlikely to occur; not common</td>
<td>10.0%</td>
</tr>
<tr>
<td>Indifferent between &quot;Very Unlikely&quot; &amp; &quot;Even chance&quot;</td>
<td></td>
<td>30.0%</td>
</tr>
<tr>
<td>Even Chance</td>
<td>50/50 chance of being higher or lower</td>
<td>50.0%</td>
</tr>
<tr>
<td>Indifferent between &quot;Very Likely&quot; &amp; &quot;Even chance&quot;</td>
<td></td>
<td>70.0%</td>
</tr>
<tr>
<td>Very Likely</td>
<td>Highly likely to occur; common occurrence</td>
<td>90.0%</td>
</tr>
<tr>
<td>Extremely Likely</td>
<td>Nearly certain to occur; near 100% confidence</td>
<td>99.0%</td>
</tr>
<tr>
<td>Absolutely Certain</td>
<td>100% Likelihood</td>
<td>100.0%</td>
</tr>
</tbody>
</table>

Only 2 probabilities needed to be elicited in order to create a Value-Scale that has 9 categories!
EE Phase 4: DI Method (iteration 1)

4. Assessing expert’s responses (Q&A)

Based upon your experience, please answer #2 - #8:
(To assist you, refer to objective & assumptions in slide 9)

2. Describe input parameter (WBS 4): Morning commute time (in minutes)
3. What has been your Most Likely commute time in FY14? 50
4. What will be your Most Likely commute time in FY15? 55 = M
5. What will be your shortest commute time in FY15? 42 = L
6. What’s the chance an FY15 commute is < 42 minutes? Indifferent-Low
   • Discuss & document extremely rare events, unusual scenarios and/or “unknown unknowns”
7. What will be your longest commute time in FY15? 80 = H
8. What’s the chance an FY15 commute is > 80 minutes? Very Unlikely
   • Discuss & document extremely rare events, unusual scenarios and/or “unknown unknowns”

This 1st iteration tends to result in anchoring bias on M, over-confidence on L and H, and poor rationale
4. Assessing expert’s responses (Q&A)

**User-Provided Distribution for Commute Time**

Red dot depicts unadjusted point estimate. Dashed lines depict unadjusted lowest & highest

PDF created based upon Expert’s responses to Questions 2 through 8.

Take 15 min. break then build “Risk Reference Table” and start Q&A Iteration #2

Given from Expert : L=42, M=55, H=80, \( p(x<L)=0.30 \) and \( p(x>H)=0.10 \)

Calculation of ‘true’ L and H (a) : \( L = 1.56 \) and \( H = 101.15 \) … Do these #’s appear reasonable?

(a) Method to solve for L and H presented in “Beyond Beta,” Ch1 (The Triangular Distribution)
EE Phase 4: **DI Method** (prep for iteration 2)

- Prior to starting DI Method - Iteration #2, the SME and Interviewer should work together to create a “**Risk Reference Table**”
  - **Step 1**: Create an Objective Hierarchy
  - **Step 2**: Brainstorm Risk Factors
  - **Step 3**: Map Risk Factors to Objective Hierarchy
  - **Step 4**: Describe / Define the Risk Factors

**Note:** This Risk Reference Table is also used for SB-RRW
Create Risk Reference Table (Step 1)

Step 1: SME & Interviewer Create an Objective Hierarchy

Q: To minimize commute time, what is your primary objective?
A: Maximize average driving speed

Q: What are primary factors that can impact driving speed?
A: Route Conditions, # of Vehicles on Roads, Mandatory Stops & Driving Efficiency

Q: Is it possible that other factors can impact driving speed?
A: Yes … (but SME cannot specify them at the moment)

<table>
<thead>
<tr>
<th>Objective</th>
<th>Means</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximize Average Driving Speed</td>
<td>These are Primary Factors</td>
</tr>
<tr>
<td></td>
<td>that can impact Objective</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Route Conditions</td>
<td></td>
</tr>
<tr>
<td># of Vehicles on Roads</td>
<td></td>
</tr>
<tr>
<td>Mandatory Stops</td>
<td></td>
</tr>
<tr>
<td>Driving Efficiency</td>
<td></td>
</tr>
<tr>
<td>Undefined</td>
<td></td>
</tr>
</tbody>
</table>

The utility of this Objective Hierarchy is to aid the Expert in:

(a) Establishing a Framework from which to elicit most risk factors,

(b) Describing the relative importance of each risk factor with respect to means & objective, and

(c) Creating specific risk scenarios
Step 2: SME & Interviewer Brainstorm Risk Factors

Using the Objective Hierarchy as a guide, the SME answers the following:

Q: What are some factors that could degrade route conditions?
A: Weather, Road Construction, and Accidents

Q: What influences the # of vehicles on the road in any given morning?
A: Departure time, Day of the Work Week, and Time of Season (incl. Holiday Season)

Q: What is meant by Mandatory Stops?
A: By law, need to stop for Red Lights, Emergency Vehicles and School Bus Signals

Q: What can reduce Driving Efficiency?
A: Picking the “Slow Lane”, Talking on the Cell Phone and Driving Below Speed Limit

<table>
<thead>
<tr>
<th>Objective</th>
<th>Means</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximize Average Driving Speed</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Route Conditions</td>
</tr>
<tr>
<td></td>
<td># of Vehicles on Roads</td>
</tr>
<tr>
<td></td>
<td>Mandatory Stops</td>
</tr>
<tr>
<td></td>
<td>Driving Efficiency</td>
</tr>
<tr>
<td></td>
<td>Undefined</td>
</tr>
</tbody>
</table>

These are Primary Factors that can impact Objective.
### Create Risk Reference Table (Steps 3 & 4)

**Step 3:** SME & Interviewer **Map Risk Factors** to the Objective Hierarchy

**Step 4:** SME & Interviewer work together to **Describe Risk Factors**

<table>
<thead>
<tr>
<th>Objective</th>
<th>Means</th>
<th>Risk Factors</th>
<th>Description (can include examples)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximize Average Driving Speed</td>
<td>Route Conditions</td>
<td>Weather</td>
<td>Rain, snow or icy conditions. Drive into direct sun.</td>
</tr>
<tr>
<td># of Vehicles on Roads</td>
<td>Accidents</td>
<td>Vehicle accidents on either side of highway.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Road Construction</td>
<td>Lane closures, bridge work, etc.</td>
<td></td>
</tr>
<tr>
<td>Mandatory Stops</td>
<td>Departure Time</td>
<td>SME departure time varies from 6:00AM to 9:00AM</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Day of Work Week</td>
<td>Driving densities seem to vary with day of week</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Season &amp; Holidays</td>
<td>Summer vs. Fall, Holiday weekends</td>
<td></td>
</tr>
<tr>
<td>Driving Efficiency</td>
<td>Red Lights</td>
<td>Approx 8 traffic intersections; some with long lights</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Emergency Vehicles</td>
<td>Incl. police, firetrucks, ambulances &amp; secret service</td>
<td></td>
</tr>
<tr>
<td></td>
<td>School Bus Signals</td>
<td>School buses stopping to pick up / drop off</td>
<td></td>
</tr>
<tr>
<td>Undefined</td>
<td>Pick Slow Lane</td>
<td>Just check out opening scene of &quot;Office Space&quot; :)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Talking on Cellphone</td>
<td>On rare occasion, will call someone during commute</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Driving below Speed Limit</td>
<td>Can be due to less work pressure or not feeling well</td>
<td></td>
</tr>
<tr>
<td>Undefined</td>
<td>Undefined</td>
<td>It’s possible for SME to exclude some risk factors</td>
<td></td>
</tr>
</tbody>
</table>

This is the most time-intensive part of the interview process.

It serves as the reference for Iteration #2 and SB-RRW.
EE Phase 4: DI Method (iteration 2)

4. Assessing expert’s responses (Q&A)

Based upon your experience & iteration #1, please answer #1- #8:
(To assist you, refer to objective & assumptions in slide 9 and Risk Reference Table)

1. Do you need to modify the probability value scale? No
2. Do you need to re-characterize the input parameter? No
3. Do you want to adjust your Most Likely commute time? No
4. What will be your Most Likely commute time in FY15? 55 = M
5. What will be your shortest commute time in FY15? 40 = L
6. What’s the chance an FY15 commute is < 40 minutes? Extremely Unlikely
   • Use risk factors in Risk Reference Table to characterize best-case scenarios that could < 40min
7. What will be your longest commute time in FY15? 90 = H
8. What’s the chance an FY15 commute is > 90 minutes? Indifferent-Low
   • Use risk factors in Risk Reference Table to characterize worst-case scenarios that could > 90min

Iteration #1 and Risk Reference Table help improve basis of inputs
EE Phase 4: DI Method (iteration 2)

4. Assessing expert’s responses (Q&A)

User-Provided Distribution for Commute Time

Red dot depicts unadjusted point estimate. Dashed lines depict unadjusted lowest & highest

Given from Expert: L=40, M=55, H=90, p(x<L)=0.10 and p(x>H)=0.30

Calculation of ‘true’ L and H: L = 35.44 and H = 143.92 … Do these #’s appear reasonable?

2nd iteration helps “condition” expert to reduce anchoring bias on M, counter over-confidence on L & H, calibrate ‘values’ & improve rationale.

(a) Method to solve for L and H presented in “Beyond Beta,” Ch1 (The Triangular Distribution)
5. Verifying encoded values & documentation

The 2\textsuperscript{nd} iteration helped elicit a Min that seems feasible and a Max that accounts for worst-case risk factors.


**Estimating Min & Max with **SB-RRW** Method**

**Scenario Based Relative Risk Weighting (SB-RRW) Method.**

The **SB-RRW** Method elicits “risk scenarios” from a subject matter expert (SME) to enable her to describe risks & risk intensities that occur in typical, optimistic & pessimistic scenarios.

**Pros:**
- a) SME is not required to provide initial estimates of high & low values
- b) Enables SME to iterate using graphics, risk factors and risk scenarios
- c) Provides descriptive risk factors that contribute to the uncertainty
- d) Provides a means to estimate to what extent each risk factor drives the uncertainty in order to estimate Minimum & Maximum values

**Cons:**
- a) Takes time to set up “risk reference” table
- b) Takes time to perform pairwise comparisons (based upon risks)
- c) Takes time to develop intensity scale
- d) Typically captures only significant known risks

---

Q: What are the top 6 risk factors that impact your commute time?

A: Top 3 are … #1. **Accidents** , #2. **Weather** and #3. **Road Construction**

Next 3 are … #4. **Departure Time** , #5. **Red Lights** and #6. **Seasons & Holidays**

Through the use of a simple Pairwise Comparison technique, the Expert can provide relative importance of each risk factor.

Because 6 Risk Factors = 15 pairs, use of Visual Aids is recommended (see examples below):

<table>
<thead>
<tr>
<th>Pair #1</th>
<th>Pairwise Comparison wrt IMPACTS on Average Driving Speed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk Factor</td>
<td>Weather</td>
</tr>
<tr>
<td>LHS is More Important</td>
<td>RHS is More Important</td>
</tr>
<tr>
<td>9</td>
<td>8</td>
</tr>
</tbody>
</table>

Q1 Equal? No (If No, then answer Q2)
Q2 More Important? Accidents
Q3 Likert Score = 1.5

<table>
<thead>
<tr>
<th>Pair #11</th>
<th>Pairwise Comparison wrt IMPACTS on Average Driving Speed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk Factor</td>
<td>Road Construction</td>
</tr>
<tr>
<td>LHS is More Important</td>
<td>RHS is More Important</td>
</tr>
<tr>
<td>9</td>
<td>8</td>
</tr>
</tbody>
</table>

Q1 Equal? No (If No, then answer Q2)
Q2 More Important? Road Construction
Q3 Likert Score = 4
Pairwise comparison of risk factors results in the following raw values:

<table>
<thead>
<tr>
<th>Raw P/W Weighting</th>
<th>Weather</th>
<th>Accidents</th>
<th>Road Construction</th>
<th>Departure Time</th>
<th>Red Lights</th>
<th>Season &amp; Holidays</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weather</td>
<td>1</td>
<td>2/3</td>
<td>1 1/2</td>
<td>2</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>Accidents</td>
<td>1 1/2</td>
<td>1</td>
<td>2 2/3</td>
<td>2 1/2</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>Road Construction</td>
<td>2/3</td>
<td>1/2</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>Departure Time</td>
<td>1/2</td>
<td>2/5</td>
<td>1 1/2</td>
<td>2</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Red Lights</td>
<td>1/4</td>
<td>1/6</td>
<td>1/4</td>
<td>1/2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Season &amp; Holidays</td>
<td>1/8</td>
<td>1/9</td>
<td>1/7</td>
<td>1/5</td>
<td>1/2</td>
<td>1</td>
</tr>
<tr>
<td><strong>Sum</strong></td>
<td>4.0</td>
<td>2.8</td>
<td>5.4</td>
<td>8.2</td>
<td>17.5</td>
<td>32.0</td>
</tr>
<tr>
<td><strong>Rank</strong></td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>

The raw values are normalized to a 100% scale, then summed to Weights per Risk Factor:

<table>
<thead>
<tr>
<th>Normalized Matrix</th>
<th>Weather</th>
<th>Accidents</th>
<th>Road Construction</th>
<th>Departure Time</th>
<th>Red Lights</th>
<th>Season &amp; Holidays</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weather</td>
<td>0.247</td>
<td>0.234</td>
<td>0.278</td>
<td>0.244</td>
<td>0.229</td>
<td>0.250</td>
</tr>
<tr>
<td>Accidents</td>
<td>0.371</td>
<td>0.352</td>
<td>0.371</td>
<td>0.305</td>
<td>0.343</td>
<td>0.281</td>
</tr>
<tr>
<td>Road Construction</td>
<td>0.165</td>
<td>0.176</td>
<td>0.185</td>
<td>0.244</td>
<td>0.229</td>
<td>0.219</td>
</tr>
<tr>
<td>Departure Time</td>
<td>0.124</td>
<td>0.141</td>
<td>0.093</td>
<td>0.122</td>
<td>0.114</td>
<td>0.156</td>
</tr>
<tr>
<td>Red Lights</td>
<td>0.062</td>
<td>0.059</td>
<td>0.046</td>
<td>0.061</td>
<td>0.057</td>
<td>0.063</td>
</tr>
<tr>
<td>Season &amp; Holidays</td>
<td>0.031</td>
<td>0.039</td>
<td>0.026</td>
<td>0.024</td>
<td>0.029</td>
<td>0.031</td>
</tr>
<tr>
<td><strong>Sum</strong></td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Accidents have the biggest impact (34%) on commute time uncertainty.

If Expert is not comfortable with calculated Weights, need to revisit (a) selection of her top 6 risk factors and/or (b) expert-provided Pairwise Comparisons.
Create **Intensity Scale** for 6 risk factors that impact commute time

<table>
<thead>
<tr>
<th><strong>Intensity Scale</strong></th>
<th><strong>Weather</strong></th>
<th><strong>Accidents</strong></th>
<th><strong>Road Construction</strong></th>
<th><strong>Departure Time</strong></th>
<th><strong>Red Lights</strong></th>
<th><strong>Season &amp; Holidays</strong></th>
<th><strong>Value</strong></th>
<th><strong>Normalized</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Low</strong></td>
<td>Perfect</td>
<td>None</td>
<td>None</td>
<td>&lt; 7:00AM</td>
<td>No lights</td>
<td>Never</td>
<td>1</td>
<td>0.061</td>
</tr>
<tr>
<td><strong>Medium-Low</strong></td>
<td>Some wind</td>
<td>Evacuated car on side of road</td>
<td>Shoulder work at 1 location</td>
<td>7:15AM</td>
<td>1 light</td>
<td>Rarely</td>
<td>1.5</td>
<td>0.091</td>
</tr>
<tr>
<td><strong>Medium</strong></td>
<td>Some rain</td>
<td>1 accident on shoulder</td>
<td>Shoulder work at 2 locations</td>
<td>7:30AM</td>
<td>2 lights</td>
<td>Half of commutes</td>
<td>2</td>
<td>0.121</td>
</tr>
<tr>
<td><strong>Medium-High</strong></td>
<td>Rain &amp; Wind</td>
<td>2 accidents on shoulder</td>
<td>1 of 3 lane closures</td>
<td>8:00AM</td>
<td>3 lights</td>
<td>More than half of commutes</td>
<td>3</td>
<td>0.182</td>
</tr>
<tr>
<td><strong>High</strong></td>
<td>Rain &amp; Snow</td>
<td>Accident shutting 1 lane</td>
<td>2 of 3 lane closures</td>
<td>8:15AM</td>
<td>4 lights</td>
<td>&gt;75% of commutes</td>
<td>4</td>
<td>0.242</td>
</tr>
<tr>
<td><strong>Very High</strong></td>
<td>Snow &amp; Wind</td>
<td>Accident shutting 2 lanes</td>
<td>Temporary road closure</td>
<td>8:30AM</td>
<td>&gt; 4 lights</td>
<td>Nearly Always</td>
<td>5</td>
<td>0.303</td>
</tr>
</tbody>
</table>

Expert provides “intensity” levels for each risk factor in each scenario

<table>
<thead>
<tr>
<th><strong>Scenario Intensities</strong></th>
<th><strong>Weather</strong></th>
<th><strong>Accidents</strong></th>
<th><strong>Road Construction</strong></th>
<th><strong>Departure Time</strong></th>
<th><strong>Red Lights</strong></th>
<th><strong>Season &amp; Holidays</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Most Likely Intensities</strong></td>
<td>Medium-Low</td>
<td>Low</td>
<td>Medium-Low</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium-Low</td>
</tr>
<tr>
<td><strong>Optimistic Intensities</strong></td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Medium-Low</td>
<td>Medium-Low</td>
<td>Low</td>
</tr>
<tr>
<td><strong>Pessimistic Intensities</strong></td>
<td>Very High</td>
<td>Very High</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
</tr>
</tbody>
</table>

- Typical commute
- Best case commute
- Worst case commute
Using the intensity scale from previous slide, the following inputs …

<table>
<thead>
<tr>
<th>Scenario Intensities</th>
<th>Weather</th>
<th>Accidents</th>
<th>Road Construction</th>
<th>Departure Time</th>
<th>Red Lights</th>
<th>Season &amp; Holidays</th>
</tr>
</thead>
<tbody>
<tr>
<td>Most Likely Intensities</td>
<td>Medium-Low</td>
<td>Low</td>
<td>Medium-Low</td>
<td>Medium</td>
<td>Medium</td>
<td>Medium-Low</td>
</tr>
<tr>
<td>Optimistic Intensities</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Medium-Low</td>
<td>Medium-Low</td>
<td>Low</td>
</tr>
<tr>
<td>Pessimistic Intensities</td>
<td>Very High</td>
<td>Very High</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>High</td>
</tr>
</tbody>
</table>

… are replaced with respective normalized values from intensity scale, then multiplied by respective risk factor weights (ref. slide 13) to produce a “Score” for each Scenario …

<table>
<thead>
<tr>
<th>Scenario Intensities</th>
<th>Risk Factor Weights:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.2471</td>
</tr>
<tr>
<td>Most Likely Intensities</td>
<td>0.091</td>
</tr>
<tr>
<td>Optimistic Intensities</td>
<td>0.061</td>
</tr>
<tr>
<td>Pessimistic Intensities</td>
<td>0.303</td>
</tr>
</tbody>
</table>

**SCORE**

- Sum: 0.0862
- Product: 0.0661
- Total: 0.2778
Use Scores from the 3 scenarios to calculate Ratios wrt Most Likely Score

- Optimistic Score / Most-Likely Score = 0.0661 / 0.0862 = 0.7671
- Pessimistic Score / Most-Likely Score = 0.2778 / 0.0862 = 3.2218

Given a Most Likely Commute of 55 minutes, apply these Ratios to get:
- Minimum Commute Time = 0.7671 x 55 = 42.2 minutes
- Maximum Commute Time = 3.2218 x 55 = 177.2 minutes

If Expert is not comfortable with Min & Max values, need to revisit (a) Intensity scale content and/or (b) expert-provided Intensities.
Using weights (slide 13), “Accidents” contribute most to dispersion (46 minutes)

But this is not accounting for impact of “undefined” risk factor.

Therefore, Interviewer must ask the Expert:

Q: Suppose you knew the state of all 6 risk factors just prior to your commute. On average within a spread of how many minutes could you estimate your commute time?

A: About 15 minutes
Suggested Use of **DI & SB-RRW Methods in Practice**

The most critical effort is to create a “**Risk Reference Table**”
   And it will only serve schedule / cost elements that share these risks & objective
   Can take >2 hours to set up each, but can be used again for other estimates

**DI Method** takes little time to execute, relying on SMEs ability to recall Min and Max values, then adjust on 2\textsuperscript{nd} iteration
   Like SB-RRW, DI Method does use Risk Reference Table (on 2\textsuperscript{nd} iteration)
   However, if SME cannot sufficiently justify Min & Max, then SB-RRW is preferred

**SB-RRW Method** takes more time than DI to set-up, primarily because Intensity Scale is customized to specific risk factors
   After which the SME can efficiently select Intensities for each activity or CER (that could be affected by specific risk factors)
   The Pairwise Comparison only needs to be completed one time to get Weights

One method could be used to calibrate results other method
   Example: After applying DI Method to 10 WBS elements, apply SB-RRW on 1 or 2 of these WBS that have largest spread. Then calibrate DI Method using SB-RRW results.
This presentation demonstrated elicitation methods that ...

1. Modeled expert’s inputs as a triangular distribution
   - Direct Input (DI) Method
     • Q&A to elicit Min, Most-Likely & Max from expert, and then adjust for expert bias.
   - Scenario Based Relative Risk Weighting (SB-RRW) Method
     • Expert-derived scenario based factors applied to Most-Likely to estimate Min & Max.

2. Incorporated techniques to account for expert bias
   - DI: Q&A elicits likelihood to be below Min & above Max
   - SB-RRW: Use of pairwise comparison helps prevent ‘gaming’ the outcome
   - For both methods, use of visual aids helps expert calibrate original inputs

3. Were structured in a way to justify expert inputs
   - DI: Each response to each question requires a rationale from the expert
   - SB-RRW: Output provides each risk factor’s contribution to uncertainty

So … hopefully … this adds to the conversation on how best to leverage expert judgment in the cost community.
Questions?

Marc Greenberg
202.358.1025
marc.w.greenberg@nasa.gov
Backup Slides
Potential Improvements / Future Work

• General
  – Develop standardized NASA system “Risk Reference Tables”
    • Example: One for Satellites, One for Rockets, One for Aircraft, etc.
    • Note: A system’s objective hierarchy may have 2 or more risk factor sets depending on estimate type
  – Develop step-by-step templates for each method (i.e. automate like 1040EZ)
  – Explore other distributions, e.g. Weibull & LogNormal
  – Provide criteria when to elicit mean or median (vs mode)
  – Incorporate methods to combine expert judgments

• DI Method
  – Add questions to enable better “training” of the SME
  – Add questions to help create a Modified Beta-PERT (vs. triangular)
  – Have a way to convert best case & worst case scenarios into probabilities

• SB-RRW Method
  – Develop alternative methods of weighting risk factors
  – Improve intensity tables that depict expert judgment
    • Example: Make less subjective using pairwise comparison method similar to one used to weight risk factors
  – See how SB-RRW may add insight into risks associated w/data-driven CERs
Expert Judgment Elicitation (EE) Procedure

EXPERT JUDGEMENT ELICITATION PROCEDURE

STRUCTURED APPROACH TO CAPTURING AN EXPERTS KNOWLEDGE BASE AND CONVERT HIS\HER KNOWLEDGE BASE INTO QUANTITATIVE ASSESSMENTS.

MODELERS SKILLED IN DECOMPOSITION AND AGGREGATION OF ASSESSMENTS

NORMATIVE

EXPERTS

SUBSTANTIVE

KNOWLEDGABLE ABOUT THE SUBJECT MATTER AND EXTENSIVE EXPERIENCE

ELICITATION PROCESS = MULTIPLE CYCLES (AT LEAST 2)

1. DECOMPOSITION OF EVENT OF INTEREST TO A MEANINGFULL LEVEL FOR SUBSTANTIVE EXPERT

2. ELICITATION OF JUDGMENT OF SUBSTANTIVE EXPERT FACILITATED BY NORMATIVE EXPERT

3. AGGREGATION OF JUDGEMENTS BY NORMATIVE EXPERT

Source: Making Hard Decisions, An Introduction to Decision Analysis by R.T. Clemen
Reasons For & Against Conducting EE

Reasons for Conducting an Expert Elicitation
• The problem is complex and more technical than political
• Adequate data (of suitable quality and relevance) are unavailable or unobtainable in the decision time framework
• Reliable evidence or legitimate models are in conflict
• Qualified experts are available & EE can be completed within decision timeframe
• Finances and expertise are sufficient to conduct a robust & defensible EE

Reasons Against Conducting and Expert Elicitation
• The problem is more political than technical
• A large body of empirical data exists with a high degree of consensus
• Findings of an EE will not be considered legitimate or acceptable by stakeholders
• Information that EE could provide is not critical to the assessment or decision
• Cost of obtaining EE info is not commensurate with its value in decision-making
• Finances and/or expertise are insufficient to conduct a robust & defensible EE
• Other acceptable methods or approaches are available for obtaining the needed information that are less intensive and expensive
Some Common Cognitive Biases

• Availability
  – Base judgments on outcomes that are more easily remembered

• Representativeness
  – Base judgments on similar yet limited data and experience. Not fully considering other relevant, accessible and/or newer evidence

• Anchoring and adjustment
  – Fixate on particular value in a range and making insufficient adjustments away from it in constructing an uncertainty estimate

• Overconfidence (sometimes referred to as Optimistic bias)
  – Strong tendency to be more certain about one’s judgments and conclusions than one has reason. Tends to produce optimistic bias.

• Control (or “Illusion of Control”)
  – SME believes he/she can control or had control over outcomes related to an issue at hand; tendency of people to act as if they can influence a situation over which they actually have no control.
Four Categories of Uncertainty

- Aleatoric (not knowable)
  - Single point failures 'will the mug break'
- Epistemic (can learn more)
  - Environmental uncertainties 'the weather'
- Typical accidents 'variable outcomes'
- Most project estimates 'cost & time'

Incident (will / won't)
Variable (range)
Probability Distributions

**Bounded**
- Triangular & Uniform
- Histogram
- Discrete & Cumulative
- Beta & Beta-PERT

**Unbounded**
- Normal & Student-t
- Logistic

**Left bounded**
- Lognormal
- Weibull & Gamma
- Exponential
- Chi-square

**Non-Parametric Distributions:** Mathematics defined by the shape that is required. Empirical, intuitive and easy to understand.

**Parametric Distributions:** Shape is born of the mathematics describing theoretical problem. Model-based. Not usually intuitive.

Of the many probability distributions out there, Triangular & Beta-PERT are among the most popular used for expert elicitation.
Triangular Distribution

* Used in situations were there is little or no data
  
  - Just requires the lowest \((L)\), highest \((H)\) and most likely values \((M)\)

Each \(x\)-value has a respective \(f(x)\), sometimes called “Intensity” that forms the following PDF:
\[
f(x) = \begin{cases} 
\frac{2(x - L)}{(M - L)(H - L)}, & L \leq x < M \\
\frac{2(H - x)}{(H - M)(H - L)}, & M \leq x < H \\
0, & \text{otherwise}
\end{cases}
\]

\(L, M & H\) are all that’s needed to calculate the Mean and Standard Deviation:
\[
\mu = \frac{(L + M + H)}{3}
\]
\[
\sigma = \sqrt{\frac{\left(L^2 + M^2 + H^2 - L*M - L*H - M*H\right)}{18}}
\]
Beta Distribution

Bounded on [0,1] interval, scale to any interval & very flexible shape

\[ f(x) = \frac{1}{(H-L)} \left( \frac{x-L}{H-L} \right)^{\alpha-1} \left( \frac{H-x}{H-L} \right)^{\beta-1} \quad L < x < H \]

= 0 otherwise

Shape Parameters: \( \alpha > 0, \beta > 0 \)

\[ \Gamma(\alpha) = \text{EXP}[\text{GAMMALN}(\alpha)] \]

\[ \Gamma(\beta) = \text{EXP}[\text{GAMMALN}(\beta)] \]

Calculated Gamma values using Excel’s GAMMALN function:

\( \beta > \alpha > 1 \), distribution is right skewed

Most schedule or cost estimates follow right skewed pattern. But how do we know \( \alpha \) and \( \beta \)? Answer: Beta-PERT Distribution.

Sources:
2. LaserLight Networks, Inc, “Beta Modeled PERT Schedules”
Beta-PERT Distribution

Requires lowest ($L$), highest ($H$) & most likely values ($M$)

Use $L$, $M$ and $H$ to calculate mean ($\mu$) and standard deviation ($\sigma$):

$$\mu = \frac{(L + \lambda \cdot M + H)}{\lambda + 2}$$

$$\sigma = \frac{(H - L)}{6}$$

Use $L$, $H$, $\mu$ and $\sigma$ to calculate shape parameters, $\alpha$ & $\beta$:

$$\alpha = \left(\frac{\mu - L}{H - L}\right) \cdot \frac{(\mu - L)(H - \mu)}{\sigma^2} - 1$$

$$\beta = \left(\frac{H - \mu}{\mu - L}\right) \cdot \alpha$$

where $\alpha > 0$, $\beta > 0$

$\alpha$ and $\beta$ are needed to define the Beta Function and compute the Beta Probability Density:

$$f(x) = \left(\frac{1}{H - L}\right) \cdot \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot \left(\frac{x - L}{H - L}\right)^{\alpha - 1} \cdot \left(\frac{H - x}{H - L}\right)^{\beta - 1}, \quad L < x < H$$

Calculated Gamma values using Excel’s GAMMALN function:

$$\Gamma(\alpha + \beta) = \text{EXP}[\text{GAMMALN}(\alpha + \beta)]$$

$$\Gamma(\alpha) = \text{EXP}[\text{GAMMALN}(\alpha)]$$

$$\Gamma(\beta) = \text{EXP}[\text{GAMMALN}(\beta)]$$

Sources:
2. LaserLight Networks, Inc, “Beta Modeled PERT Schedules”
Results (Triangular & Beta-PERT)

- **In most cases, Beta-PERT is preferred (vs triangular)**
  - Beta-PERT’s mean is only slightly greater than its mode

- **However, triangular would be preferred (vs Beta-PERT) if elicited data seems to depict over-confidence (e.g. H value is optimistic)**
  - Triangular PDF compensates for this by ‘exaggerating’ the mean value
### EE Phase 3: Commute Time (cont’d)

#### Training the expert (continued)

For 2 Questions, you’ll need to provide your assessment of likelihood:

<table>
<thead>
<tr>
<th>Descriptor</th>
<th>Explanation</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absolutely Impossible</td>
<td>No possibility of occurrence</td>
<td></td>
</tr>
<tr>
<td>Extremely Unlikely</td>
<td>Nearly impossible to occur; very rare</td>
<td></td>
</tr>
<tr>
<td>Very Unlikely</td>
<td>Highly unlikely to occur; not common</td>
<td></td>
</tr>
<tr>
<td>Indifferent between &quot;Very Unlikely&quot; &amp; &quot;Even chance&quot;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Even Chance</td>
<td>50/50 chance of being higher or lower</td>
<td></td>
</tr>
<tr>
<td>Indifferent between &quot;Very Likely&quot; &amp; &quot;Even chance&quot;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Very Likely</td>
<td>Highly likely to occur; common occurrence</td>
<td></td>
</tr>
<tr>
<td>Extremely Likely</td>
<td>Nearly certain to occur; near 100% confidence</td>
<td></td>
</tr>
<tr>
<td>Absolutely Certain</td>
<td>100% Likelihood</td>
<td></td>
</tr>
</tbody>
</table>

**Example:** Assume you estimated a "LOWEST" commute time of 20 minutes.

Your place a value = 10.0% as the probability associated with "Very Unlikely."

Therefore:

- a) You believe it’s "VERY UNLIKELY" your commute time will be less than 20 minutes, and
- b) This is equal to a 10.0% chance that your commute time would be less than 20 min.