Science of Estimation

Computationally intensive estimation methods that cannot be done by hand, or with a calculator

1. Simulate project outcomes
2. Account for diseconomies of scale
3. Account for creeping req’ts
4. Estimate less common SW issues
5. Calculate planning options & integration with planning tools
6. Allow for What-if-analysis
7. Referee for unrealistic project expectations
8. Act as an objective authority when revising estimation assumptions
9. Sanity-checking estimates created with the art of estimation
10. Estimating large projects (the larger the project the less appropriate to rely on the *art* of estimation)
1,000 Simulated Project Outcomes

- Solid lines represent the median
  Effort on the vertical axis
  Schedule on the horizontal axis

- Dashed lines represent 25th percentile and 75th percentile
  25% of the schedules are 8.5 months or less
  75% of the Effort was 45 months or less

Figure 14-1  A tool-generated simulation of 1,000 project outcomes. Output from Construx Estimate.
Estimation Software

Simulation can account for variability from several sources:

• Variation in productivity
• Variation in program size, possibly decomposed into multiple modules
• Variation in rates of staff buildup

Monte Carlo Simulation

The scatter diagram – each point represents a simulated project (1,000 simulations)
Probability Analysis

- Estimates generated using an estimation tool
  Calibrated with historical data
  Large increases in effort to move from 70% to 80%...

<table>
<thead>
<tr>
<th>Probability</th>
<th>Effort will be less Than or Equal to</th>
<th>Difference from Nominal Effort Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>10%</td>
<td>7</td>
<td>-65%</td>
</tr>
<tr>
<td>20%</td>
<td>10</td>
<td>-50%</td>
</tr>
<tr>
<td>30%</td>
<td>13</td>
<td>-35%</td>
</tr>
<tr>
<td>40%</td>
<td>16</td>
<td>-20%</td>
</tr>
<tr>
<td>50%</td>
<td>20</td>
<td>0%</td>
</tr>
<tr>
<td>60%</td>
<td>26</td>
<td>30%</td>
</tr>
<tr>
<td>70%</td>
<td>37</td>
<td>85%</td>
</tr>
<tr>
<td>80%</td>
<td>58</td>
<td>190%</td>
</tr>
<tr>
<td>90%</td>
<td>142</td>
<td>610%</td>
</tr>
</tbody>
</table>

% Difference = \( \frac{(50\% \text{effort} - \text{effort}_i)}{50\% \text{effort}} \)
Table 14-1 Graphically

- 80% of the projects require roughly 57 staff months or less
- Probability that the project effort will be roughly 57 staff months or less is 0.80
Estimation Tools

1. Account for diseconomies of scale
   Account for differences in project size & size effects on productivity
2. Allow for the effect of creeping requirements
3. Estimate less common software issues
   Size of req’ts & design docs, # of test cases, mean-time-to-failure, etc.
4. Calculate planning options & integration with planning tools
   (support dividing project into iterations, phases, etc.)
5. What-if analysis
6. Referee for unrealistic project expectations (*next slide*)
7. Act as an objective authority in arbitrating proposed revisions
   The tool can be used to assess the revisions (serving at times as the “bad cop”)
Tradeoffs between project effort & schedule

Fig. 14-3 Using Simulation data

“Boss” wants project completed in 11 months with 50 staff months of effort!

Only 8 of the 1,000 outcomes fall within the desired combination of cost and schedule
8. Act as an objective authority when revising estimation assumptions

Shows tradeoffs between project effort & schedule

- Increased staff needed to shorten a schedule or
- Savings in effort if the schedule can be lengthened

*Figure 14-4* Calculated effect of shortening or lengthening a schedule.
9. Sanity-checking estimates created with the art of estimation
   Rule: use multiple estimation approaches and look for convergence (use of a tool can give you one such estimate)

10. Estimating large projects (the larger the project the less appropriate to rely on the *art* of estimation)
    With larger projects, not a good idea to rely only on the “art of estimation”.
    Use commercial SW estimation tool for at least one estimate
Data you need to calibrate the tools

- Effort, in staff months
- Schedule, in elapsed months
- Size, in Lines Of Code

One thing you should not do with a tool:
“… treat the output of a software estimation tool as divine revelation.
Sanity check estimation tool outputs just as you would other estimates”
Use of Multiple Approaches
Chapter 15

• No single estimation technique is perfect.
• “Convergence among estimates tells you that you probably have a good estimate.”

Author’s Example:
Estimated number of pages for book
Divergent estimates… analyzed the reason for the difference

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Estimate #1 Original Whole-Book &quot;Gut Feel&quot; Estimate</th>
<th>Estimate #2 Expert Judgement with Decomposition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preface</td>
<td>-</td>
<td>4</td>
</tr>
<tr>
<td>Welcome</td>
<td>-</td>
<td>5</td>
</tr>
<tr>
<td>Metaphors</td>
<td>-</td>
<td>11</td>
</tr>
<tr>
<td>Prerequisites</td>
<td>-</td>
<td>52</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Character</td>
<td>-</td>
<td>20</td>
</tr>
<tr>
<td>Review of themes</td>
<td>-</td>
<td>20</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>250-300</strong></td>
<td><strong>802</strong></td>
</tr>
</tbody>
</table>
Used a 3\textsuperscript{rd} estimate

- Further decomposition, using a detailed outline

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Estimate #1 Original Whole-Book &quot;Gut Feel&quot; Estimate</th>
<th>Estimate #2 Expert Judgement with Decomposition</th>
<th>Estimate #3 Outline Points and Historical data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preface</td>
<td>-</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Welcome</td>
<td>-</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Metaphors</td>
<td>-</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>Prerequisites</td>
<td>-</td>
<td>52</td>
<td>52</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Character</td>
<td>-</td>
<td>20</td>
<td>16</td>
</tr>
<tr>
<td>Review of themes</td>
<td>-</td>
<td>20</td>
<td>21</td>
</tr>
<tr>
<td>TOTAL</td>
<td>250-300</td>
<td>802</td>
<td>759</td>
</tr>
</tbody>
</table>

- 3\textsuperscript{rd} estimate within 5\% of the 2\textsuperscript{nd} estimate
- Convergent estimates (spread between the results narrowed)
“People form ideas about possible project costs, durations, and features that are based on nothing in particular. They will keep those preconceived ideas until someone presents them with enough data to dislodge their preconceptions.”

“… The fact that two completely different approaches had produced similar estimates increased my confidence in those estimates.”

Further note:

“Even in my decomposed estimate I committed the classic estimating mistake of doing a good job of estimating the things I knew about but forgetting to estimate certain significant parts of the project.”
**Author’s experience estimating**

- Size of + sign represents author’s confidence in the estimate
- Triangle was his most accurate estimate (12, 75)
- Estimates converged within 5% of the most accurate
- Client’s business *target* was 5, 25, which was used
- Actual: (14, 80) that is, 14 months and 80 staff hours