COCOMO II

Brad Clark
Software Metrics, Inc.
bradclark@erols.com

Outline

» What is a parametric model?
• How is one used?
• Parametric models need to be calibrated
• Collecting data
Parametric Model

• Mathematical representation of “idealized” real-world relationships
• The mathematical formula has a number of “parameters”
• Mathematical formula to estimate development effort:

$$Effort = A \cdot (Size)^B \cdot C$$

Parametric Model

• Parameters can be:
  – Quantitative such as size, number of defects, months
  – Qualitative such as complexity, required reliability, tool usage, analyst capability
• Models are used for analysis and estimating (forecasting)
• COCOMO II is an example of a parametric model for estimating effort and schedule from size and other factors
COCOMO II

\[ PM = A \cdot (KSLOC)^B \cdot \prod_{i=1}^{17} EM_i \]

- A is a constant
- KSLOC is thousands of source lines of code
- EM are effort multipliers, parameters that effect effort the same amount regardless of project size
- SF are scale factors, parameters that have large influence on big projects and small influence on small projects

COCOMO II Parameters

- **EM Example: Application Experience**

<table>
<thead>
<tr>
<th>Criteria</th>
<th>6 months</th>
<th>1 year</th>
<th>3 years</th>
<th>6 years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rating</td>
<td>Low</td>
<td>Nominal</td>
<td>High</td>
<td>Very High</td>
</tr>
<tr>
<td>Value</td>
<td>1.10</td>
<td>1.00</td>
<td>0.88</td>
<td>0.81</td>
</tr>
</tbody>
</table>

- **SF Example: Process Maturity**

<table>
<thead>
<tr>
<th>Criteria</th>
<th>CMM 1 Lower</th>
<th>CMM 1 Upper</th>
<th>CMM 2</th>
<th>CMM 3</th>
<th>CMM 4</th>
<th>CMM 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rating</td>
<td>Very Low</td>
<td>Low</td>
<td>Nominal</td>
<td>High</td>
<td>Very High</td>
<td>Extra High</td>
</tr>
<tr>
<td>Value</td>
<td>0.78</td>
<td>0.62</td>
<td>0.47</td>
<td>0.31</td>
<td>0.16</td>
<td>0.00</td>
</tr>
</tbody>
</table>
Differences between COCOMO 81 and COCOMO II

- Three modes replaced by five Scale Factors (SF)
- Reuse of code is non-linear in COCOMO II
- Requirements Volatility replaced by Breakage (BRAK)
- Added DOCU, RUSE, PVOL, PEXP, LTEX, PCON, SITE
- Removed VIRT, TURN, VEXP, LEXP, MODP

Outline

- What is a parametric model?
  - How is one used?
- Parametric models need to be calibrated
- Collecting data
Analysis

• Analysis
  – Parameter influence

<table>
<thead>
<tr>
<th></th>
<th>VL</th>
<th>L</th>
<th>N</th>
<th>H</th>
<th>VH</th>
<th>XH</th>
<th>Influence</th>
</tr>
</thead>
<tbody>
<tr>
<td>RELY</td>
<td>0.82</td>
<td>0.92</td>
<td>1.00</td>
<td>1.10</td>
<td>1.26</td>
<td></td>
<td>1.54</td>
</tr>
<tr>
<td>CPLX</td>
<td>0.73</td>
<td>0.87</td>
<td>1.00</td>
<td>1.17</td>
<td>1.34</td>
<td>1.74</td>
<td>2.38</td>
</tr>
<tr>
<td>AEXP</td>
<td>1.22</td>
<td>1.10</td>
<td>1.00</td>
<td>0.88</td>
<td>0.81</td>
<td></td>
<td>1.51</td>
</tr>
</tbody>
</table>

  – Sensitivity analysis (change a parameter one increment)
  – Risk identification (compare parameters)

Estimation

• Effort and schedule are estimated for a specific span of phases in the development cycle
  – Which phases?
  – What labor categories are included in the effort?
  – How does software reuse and COTS effect the overall effort
  – What about critical resources, facilities, etc.?
Outline

• What is a parametric model?
• How is one used?
  » Parametric models need to be calibrated
• Collecting data

Calibration

• Parametric models are built on data - usually someone else’s
• Best results are obtained if model is calibrated to local development environment.
• Usually, only the model constants are adjusted
  – Reduces the amount of data needed
COCOMO II Calibration

\[ PM = A \cdot (KSLOC)^B \cdot \prod_{i=1}^{17} EM_i \]
\[ B = 1.01 + \sum_{j=1}^{n} SF_j \]

- Calibrate constant A
- Calibrate constant A and fixed exponent

COCOMO II 1997 Accuracy Results

<table>
<thead>
<tr>
<th>Effort Prediction</th>
<th>Before Stratification By Organization</th>
<th>After Stratification By Organization</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRED(.20)</td>
<td>46%</td>
<td>49%</td>
</tr>
<tr>
<td>PRED(.25)</td>
<td>49%</td>
<td>55%</td>
</tr>
<tr>
<td>PRED(.30)</td>
<td>52%</td>
<td>64%</td>
</tr>
</tbody>
</table>

- Stratification means model constant, A, was calibrated for each organization
Outline

- What is a parametric model?
- How is one used?
- Parametric models need to be calibrated
  » Collecting data

Data Collection

- Collect all of the model parameters on complete project data (size, application experience, complexity, etc.)
- Collect the project result values for the model output (effort, schedule)
- You calibrate the model inputs to the known project results
Data Collection

• Collection can be done using paper forms
  – Interview project personnel
  – Sift through existing data
• Collection can be done with Software Process Database
  – Project historical estimates
  – Interview project personnel
• Collection can be done with the PSM Insight tool
  – Add model parameters to Insight using customization feature

Summary

• Parametric models offer both analysis and estimation capability; they are worth learning
• Parametric models need to be calibrated to local conditions; collect data
• COCOMO II is in the public domain and is free available at:
  http://sunset.usc.edu/COCOMOII/Cocomo.html