Planning III-B: Estimating Software Size - The PROBE Method

Outline

- PROBE
- Object categories
- Ways to develop & improve your estimating skills
- Homework #4

PROBE (cf. Humphrey, 1995, 117-125)

PROBE = PROxy-Based Estimating
- Uses objects (or functions) as proxies
- See diagram below, excerpt from p. 679 & 680
- Walk through example on p. 120.

Object Categories (cf. Humphrey, 1995, 125-134)

- In order to yield the most useful information, your historical database must eventually be categorized according to type and size of object.
- Use basic “fuzzy-logic” approach and create categories and size ranges based on assumed normally distributed sizes.
- Note that $s_2$ and $s_3$ should be calculated with $n-1$ rather than $n$ as is done in the example ($n<30$).
- Natural log (ln) can be used to create normally-distributed LOC data from which LOC category ranges can be more effectively calculated.
- Walk through example on p. 126-134.
- Note incorrect calculations.


- Estimate many small pieces and sum them to get a total estimate - the sum of the variances is probably smaller than the variance of a single large estimate.
- Over time your $b_0$ and $b_1$ values will stabilize. Then you do not need to recalculate them every time you do an estimate.
- If $b_0$ and $b_1$ appear to be unreasonable / $b_0$ larger than the smallest code sizes and $b_1$ significantly larger or smaller than 1.0 then recheck your calculations, and you may need to use historical averages instead of the $b$-weights. A ratio based on averages can be calculated from the proportion of estimated object LOC to the estimated total new & changed LOC.
- Until you have sufficient data (> 2 estimated programs), you will need to calculate the $b$’s from actual program data and make estimates based on averages of the actual program data as described above.
- Make revised estimates (in large projects) at various phases when you have additional information.
- Don’t try to correct estimating errors every time - statistical variation is natural and OK. Relying on historical data will eventually help correct errors as this database grows larger over time.

Homework #4

- Program 4A
  - Linear regression parameters
  - See p. 705-707, 543-547, and Assignment Kit #4