PSP℠

A Self-Improvement Process for Software Engineers

Watts S. Humphrey

PSP℠

TABLE 6.3 (continued)

Student: Student 12 Program: 8

Reused Parts		Estimated Size	Actual Size
LinkedList		73	73
DataEntry		96	96
	Total **R**	169	169

PROBE Calculation Worksheet (Added and Modified)		**Size**	**Time**
Added size (A):	A = BA + PA	361	
Estimated proxy size (E):	E = BA + PA + M	366	
PROBE estimating basis used:	(A, B, C, or D)	A	A
Correlation:	(R^2)	0.62	0.58
Regression Parameters:	β_0 Size and Time	62	108
Regression Parameters:	β_1 Size and Time	1.3	2.95
Projected Added and Modified Size (P):	$P = \beta_{0\,size} + \beta_{1\,size} * E$	538	
Estimated Total Size (T):	T = P + B – D – M + R	1397	
Estimated Total New Reusable (NR):	sum of * items	49	
Estimated Total Development Time:	$Time = \beta_{0\,time} + \beta_{1\,time} * E$		1186
Prediction Range:	Range	235	431
Upper Prediction Interval:	UPI = P + Range	773	1617
Lower Prediction Interval:	LPI = P – Range	303	755
Prediction Interval Percent:		90	90

shows the completed PSP Size Estimating Template with the added and reused parts, the base parts, and the planned modifications. Student 12 identified 695 LOC of base code, 5 of which she planned to modify. She actually did modify 18 LOC.

Although developers occasionally get to develop entirely new programs, most of our work is enhancing existing programs. Here, base program size is the total size of the unmodified program before development. Base additions are for enhancements you make to the base program. If you know enough about the base product to do so, estimate base additions the same way you estimate added parts.

The reused category is only for unmodified parts. When modifying existing programs, the unmodified program is the base, and you estimate its additions, changes, and deletions. Even if the base program is unmodified, it is not considered reused unless it was specifically intended for reuse. The reused category is only for parts that come directly from the reuse library without *any* modification.

The final part of step 3 is to add all of the parts estimates to obtain the estimated proxy size (E), to be entered in the Size Estimating Template in Table 6.3. E is the number used by PROBE to make the size and time projections. In the example, Student 12 had a total of 361 LOC of parts additions, 5 LOC of modifications, and 0 LOC of base additions, resulting in a value of 361 + 5 + 0 = 366 LOC for E. This number is entered in the Estimated proxy size (E) space on line 2 of the PROBE Calculation Worksheet.

PROBE Step 4: The Size Estimating Procedure

In PROBE step 4, check your data to determine whether you can use PROBE method A. After you have used the PSP for a while, you will normally use method A to make size and time estimates. Later in the chapter, after examining the basic PROBE method A, you will learn how to use PROBE methods B, C, and D to make estimates with limited data.

With estimated proxy size E, you can calculate the projected program size P and total estimated development time. If, for example, historical data showed that your finished programs were generally about 25% larger than estimated, you might add a 25% fudge factor to each estimate. The PROBE method essentially does this but in a statistically sound way. As noted before, the method is called linear regression. Although these calculations are a little more complex than taking simple averages, they use historical data to produce a statistically sound estimate. The parameters β_0 and β_1 are used in the following equation to calculate projected added and modified size:

$$\text{Projected Added and Modified Size (P)} = \beta_0 + \beta_1 * \text{E}$$

When two sets of data are strongly related, you can use the linear regression method to represent that relationship. As Figure 5.2 (see p. 73) and Figure 5.3 (see p. 75) show, estimated part size and actual added and modified size are often closely correlated. This means that linear regression is often appropriate. The parameters β_0 and β_1 are calculated from your historical data. These calculations are described later with an example.

Finished programs usually contain more than just the parts specified in the conceptual design. For example, they will likely contain declaration and header code that is not included in the parts estimates. To account for this and for any other such code, you must use a factor that is based on your historical experience. Fortunately, the PROBE method accounts for this factor when it calculates the β_0 and β_1 parameters.

In the example shown in Table 6.3, the fact that $\beta_1 = 1.3$ indicates that Student 12's finished programs have historically been 30% bigger than the total of the

estimated parts and modifications. In addition, the value $\beta_0 = 62$ indicates that, on average, she underestimated by 62 LOC. As shown in Table 6.3, the regression calculations result in a total of 538 projected added and modified LOC. In this case, because added and modified code were used in making the size estimate, total program size is calculated by adding the 695 LOC of base code and the 169 LOC of reused code, and subtracting the 5 LOC of modified code. The total estimate is then 1,397 LOC for the finished program. The modified LOC are subtracted because they would otherwise be counted twice: once in the 695 LOC of base code and again in the estimated proxy size (E). As noted in size accounting in Chapter 3 (see p. 42), a line of modified code is treated as both an added line and a deleted line.

PROBE Step 5: The Time Estimating Procedure

In PROBE step 5, again check your data to determine whether you can use PROBE method A for the time estimates. Although method A is also the preferred method for time estimates, you can also use alternative methods B, C, and D, which are described later in the chapter. Once you have obtained the β values for time, use them and the estimated proxy size (E) to calculate the estimated development time. Then enter the size and time estimates in the plan spaces in the Project Plan Summary shown in Table 6.4.

PROBE Step 6: The Prediction Interval

The final calculations on the PROBE Size Estimating Template are for the **prediction interval**. The prediction interval is a statistically determined range around your size or time estimate within which the actual value is likely to fall. For a 70% prediction interval, you would expect the actual time and size values to fall outside of this range about 30% of the time. The prediction interval is described in Chapter 7.

TABLE 6.4 PSP1 PROJECT PLAN SUMMARY

Student	Student 12	Date	5/1
Program	Multiple Regression	Program #	8
Instructor	Humphrey	Language	C++

Summary	***Plan***	***Actual***	***To Date***
Size/Hour	27.4	30.9	23.4

Program Size	***Plan***	**Actual**	**To Date**
Base (B)	695 (Measured)	695 (Measured)	
Deleted (D)	0 (Estimated)	0 (Counted)	
Modified (M)	5 (Estimated)	18 (Counted)	
Added (A)	533 (A&M – M)	464 (T – B + D – R)	
Reused (R)	169 (Estimated)	169 (Counted)	
Added and Modified (A&M)	538 (Projected)	482 (A + M)	2081
Total Size (T)	1397 (A&M + B – M – D + R)	1328 (Measured)	5161
Total New Reusable	49	54	294

Estimated Proxy Size (E) 366

Time in Phase (min.)	**Plan**	**Actual**	**To Date**	**To Date %**
Planning	146	166	710	13.3
Design	425	332	1220	22.8
Code	390	333	1639	30.6
Compile	30	16	288	5.4
Test	105	31	749	14.1
Postmortem	90	58	737	13.8
Total	1186	936	5343	100

Defects Injected	**Actual**	**To Date**	**To Date %**
Planning	2	6	3.5
Design	11	38	22.0
Code	26	129	74.5
Compile	0	0	0
Test	0	0	0
Total Development	39	173	100

Defects Removed	**Actual**	**To Date**	**To Date %**
Planning	0	0	0
Design	10	10	5.8
Code	19	72	41.6
Compile	10	61	35.3
Test	0	30	17.3
Total Development	39	173	100
After Development			

6.3 Estimating with Limited Data

Table 6.5 shows the four PROBE methods, the conditions for using them, and how they are used. These methods are also described in the PROBE Estimating Script in Table 6.6, the Size Estimating Template Instructions in Table 6.7, and the PROBE Calculation Worksheet Instructions in Table 6.8.

By choosing one of the four PROBE size estimating procedures, you decide how to calculate the β parameters. Base this choice on the quality of your data. Method A should be your first choice but it requires at least three, and preferably four, data points of *estimated proxy size* (E) and *actual added and modified size* that correlate with an $r >= 0.7$. If you can't use method A, try to use method B. This method uses *plan added and modified size* and *actual added and modified size*. Again, you must have at least three, and preferably four, data points that correlate with an $r >= 0.7$. If the data are not adequate for methods A and B, then use method C if you have at least some data on plan and actual added and modified size. If you don't have any data, you must use method D. With method D, you are not actually making a projection but merely guessing at a value to enter as

TABLE 6.5 THE FOUR ALTERNATE PROBE CALCULATION METHODS

Method	Data Used for Beta Values		Data Requirements
	Size	Time	
A	Estimated proxy size and actual program size	Estimated proxy size and actual development time	The data must correlate with $r >= 0.7$
B	Planned program size and actual program size	Planned program size and actual development time	The data must correlate with $r >= 0.7$
C	Planned program size if available and actual program size. Set β_0 = 0 and β_1 = to-date actual size/to-date planned size. If planned data not available, set β_1 = 1.0	Planned program size if available and actual development time. If planned size data not available, use actual size. Set β_0 = 0 and β_1 = to-date actual time/to-date planned size. If planned data not available, set β_1 = to-date actual time/to-date actual size.	Some actual size and time data
D			No data

TABLE 6.6 PROBE ESTIMATING SCRIPT

Purpose		To guide the size and time estimating process using the PROBE method
Entry Criteria		• Requirements statement • Size Estimating template and instructions • Size per item data for part types • Time Recording log • Historical size and time data
General		• This script assumes that you are using added and modified size data as the size-accounting types for making size and time estimates. • If you choose some other size-accounting types, replace every "added and modified" in this script with the size-accounting types of your choice.
Step	**Activities**	**Description**
1	Conceptual Design	Review the requirements and produce a conceptual design.
2	Parts Additions	Follow the Size Estimating template instructions to estimate the parts additions and the new reusable parts sizes.
3	Base Parts and Reused Parts	• For the base program, estimate the size of the base, deleted, modified, and added code. • Measure and/or estimate the side of the parts to be reused.
4	Size Estimating Procedure	• If you have sufficient estimated proxy size and actual added and modified size data (three or more points that correlate), use procedure 4A. • If you do not have sufficient estimated data but have sufficient plan added and modified and actual added and modified size data (three or more points that correlate), use procedure 4B. • If you have insufficient data or they do not correlate, use procedure 4C. • If you have no historical data, use procedure 4D.
4A	Size Estimating Procedure 4A	• Using the linear-regression method, calculate the β_0 and β_1 parameters from the estimated proxy size and actual added and modified size data. • If the absolute value of β_0 is not near 0 (less than about 25% of the expected size of the new program), or β_1 is not near 1.0 (between about 0.5 and 2.0), use procedure 4B.
4B	Size Estimating Procedure 4B	• Using the linear-regression method, calculate the β_0 and β_1 parameters from the plan added and modified size and actual added and modified size data. • If the absolute value of β_0 is not near 0 (less than about 25% of the expected size of the new program), or β_1 is not near 1.0 (between about 0.5 and 2.0), use procedure 4C.
4C	Size Estimating Procedure 4C	If you have any data on plan added and modified size and actual added and modified size, set $\beta_0 = 0$ and β_1 = (actual total added and modified size to date/plan total added and modified size to date).
4D	Size Estimating Procedure 4D	If you have no historical data, use your judgment to estimate added and modified size.