

This tutorial will go into the details of how PROBE Methods are evaluated for best fit based on the individual's historical data.

PSP Advanced: Tutorial: PROBE Methods A & B with the Student Workbook

Lecture Topics

- Using PROBE to make size and time estimates
- Overview of PROBE Methods A & B
- Method selection criteria for "size" and "time" in methods A & B
- A review of
 - PROBE Size Estimating Template
 - Derived Measures
- PROBE Historical Data
- PROBE Charts

Please make a copy of your student database,
as you will be exploring the tool and experimenting with the options

© 2012 Carnegie Mellon University 2

We will provide a refresher for the class and go over the PSP Size Estimating Template in the student tool.

It may be a good idea to have the students bring up their student tool.

Engaging the class to look at their tool will also help keep the focus on the tutorial and eliminate internet surfing and email correspondence during the lecture.



Using PROBE

PROBE is used to make continuously improving Size and Time estimates by

- collecting historical data
- calculating regression parameters on historical data
- adjusting the estimates for trends

PROBE methods A & B are introduced this week.

What differentiates methods A & B from method C, introduced in PSP Fundamentals, is that methods A & B use **linear regression** on your historical data to calculate the projected size and time estimates.

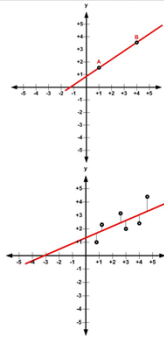


Minimum Data for Using Regression

The regression line represents the best fit (or smallest possible distance from the regression line to the data points).

When we have only 2 data points, the regression line goes right through them for a perfect fit, but it is not meaningful.

For the regression line to act as the predictor, we need more data points (at least 3).

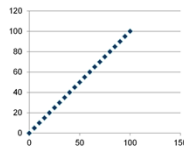


© 2012 Carnegie Mellon University 4

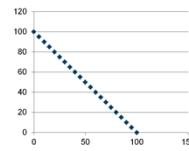


Correlation

Correlation (r) (a.k.a. Correlation Coefficient): Degree to which two sets of data are linearly related.



Positive Relationship ($r = +1$)



Negative Relationship ($r = -1$)

Correlation squared (r^2): Descriptive measure between zero and one, indicating how good one term is at predicting another.

An r^2 of 1 is considered a perfect fit!

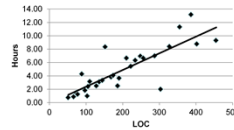
To use methods A & B, an r^2 of 0.5 or higher is required.



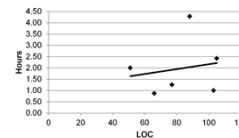
Correlation & Significance

Significance is a measure that indicates the probability that the correlation occurred by chance.

When the number of items in the data set grows, correlation tends to be reliable. When this is the case, pay attention to correlation!



When the number of data points is small, pay attention to significance!



© 2012 Carnegie Mellon University 6

Note, there is always the possibility that a high correlation was the result of chance (just a random event) and there is not a real cause and effect relationship between the two data sets. Significance is measure of that probability.

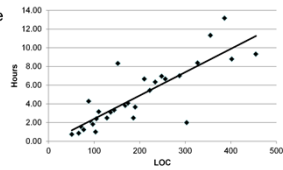
The point here is that when you have only a small number of data points in the regression and you get a good correlation, you should also check the significance measure to determine if the relation is likely due to chance and is not a true cause/effect relationship.

The students will implement the significance calculation in assignment on day 3 of this course.



Adjusting estimates with linear regression

PROBE methods A & B calculate the trend line with linear regression:



In PROBE " $y = mx + b$ " equates to:

$$y \text{ (adjusted estimate)} = b (\beta_0) + m (\beta_1) * x \text{ (Estimated Proxy Size)}$$

(P)rojected Added and Modified Size = $\beta_0 + \beta_1 * \text{Estimated Proxy Size}$

β_0 is a constant bias in the size, e.g., overhead LOC and/or constant estimating bias
 β_1 is the factor for the ratio of under-estimates and over-estimates

Estimated Development Time = $\beta_0 + \beta_1 * \text{Estimated Proxy Size}$

β_0 is a constant bias in the estimated time, e.g., overhead time and/or constant estimating bias
 β_1 is the average additional time to develop one line of code

© 2012 Carnegie Mellon University 7

Note that beta0 is not strictly overhead (LOC or time), it also accounts for a consistent constant estimating bias.



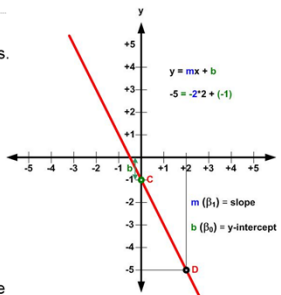
What do negative β_0 and β_1 mean?

The y-intercept (β_0) represents overhead and constant estimating bias.

While overhead can only be positive, constant estimating bias can be negative due to a strong bias towards over-estimating.

Only small, relative to the final estimate, negative values of β_0 are acceptable.

Negative slope (β_1) means that as the proxy size estimate increases, the size and/or effort decreases! This should not be true, thus **negative values for β_1 are not acceptable.**



© 2012 Carnegie Mellon University 8

It's important to note that negative values of β_0 are valid. They can arise due to a constant estimating bias to overestimate. The regression compensates for this tendency to overestimate by a constant amount with a negative β_0 .

Note that the absolute value of B_0 is used in the PROBE method selection scripts. A small negative B_0 is acceptable because it will not make the estimate unreasonable. The script guidelines work well enough for planning purposes.



PROBE Methods A & B Overview

PROBE Method A is the most preferred method.

Method A requires a minimum of 3 data points and uses:

- "E"stimated Proxy Size and Actual A&M for "size" estimation.
- "E"stimated Proxy Size and Actual development time for "time" estimation.

If the correlation and regression parameter values are within the selection criteria Method A is used.

PROBE Method B is the next preferred method in the hierarchy.

Method B requires a minimum of 3 data points and uses:

- Planned A&M (planned program size) and Actual A&M (actual program size) for "size" estimation.
- Planned A&M (planned program size) and Actual Development Time for "time" estimation.

If the correlation and regression parameter values are within the selection criteria Method B is used.

When PROBE Methods A or B do not fit the selection criteria, use

Method C if you have some historical data

- Method C uses averaging and does not look at correlation

Method D if you have no historical data

- Method D uses "engineer's best guess!"



Rules for Method Selection regarding Size

Check if method A can be used.

- You have three or more data points (*estimated E* and *actual A&M*) that correlate ($r^2 \geq 0.50$).
- The absolute value of β_0 is less than about 25% of the expected size of the new program
- β_1 is between 0.5 and 2.

If method A cannot be used, check method B.

- You have three or more data points (*plan A&M* and *actual A&M*) that correlate ($r^2 \geq 0.50$).
- The absolute value of β_0 is less than about 25% of the expected size of the new program.
- β_1 is between 0.5 and 2.

If method B cannot be used and you have historical data, use method C.

If you have no historical data, use method D.



Rules for Method Selection regarding Time

Check if method A can be used.

- You have three or more data points (**estimated E** and **actual time**) that correlate ($r^2 \geq 0.50$).
- β_0 should be near 0 (substantially smaller than the expected development time for the new program).
- β_1 should be within 50% of 1/(historical productivity)

If method A cannot be used, check method B.

- You have three or more data points (**plan A&M** and **actual time**) that correlate ($r^2 \geq 0.50$).
- β_0 should be near 0 (substantially smaller than the expected development time for the new program).
- β_1 should be within 50% of 1/(historical productivity)

If method B cannot be used and you have historical data, use method C.

If you have no historical data, use method D.



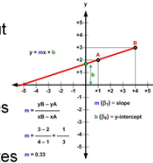
Where do the Size guidelines come from?

Rule for **Size** β_0

- Having the overhead and constant bias (β_0) almost as big as the finished product would not make sense. It would mean estimated proxy size has little influence on the estimate size.
- The absolute value of β_0 should be less than about 25% of the expected size of the finished product.

Rule for **Size** β_1

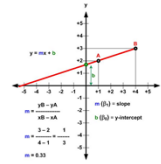
- Normally, the slope (β_1) should be near 1.0. Values between 0.5 and 2.0 are acceptable.
- A value of less than 0.5 or greater than 2.0 indicates a large proportional over or under bias and should not be used unless your instructor agrees.



Where do the Time guidelines come from? - 1

Rule for **Time** β_0

- Having the overhead and constant bias (β_0) almost as large as the final estimated time would not make sense. It would mean estimated proxy size has little influence on the estimated time.
- The absolute value of β_0 should be significantly smaller than the estimated development time.



Where do the Time guidelines come from? - 2

Rule for *Time* β_1

- "Time" β_1 needs to be within 50% of (1/Historical Productivity)

Example:

How long would it take to code one LOC if historical average productivity is at 30 LOC per hour?

$$\begin{array}{rcl} 30 \text{ LOC} & & 1 \text{ hr} \\ 1 \text{ LOC} & & x \text{ hrs} \\ \hline 30x = 1 & \rightarrow & x = 1/30 = 2 \text{ minutes per LOC} \end{array}$$

For this example, if β_1 is not between 1 and 3 minutes per LOC (within 50%), the correction applied is too great and should not be used.



Follow Along ...

- Download the PROBE methods A & B.mdb database and import it into a copy of your PSP Student Workbook.



PSP Advanced: Tutorial: PROBE Methods A & B with the Student Workbook

Historical Data - PROBE Calculation Worksheet

This is the lower part of the PSP Size Estimating Template where the analysis of data and selection of the PROBE methods is performed.

This screen is reached through the PSP2.1 window's drop down list:

- PSP2.1
 - PSP2.1 Forms...
 - Time Log
 - Defect Log
 - Defect Type Standard
 - PIP
 - Size Estimating Template
 - Test Report
 - Task and Schedule Plan

© 2012 Carnegie Mellon University 16

We display the lower portion of the PSP Size Estimating Template first as our journey will start in the lower portion.

This slide introduces the parts of the screen with red arrows to indicate the main topics of discussion,

While we will go into the details later, it may be a good idea here to show (with the green arrows) that Method B&C's historical data starts with Program-2 where they did estimate time for each phase for the first time, and that Method A's historical data starts with Program-3 where they started proxy based estimating. Note the first orange arrow on Program-2 and point out that there is no "E" and then point out the second orange arrow on Program-3 where there is an "E" estimate and that with Program-3 we started building historical data on "E"estimated Proxy Size.

Since both Method A and Method B require a minimum of 3 data points, Method B has data starting from Program-2, and Method-A from Program-3.

Method A uses "E"estimated Proxy Size and Actual A&M for "size", and "E"estimated Proxy Size and Actual development time for "time" (The double sided red arrows in PROBE Historical Data – Method A section).

Method B uses Planned program size and Actual program size for "size", and Planned program size and Actual development time for "time" (The double sided red arrows in PROBE Historical Data – Method B and C section).

Point to blue arrows and remind that they can scroll through and view the historical data for all the programs they have developed so far.

Have the class go into their student tools and acknowledge that the entire class is now familiar with this part of the PSP Size Estimating Template.



PSP Advanced: Tutorial: PROBE Methods A & B with the Student Workbook

PSP Size Estimating Template

This is the upper part of the PSP Size Estimating Template where the detailed size estimating is performed.

We are interested in:

Added Size =
Base Additions + Part Additions

Estimated Added & Modified =
Base Additions + Part Additions
+ Modifications

No effort is spent in development on inclusion of **Reused** parts so they are not included in **A** or **E**.

This slide is a refresher to get everyone's understanding on the same page.

We are interested in two derived measures: $A = BA + PA$ and $E = BA + PA + M$.

It is a good idea to remove any confusion about "base" and "reused" here, if there is any!

Also remind the students that if most of the newly added parts are of medium size it is indicative of a good conceptual design, not too detailed where many parts are too small, and not too high level where many parts are too large.



PSP Advanced: Tutorial: PROBE Methods A & B with the Student Workbook

PROBE Method A

PROBE Historical Data - Method A

ID	Name	End Date	Time	Plan A&M	E	Time	Actual A&M	Outlier
402	Program 3	09/12/07	119.536	76.9259	98.94	169.467	126	<input type="checkbox"/>
403	Program 4	09/13/07	138.601	89.1948	114.72	110.35	161	<input type="checkbox"/>
404	Program 5	10/15/07	181.311	185.670	96.04	219.45	62	<input type="checkbox"/>

PROBE Method Parameters - Method A

R ² Size	Beta0 Size	Beta1 Size	R ² Time	Beta0 Time	Beta1 Time
Method A	0.162327	37.788191	0.8797407	0.032137	215.17532

Method A uses:

- "E"stimated Proxy Size and Actual A&M for "size" estimation,
- "E"stimated Proxy Size and Actual development time for "time" estimation.

PROxy Based Estimation started with the 3rd programming assignment!

Outliers are best visible on the chart. They should be marked as such after careful evaluation.

We will visit Method A's chart next by clicking on the **Chart** button.

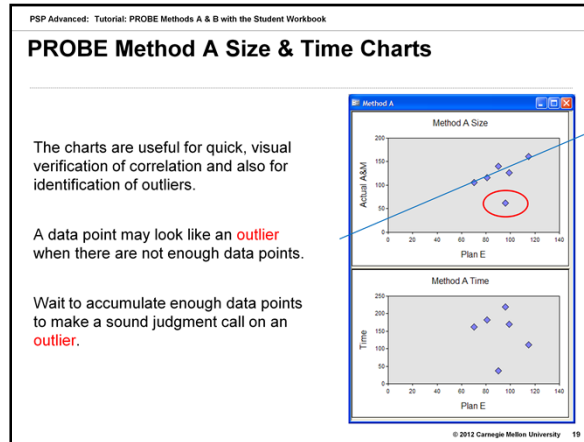
© 2012 Carnegie Mellon University 18

We show on the slide where the data is coming from for use in Method A.

Also mention that it is discouraged to mark a data point as an outlier as fluctuation is normal and evaluation should be deferred until substantial amount of data is collected.

Method A's historical data starts with Program-3 where they started proxy based estimating.





Note that based on an individual's perception a visually drawn regression line may or may not be accurate!

On this example explain that when the data point circled in red is marked as an outlier we get a pretty good regression, but this would be the wrong approach as there are too few data points to make a sound judgment call!

Note: Student tool does not draw a **regression line**. It is best to go back to PROBE Size Estimating Template and look at the calculated correlation value and regression parameters!



PSP Advanced: Tutorial: PROBE Methods A & B with the Student Workbook

PROBE Method A - Outlier

PROBE Historical Data - Method A				Plan		Actual		
ID	Name	End Date	Time	A&M	E	Time	A&M	Outlier
403	Program 4	09/13/07	138.601	89.1948	114.72	110.35	161	<input type="checkbox"/>
404	Program 5	10/15/07	181.311	185.670	96.04	219.45	62	<input checked="" type="checkbox"/> Before
405	Program 6	10/16/07	111.825	66.2187	70.15	162.15	105	<input type="checkbox"/>

PROBE Method Parameters - Method A						
	R ² Size	Beta0 Size	Beta1 Size	R ² Time	Beta0 Time	Beta1 Time
Method A	0.162327	17.788191	0.8797487	0.032137	215.17532	-0.7470015

PROBE Historical Data - Method A				Plan		Actual		
ID	Name	End Date	Time	A&M	E	Time	A&M	Outlier
403	Program 4	09/13/07	138.601	89.1948	114.72	110.35	161	<input type="checkbox"/>
404	Program 5	10/15/07	181.311	185.670	96.04	219.45	62	<input checked="" type="checkbox"/> After
405	Program 6	10/16/07	111.825	66.2187	70.15	162.15	105	<input type="checkbox"/>

PROBE Method Parameters - Method A						
	R ² Size	Beta0 Size	Beta1 Size	R ² Time	Beta0 Time	Beta1 Time
Method A	0.623972	25.691613	1.1454832	0.095584	230.43022	-1.0621272

In your student tool, mark a program as an outlier and see how PROBE reacts to it with new calculation. Review displayed data.

© 2012 Carnegie Mellon University 20

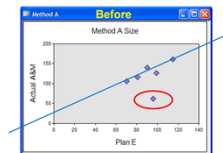
Explain that this is the “advanced” course and the students should feel pretty comfortable exploring the options on the tool and get familiar with them.

Note: **Correlation changed drastically in this example!**

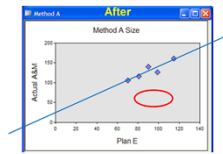


PROBE Method A Size Charts with(out) Outlier

The before chart was built based on all the data.



Once one or more data points get marked as outliers they no longer appear on the charts and they are also excluded from the calculations!



© 2012 Carnegie Mellon University 21

Tell the students that marking a data point as outlier does not delete the historical data on that program, the PROBE will be blind to that program's data.



PSP Advanced: Tutorial: PROBE Methods A & B with the Student Workbook

PROBE Method A Selection Criteria for Size - 1

PROBE Method Parameters - Method A						
	R ² Size	Beta0 Size	Beta1 Size	R ² Time	Beta0 Time	Beta1 Time
Method A	0.829723	25.691613	1.1454832	0.095584	230.43022	-1.0821272

Size selection criteria (Step-1)

- Is the correlation (r^2) equal to or greater than +0.50?
- In this example r^2 meets the requirement (0.829723)

Reminder:
Method A uses "E"stimated Proxy Size and Actual A&M for "size" and
"E"stimated Proxy Size and Actual development time for "time" estimation.

© 2012 Carnegie Mellon University 22

Note that the correlation should be a positive number. If there is a negative correlation for example, between size and time, then it would mean that it takes less time to write larger programs!



PSP Advanced: Tutorial: PROBE Methods A & B with the Student Workbook

PROBE Method A Selection Criteria for Size - 2

PROBE Method Parameters - Method A						
	R ² Size	Beta0 Size	Beta1 Size	R ² Time	Beta0 Time	Beta1 Time
Method A	0.829723	25.691613	1.1454832	0.095584	230.43822	-1.0821272

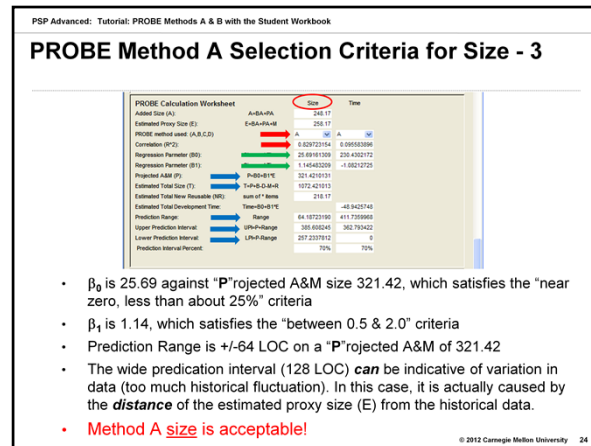
Size selection criteria (Step-2)

- Absolute value of β_0 should be near 0, less than about 25% of the expected size of the new program. β_0 represents overhead and/or constant estimating bias, so it should not be more than about 25% of projected A & M size!
- β_1 represents proportional estimating bias and proportional overhead. It should be between 0.5 and 2.0
- Now we will go into PROBE Calculation Worksheet (by scrolling up from this point) to evaluate if the regression parameters satisfy the selection criteria.

© 2012 Carnegie Mellon University 23

β_1 represents proportional estimating bias and is used to correct the estimated LOC. If we are correcting it by more than 100% as is the case with 2.0, then we should not be using these regression parameters as the tool thinks that the finished product will be two times the size of our estimate. In contrast if we are reducing our estimated LOC by more than 50% as is the case with 0.5, then again we should not be using these regression parameters as the tool thinks we are being overly cautious.





Method A fits the selection criteria, so it should be selected.

A couple things to note about the prediction interval

- It is not part of the selection criteria for the method choice.
- The magnitude of the prediction interval is mainly affected by 2 items
 - the variation of the historical data about the regression line
 - The "distance" of the new estimated proxy size (E) from the historical data
- The larger either of these is, the larger the prediction interval is
- In this example, the variation of the historical data about the regression line is "small" – see the earlier charts. But the "distance" of the new estimated proxy size from the historical data is "large". So although the historical data has little fluctuation, because the new estimated proxy size (e) is well outside the range of the existing data, the prediction interval is large.



PSP Advanced: Tutorial: PROBE Methods A & B with the Student Workbook

PROBE Method A Selection Criteria for Time - 1

PROBE Method Parameters - Method A						
	R ² Size	Beta0 Size	Beta1 Size	R ² Time	Beta0 Time	Beta1 Time
Method A	0.829723	25.691613	1.1454632	0.095584	230.43022	-1.0821272

Time selection criteria (Step-1)

- Is the correlation (r^2) equal to or greater than +0.50?
- In this example r^2 does NOT meet the requirement (0.096)
- Let's suppose that correlation was good and check the criteria for β_0 and β_1

Reminder:
Method A uses "E"stimated Proxy Size and Actual A&M for "size", and
"E"stimated Proxy Size and Actual development time for "time" estimation.

© 2012 Carnegie Mellon University 25

Note that the slope (β_1) should be a positive number. For example, if the slope, between size and time, is negative, then it would mean that it takes less time to write larger programs!

In this case, the slope (b_1), is negative so we could immediately conclude that Method A is not usable for time!



PSP Advanced: Tutorial: PROBE Methods A & B with the Student Workbook

PROBE Method A Selection Criteria for Time - 2

PROBE Method Parameters - Method A						
	R ² Size	Beta0 Size	Beta1 Size	R ² Time	Beta0 Time	Beta1 Time
Method A	0.823723	25.691613	1.1454832	0.095584	230.43022	-1.0821272

Time selection criteria (Step-2)

- β_0 should be near 0 (substantially smaller than the expected development time for the new program).
- β_1 should be within 50% of $\left(\frac{1}{\text{Historical Productivity}} \right)$

$$\left(0.5 * \left(\frac{1}{\text{Historical Productivity}} \right) \right) < \beta_1 < \left(\frac{1}{\text{Historical Productivity}} \right) + \left(0.5 * \left(\frac{1}{\text{Historical Productivity}} \right) \right)$$

- Now we will go into PROBE Calculation Worksheet (by scrolling up from this point) to evaluate if the regression parameters satisfy the selection criteria.

© 2012 Carnegie Mellon University 26

β_0 and β_1 should be explained here as to what they mean regarding “time”. In LOC/Size, β_0 , represented overhead such as declarations, setting up development environment and/or constant estimating bias.

In “time” β_0 represents time spent that isn’t related to size (possibly planning and/or postmortem and/or overhead items) and constant estimating bias.

β_1 should be within half of $1/\text{“Historical Productivity”}$. The rational for this is that since the B0 should be substantially smaller than the expected development time and that the Projected A&M should be somewhere between $\frac{1}{2}$ and $2x$ the estimated proxy size, b1, which is the main determinate of the time estimate, should be within 0.5 and 1.5 the reciprocal of the historical productivity expressed in minutes/LOC.

NOTE, however, this is an academic exercise in this case since b1 is negative, we can immediately conclude that method A is not usable for time!

NOTE: The historical productivity needs to be converted to minutes per LOC instead of Hours per LOC before making the comparison.



PROBE Method A Selection Criteria for Time - 3

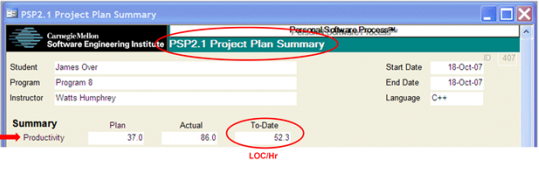
PROBE Calculation Worksheet		Size	Time
Actual Size (S)		248.17	
Estimated Proxy Size (E)		248.17	
Model without cost (A,B,C,D)	A	36	5
Correlation (P2)		0.92922154	0.99992388
Regression Parameter (B1)	Size and Time	0.00187180	230.432112
Regression Parameter (B1)	Size and Time	1.14546228	-1.0212725
Projected Label (P)		251.421811	
Estimated Total Size (T)	Size and Time	337.431913	
Estimated Total Proxy Size (P2)		248.17	
Estimated Total Proxy Residual (R2)	Size and Time	-48.9421745	
Estimated Total Development Time			
Problems Range	Range	64.1972138	411.733888
Upper Prediction Interval	Upper Range	185.65454	362.79432
Lower Prediction Interval	Lower Range	237.237812	0
Prediction Interval Percent		75%	75%

- β_1 is -1.08. A **negative** slope is not acceptable! We can immediately conclude that method A is not usable for time.
- β_0 is 230.43 against total development "Time" of -48.94, which is totally outside of acceptable parameters
- β_0 also does NOT satisfy the selection criteria
- "Time" is -48.94 minutes, which is meaningless!
- If β_1 was good, how do we check the selection criteria for β_1 ?



PSP Advanced: Tutorial: PROBE Methods A & B with the Student Workbook

PROBE Method A Selection Criteria for Time - 4



Summary

	Plan	Actual	To-Date
Productivity	37.0	86.0	52.3

LOC/Hr

- If β_1 were positive, you would then check if it met the criteria.
reciprocal of productivity = $(1/52.3) * 60 = 1.14 \text{ min/LOC}$
 $1.14 - (1.14/2) < \beta_1 < 1.14 + (1.14/2)$
 $0.57 < \beta_1 < 1.71$
(From previous slide) β_1 does not satisfy Method A's selection criteria for Time!

© 2012 Carnegie Mellon University 28

Note that historical productivity is LOC/Hr, so we multiply the result by 60 to get the per minute value!



PROBE Method A Selection Criteria for Time - 5

PROBE Calculation Worksheet		Size	Time
Added Size (A)	A-B-A/A	248.17	
Estimated Proxy Size (E)	E-B-A/B-A	255.17	
PROBE method used (A,B,C,D)			
Correlation (R ²)		0.82972154	0.98522399
Regression Parameter (B)	Size and Time	25.8916109	230.4302172
Regression Parameter (B')	Size and Time	1.14548209	-1.68212729
Projected A/B (P)	P-B-A/B	321.421623	
Estimated Total Size (T)	T-A-B-B-A/B	1072.421613	
Estimated Total New Research (N)	sum of "Items"	216.17	
Estimated Total Development Time	Time-B-B-A/B	-48.9435740	
Prediction Range	Range	64.1872190	411.7225962
Upper Prediction Interval	UP-Pr-Range	385.658245	362.783452
Lower Prediction Interval	LP-Pr-Range	297.233812	-
Prediction Interval Percent		75%	75%

- Range is 411.74 minutes against estimated development "Time" of -48.94, which is totally meaningless,
- Upper and Lower prediction intervals therefore are meaningless too.
- Our next option for time is Method B.



PSP Advanced: Tutorial: PROBE Methods A & B with the Student Workbook

PROBE Methods B&C

The screenshot displays the PROBE Methods B&C interface. It features a table of historical data for four programs, with columns for ID, Name, End Date, Time, A&M, E, Actual Time, Actual A&M, and Outlier. Below the table, there are sections for PROBE Method Parameters - Methods B and C, and a list of data sources used by Method B.

ID	Name	End Date	Time	A&M	E	Actual Time	Actual A&M	Outlier
451	Program 2	09/11/07	200	400	0	483.267	262	<input type="checkbox"/>
452	Program 3	09/12/07	119.536	76.9259	98.94	169.467	126	<input type="checkbox"/>
453	Program 4	09/13/07	138.601	89.1948	114.72	110.35	161	<input type="checkbox"/>
454	Program 5	10/15/07	181.311	185.670	96.04	219.45	62	<input type="checkbox"/>

PROBE Method Parameters - Methods B and C

Method	R ² Size	Beta0 Size	Beta1 Size	R ² Time	Beta0 Time	Beta1 Time
Method C	N/A	0	0.9899326	N/A	0	1.38543737
Method B	0.470499	89.711109	0.3506719	0.90461	41.149897	1.09267135

Method B uses:

- Planned A&M (planned program size) and Actual A&M (actual program size) for "size" estimation.
- Planned A&M (planned program size) and Actual Development Time for "time" estimation.

Method C uses averaging and does not look at correlation

Other than using a different set of data, steps in Method B are identical to Method A.

Let's look at PROBE Calculation worksheet next and see if Method B provides a better fit!

© 2012 Carnegie Mellon University 30

We show on the slide where the data is coming from for use in Method B.

Also mention that it is discouraged to mark a data point as an outlier as fluctuation is normal and evaluation should be deferred until substantial amount of data is collected.



PSP Advanced: Tutorial: PROBE Methods A & B with the Student Workbook

PROBE Method B Selection Criteria for Time

PROBE Calculation Worksheet		Size	Time
Added Size (A)		246.17	
Estimated Error Size (E)		246.17	
PROBE method used (A,B,C,D)		A	B
Correlation (PC2)		0.82972154	0.910211159
Regression Parameter (B1)	Size and Time	25.69161509	43.20540403
Regression Parameter (B2)	Size and Time	1.145402209	1.107050629
Proposed Addn (P)		251.4216133	
Estimated Total Size (T)		1072.421613	
Estimated Total New Requisite (NR)	sum of 7 items	216.17	
Estimated Total Development Time	Time-B1-B2*E	329.20339100	
Prediction Range	Range	64.18723190	75.78012810
Upper Prediction Interval	UPH-R-Range	385.605245	398.8040447
Lower Prediction Interval	LPH-R-Range	217.2337812	258.4237962
Prediction Interval Percent		70%	70%

R^2 of 0.91 is equal to or greater than 0.5 ✓
 β_1 of 1.11 is positive and between 0.57 and 1.71 ✓
 β_0 of 43 is substantially smaller than the expected development time of 329 ✓

Method B for time estimating is acceptable!

© 2012 Carnegie Mellon University 31

In this example, the Time selection criteria was not met by Method A, then we came into Method B and found it to be acceptable





Messages to Remember

PROBE offers four methods to arrive at size and time estimates

- Method A is the most preferred choice
 - requires a minimum of 3 data points
 - uses "E"stimated Proxy Size and Actual A&M for "size" estimation
 - uses "E"stimated Proxy Size and Actual development time for "time" estimation
 - evaluates correlation along with regression parameters
- Method B uses:
 - requires a minimum of 3 data points
 - Planned A&M (planned program size) and Actual A&M (actual program size) for "size" estimation
 - Planned A&M (planned program size) and Actual Development Time for "time" estimation
 - evaluates correlation along with regression parameters
- Method C uses averaging and does not look at correlation
- Method D uses "engineer's best guess!"



