Disciplined Software Engineering
Lecture #6

Software Engineering Institute
Carnegie Mellon University
Pittsburgh, PA 15213

Sponsored by the U.S. Department of Defense

Process Measurement
Principles of process measurement

The goal-question-metric (GQM) paradigm

Basic process measures

Gathering and using data

The process baseline
Process Measurement Principles

To be useful, measurements should be
• gathered for a specific purpose
• explicitly defined
• properly managed
• properly used

Measuring your process will not improve it. You must make process changes to achieve lasting improvement.

Process Measurement Purposes

We measure for the following reasons
• to understand and manage change
• to predict or plan for the future
• to compare one product, process, or organization with another
• to determine adherence to standards
• to provide a basis for control
Measurement Objectives

Measurements only produce numbers.

To be useful, they must
- relate to business objectives
- be properly interpreted
- lead to appropriate action

If the business purposes for the measurements are not understood
- the wrong data may be gathered
- the data may not be properly used

Types of Measurements - 1

Measurements may be
- objective/subjective
- absolute/relative
- explicit/derived
- dynamic/static
- predictive/explanatory

We generally seek objective, absolute, and explicit measures

They may be dynamic or static, predictive or explanatory
Types of Measurements - 2

We seek useful relationships that correlate, for example
- program size versus development hours
- cost distributions
- defect densities

We also seek controlling or predictive parameters, for example
- actions to reduce test defects
- steps to improve review quality
- means to improve productivity

The Goal-Question-Metric (GQM) Paradigm

The GQM paradigm establishes a framework for gathering and using data.

It starts with an explicit statement of data gathering goals

Next it defines the questions the data are to answer

And then defines the data metrics
Measurement Goals

What are the goals for which data are required?
• these may be personal, project, or business
• they should be explicit

Be clear on whose goals these are
• yours, your project’s, your management’s
• try to understand the goals at all these levels

Relating your goals to other’s goals provides
• a clearer context for the measurements
• more likely support for your work

Measurement Questions

What will it take to meet these goals?
• plan the actions required
• who must take these actions?

Do people need to be convinced to act?
• are data needed to convince them?
• how will these data be used?

What is needed to implement the actions?
• what data are required?
• how will these data be used?
Measurement Metrics

Precisely what data are needed?
• define the data so others could gather them
• establish specific fields and formats

How will these data be gathered?
• provide data gathering forms
• define the data gathering process

How will these data be retained and used?
• define the specific calculations/analyses
• work through some test cases

Making the Measurements

You are now ready to start the data gathering process

First
• communicate the entire GQM framework
• ensure that the process is understood
• conduct a test if necessary

Start data gathering
• monitor data quality
• provide feedback
GQM for the PSP

The PSP has the basic goal of assisting you in improving your personal performance.

This suggests some likely personal goals to:
- understand your personal process
- determine steps to improve product quality
- understand your personal productivity
- establish benchmarks for measuring improvement
- make better plans

Some General PSP Questions

How good is my process?

Where can it be improved?

What is the most important improvement I should make now?

What are others doing that works better?

How can I learn from them?
How Good is my Process?

The principal dimensions of process quality are
• product quality
• predictability
• productivity
• cycle time

Since improvement must start from the current process, the first step is to establish measures of current process quality.

You can then ask more informed questions.

Measuring Process Quality

Start with building a basic understanding of your process
• what do you do?
• how much time do you spend?
• how is this time distributed?
• how predictable is your work?
• how stable is your work?
• what is the quality of the products you produce?

The PSP addresses these issues by gathering basic process data.
The PSP Measurements

The basic PSP data are
• program size
• time spent by phase
• defects found and injected by phase

Both actual and estimated data are gathered on every item

The derived measures then establish
• data to support planning
• measures of process quality

PSP Size Measures

The goals of the PSP size measures are to
• define a consistent size measure
• establish a basis for normalizing time and defect data
• help make better size estimates

The questions asked are
• what size program did I plan to develop?
• how good was my size estimate?
• what was the complete size description of the finished program?
PSP Time Measures

The goals of the PSP time measures are to
• determine how much time you spend in each PSP phase
• help you to make better time estimates

The questions asked are
• how much time did you spend by PSP phase?
• how much time did you plan to spend by PSP phase?

PSP Defect Measures

The goals of the PSP defect measures are to
• provide an historical baseline of defect data
• understand the numbers and types of defects injected
• understand the relative costs of removing defects in each PSP phase

The questions asked are
• how many defects did I make in each phase?
• how many defects did I inject in each phase?
• how much time did it take to find and fix each defect?
PSP Derived Measures

The PSP1.1 derived measures are
• To Date and To Date %
• LOC/hour
• CPI
• % Reuse and % New Reuse

To Date and To Date %

Goal
• provide a basis for projecting the time distribution of a new project based on the time distribution of previous projects

The questions are
• how much time have I spent in each PSP phase for all projects to date?
• what is the % distribution of this time?

Note - you may wish to restart these To Date data when your process changes
LOC/hour

Goals
- provide a convenient basis for comparing plans with historical performance
- provide a fall-back in case the regression method does not produce a reasonable result

The questions are
- how many LOC per hour have I produced on the most recent project?
- how many LOC per hour did I plan?
- what has been my average LOC/hour on prior projects?

CPI

Goal
- plan and manage projects so they are generally completed at or near to the committed plan

The question is
- what is the ratio of the actual to the planned time for all my PSP projects to date?
% Reused and % New Reused

Goals
• understand the degree to which previously developed code are reused
• understand the degree to which new code are added to the reuse library

The questions are
• how much code was reused?
• how much was planned for reuse?
• what percent of new development contributed to the reuse library?
• what are the percents to date for these data?

Some Other Sources of Questions

Your peers and coworkers

Project management

Senior management
Your Peers and Coworkers

If my work depends on yours, will you be ready in time?

Can you provide support in areas that I need?

If I need your support
  • when can you provide it?
  • what is the likely quality of this support?

Are you doing something that would help me to do a better job?

Project Management Questions

What is the project’s status?

How does this status compare with the plan?

Will we meet the schedule?

Where are we in trouble?

Where are we ahead?
General Management Questions

Is productivity improving?
• where can we cut costs?
• what can we do to reduce cycle time?
• how do we compare with competition?

Is quality improving, by how much, and how does it compare to competition?

What has changed since the last report and is it an improvement?

Data Gathering Considerations

Automation

A personal process notebook

Forms and templates

Databases and spreadsheets

Summary project report
Establishing a Process Baseline

Are you getting better?
- you need an objective basis to tell
- measuring alone will not cause improvement

Statistical variation
- your performance will fluctuate
- improvement is a statistical question

Bolstering

Clutching

Data Distributions

Data analysis requires knowledge of the distribution of the data

There are many types of distributions
- normal, log-normal
- Poisson
- Chi-squared
- student’s t
- F

The PSP analyses assume the data are normal or log-normal
The Normal Distribution - 1

Often occurs in nature
  • heights or weights of people
  • measurement errors

Properties
  • symmetrical
  • ranges from minus to plus infinity
  • the median equals the mean

The Normal Distribution - 2

In adding independent normal distributions
  • the mean of the sums = the sum of the means
  • the variance of the sums = the sums of the variances

Error calculations
  • standard deviation = square root of variance
  • about 68% of values are within +/- 1 standard deviation of the mean
  • about 95% are within 2 standard deviations of the mean
Assignment 6

Read chapter 7 of the text.

Using PSP1.1, write program 6A to calculate the 70% and 90% prediction intervals for your size and time estimates.

Calculate the 70% and 90% prediction intervals for your size and time estimates for program 6A.

Follow the program, assignment, and process specifications in Appendices C and D.

Prediction Interval - 1

The formula for calculating the prediction range is

$$\text{Range} = t(\alpha / 2, n-2)\sigma \sqrt{1 + \frac{1}{n} + \frac{(x_k - x_{\text{avg}})^2}{\sum_{i=1}^{n}(x_i - x_{\text{avg}})^2}}$$

Where

• $x$ is your historical object size data
• $n$ is the number of members in $x$
• $t$ is the two-sided $t$ distribution for $n-2$ degrees of freedom
Prediction Interval - 2

The formulas for calculating the variance and standard deviation terms are

\[ \sigma^2 = \left( \frac{1}{n-2} \right) \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2 \]

\[ \sigma = \sqrt{\sigma^2} \]

The upper and lower prediction interval values are then

UPI = estimate + range
LPI = estimate - range

Calculating the Value of t - 1

To calculate the value of t, using program 5A, start with a known value of p and find the value of t that produces that p value.

Make successive trial integrations from 0 to t until you obtain a proper p value
• for 70%, integrate to get 0.35 (0.85 - 0.5)
• for 90%, integrate to get 0.45 (0.95 - 0.5)

Continue iterating until the value is within an acceptable error, say 0.0001.
Calculating the Value of t - 2

One way to make this calculation is
1. start with a trial value of t, say 1.0
2. make an initial integral and test to see if it
gives the proper value of p, if not, continue
3. if it is too low, add d = 1 to trial t
4. if it is too high, subtract d = 0.5 from trial t
6. integrate again and test if the result is
   within an acceptable error, if not continue
7. if too low, subtract d/2 from trial t
8. if too high, add d/2 to trial t
9. repeat at 6

Messages to Remember from Lecture 6

1. To be useful, measures should be:
   related to a goal
   address specific questions
   be precisely defined

2. The PSP provides examples of useful
   process measures.

3. Once you define your process and establish
goals, you can find many useful measures.