DCC / ICEx / UFMG

Measurement

Eduardo Figueiredo

http://www.dcc.ufmg.br/~figueiredo

Measurement in Experiments

- Measurement is a central part in empirical studies
- Software measurement is crutial to enable control of projects, products and processes

Evaluation is Control

"You can't control what you can't measure" *Tom DeMarco*

- Control comes from being able to evaluate new methods, techniques, languages and tools

Definitions

- Measurement is the process by which numbers or symbols are assigned to attributes of entities
  - They describe these attributes according to clearly defined rules
- A measure is a mapping from the attribute of an entity to a value
  - Usually a numeric value
- Entities are objects in the real world

Measure vs. Metric

- A measure is the number or symbol assigned to an entity to characterize an attribute
- Software metrics denote the field of measurement in software engineering
- Product metric denotes a property which is measured
  - Example: lines of code (LOC)

Measure Validation

- To use a measure, it must be valid
  - It must not violate any necessary property of the measured attribute
  - It must be a proper mathematical characterization of the attribute
- Different objects can have the same measurement value
  - The measure must preserve our intuitive notion about the attribute

Analytical and Empirical Validation

- Analytical validity of a measure relates to its ability to capture accurately and reliably the item of interest
- Empirical validity describes how well a score correlates to something measured in another context

Scale and Scale Type

Measurement Scale

- Scale defines the mapping from an attribute to a measurement
  - Meters, centimetres and inches are different scales of length
- We can transform the measure from one scale to another
  - If the transformation preserves the relationship among the objects, it is called *admissible transformation* or *rescaling*

Meaningful Statement

- Meaningful statements are true after rescaling
- Example
  - Object A is 1 meter long
  - Object B is 2 meters long
- Statement: B is twice as long as A
  - Statement is true even if we rescale to centimetres or inches

Meaningless Statement

- Meaningless statements are not true after rescaling
- Example
  - Room A is 10º C
  - Room B is 20º C
- Statement: B is twice as warm as A
  - The statement is no longer true if we rescale to Fahrenheit (50º F and 68º F)

Scale Types

- Scales belonging to the same scale type share the same properties
- The most common scale types are
  - Nominal
  - Ordinal
  - Interval
  - Ratio

Nominal Scale

- The nominal scale is the least powerful of the scale types
- It maps the attribute into a name or symbol
- Transformations preserve the one-to-one mapping
- Example: Defect type
  - Requirements, Design, Coding, etc.

Ordinal Scale

- The ordinal scale ranks entities after an ordering criterion
  - Greater than, better than, more complex…
- Transformations preserve the order of the entities
- Example
  - Sorting systems by their complexity

Interval Scale

- In the interval scale, the difference between two measures are meaningful
  - The value itself is not meaningful
- It orders the values as the ordinal scale
- There is a notion of relative distance between two entities
- Example
  - Temperature in Celsius and Fahrenheit

Ratio Scale

- It is the most powerful scale type
- In a ratio scale
  - There is a meaningful zero
  - Ratio between two measures is meaningful
- Example
  - Length of an array
  - Duration of a development phase

Qualitative vs. Quantitative

- Qualitative research is concerned with measurement mainly in the nominal and ordinal scales
- Quantitative research treats measurement on the interval and ratio scales

Types of Measures

Objective and Subjective

Direct and Indirect

Objective Measure

- There is no judgement in the measurement value
- It depends only on the measured object
- It can be measured by different people and the same value is obtained
  - Within the measurement error
- Example of objective measure
  - Number of Classes

Subjective Measure

- The person measuring makes some judgement
- The measure depends on both the object and the viewpoint
- Its value can vary if the object is measured by different people
- Example of subjective measure
  - Programming skill of a developer

Direct and Indirect Measures

- Direct Measure
  - It does not involve measurements of other attributes
  - Example: Defects found in tests
- Indirect Measure
  - It is derived from other measures
  - Example: Defect Density (defects divided by lines of code)

Measurement in Software Engineering

Object of Measurement

- The object can be divided into process, product, and resources
  - Process: describes which activities are needed to produce software
  - Product: the artefacts, deliverables or documents that results from an activity
  - Resources: are personnel, hardware or software needed for a process activity

Internal and External Attributes

- Internal attribute can be measured purely in terms of the object
- External attribute can only be measured with respect to how objects relates to other objects
  - External attributes are mostly indirect measures and must be derived from internal attributes

Examples of Measures

Class	Object	Attribute	Measure
Process	Testing	Internal	Effort
		External	Cost
Product	Code	Internal	Size
		External	Reliability
Resource	Personnel	Internal	Age
		External	Productivity

Challenges in SE

- It is often sometimes hard to define an attribute in a measurable way
- Validation of indirect metrics is difficult
  - We have to validate the underlying model and the direct metrics
- It is difficult to prove that the measures are in a ratio scale (most powerful one)
  - Statistical analysis depends on the scale type

Metric Collection

- Metrics should not require too much effort to be collected
- How to collect metrics?
  - Manual: e.g., subjects fill out forms
  - Automatic: e.g., development environment is instrumented
- The quality of the collected metrics is the basis for further analysis

Analysis of Metrics

- It is important to understand what kind of metrics were collected
- Example of issues
  - Scale type: the most powerful statistical analyses require the ratio scale
  - Distribution: it is important to verify if measures are normally distributed
- Ratio of two metrics (A/B) should be used carefully

Bibliography

- C. Wohlin et al. **Experimentation in Software Engineering**, Springer. 2012.
  - Chapter 3 - Measurement