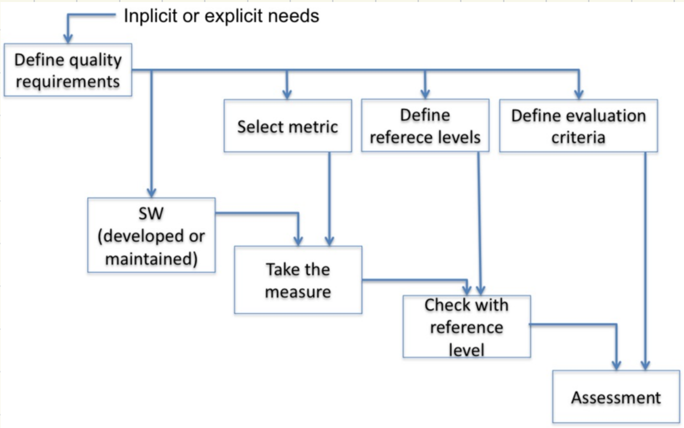
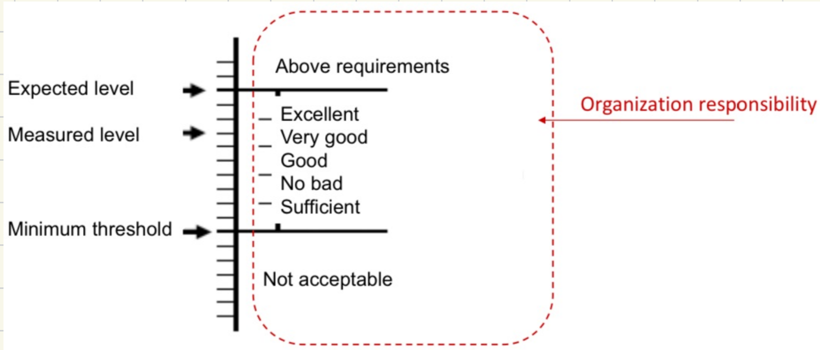
THEORY OF MEASURE

WE MEASURE FOR MONITORING AND MAKING DECISIONS.
MEASUREMENT IN SW ENGINEERING HELPS TO:

- VERIFYING HOW FAR QUALITY PARAMETERS ARE FROM REFERENCE VALUES.
- IDENTIFYING DEVIATIONS FROM TEMPORAL AND RESOURCE ALLOCATION PLANNING.
- IDENTIFYING PRODUCTIVITY INDICATORS.
- VALIDATING THE EFFECT OF STRATEGIES AIMED AT IMPROVING THE DEVELOPMENT PROCESS.





MEASUREMENT SCALES

- **NOMINAL SCALE**: CLASSIFIES PERSONS OR OBJECTS INTO TWO OR MORE CATEGORIES WITHOUT ORDER. IT DOESN'T MAKE SENSE TO CALCULATE AVERAGES. {=, !=}
- **ORDINAL SCALE**: CLASSIFIES SUBJECTS AND RANK THEM IN TERMS OF HOW THEY POSSESS THE CHARACTERISTIC OF INTEREST. {=, !=, >, <}
- **INTERVAL SCALE**: ALLOW US TO RANK THE ORDER OF THE ITEMS THAT ARE MEASURED, AND TO QUANTIFY AND COMPARE THE SIZE OF DIFFERENCES BETWEEN THEM. MEASURES DIFFERENCES WITH EQUAL INTERVALS, BUT WITH ARBITRARY ZERO. {=, !=, <, >, +, -}
- **RATIO SCALE**: SIMILIAR TO INTERVAL SCALE, BUT WITH ABSOLUTE ZERO. {=, !=, <, >, +, -, *, /}
- **ABSOLUTE SCALE**: BASED ON DISCRETE ENTITY COUNTS (LOC)

Scale types	Admissible transformations	Basic empirical operation	Appropriate statistical indexes	Appropriate statistical tests	EXAMPLES
NOMINAL	any one-to-one transformation	equality test	Mode Frequency	not parametric	labeling classify
ORDINAL	M(x)≥M(Y) implies that M'(x)≥M'(Y)	equality test and > <	Median Percentiles Spearman *r* Kendall W Kendall T	not parametric	preferences ordering di entità
INTERVALS	M'=aM+n(a>0) [*positive, linear*]	equality test and > < + and -	Aritmetic mean Standard deviation Pearson correlation Multiple correlation	not parametric	Fahrenheit o Celsius date time
RATIO	M'=aM (a>0) [*similarity transformation*]	equality test and > < + and - * and /	Geometric mean Armonic mean Coefficiente di variazione Percentage variation Correlation index	not parametrico and parametric	time intervals Kelvin lenghts
ABSOLUTE	M'=M [*identity*]				entity count

TYPES OF MEASURES

- **RATIO**: THE RESULT OF A DIVISION BETWEEN TWO VALUES THAT COME FROM TWO DIFFERENT AND DISJOINT DOMAINS.

  (LINES OF COMMENTS/LOC)* 100

- **PROPORTION**: THE RESULT OF A DIVISION BETWEEN TWO VALUES WHERE THE DIVIDEND CONTRIBUTES TO THE DIVISOR.

  SATISFIED USERS/NUMBER OF USERS

- **PERCENTAGE**: A PROPORTION OR FRACTION EXPRESSED BY NORMALIZING THE DIVISOR TO 100.

- **RATE**: MEASURES THE CHANGE OF A QUANTITY (y) WITH RESPECT TO ANOTHER QUANTITY (x) ON WHICH IT DEPENDS.

INFERENTIAL STATISTICS

WE WANT TO INFER PROPERTIES BY USING A SAMPLE OF THE DATA.

A CONFIDENCE INTERVAL ALLOWS FOR ESTIMATING A POPULATION PARAMETER. THE SIZE OF THE INTERVAL INDICATES THE RELIABILITY OF AN ESTIMATE.

EX:

Exercise 3 (7 points) The quality manager of the ACME software house is going to assess the maintainability of the SW house software (50.000 Java classes). In order to do that s/he designs the metric

USEFUL=number of useful comments/number of comments

that is measured through manual inspection on seven classes, randomly chosen.

The collected measures are [0.8, 0.92, 0.75, 0.95, 0.82, 0.88, 0.93].

The candidate has to:

a) indicate the scales of dividend and divisor;

b) indicate the measure type (according to the theory of measure and ISO 25010)

c) indicate the confidence interval for confidence level of 95%, using the following formulae

d) explain what confidence value of 95% means

e) compute how many samples are required to reduce the confidence interval to half (assuming the same value for σ)

$2.77*\sigma/N^{1/2}$ per N=5

$2.26*\sigma/N^{1/2}$ per N=10

$2.09*\sigma/N^{1/2}$ per N=40

$1.96*\sigma/N^{1/2}$ per N "big"

a) SCALE = ABSOLUTE/ABSOLUTE

b) BEING A RATIO BETWEE MEASURABLE AND ABSOLUTE QUANTITIES, IT'S A PROPORTION.

c) MEAN: $\dfrac{0.8+0.92+0.75+0.95+0.82+0.88+0.93}{7} = 0.86$

$$\text{VAR} : \frac{(0.8-0.86)^2+(0.92-0.86)^2+(0.75-0.86)^2+(0.95-0.86)^2+(0.82-0.86)^2+(0.88-0.86)^2+(0.93-0.86)^2}{n-1} = 0.005717$$

$\sigma = \text{VAR}^{1/2} = 0.076$ INTERVAL: $z\ \sigma/N^{1/2} = 2.77 \cdot 0.075/\sqrt{5} = 0.079$

$[0.86-0.079,\ 0.86+0.079] = [0.79, 0.94]$

d) A 95% CONFIDENCE LEVEL MEANS THAT, IF YOU REPEAT THE SAME MEASUREMENT ON DIFFERENT SAMPLES, 95% OF THE TIME THE CALCULATED INTERVAL WILL CONTAIN THE TRUE MEAN VALUE OF THE POPULATION

e) $2.26 \cdot \sigma / \sqrt{x}$ $\qquad 0.079 / 2 = 0.0395$

$0.0395 = 2.26 \cdot 0.075 / \sqrt{x} \rightarrow \sqrt{x} = \frac{2.26 \cdot 0.075}{0.0395} = 4.29 \rightarrow x = 4.29^2 = 17$

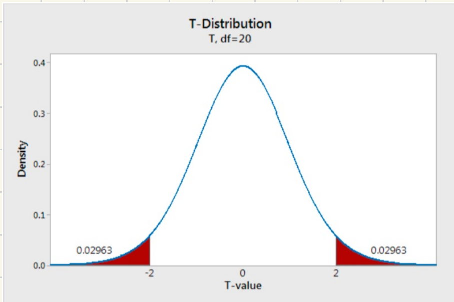
17 SAMPLES ARE NEEDED TO REDUCE THE CONFIDENCE INTERVAL IN HALF

HYPOTHESIS VERIFICATION

OFTEN WE NEED TO COMPARE DIFFERENT REPEATED MEASURES.
A STATISTIC TEST CONSIST OF CHALLENGING THE HYPOTHESIS THAT THE MEANS OF DIFFERENT SAMPLES ARE THE SAME.

T-TEST: A TECHNIQUE THAT ALLOWS TO COMPARE THE DIFFERENCE BETWEEN THE MEAN VALUES OF TWO SAMPLES. IT EXPLOITS A COMPARISON BETWEEN MEANS AND STANDARD DEVIATION.

$$t = \frac{\mu_a - \mu_b}{\sqrt{\frac{\sigma_a^2 + \sigma_b^2}{n}}}$$



P-VALUE: PROBABILITY OF OBTAINING THE SAME OR MORE EXTREME RESULTS IF THE NULL HYPOTHESIS IS TRUE.

- **FALSE POSITIVE**: REJECTING A TRUE HYPOTHESIS.
- **FALSE NEGATIVE**: ACCEPTING A FALSE HYPOTHESIS.

ANALYSIS OF VARIANCE (ANOVA)

ANOVA COMPARE MULTIPLE AVERAGES AT ONCE AND CHECK WHETHER AT LEAST ONE MEAN IS DIFFERENT (ALTERNATIVE HYPOTHESIS) OR WHETHER ALL MEANS ARE THE SAME (NULL HYPOTHESIS).

USE RANDOM VARIABLE SNEDECOR'S F TO COMPARE WITHIN-GROUP VARIABILITY WITH THAT BETWEEN GROUPS.

$$F = \frac{SS_B / (I-1)}{SS_W / [I(J-1)]}$$

$$SS_B = J \sum_{i=1}^{I} (\mu_i - \mu)^2$$

$$SS_W = \sum_{i=1}^{I} \sum_{j=1}^{J} (Y_{ij} - \mu_i)^2$$