SC.203 - Scientific Method
Lecture 4 -
Observational and Experimental Methods

Prof. DUONG Nguyen Vu
Professor, Director
John von Neumann Institute, Vietnam National University HCM

Assoc. Prof. TRAN Minh-Triet
Faculty of Information Technology
University of Science - Vietnam National University HCM

Acknowledgements

- This part of the lecture is adapted from:
  - Vu Duong (2007) "*Scientific Research: Methods, Practices and Models.*" University of Technology & University of Science - Vietnam National University.

Research Problems in CSE

- Studies vs. Research in CSE:
  - Study in Computer Science and Engineering (CSE) tends to solve a "problem" i.e. "How To Design a System that Accomplishes an Automation Task."
  - Research in CSE aims at developing or contributing to a "Theory" that makes an automation task ***better***. It tends to answer to ***Why*** such automation solution is suggested.

Research Problems in CSE

- Human is user of "automated system" hence in many cases, validation requires human testing of its uses. Research includes, in such cases, evaluation of:
  - Cognitive Tasks,
  - Subjective Validations,
  - Human Performance,

- Validation consists in conducting experimentations (with or without human) looking for the evidence that a research topic reckons, positively or negatively.

Observational Method

- Observational Method:
  - Observations in Science and Technology,
  - Characteristics of the Method
    - Observation
    - Selection of samples
      - Arithmetic method to generate random numbers,
      - Physics-based method to generate random numbers,
    - Size of Samples
      - Histograms
      - Distribution of Probabilities

Why Observation is Important in CSE Research?

- Computational Mathematics:
  - Physics-based models (where observation is sometimes required),
  - Mathematical models,
  - Discretization models,
  - Programming models.
- Models are the basis of Simulations,
- Simulations show the needed "evidence" to demonstrate the applicability, extensibility, reliability, etc. of the concerned theory.

Observation Method

- Observe
- Note
- Observed population
- Analyzing relations

Observation

- Another example: "Trees - observed growth in *n* years."
  - How to photograph these changes?
    - Selection representative samples.
    - How to select each member of the samples?
      - Arithmetic
      - Physics
    - How to determine the size of the samples?
  - Observation:
    - We are capable to describe the variations and evolutions, but we could all forget after a certain time.
    - Must note the changes.

Sampling

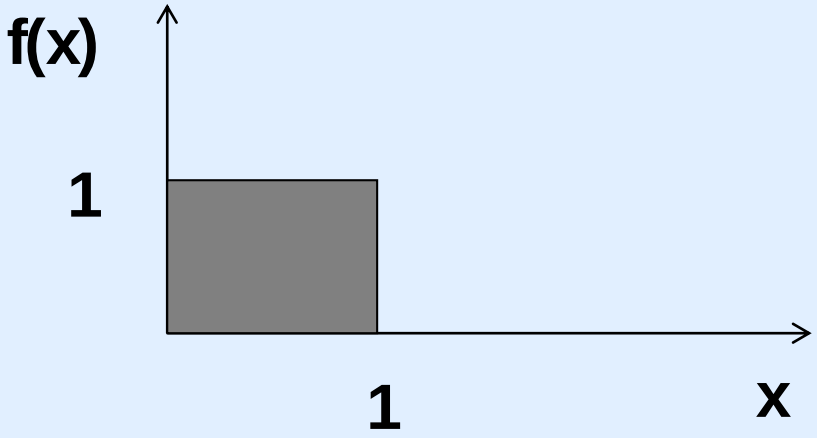
- Supposing that there exists a large population of objects of which samples are required for a manageable description.
  - If this population is homogeneous, each element of the set of sample is representative. There's no problem because any selected sample can represent the characteristics and behaviors of the whole population.
  - If this population is heterogeneous, a separation method must be applied to the set to obtain different types (or different pseudo-homogeneous subsets) before the sampling process.

  The typical procedure consists in selecting the samples of the general population in an *random* manner.

  So, what is the method to generate random numbers or random events that are uniformly distributed?

Sampling

- This procedure aims at obtaining random numbers that are uniformly distributed with the following probability function:



Sampling

Two methods to obtain the desired probability:

- Mathematically or Algorithmically
- Physically-based

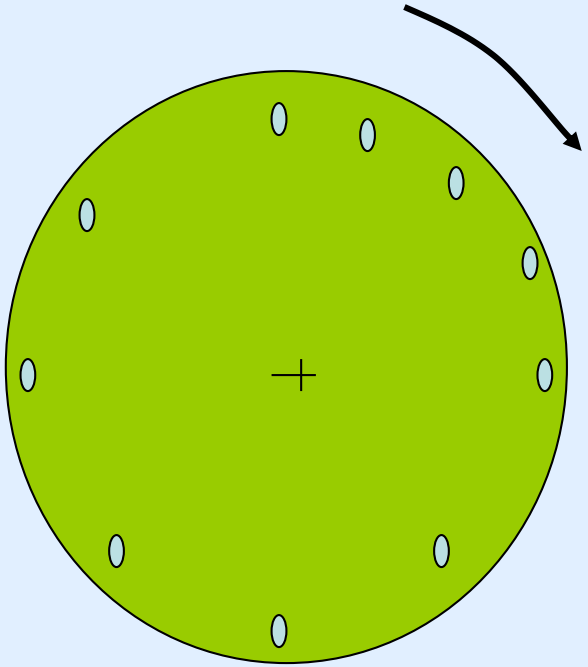
- **Mathematic / Algorithmic Method**
  - Simple program to calculate random numbers in a regular manner:

  x = random (n)

    - We can obtain pseudo-random numbers.
    - To test the uniform probability function, we can generate a series or pseudo-random numbers an then study their distribution.

Sampling

- **Physical Method:**
  - Example: «Wheel of Fortune»
  - Equal chances for all numbers.
  - Each of the N numbers corresponds to a member of the whole population that must be sorted *a priori.*
  - This method can only be applied when N is relatively small, i.e. < 100



Size of Samples

- Intuitive rule: more data better precision.
  - Size of samples?
  - In the practice, there's economical (and/or other) constraints therefore a boundary condition is required..
- Basic concept: Notion of uncertainty.
  - Supposing that we want to measure the quantity x of the observed objects, e.g. the length. The measurements obtained are $(x_1, x_2, \ldots, x_i, \ldots, x_n)$
  - Mean Value

    is the estimates of real value of the whole population.

    we can show that it's the best estimates.

Sampling

- Errors are better obtained with the use of histograms.
  - abscissa orders measured values,
  - Axis orders the number of samplings performed.
  - The shape of the histogram often looks like a Gaussian with two parameters: :
    - Base R represents the set of obtained measurements, usually $R = 4r$
    - Larger $2r$ is defined by $r^2/\Delta^2 = n = R^2/(4\Delta)^2$.
    - n is the number of sampling members.
    - Δ is the standard deviation, or admitted error.
    - Delimiting equation is $\mu = x \pm \Delta$
  - ***Exercise:*** we are interested in the measurements of the length of a mechanical component. Le number of component is high. The base R of the population is approximately 2 microns. On how many components shall we measure to obtain a sampling value within the limit of Δ = 0.1 micron?

Naturalistic Observation

- Also referred to as *field observation,* involves observing *subjects* in their natural habitats for their natural *behaviors.*
- Realistic conditions can lead to higher *ecological validity,* but reproduction of experimental conditions can be difficult.

Ecological validity.

The extend to which research can be generalized to real-life situations (Aronson & Carlsmith, 1968)

Options

- **Undisguised observation.** Studies in which the subjects are aware that the researcher is observing their behavior.
- **Participant observation.** Studies in which the researcher does actively take part in the situation in which the research subjects are involved.
- **Nonparticipant observation.** Studies in which the researcher does not take part in the situation in which the research subjects are involved.

Example: Jane Goodall and Dian Fossey in studies of chimpanzees and gorillas,

Expectancy effects. The influence of the researcher's expectations on the outcome of the study.

Laboratory Observation

- Sometimes referred to as *systematic observation.* Involves observing behavior in a more contrived setting, usually a laboratory, and focusing on a small number of carefully defined behaviors.
- Less costly and less time consuming and affords more control, but could be less realistic than naturalistic observation.

Options

- Subjects are likely to be aware of the observation. However, as with naturalistic observation, the researcher can be either a participant or a non-participant and either disguised or undisguised.
- Subjects can be disguised and undisguised (example of mother and child)
- In lab observation, we are concerned with reactivity and expectancy effects. One way of attempting to control reactivity effects is by using a disguised type of design.
- Notice that increasing controllability can lead to decreasing flexibility.

Reactivity effects. The influence of the subject's reactions to being in the laboratory conditions.

Data Collection in Observational Methods

- Because observational research involves observing and recording behaviors, data are often collected through the use of behavioral measures that can be taken in a direct manner (at the time the behavior occurs) or in an indirect manner (via audio- or videotape).
- Researchers using observational technique can collect data using *narrative records* or *checklist.*

Narrative records. Full narrative descriptions of a subject's behavior. To capture completely everything the subject acts or reacts during a specified period of time.

Checklist. A tally sheet in which the researcher records attributes of the subjects and whether the subject's behaviors were observed. A ***static*** *checklist* records attributes that do not change, while an ***action*** *checklist* is employed to note the presence or absence of behaviors.

Case Study Method

- One of the oldest methods. An in-depth study of one or more individual subjects in the hope of revealing things that are true to all cases.
- Characteristics:
  - One advantage of case study research is that it often suggests hypothesis for future studies.
  - The individual subject under study can have particularities that can not be generalized or the generalization could be erroneous.
  - Expectancy effects: researchers may be biased in their interpretations of their observations or data collection, paying more attention to data that support their theory and ignoring data that present problems for it.
  - Case study research should be used with caution.

Archival Method

- Is a descriptive research method that involves describing data that existed before the time of the investigation.
- Characteristics:
  - Primary data were not generated as part of the study.
  - Problem of reactivity is minimized.
  - Less time consuming than most other research methods because the data already exists, the researcher doesn't interact with the subjects.
  - Not an experimental study but descriptive research.
  - Data was collected a priori for other purposes, and may not fully fit with the objectives of the research.
  - No control on how the data was to be collected.

Qualitative Methods

- Qualitative research focuses on phenomena that occur in natural settings, and the data are analyzed without the use of statistics. Researchers are typically not interested in simplifying, objectifying, or quantifying what they observe. Instead they are more interested in interpreting and making sense of their observations.
- Characteristics:
  - Qualitative research entails observation and/or unstructured interviewing in natural settings. The data are collected in a spontaneous and open-ended fashion.
  - Flexibility over control. Data collection is not necessarily structured and can be sometimes opportunistic.
  - Commonly used by social researchers such as sociologists, anthropologists, and are eventually interesting for psychologists.

In-Class Exercise

- Imagine that you want to study the use of mobile-phone by car drivers. Discuss:
  - The research method to be used?
  - Data collection: which data and how?
  - How would you recommend conducting this study?
  - What concerns do you need to take into consideration?

"In-bus" Exercises

- Explain the difference between participant and non-participant observation and disguised and undisguised observation.
- How does using a narrative record differ from using a checklist?
- Explain the archival method.

Survey Methods

- An observational method where data are collected via questionnaires or survey instruments.
- For the collected data to be both reliable and valid, the construction/design of the questionnaires or any survey instruments shall be done with careful thoughts.
- The type of questions used and the order in which they appear may vary depending on how the survey is ultimately administered (e.g., mail survey vs. telephone survey).

Constructing the Survey (1)

- Writing the questions
  - Should be in clear, simple language to minimize confusion. Example:
    - How long have you lived in Budapest?
      - Since I was 15 years old / Right after I met with my girlfriend!!
    - How many years have you lived in Budapest?
      - 5, 10, 20
  - Whether the question is open-ended, close-ended, partially open-ended, or rating scale questions.
    - **Open-ended questions** ask respondents to formulate their own responses.
      - We can control the length to some extent by the amount of room on the questionnaire.
    - **Close-ended questions** ask the respondents to choose a limited number of alternatives, and to check on the best fits.
      - Nominal or Likert rating scale (1 …. 5 ; Strongly disagree …. Strongly agree).
    - **Partially open-ended questions** are similar to close-ended but one alternative is "**Other**" with a **blank space** next to it for an open-ended short explanation.

Constructing the Survey (2)

- Examples of types of survey questions:
  - **Open-ended**
    Has this lecture responded to your expectations?
    _____
  - **Close-ended**
    Has this lecture responded to your expectations? Yes ____ No ____
  - **Partially-opened**
    Which of the following elements do you find responded to your expectations?
    Lecturer____ Course Note _____ Course Schedule _____ Class room _____
    Lecture 1___ Lecture 2____ Lecture 3____ Other _____
  - **Likert rating scale** (a type of numerical rating scale developed by Rensis Likert, 1932)

This lecture responded to my expectation.

1	2	3	4	5
Strongly disagree		Disagree	Neutral	Agree

Strongly Agree

Likert scale from Wikipedia©

- Likert scale (pronunciation in the field of Psychology varies between /lɪkɜrt/ 'lick-urt' and /laɪkɜrt/ 'lie-kurt', is a type of psychometric response scale often used in questionnaires, and is the most widely used scale in survey research. When responding to a Likert questionnaire item, respondents specify their level of agreement to a statement. The scale is named after Rensis Likert, who published a report describing its use (Likert, 1932).

- A typical test item in a Likert scale is a statement. The respondent is asked to indicate his or her degree of agreement with the statement or any kind of subjective or objective evaluation of the statement. Traditionally a five-point scale is used, however many psychometricians advocate using a seven or nine point scale. A recent empirical study showed that data from 5-point, 7-point and 10-point scales showed very similar characteristics in terms of mean, variance, skewness and kurtosis after a simple re-scaling was applied (although as discussed below, it is inappropriate to apply such statistics to Likert-scale data).
- Likert scaling is a bipolar scaling method, measuring either positive or negative response to a statement. Sometimes a four-point scale is used; this is a forced choice method since the middle option of "Neither agree nor disagree" is not available. Likert scales may be subject to distortion from several causes. Respondents may avoid using extreme response categories (*central tendency bias*); agree with statements as presented (*acquiescence bias*); or try to portray themselves or their organization in a more favorable light (*social desirability bias*).
- Scoring and analysis After the questionnaire is completed, each item may be analyzed separately or in some cases item responses may be summed to create a score for a group of items. Hence, Likert scales are often called summative scales.Responses to a single Likert item are normally treated as ordinal data, because, especially when using only five levels, one cannot assume that respondents perceive the difference between adjacent levels as equidistant. When treated as ordinal data, Likert responses can be collated into bar charts, central tendency summarised by the median or the mode (but not the mean), dispersion summarised by the range across quartiles (but not the standard deviation), or analyzed using non-parametric tests, e.g. Chi-square test, Mann-Whitney test, Wilcoxon signed-rank test, or Kruskal-Wallis test.Responses to several Likert questions may be summed, providing that all questions use the same Likert scale and that the scale is a defendable approximation to an interval scale, in which case they may be treated as interval data measuring a latent variable. If the summed responses fulfils these assumptions, parametric statistical tests such as the analysis of variance can be applied. These can be applied only when the components are more than 5.Data from Likert scales are sometimes reduced to the nominal level by combining all agree and disagree responses into two categories of "accept" and "reject". The Chi-Square, Cochran Q, or McNemar-Test are common statistical procedures used after this transformation.Consensus based assessment (CBA) can be used to create an objective standard for Likert scales in domains where no generally accepted standard or objective standard exists. Consensus based assessment (CBA) can be used to refine or even validate generally accepted standards.

Constructing the Survey (3)

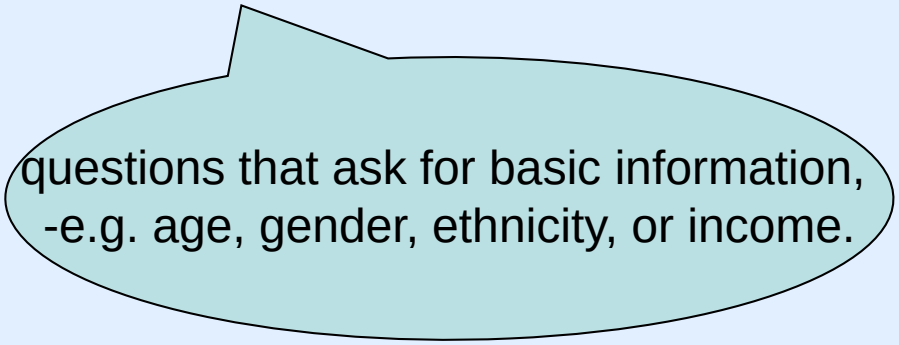
(Sherri L. Jackson, 2008 - Research Methods: A Modular Approach, Thomson Wadsworth, USA)

- Knowing exactly what you want. Don't use:
  - **Loaded question -** a question that includes non-neutral or emotionally laden terms. Example: Do you believe that radical extremists should be allowed to burn flags?
  - **Leading question -** a question that sways the respondent to answer in a desired manner. Example: Most people agree that conserving energy is important - do you agree?
  - **Double-barreled question -** a question that asks more than one thing, often includes **"and"** or **"or."** Example: Do you find using a cell phone to be convenient and time-saving?
- Knowing about the participants/respondents. Care about:
  - **Response bias -** the tendency to consistently give the same answer to almost all of the items on the survey. This is often referred tp as *yea-saying* or *nay-saying.*

    Use standard and reversed formats to detect response bias. Example: "I frequently feel sad" vs. "I am happy almost all the time."

Constructing the Survey (4)

- **Arranging the questions** (Dillman. *Mail and Telephone Surveys: the Total Design Method*. Wiley & Sons, 1978, NY, USA.)
  1. Present related questions in subsets to focus on one issue at a time;
  2. Place questions that dead with sensitive topics (eg. Drug use or sexual experiences) at the end of the applicable subset of questions;
  3. Last, to prevent participants from losing interest in the survey, place the **demographic questions** at the end of the survey.



Administering the Survey by Mail

- Mail Surveys
  - Are written surveys that are self-administered. It is important that mail survey be clearly written and self-explanatory.
  - Advantages:

- can eliminate the problem of **interviewer bias** - the tendency to be influenced
  - can allow the collection of sensitive information
  - usually less expensive than phone survey or personal interviews
  - answers on mail survey are sometimes more complete
  - no pressure to the participants.

  - Potential Problems:

- low response rate (usually 25%-30%).
  - must be carefully analyzed to avoid concentration of responses from few groups, and
  - case of self-selection ie. you only answer when you are or have been in such situation.
  - attention to **sampling bias** - a tendency for one group to be over-represented

Administering the Survey by Phone

- Phone Surveys
  - Involves telephoning the participants and reading questions to them.
  - Advantages:

    - representative samples even with random-digit dialing technique

    - respondents can ask for clarification if needed

    - researcher can ask follow-up questions if needed (for more reliable data)

    - generally higher response rate, 59 to 70% (Groves & Kahn, 1979)
  - Potential Problems:

    - response rate may decrease because of increase in telemarketing

    - more time consuming and more costly

    - can be interviewer bias

    - participants can give so

respondent believes it is deemed appropriate by society

Administering the Survey by Interviews

- **Personal Interviews**
  - A survey in which the questions are asked face to face.
  - Advantages:

    - Can be conducted anywhere, e.g. home, street, shopping mall, university, etc.

    - Allow research t record not only verbal responses but also "body languages"

    - participants usually devote more time

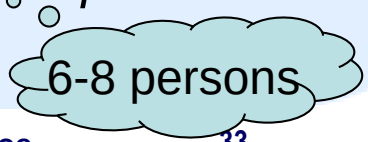
    - clarification of questions and responses possible

    - fairly high response rate, typically around 80% (Dillman, 1978; Erdos 1983)
  - Potential Problems:

    - similar to phone: interviewer bias, socially desirable questions

    - lack of anonymity in personal interviews may affect the responses

A variation on personal interview is the *focus group interviews*



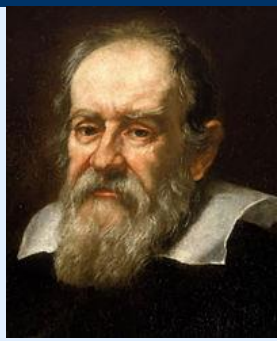
Predictive (Relational) Methods

- Correlational Research - determine whether two naturally occurring variables are related to each other.
- Quaisi-Experimental Designs - some but not all of the features are experimental.

Experimental Method

- Experimental method:
  - Characteristics
    - Basic Steps of an Experiment:
      - Determine relevant factors,
      - Experiment with two variables,
      - Analyze primary data,
      - Analyze secondary (processed) data,
      - Investigate abnormal data,
      - Test baseline hypothesis,
      - Experiment with remaining variables,
      - Link the experimental results to a model.
    - Experimental Conditions :
      - Availability of the system,
      - Quantification of the problem and accountability of the results,
      - Absence of external restrictions on the used methods.

Experimental Method



- Characteristics:
  - Basic steps in an experiment,
  - Experimental conditions,

  - Illustration - case study developed by Galilee Galileo: «motion of a rigid body along an axis.»
    - Rigid body = vehicle,
    - Axis = horizontal direction
    - Staring point = A
    - Goal = Find the traveling time needed to attaint a certain distance.

Step 1

- Determine relevant factors:
  - What are the important factors or variables to describe the motion of the vehicle?
    - Traveling time,
    - Traveling distance,
    - Quantity of available gasoline (no gas to move!)

Step 2

- Perform an experiment with two variables:
  - Simplification: consumption constant.
  - Experimentation: investigate the variation of ***distance*** as a function of traveling ***time***.
  - Organization of the experiment:
    - Clocks are placed on the trajectory at known distances. These clocks are synchronized and reset to zero at the departure of the vehicle.
    - Perform several runs and measure the distance and time. These data are calculated and tabulated. Each value is represented by a mean value and a delta.

measures

Distance (m)	Time (sec)
00.0	0.0
10.00 ± 0.01	2.0 ± 0.2
20.01 ± 0.02	2.8 ± 0.1
30.00 ± 0.01	3.5 ± 0.1
40.01 ± 0.01	4.1 ± 0.3
50.00 ± 0.01	4.8 ± 0.2

Step 3

- Analyze primary data:
  - Detect the possible relations between variables. The easiest is to examine the linearity between the two variables with help of graphical description.
  - For our case study: the graphic depicts that measurement points are not aligned on a straight line. Noticeably the ones at the departure and at the end of the trajectory.
  - This result invalidates our initial hypothesis that traveling time and distance are proportionally linear, in other words $d \neq k.t$ (k is a constant)

Step 4

- Analyze secondary (processed) data:
  - Observe and analyze the relations between variables.
  - Interrogate on deviations if there's any.
  - Determine domains of values on which the relations between variables change.
  - Example of case study:
    - The relationship between distance and time is not constant.
    - In using the least-square fitting method, a straight line fits to the measured data up to the distance of 50m. There's a deviation that falsifies the linearity between time and distance.
    - A specific analysis is required if another test shows the same phenomenon (confirmation of the deviation, and not a hazard / fluctuation).
    - New hypothesis: «effect of resistance of the air?»
    - Note that the resistance increases at the square of the speed.

Step 5

- Treat specific points (anomalies)
  - Pose hypothesis on the anomaly under interrogative form.
  - Establish scenario to experiment with new hypothesis.
  - Example of our case:
    - «air resistance?»
    - How to measure or confirm?
      - Continue experimentation to verify the existence of air resistance as a function of speed.

Step 6

- Confirm performed analyses:
  - Test the obtained linearity,
  - Test hypothesis about the resistance of the air.
  - A physical model may be needed.
  - Physical models in laboratory experiments often require:
    - Techniques in dimensional analysis,
    - Analogical models.

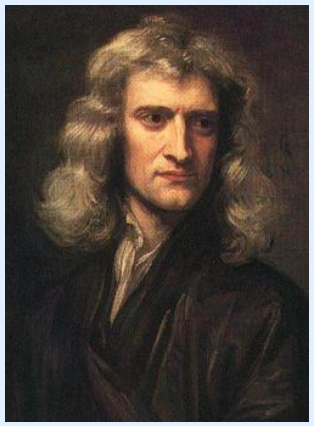
Step 7

- Experiment with remaining variables
  - Same approach to show evidence the relationships between another variable and those experimented.
  - Through linearity relations, identify the tendency between variations
  - Techniques:
    - Dimensional Analysis
    - Reduced Models
    - Analogical Models
    - Qualitative Physique

Step 8

- Link experimental results to a model:
  - Ultimate goal of an experimentation,
  - Conclude performed analyses.

  - Example:
    - Obtained model is a mathematical model known as "Newton Mechanics"
    - If the resistance of the air is neglected, the relation between time and distance can be written $d = 1/2\, at^2$; a is the acceleration.



Experimental Conditions

- Availability of the “system”,
- Quantifiability of the problem,
- Absence of restrictions on the use of the experimental method

Experimental Method

- Difference between exercise and research:
  - Exercise :
    - Instruments
    - Time
    - No feedback
    - Routine
  - Research :
    - Instrument ++
    - Time ++
    - Precision
    - New discoveries

Experimental Method

- Experimental Design and Optimization
  - Design of experiments look for optimal conditions for an investigation;
  - Parameters to identify in experimentations:
    - *cost function* of the operation,
    - Precision, expressed in terms of uncertainty or errors, Δ
    - *damage function*
  - Optimization :
    - Minimization of the cost, and damages,
    - Maximization of the precision or minimization of the errors or uncertainty.