MUNI
SPORT

Methodology and Statistics in Sports Sciences

e059

Masaryk University
Faculty of Sports
Studies

MaSiSS

Guarantor and Teacher subject

Mgr. Michal Bozděch, Ph.D.

- **Email:** Michal.bozdech@fsps.muni.cz
- **Phone number:** 549 49 6863
- Department of Kinesiology
- **Room:** D33/333

MUNI SPORT

MaSiSS

Research Methodology

1. A review of the Fundamentals
2. Type of Research
   1. Selected Types of Qualitative Research
   2. Selected Types of Quantitative Research
3. Research Project
4. Research Ethics
5. The Literature Review
6. The Research Hypotheses
7. Data Collection Methods
   1. Collecting a Primary and Secondary Data
   2. Collecting a Qualitative and Quantitative Data

Statistica Analysis

1. Characteristics of a Good Test
2. Variable Types and Scaling
   1. Research Variables
   2. Levels of Measurement (Scale of Measure)
3. Descriptive Statistic
   1. Measures of Distribution or Frequency
   2. Measures of Central Tendency
   3. Measures of Dispersion or Variation
   4. Measures of Position
4. Inferential or Analytical Statistic
   1. Hypothesis (or Predictions) Testing
   2. Choosing a Comparison Test
   3. Choosing a Correlation Tests
   4. Choosing a Regression Tests
5. The effect size

MaSiSS

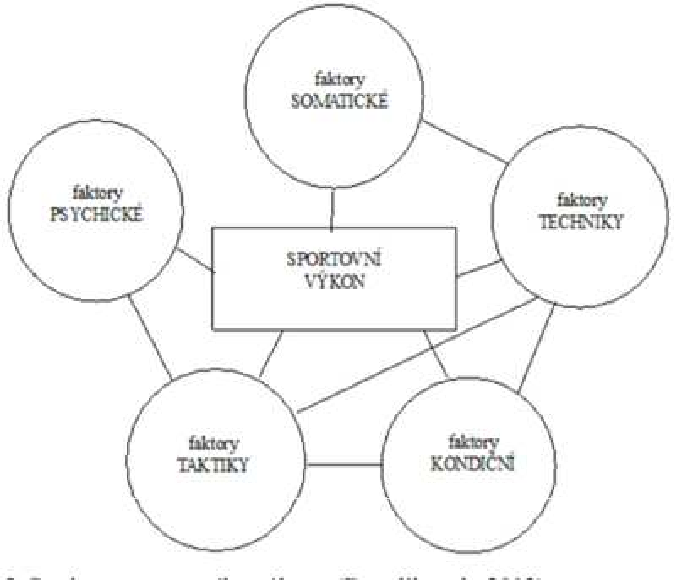
Research Methodology

Science (Benešová, 1999) - is a *systematic*, *critical* and *methodological* pursuit of *true* and common knowledge in a defined area of reality

Science (Ferjenčík, 2000) - is a comprehensive system of information obtained by a *scientific method*. Science provides guidelines for examination (methods) and an explanation of collected information (scientific theory)

Theory (Zháněl, 2014) - The primary goal of science is , i.e. an attempt to find *general explanations* of *natural phenomena* (via science research)

Examples of theory



(Dovalil et al., 2012)

MUNI SPORT

MaSiSS

Research Methodology

Sport science – Kinesiology, Kinanthropology, Sportwissenschaft - is multidisciplinary science about human (voluntary) movement.

Paradigm - a *fundamental concept* of a certain scientific discipline, which is considered to be exemplary. Defines proper procedures and methods, according to which rules and conventions.

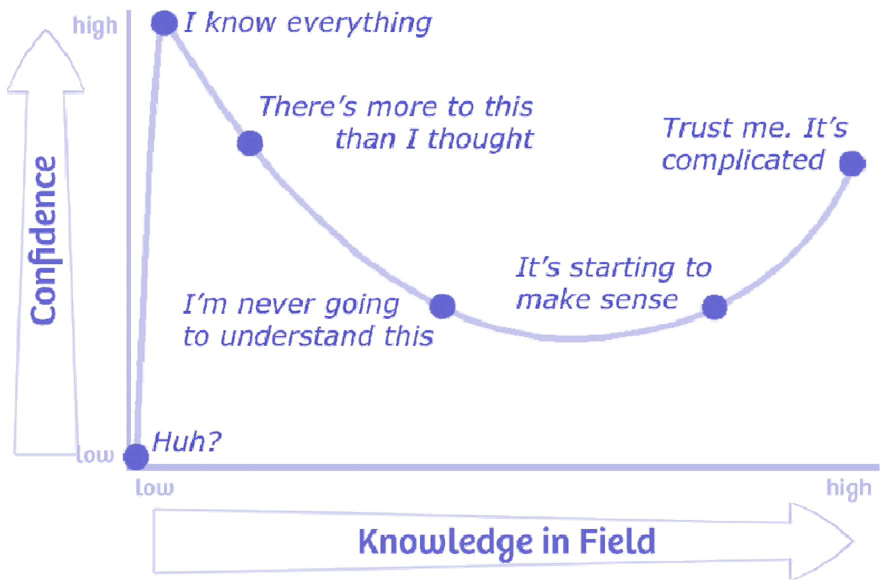
Methodology vs methods (Gabriel, 2011) – method is a *research tool* (interwiew in qualitatice study) and methodology is the *justification* for using this method.

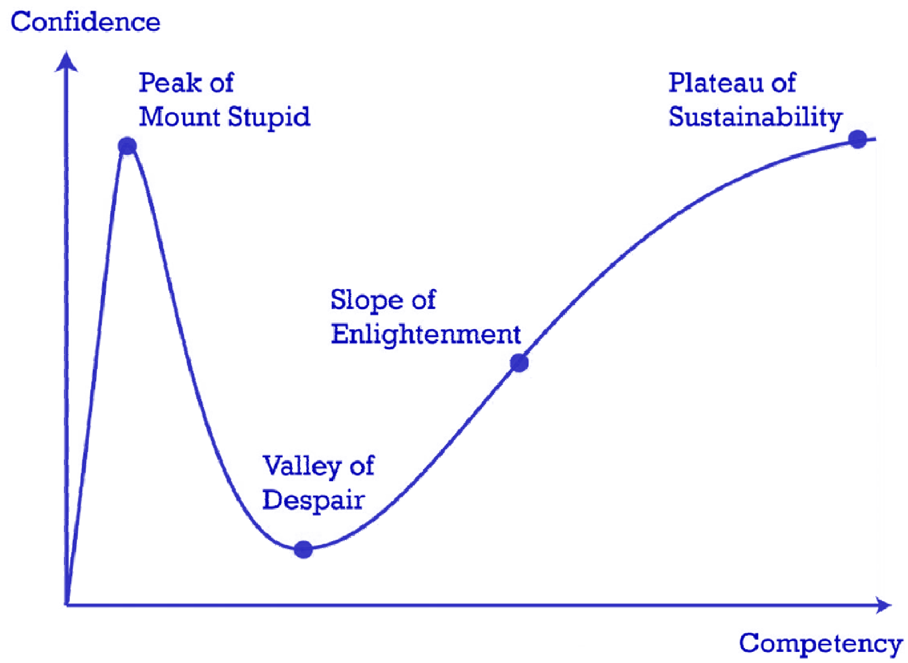
- **Methodology** – the *summary* and study of methods, or the science of methods.
- **Method** – a tool or process for monitoring, researching, learning, exploring and achieving certain goals.
- **Methodical guidelines** – specific instructions for carrying out specific activities.

MUNI SPORT

MaSiSS

Dunning-Kruger effect



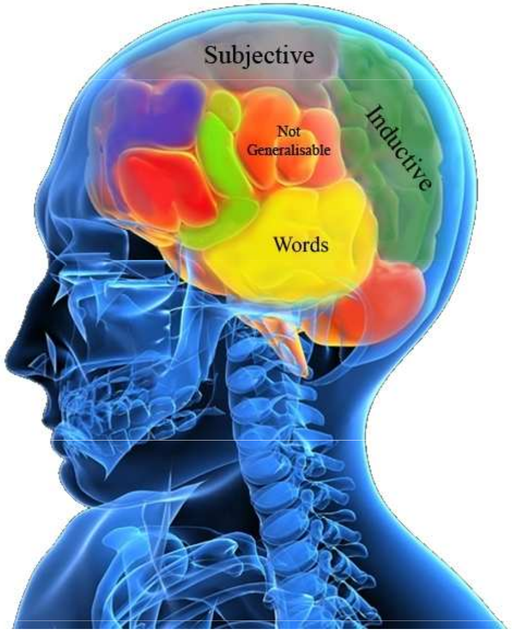


MUNI SPORT

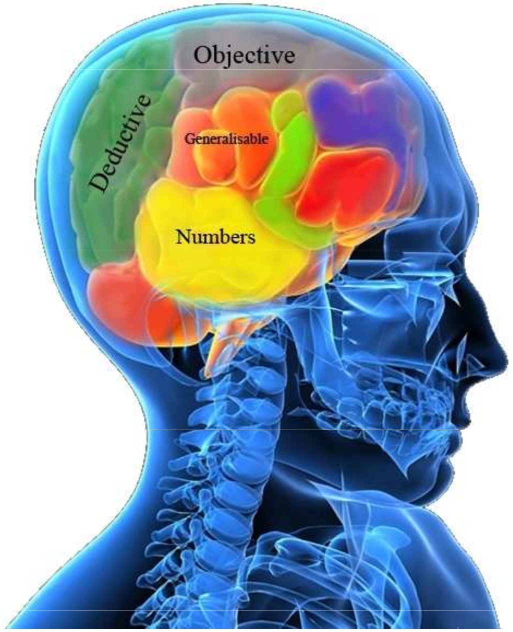
MaSiSS

Type of Research

Qualitative research



Quantitative research



MUNI
SPORT

MaSiSS

Type of Research

Qualitative research - According to Švaříček & Šeďová et al. (2007) it is a process of examining phenomena and problems in an *authentic environment* in order to obtain a *comprehensive picture* of these *phenomena* based on *deep data* and a *specific relationship* between a researcher and a research participant. The aim of a qualitative research is to uncover and represent how people *understand*, *experience* and *create* social reality. It focuses on how individuals and/or groups *view*, *understand* and *interpret* the World (Zháněl, 2014), via *non-numerical* data.

Quantitative research - is the process of *collecting* and *analyzing numerical data*. It can be used to find *patterns* and *averages*, make *predictions*, test causal *relationships*, and *generalize* results to wider populations (Bhandari, 2021).

A common mistake: when it is relatively simple to transforming words (eg answers in a questionnaire) into numbers (occurrence) these are quantitative data

Mixed research

Hendl (2005) then talks about a combination of quantitative and qualitative methods in a single research activity when speaking of a mixed research strategy, in which the results obtained by the two strategies complement each other

MaSiSS

Type of *Qualitative* Research

Phenomenological method – how participant *experiences*, *feel* have *opinion* about a specific *event* or *activity*. It utilizes in-depth interviews, observation or survey to gather information

Grounded theory - tries to explain why a *course of action evolved* the way it did. Need large subject number to developer theoretical model based on existing data (DNA – genetic model)

Case study - *in-depth look* at one test subject (opposite of grounded theory). Various data are compiled to create a bigger conclusion

Focus groups - *group of individuals* who are asked questions about their *opinions* and *attitudes* towards certain *fenomen*

Ethnographic research - subjects are experiencing a *culture* that is unfamiliar to them

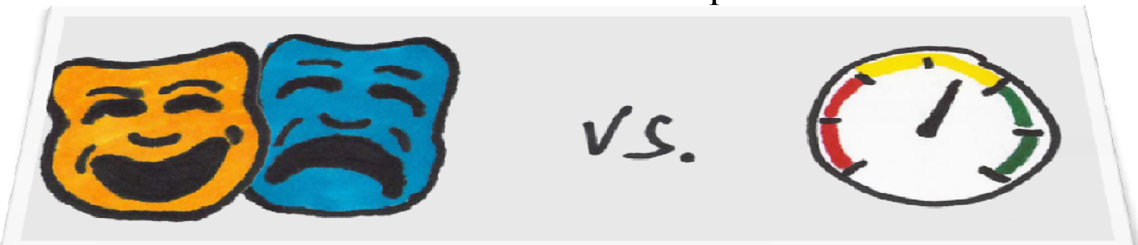
Type of *Quantitative* Research

Descriptive research - seeks to *describe the current status* of an identified variable. The researcher does not usually begin with an hypothesis, but is likely to develop one after collecting data

Correlation research - aim is to determine the extent of a *relationship* between two or more *variables* using statistical data. This type of research will recognize trends and patterns in data

Causal-Comparative/Quasi-Experimental - establish *cause-effect relationships* among the *variables*. The researcher does not randomly assign groups

Experimental Research - establish the *cause-effect relationship*. Subjects are randomly assigned to experimental



MUNI SPORT

MaSiSS

Type of Research

Types of different research (Hendl, 2016)

Methodological study, Case study, Comparison, Correlation-predictive study, Experiment, Quasi-experiment, Evaluation, Development studies, Trend analysis, attitudes, Status, Exploration, Historical study, Modelling, Proposal and demonstration, systematic review, Meta-analysis, Theoretical studies, Analytical, Qualitative study

Types of different research (Mishra & Alok, 2017)

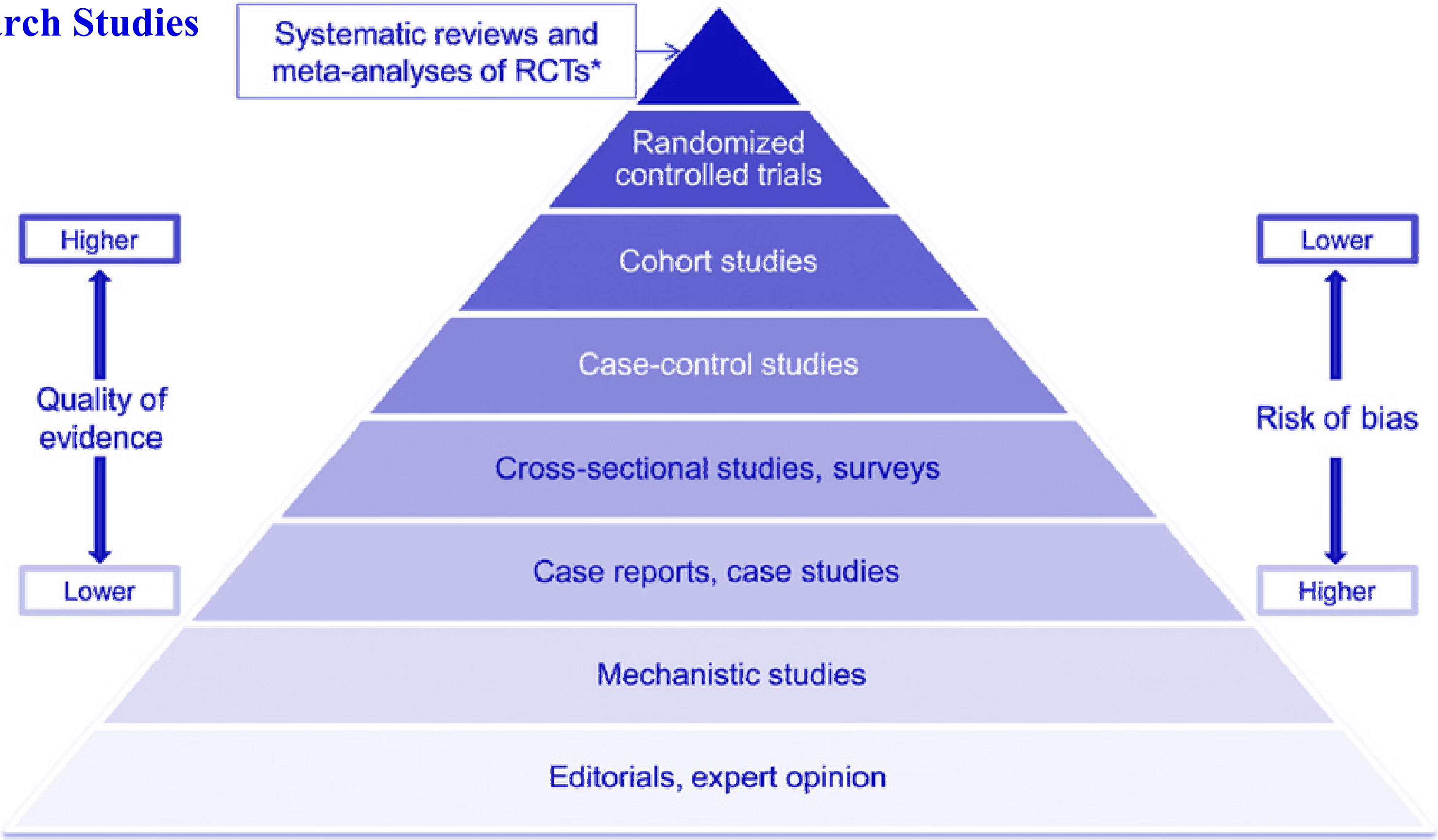
- Descriptive vs. analytical research
- Applied vs. fundamental research
- Conceptual vs. empirical research

Other types of research

- one-time research vs. longitudinal research
- field-setting research vs. laboratory research vs. model research
- clinical vs. diagnostic research
- case-study methods vs. exhaustively approaches
- exploratory vs. formalized research
- The objective of exploratory research (creation of hypotheses rather than their testing) vs. formalized research (specific hypotheses are tested)
- conclusion-oriented and decision-oriented research

MaSiSS

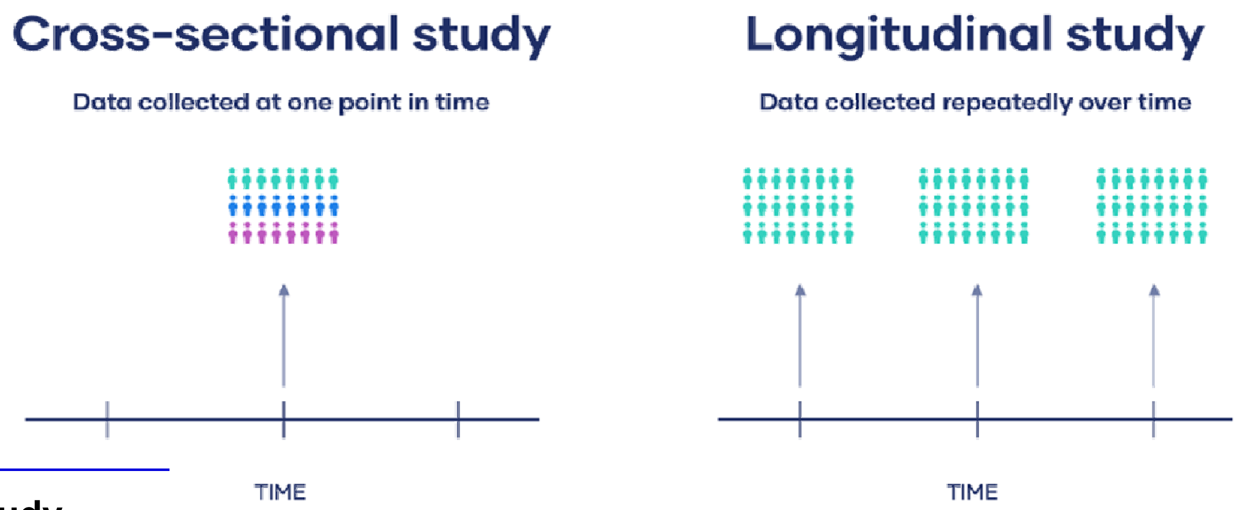
Type of Research Studies



RCT* = randomized controlled trial

MaSiSS

Type of Research Studies



Cross-sectional study	Longitudinal study
One point in time	Several point in time
Different samples	Same sample
Change at a societal level	Change at the individual level

MUNI SPORT

MaSiSS

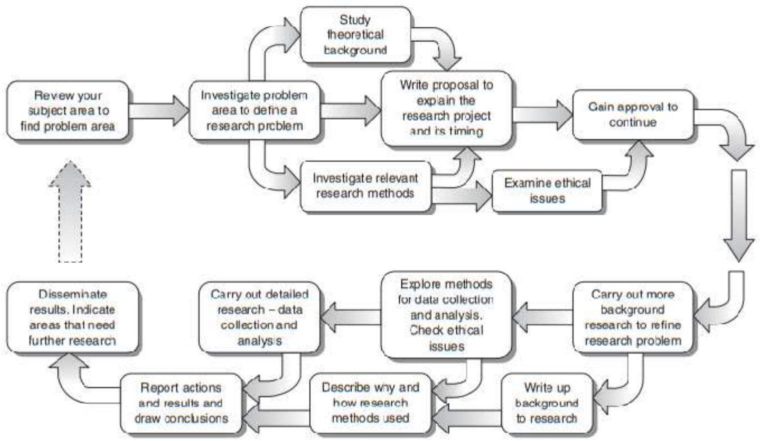
Type of Research Studies

Longitudinal study	Pseudo-longitudinal study	Semi-longitudinal study
Several point in time	One point in time	Several point in time
Same sample	Different (age) samples	Different and same samples
Change at the individual level	*Maturation*	*Until the youngest group become the oldest group*

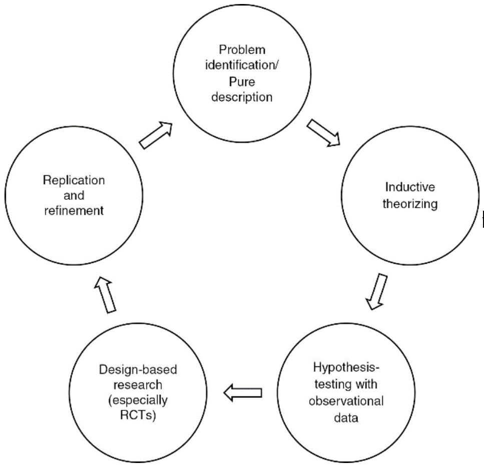
MUNI SPORT

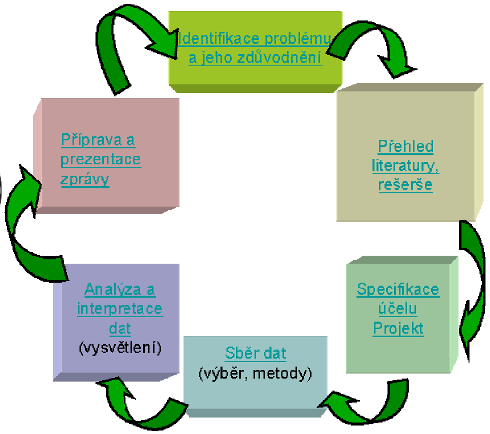
MaSiSS

Research Project



(Williman, 2011)





(Elman & Mahoney, 2020; Hendl, 2016)

MaSiSS

Research Project

Phases of the research process (Rockmann & Bömermann, 2006)

1. Formulation of a research problem
2. Research planning
3. Implementation of Research
4. Evaluation of research
5. Publication of research results

Scheme of logical thought progress of scientific work (Zháněl, 2014)

Research intention → Research problem → Research objective → Research question (hypothesis)

Principles of a good research project (Robson, 1993)

1. it says what you *want* to do and *why* do you want to do
2. is written *clearly* and without *unnecessary description* (secondary facts),
3. is clearly *organized* and *straightforward*

MaSiSS

Research Ethics

There are two aspects of ethical issues in research (Williman, 2011)

1. The individual values of the researcher relating to *honesty* and *frankness* and *personal integrity*.
2. The researcher's treatment of other people involved in the research, relating to *informed consent*, *confidentiality*, *anonymity* and *courtesy*.

Most commen problems are

- Plagiarism (or wrong used of citations)
- Non-disclosure acknowledgement
- Data collection/analysis/interpretation
- Disclosure of interest; using third-party materiál

Plagiarism - It is a simple matter to follow the clear guidelines in citation that will prevent you being accused of passing off other people's work as your own (Williman, 2011)

MUNI SPORT

MaSiSS

Research Ethics

General Principles

- The <u>*health*</u> of my patient will be my first consideration
- Physician must promote and safeguard the health, well-being and rights of patients
- … never take precedence over the rights and interests of individual research subjects
- researcher protect the *life*, *health*, *dignity*, *integrity*, right to *self-determination*, *privacy*, and *confidentiality* of personal information of research subjects
- …

Helsinki Declaration

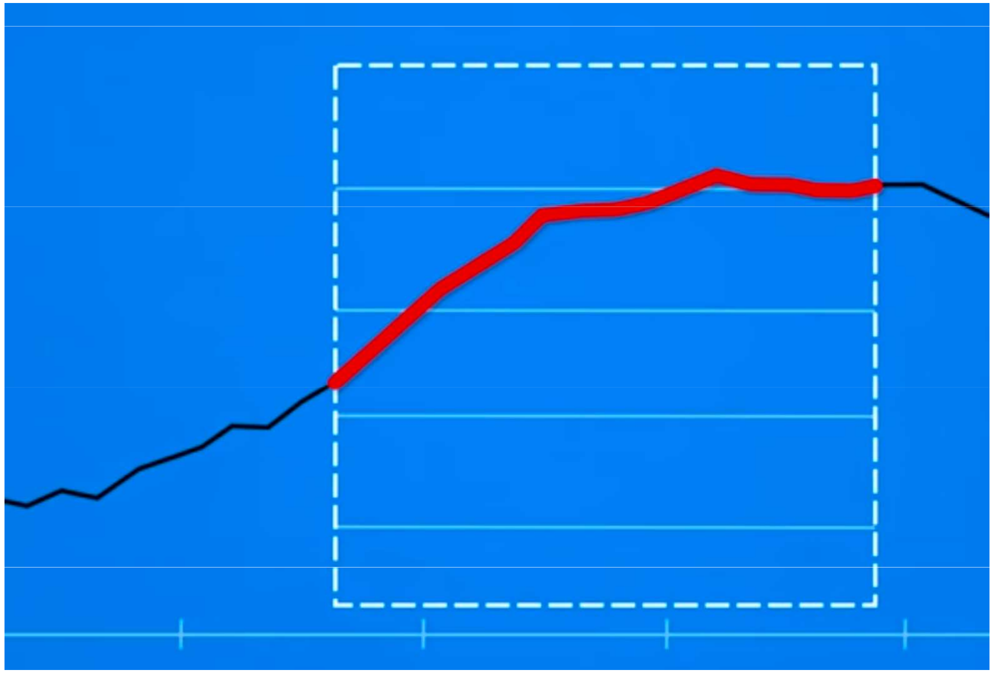
- Adopted by the 18th The World Medical Association (WMA) General Assembly, Helsinki, Finland, June 1964
- Statement of *ethical principles* for medical research involving *human subjects*, including research on identifiable *human material* and *data*

MaSiSS

Research Ethics

- Carefully *chosen* time range (on x-axis) or data scale (on y-axis)
- Picked specific data points can *hide* important changess in between
- This mean grap his not wrong, bud *leaving out relevant data* can give a *misleading impression*

Cherry pick data



MUNI SPORT

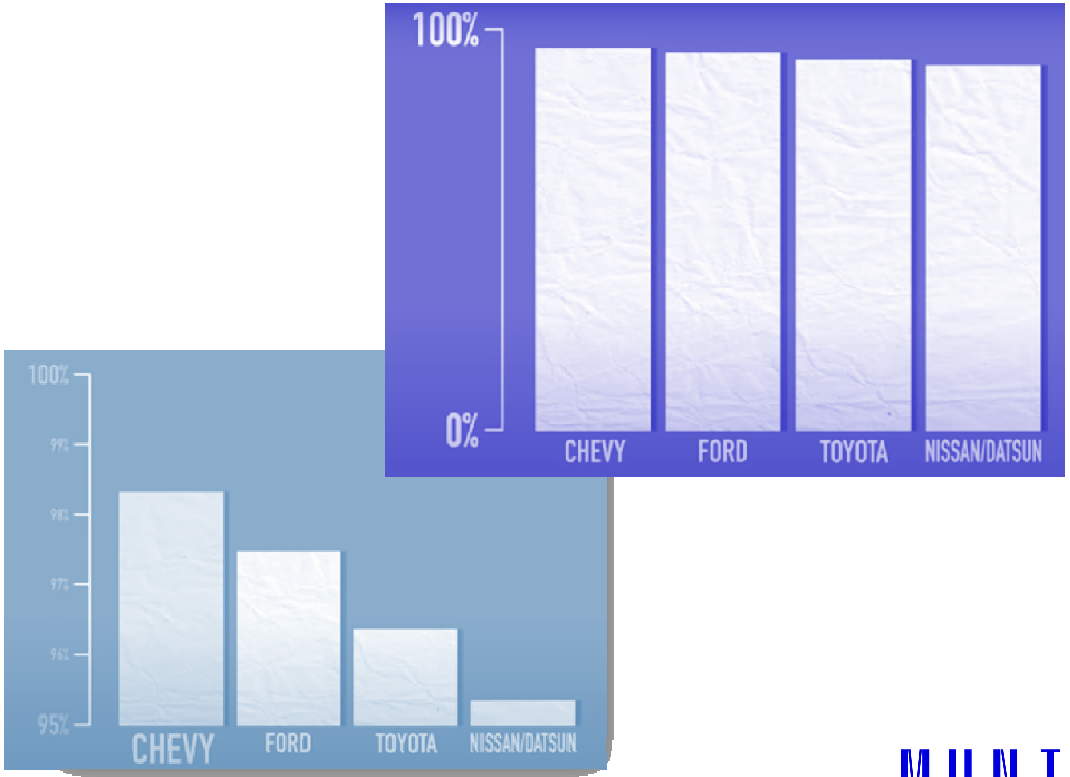
MaSiSS

Research Ethics

Cherry pick data

Example of cherry-picking (on y-axis)

- In 1992 Chevrolet car ads claims that all Chevy trunks sold in the last 10 years are still on the road. For soporting this claim there use this graph.
- The graph can also by *intepret* that Chevy can are *twice as dependable* as Toyota trucks
- Until you see *scala of graf*, which is from 95% to 100%
- Seond graf shows how graf will look on 0% to 100% scale



MaSiSS

Research Ethics

Cherry pick data

Example of cherry-picking (on x-axis)

- Usuali in line graphs showing something changing over a time
- Graph show *job loss* by quater
- But scale on x-asis is ***inconsistent***
    - September 08 to march 09 = 6 months
    - March 09 to jun 10 = 15 months

MUNI
SPORT

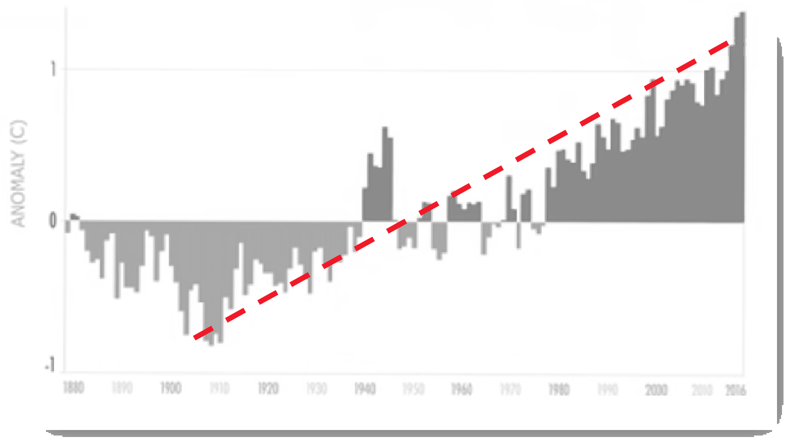
MaSiSS

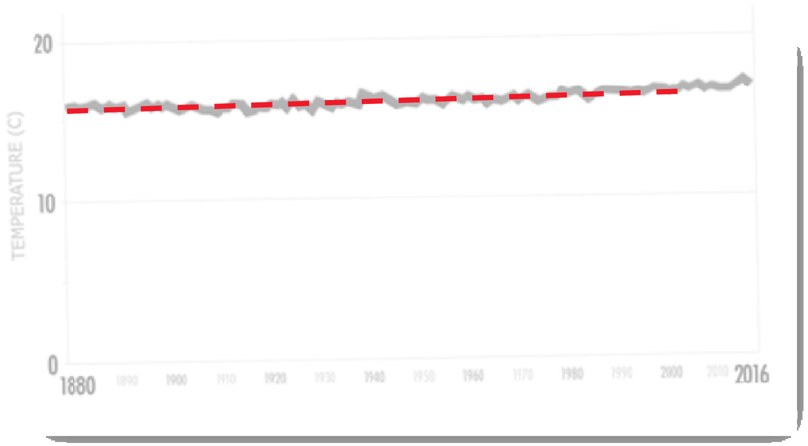
Research Ethics

Cherry pick data

Knowing the full signifikance of data that graph presented

- If boths graphs used *same data* (Annual global ocean average temperatus), whay they look so *different*?
- And if you know that even a *<u>rise of half a degree Celsius</u>* can cause massive ecological disruption, which graph you thing is more appropriate?





MUNI SPORT

MaSiSS

Research Ethics

Example of omitting variables

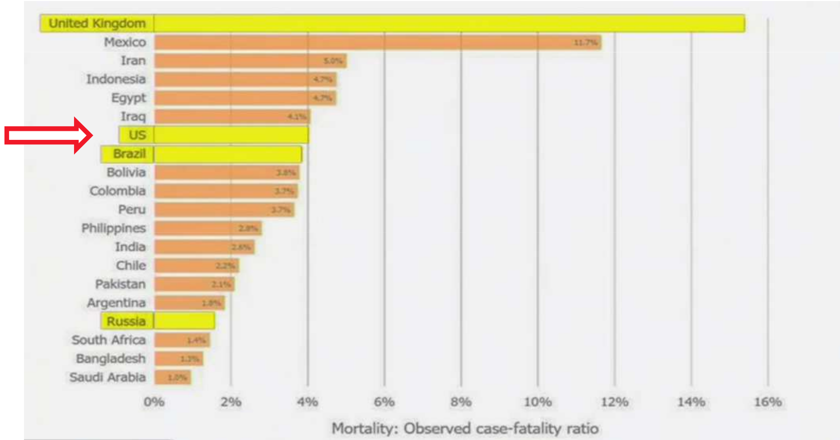
- Donald J. Trump: USA has the lowest / best mortality rate.
  - if we omit better states

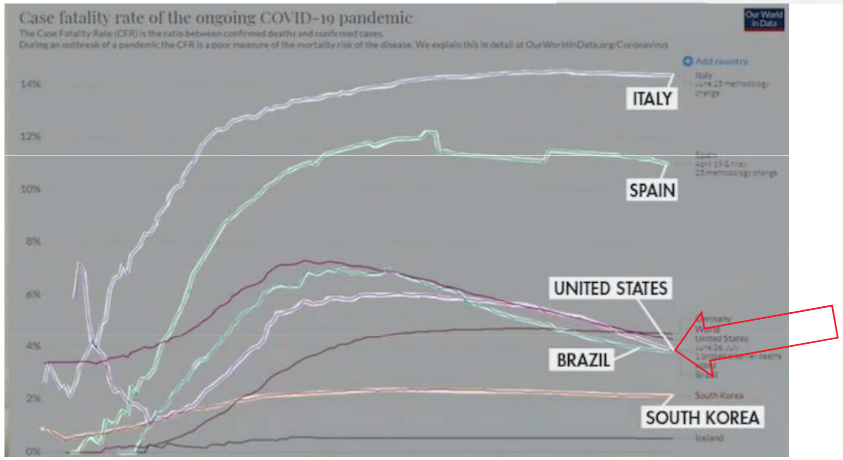
Bonus

- Donald J. Trump: "*I noticed that the more tests we do, the more we get infected*"

  = tests cause COVID-19?

Cherry pick data





MUNI SPORT

MaSiSS

Research Ethics

Hacking of p-value

Also know as *data dredging, data snooping, data fishing* or *data butchery*

Statistically significant, when in reality, there is ***no underlying effect*** (Head et al., 2015; Norman, 2014)

Academic journal prefer articles with *statistically significant results* and researchers are force to publish in high quality journals (with higher Impact Factor, IF)

In practice we often encounter so-called *p*-hacking and salami publication/slicing (splitting research data/results into several publications) with caused false **TYPE I ERROR** (accept a true null hypothesis – false positive, α)

Prevention from *p*-hacking

- Create *two sets of data* (cross-validation method)
- build strong *project* (don't change the projekt methodology during the research)
- used Bonferroni correction (for more statistical analyses),
- Scheffé's method
- False discovery rate,
- Raw data publishing

p-haking leads to false positive results = negative impact to future research field and researchers prestige

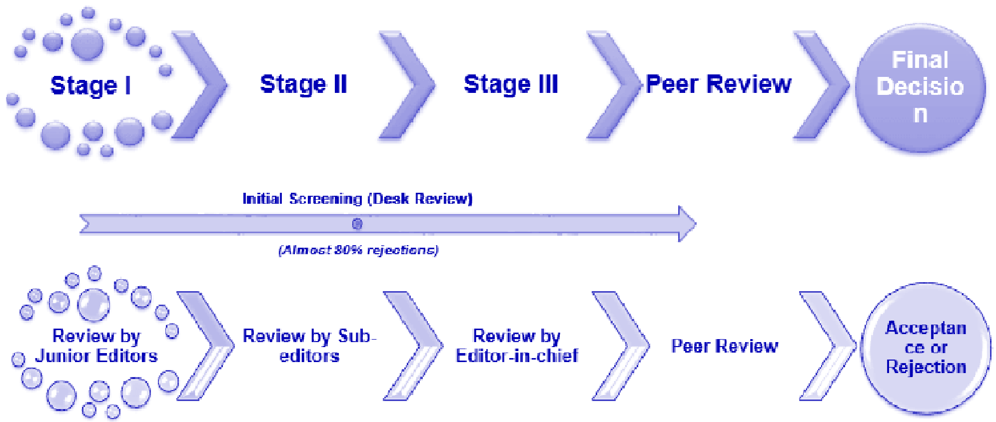
MUNI SPORT

MaSiSS

Research Ethics

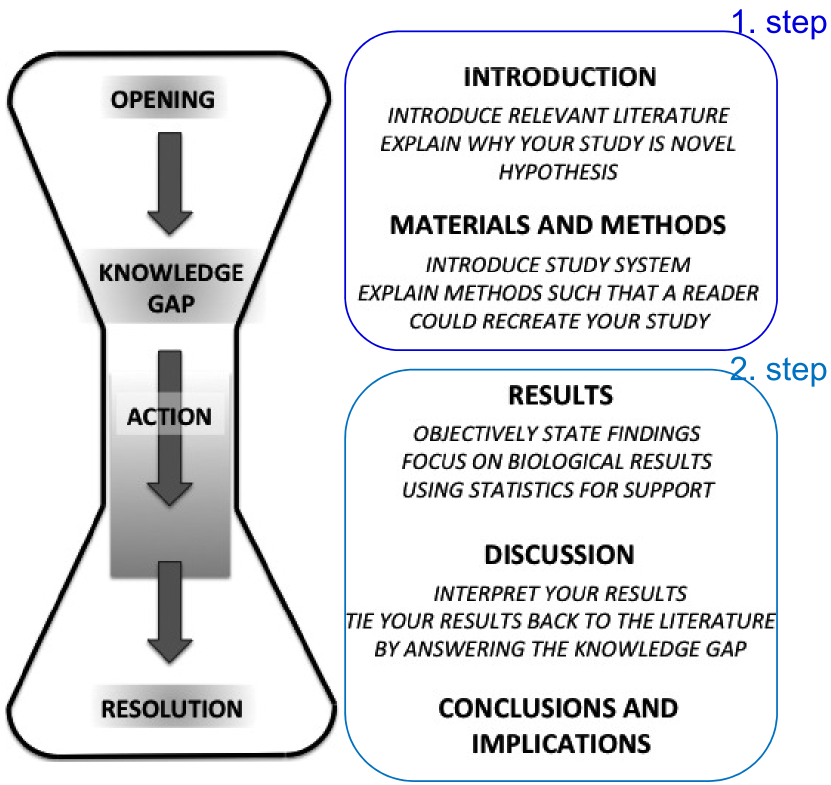
New way to **reduce *p*-hacking** or other **non-ethical manipulation** is *2-step manuscript submission*

- 1st you submit only introduction and description of your method and journal decides whether to publish before seeing your results



Blind peer review submission process

Hacking of p-value

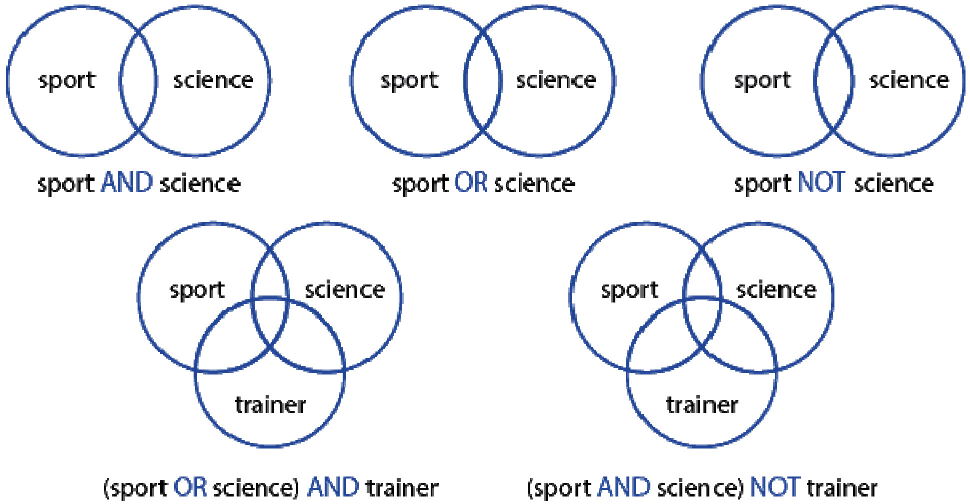


MaSiSS

The Literature Review

Step-by-step guide of literature review

1. ***Search for relevant literature on selected topic***
   a. What? – books, academic sources (journal articles)
   b. Where? Google Scholar, discovery.muni.cz; Web of science; ScienceDirect; PubMed
   c. Define your keywords and their synonyms
   d. For more relevant results use Boolean operators
2. Evaluate and select sources
3. Identify themes, debates and gaps
4. outline your project structure
5. write and rewrite what you wrote



"If you steal from one author it's plagiarism, if you steal from many it's research"

- Wilson Mizner

MaSiSS

The Literature Review

Step-by-step guide of literature review

1. Search for relevant literature on selected topic
2. ***<u>Evaluate and select sources</u>***
   a. you can't read everything that has ever been written (unless your topic is very)
   b. read just abstract, which normally consists of background, aim, methods, results, conclusion and sometimes even recommendations for practice
   c. save relevant articles (in pdf) for later purposes
   d. check cite literature to find other relevant sources
   e. tips: pay attention to the citation count and experts in the selected field

MUNI
SPORT

MaSiSS

The Literature Review

Step-by-step guide of literature review

1. search for relevant literature on selected topic
2. Evaluate and select sources
3. ***Identify themes, debates and gaps***
    a. Find connection between different sources
        i. in what they agree and disagree,
        ii. slightly or very different conclusions
        iii. trends and patterns
        iv. used process, methods, equipment and statistic tests
        v. gaps
4. ***outline your project structure***
    a. different approaches: the most used approach is from general to specific (this applies to chapters, subchapters and also to paragraphs)
5. ***write and rewrite what you wrote***

MaSiSS

The Research Hypotheses

Hypotheses are statements that relate to the existence of the relationship between variables or prediction of defined variables using other variables (Zháněl, 2014)

- Without hypothesis the research in *unfoucussed*
- Hypothesis is necessary *link* between *theory* and *investigation*

Sources of hypothesis

- Theory and studies (literature research)
- Observation (from own practice / experience)
- Intuition (less then theory and observation)
- Culture (behavior or beliefs of social, ethnic or age group)
- New trend (possible future experience)

MaSiSS

The Research Hypotheses

- *Power of prediction* – predict the future situation, not only the present situation
- *Based on observation* – If we cannot verificated a thing, which we cannot observed
- *Simplicity* – everyone should be able to understand it
- *Clarity* – It should be clear from ambiguous information
- *Testability* – it should be able to by be tested empirically
- *Limit the unninteraction*

Charakcteristic of good hypothesis

- *Relevant to problem* – A hypothesis is guidance for the identification and solution of the problem
- *Specific* – avoid generalization terms, omit unwanted factors (variables)
- *Relevant to available researchers techniques* – you must know workable techniques before formulating a hypothesis
- *Provide new suggestion/knowledge/technique/process…*- it is not a repetition of what we already know
- *Consistency and harmony* – There must be a close relationship between variables which one is dependent on other

(Mill, 1963)

MUNI SPORT

MaSiSS

The Research Hypotheses

Working hypothesis – not very specific, they can be easily modified, used with insufficient data; example: Chocolates before training ensures maximum performance

Descriptive hypothesis – variable can be situation, event, organization, person, group, object

- Relation hypothesis (describes relationship – positive, negative or casual – between two variables) example: Children from higher incomes families spend more times at leisure physical activities
- Formalised hypothesis (cause and effect relationship between independent and dependent variable) example: If families have higher incomes, than they spend more money on leisure physical activities

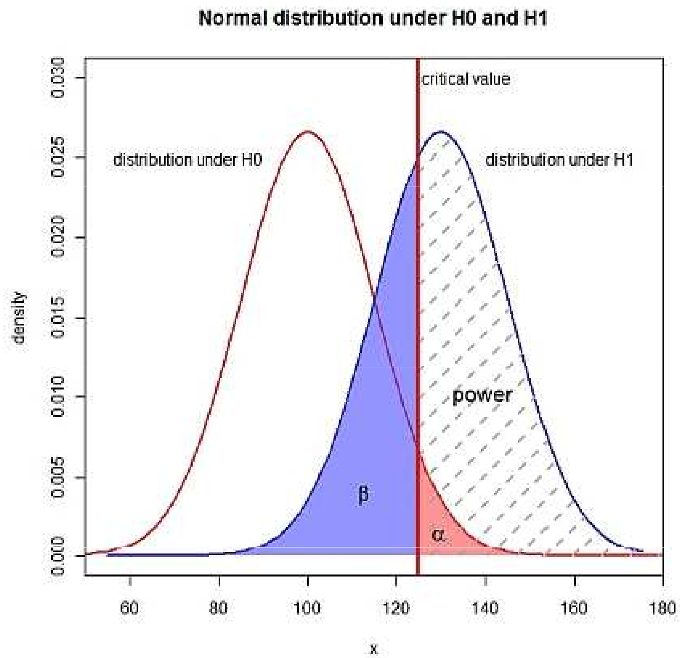
Types of research hypotheses

1. *Null hypothesis* (H_0) – predicts that there is no relationship between two variables; example: after a 3-month training program, there are no statistically significant differences in the muscular strength of the knee extensors between the experimental and control groups.
2. *Alternative hypothesis* (H_a/H_1) – the opposite statement than H_0. For acceptance of H_a, H_0 must be rejected first. Exapmle: after a 3-month training program, there are statistically significant differences in the muscular strength of the knee extensors between the experimental and control groups
   1. Directional (left or right tailed test) – hypothesis in which we can predict effect (positive/negative) of one variable on others. Example: Girls are more flexible than boys
   2. Non-directional (two tailed test) – hypothesis in which we cannot predict effect, but stat a relationship between variables (we do not know what kind of difference); example: there will be a difference in the performance of experimental and control groups

(Zháněl, 2014; Cauvery et al., 2010)

MaSiSS

The Research Hypotheses

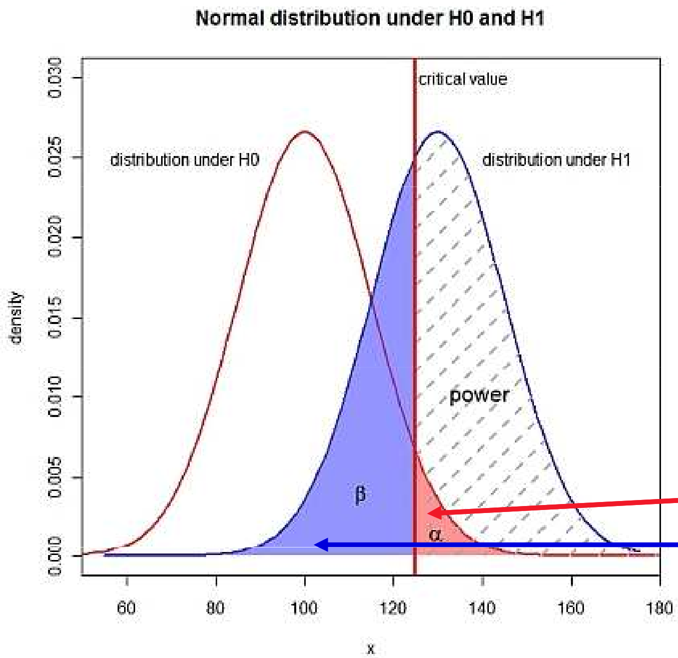


Rejecting the null hypothesis

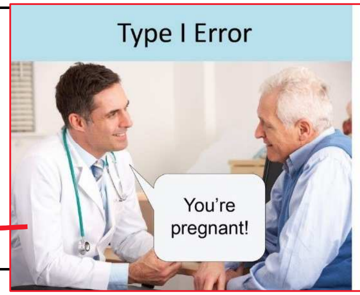
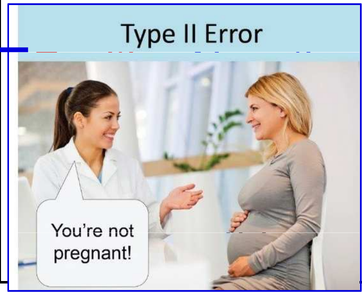
		Actual Value (reality)	
		Positive	Negative
Conclusion from hypothesis test	Positive	Positive Positive == TRUE Positive (Power, 1 – β)	Negative Positive == **FALSE Positive**, *Type I Error (α)*
	Negative	Positive Negative == **FALSE Negative** *Type II Error (β)*	Negative Negative == TRUE Negative

MaSiSS

The Research Hypotheses



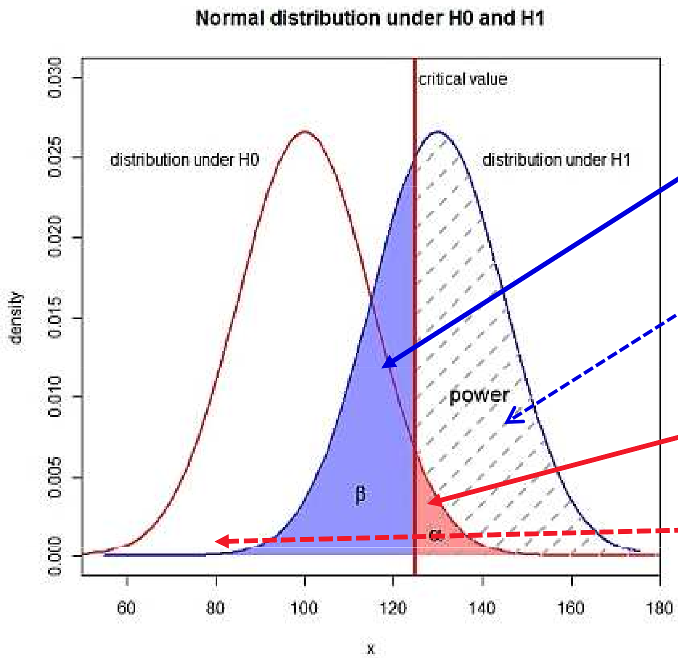
Rejecting the null hypothesis

		Actual Value (reality)	
		Positive	Negative
Conclusion from hypothesis test	Positive	Correct	
	Negative		Correct

MUNI SPORT

MaSiSS

The Research Hypotheses



Rejecting the null hypothesis

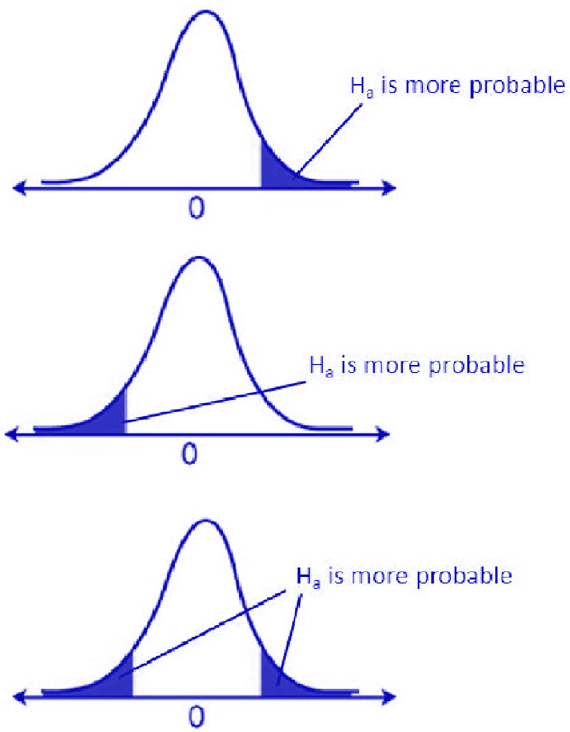
- β = probability of a Type II error,
  - known as a "false negative"
- 1-β = probability of a "true positive"
  - correctly rejecting the null hypothesis.
  - also known as the *Power* of the test.
- α = probability of a Type I error
  - known as a "false positive"
- 1-α = probability of a "true negative",
  - correctly not rejecting the null hypothesis

MUNI
SPORT

MaSiSS

The Research Hypotheses

Rejecting the null hypothesis



Right-tail test

H_a: $\mu >$ value

Left-tail test

H_a: $\mu <$ value

Two-tail test

H_a: $\mu \neq$ value

MaSiSS

Statistical Hypothesis Testing

- Null hypothesis (H_0)
  - Assumption about the outcome
    - There is no relationship between two variables (for correlation)
    - There is no difference between the means of two populations (for *t*-test)
- *p*-value
  - Propability of observing the result given that the null hypothesis is true
  - $p <= \alpha$ (0.05):
    - reject H_0, different distribution.
  - $p > \alpha$ (0.05):
    - fail to reject H_0, same distribution.

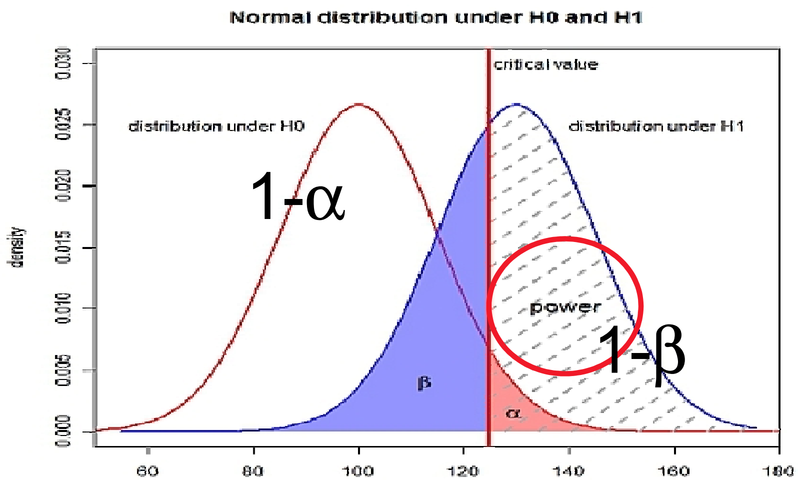
- Type I Error
  - Reject the null hypothesis when there is in fact no significant effect (false positive).
  - The p-value is optimistically small.
- Type II Error
  - Not reject the null hypothesis when there is a significant effect (false negative).
  - The p-value is pessimistically large.

MUNI
SPORT

MaSiSS

Statistical Power

- Statistical power (or the power of a hypothesis test)
  - is the probability that the test correctly rejects the null hypothesis.
- Statistical power has relevance only when the null is false.



- The higher the statistical power = the lower the probability of making a Type II error
  - (false negative)

= That is the higher the probability of detecting an effect when there is an effect

(Ellis, 2010)

MUNI SPORT

MaSiSS

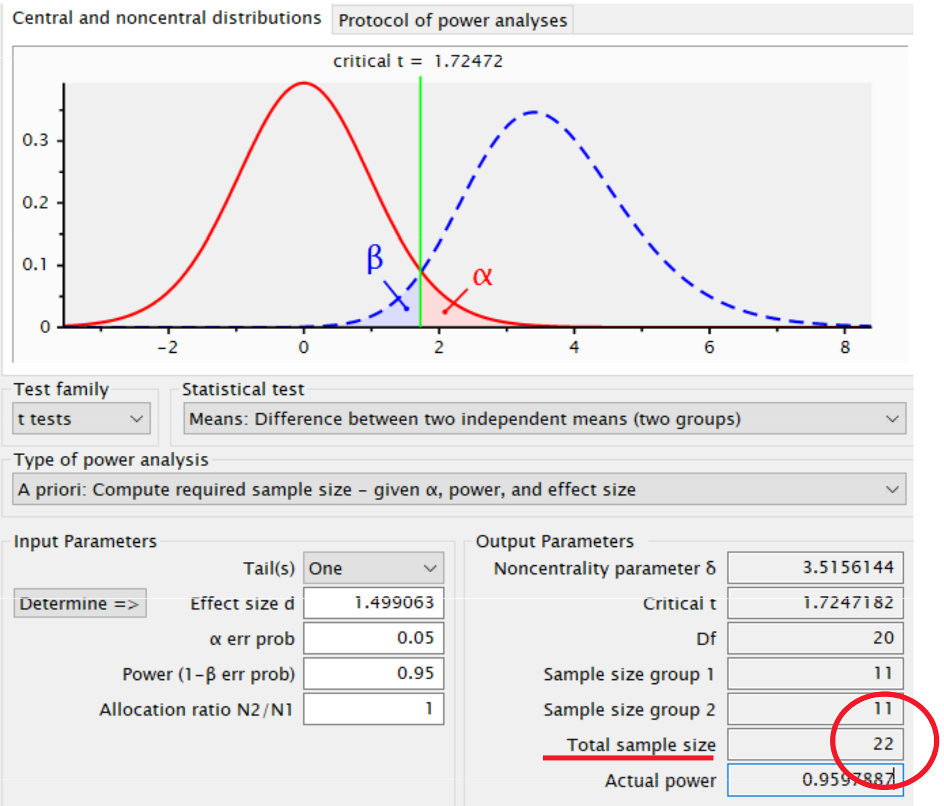
Statistical Power

- Low Statistical Power:
  - Large risk of committing Type II errors (a false negative)
- High Statistical Power:
  - Small risk of committing Type II errors
- Experimental results with too low statistical power will lead to invalid conclusions about the meaning of the results.
  - Therefore a minimum level of statistical power must be sought

- It is common to design experiments with a statistical power of 80% or better, e.g. 0.80.
  - This means a 20% probability of encountering a Type II area.
  - This different to the 5% likelihood of encountering a Type I error for the standard value for the significance level.

MaSiSS

Statistical Power



G*Power

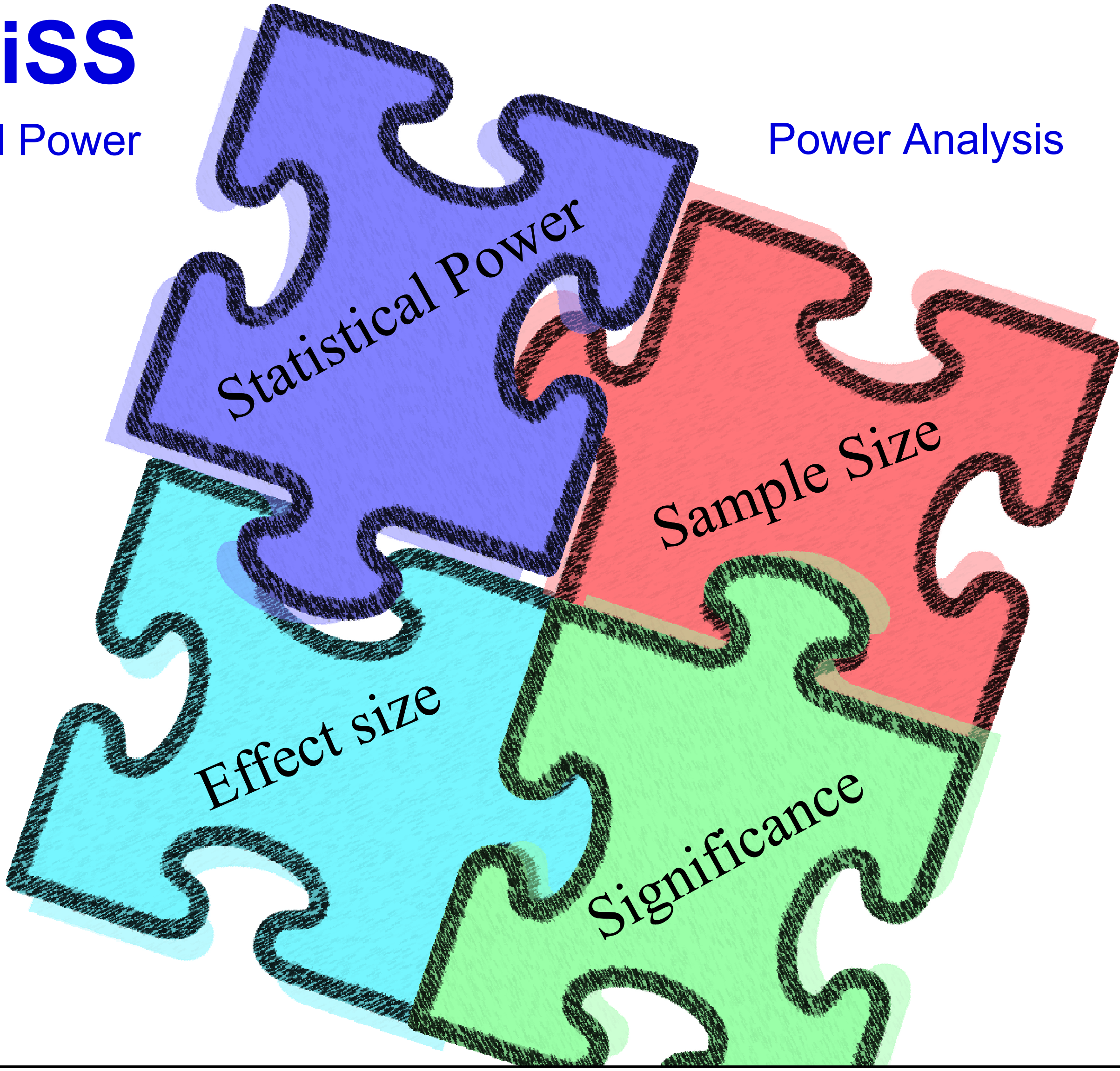
- It is common to design experiments with a statistical power of 80% or better, e.g. 0.80.
  - This means a 20% probability of encountering a Type II area.
  - This different to the 5% likelihood of encountering a Type I error for the standard value for the significance level.

MUNI SPORT

MaSiSS

Statistical Power

Power Analysis



MaSiSS

Statistical Power

- Effect Size
  - The quantified magnitude of a result present in the population
    - *Pearson's correlation coefficient* (*r*) for the relationship between variables
    - *Cohen's d* for the difference between groups
- Sample Size
  - The number of observations in the sample.

Power Analysis

- Significance
  - The significance level used in the statistical test
    - Alpha (α) = 5% or 0.05
    - Alpha (α) = 1% or 0.01
    - Alpha (α) = 0.1% or 0.001
- Statistical Power
  - The probability of accepting the alternative hypothesis if it is true.

MaSiSS

Statistical Power

Power Analysis

All variables are related:

a **larger sample** size can make an **effect easier to detect**

the statistical **power** can be increased in a test by increasing the **significance level**

MaSiSS

Statistical Power

Power Analysis

All variables are related:

The *statistical power* can be estimated given an *effect size*, *sample size* and *significance level*

The *sample size* can be estimated given different desired *levels of significance*

MaSiSS

Statistical Power

Power Analysis

Power analysis answers questions:

- How much **statistical power** does my **study have**?
- How big a **sample size** do I need?
  - Or for estimation of the minimum sample size reguired for an experiment

Post-hoc Power (Observed Power) = used as a follow-up analysis, if a finding is *non significant*

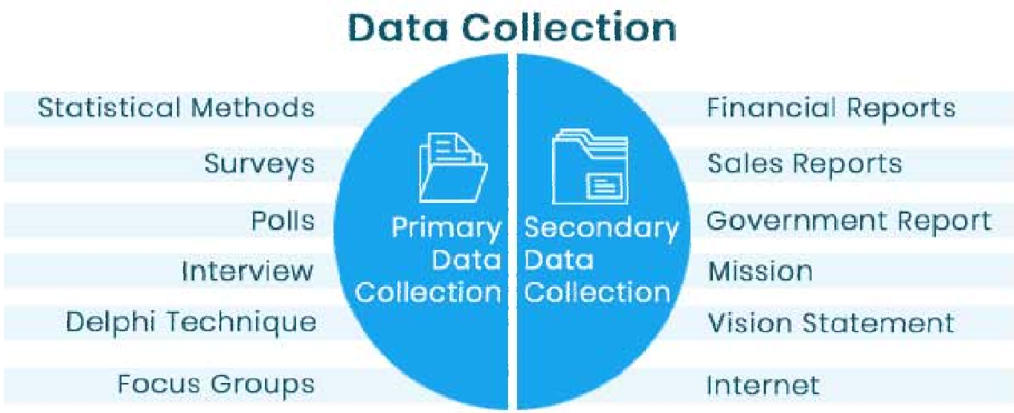
Priori Power Analysis = Power analyses made before a study is conducted

MUNI SPORT

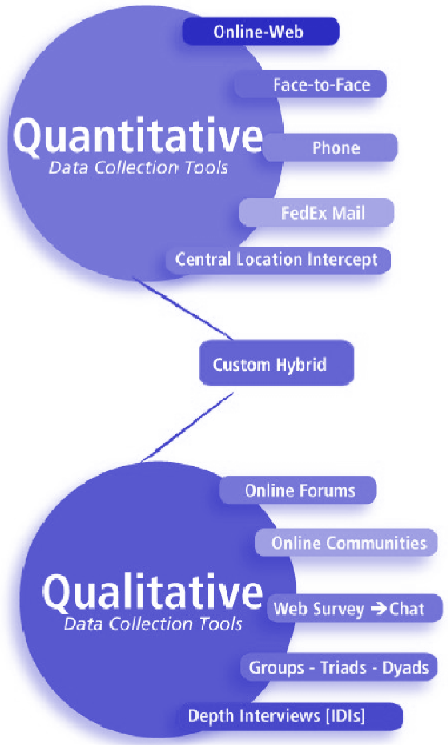
MaSiSS

Data Collection Methods

Collecting of a *Primary* and *Secondary* Data



Collecting a *Qualitative* and *Quantitative* Data



MUNI SPORT

MaSiSS

Sampling

- The study of the total population **is not possible** and it is **impracticable**.
- The practical limitation *cost*, *time*, and *other factors* which are usually operative in the situation, stand in the way of *studying the total population*
- Sampling is the process of *selecting a few* (a sample) from a bigger group (the sampling population) to become the basis for estimating or predicting the prevalence of an unknown piece of information, situation, or outcome regarding the bigger group

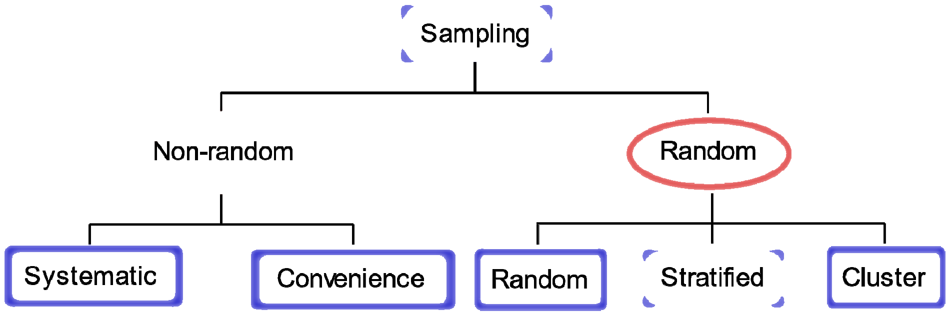
Basically we have two types of sample:

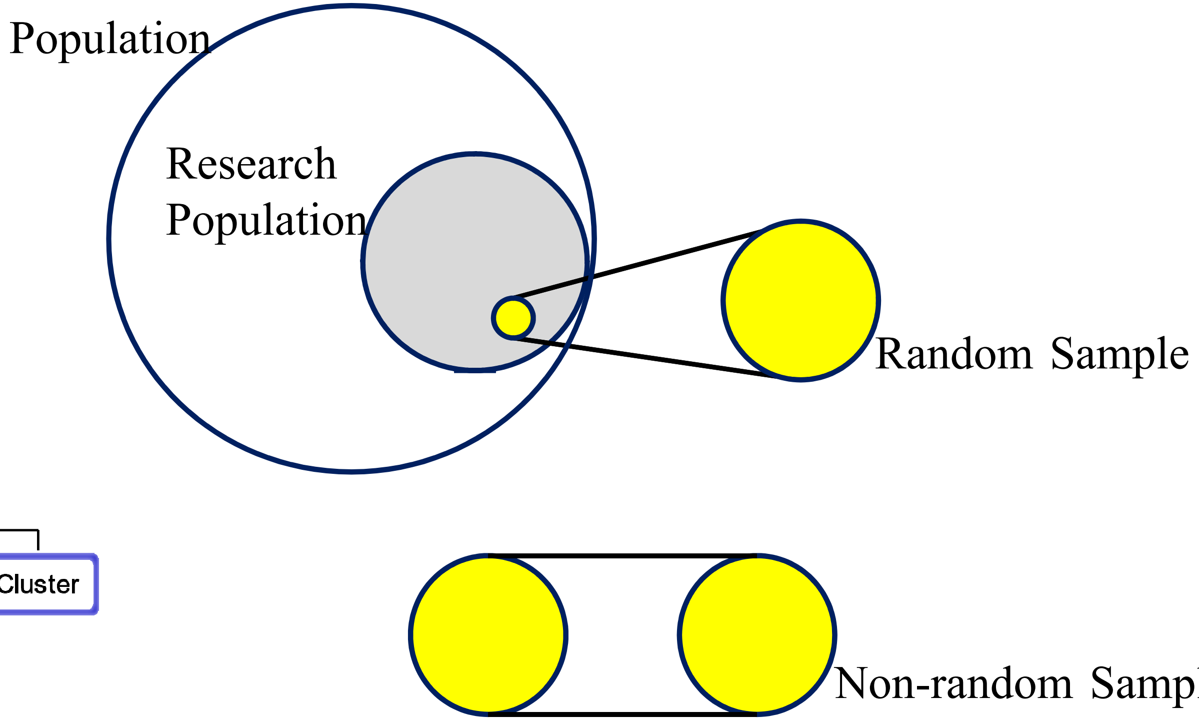
- Random sample
- Non-random sample

(Singh, 2006; Kumar, 2011)

MaSiSS

Sampling





MUNI SPORT

MaSiSS

Sampling

Characteristics of a Good Sample

- *true representative* of the population
- free from *bias*
- is an objective one
- is comprehensive in nature
- maintains *accuracy*
- it yields accurate results

(Singh, 2006)

Size of a Sample

General rule is *large sample is better* (is likely more representative of population, data are more accurate and precise with smaller standard error).

But this is not always true, because the chances of a *Type I Error* increase with the sample size (claim that something exists when it is not true).

We can reduce this risk by using the *Effect size tests* or calculate *Statistical Power* of our sample to determinate adequate sample size.

MUNI SPORT

MaSiSS

Statistical Analysis

Statistical analysis is the science of *collecting* data and uncovering *patterns* and *trends*

After collecting data you can analyse it to:

1. Summarize the data (make a pie chart)
   1. *By Measures of Distribution or frequency*
2. Find key measures of location (median, mean, …)
   1. *By Measures of Central Tendency*
3. Calculate measures of spread – find if your data are tightly or spread out cluster (R, SD, IQR)
   1. *By Measures of Variation and Position*
4. Make future prediction based on past behavior
5. Test an experiment's hypothesis (hypothesis testing)
   1. $p < 0.05$ = statistically significant results, we reject H_0 (because valid H_0 occurred in less than 5% of cases)
   2. $p > 0.05$ = statistically not significant results, we fail to reject H_0 (because valid H_0 occurred in more than 5% of cases)

MUNI SPORT

Research Methodology and Statistics

Statistical Analysis

Statistical analysis is the science of collecting data and uncovering patterns and trends

note: you <u>cannot</u> accept H_0 (only **reject** or **fail to reject**) or H_a …

… after rejecting H_0, we can add „we can accept H_a

MaSiSS

Characteristics of a Good Test

Reliability is consistency, dependence or trust

measurement reliability is the consistency with which a test yields the same result in measuring whatever it does measure

Example 1: If after first test mean score is 80, and after one week we used same test on same sample and the mean score is:

- 80 = test provided stable and dependent results
- 102 = test results are not consistent

Example 2: how two different teachers will evaluate the same test results

Reliability, Validity, Objectivity and Usability

Types of reliability testing methods:

- Test-retest method (the same test over time)
- Interrater method (the same test conducted by different people)
- Parallel forms method (different version of a test which are designed to be equivalent)
- Split-half method (same test divides into two equivalent values)
- Internal consistency method (correlation between multiple items in a test that are intended to measure the same construct)

Methods of determining reliability

- Intraclass Correlation Coefficient (ICC) - defines a measure's ability to discriminate among subjects
- Standard Error of Measurement (SEM) - quantifies error in the same units as the original measurement

(Gronlund and Linn, 1995; Ebel and Frisbie, 1991)

MUNI SPORT

MaSiSS

Characteristics of a Good Test

Reliability, Validity, Objectivity and Usability

Factors that's Affecting Reliabitity

- Length of the test
- Content of the test
- Spread of Scores
- Heterogeneity of the group
- Experience with the test
- Motivation
- testing procedure
- time limit of test
- Cheating opportunity

How Higher Should Reliability be?

Cronbach's Coefficient Alpha (α)	Reliability
0.80 to 0.95	Very good
0.70 to 0.80	Good
0.60 to 0.70	Fair
< 0.60	Poor

Note: what's wrong with this interval?

MaSiSS

Characteristics of a Good Test

- It means to what extent the **test measures that**, what the test maker **intends to measure**
  - means *truthfulness* of a test
- Validity do not have different types. It is a unitary concept, based on various types of evidence

Reliability, **Validity**, Objectivity and Usability

Factors that affecting validity

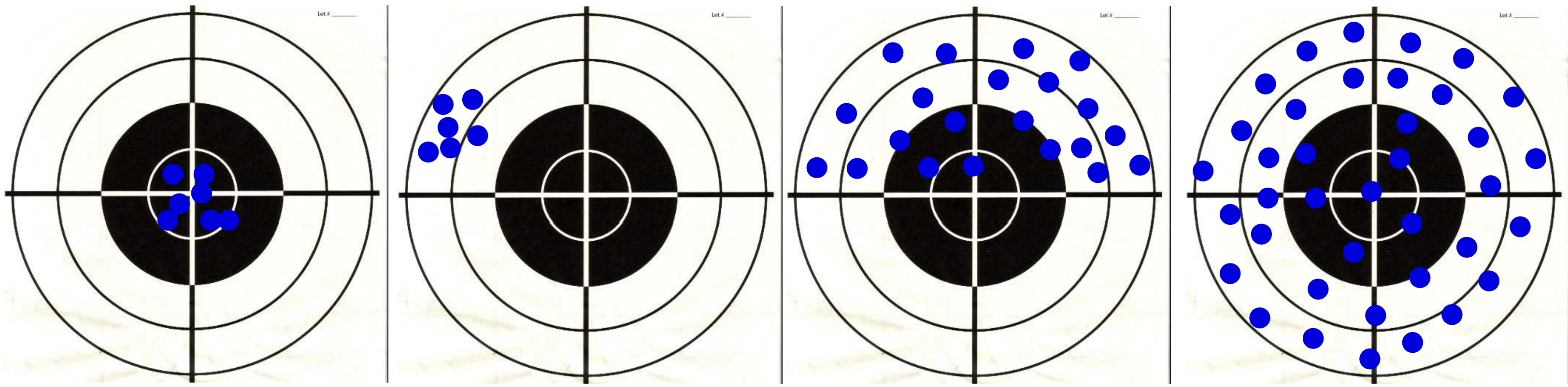
- Unclear directions to the respond of the test
- *Difficulty* of the reading vocabulary and sentence structure
- Too *easy* or too *difficult* test items
- *Ambiguous* statements in the test items
- Inappropriate test items for measuring a particular outcome
- Inadequate *time* provided to take the test
- *Length* of the test
- Unfair aid to individual students (asking for help)
- *Cheating* during testing
- Unreliable scoring
- *Anxiety*, *Physical* or *Psychological* state of the pupil
- Response set (pattern in responding)

MaSiSS

Characteristics of a Good Test

Reliability, **Validity**, Objectivity and Usability

- It means to what extent the **test measures that**, what the test maker **intends to measure**
  - means *truthfulness* of a test
- Validity do not have different types. It is a unitary concept, based on various types of evidence



Both Reliable and Valid | Reliable but not Valid | Neither reliable or valid | Not Reliable but Valid

MUNI SPORT

MaSiSS

Characteristics of a Good Test

Reliability, Validity, **Objectivity** and Usability

The extent to which the instrument is *free from personal error* (personal bias), that is subjectivity on the part of the scorer (Good, 1973)

a test is considered objective when it makes for the *elimination* of the *scorer's personal opinion* and *bias judgement*

It affects both validity and reliability of test scores

MaSiSS

Characteristics of a Good Test

Reliability, Validity, **Objectivity** and Usability

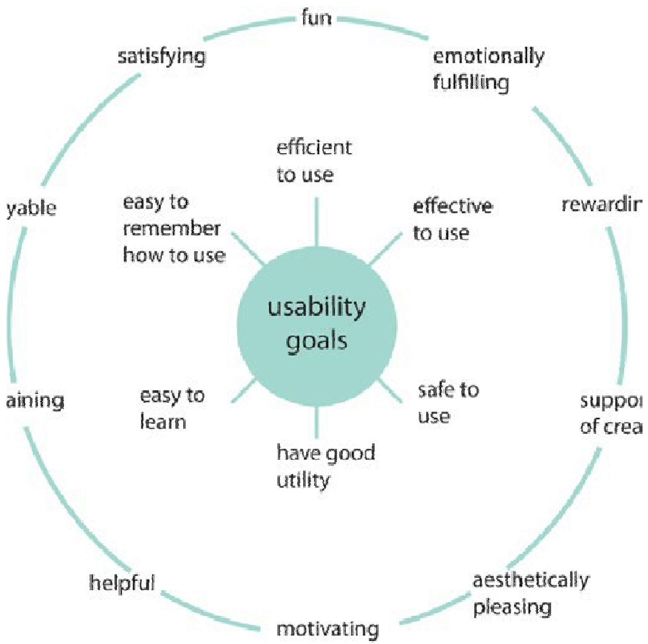
While constructing a test, two main aspects of objectivity you need to keep in mind

1. *Objectivity in scoring* - same person or different persons scoring the test at any time arrives at the same result without may chance error (personal individual judgement should not affect the test scores), The scoring procedures must be clearly defined (without doubt and ambiguity)
2. *Objectivity of test items* - the item must call for a definite single answer
    1. free from ambiguity and dual meaning sentences (it makes the test subjective)

MaSiSS

Characteristics of a Good Test

- The test must have *practical value* from time, economy, and administration point of view
- Practical considerations *cannot be neglected*



Reliability, Validity, Objectivity and <u>**Usability**</u>

While constructing or selecting a test you must be taken into account:

- *Ease of Administration* – any trained person can use it and evaluated
- *Time required for administration* – Appropriate time limit for test (20–60 minute)
- *Ease of Interpretation and Application* – If the results are misinterpreted, it is harmful and not applied results are useless
- *Availability of Equivalent Forms* – You should have available equivalent forms of the same test in terms of content, level of difficulty and other characteristics
- *Cost of Testing* - A test should be economical from preparation, administration and scoring point of view

MaSiSS

Variable Types and Scaling

Research variables are things you *measure*, *manipulate* and *control* in statistics and research

– Person, place, thing, idea, …

Other variables: Intervening (mediator) variables, Moderating (moderator) variables, Extraneous variables, Quantitative (numerical) variables, Qualitative (categorical) variables, Composite variables

Research Variables

The most common types of variable

- *Independent variable* (IV) - is a singular characteristic that the other variables in your experiment cannot change, but IV can change other variables
  - *Age* (eating or exercise habits are not changing your biological age, but you will not lift same weight as senior)
- *Dependent variable* (DV) - relies on and can be changed by other components, IV can influence DV, DV can't influence IV. Researchers goals are **determine what makes** the dependent variable change and **how**.
  - *A grade on exam* (it depends on factors as how much your slept or how long you studied, but your test does not affect the time you spent studying)
- *Control (controlling)* variables - are constant and do not change during a study, they have no effect on other variables.
  - If we are investigating how much your slept (IV) effect a grade on exam (DV), we need to control *time spent learning*, the same *level of students* and more

MUNI SPORT

MaSiSS

Variable Types and Scaling

Independent variable → Dependent variable

Independent variable (IV)

Also called:

- Exposure variable
- Control variable
- Explanatory variable
- Manipulated variable

Dependent variable (DV)

Also called:

- Outcome variable
- Controlled variable
- Explained variable
- Response variable

Research Variables

The most common types of variable

- *Independent variable* (IV) - is a singular characteristic that the other variables in your experiment cannot change, but IV can change other variables
  - *Age* (eating or exercise habits are not changing your biological age, but you will not lift same weight as senior)
  - In *t*-test/ANOVA IV = categorical variable (nomina/ordinal)
- *Dependent variable* (DV) - relies on and can be changed by other components, IV can influence DV, DV can't influence IV. Researchers goals are **determine what makes** the dependent variable change and **how**.
  - *A grade on exam* (it depends on factors as how much your slept or how long you studied, but your test does not affect the time you spent studying)
  - In *t*-test/ANOVA DV = continuous measurement (interval/ratio)

MaSiSS

Variable Types and Scaling

Research Variables

_t_-test (independent, one or two sampes)

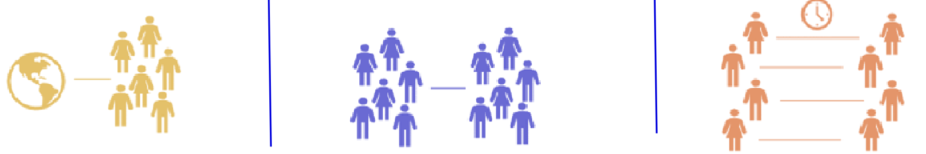
1-2 IV groups

IV (1x)	DV (1x)
Independent variable	Dependent variable
Group A/B	Value

1 IV group
One sample t-test

2 IV groups
Independent samples t-test

2 IV groups
Paired samples t-test

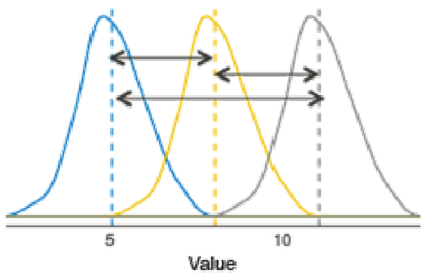


One-Way **ANOVA** (Analysis of Variance)

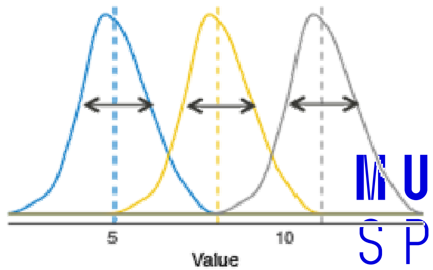
3+ IV groups

IV (1x)	DV (1x)
Independent variable *(Factor)*	Dependent variable *(Response Variable)*
Studying Technique A/B/C	Exam Score

Between-group variation
(Diff amoug group means)



Within-group variation
(Variability within each group)



MUNI SPORT

MaSiSS

Variable Types and Scaling

Research Variables

ANCOVA (Analysis of Covariance)

IV (2x) | DV (1x)

Studying Technique → Exam Score

Covariates

Current Grade → Exam Score

One-Way **ANOVA** (Analysis of Variance)

IV (1x) | DV (1x)

Independent variable *(Factor)* → Dependent variable *(Response Variable)*

3+ IV groups

Studying Technique A/B/C → Exam Score

One-Way **MANOVA** (Multivariate Analysis of Variance)

IV (1x) | DV (2x)

Level of Education → Annual Income

Level of Education → Student Loan Debt

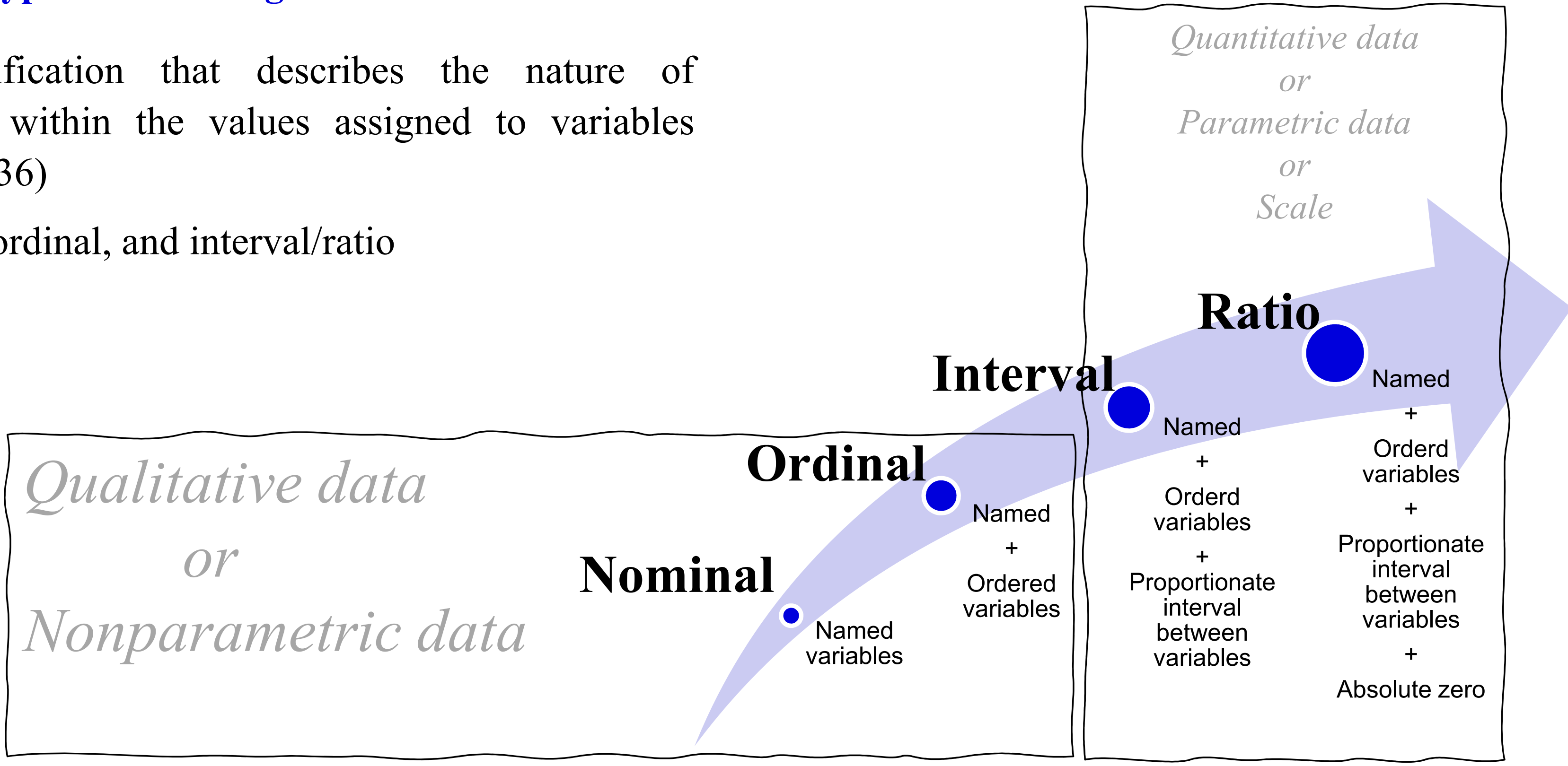
MaSiSS

Variable Types and Scaling

is a classification that describes the nature of information within the values assigned to variables (Stevens, 1936)

– nominal, ordinal, and interval/ratio

Levels of Measurement



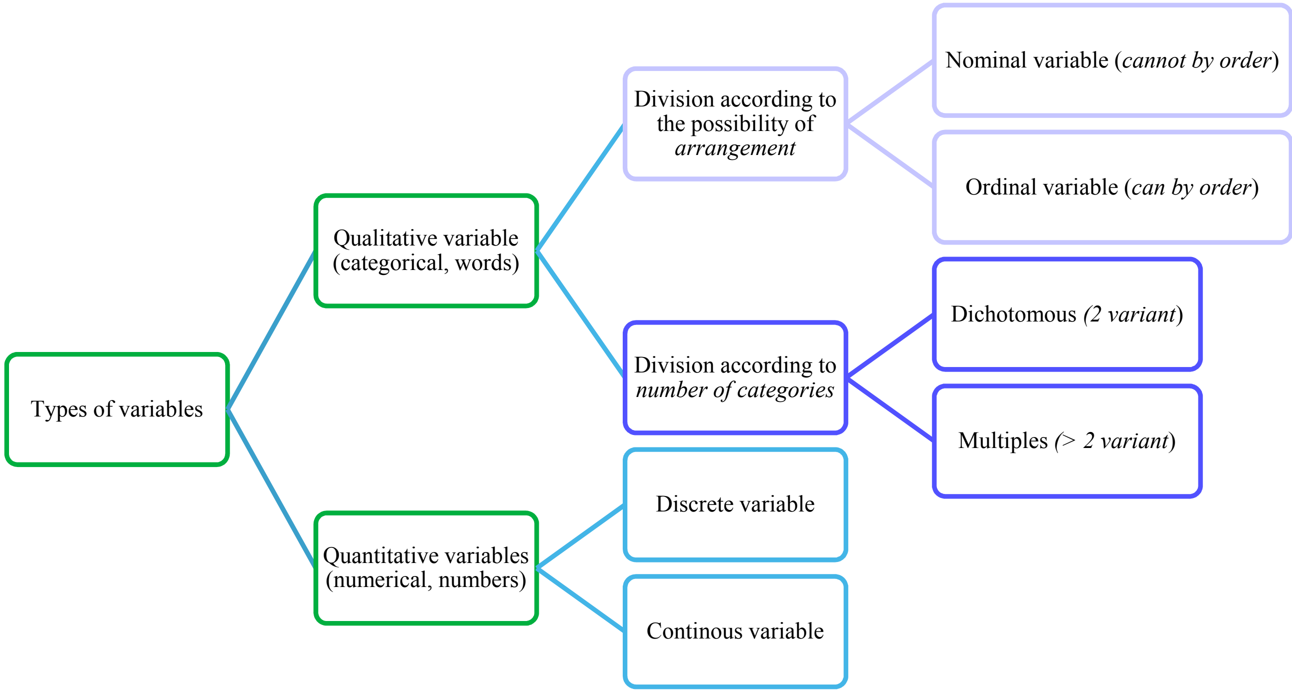
MUNI SPORT

MaSiSS

Variable Types and Scaling

Levels of Measurement

Levels of Measurement (Scale of Measure) are *Nominal, Ordinal, Interval, Ratio*



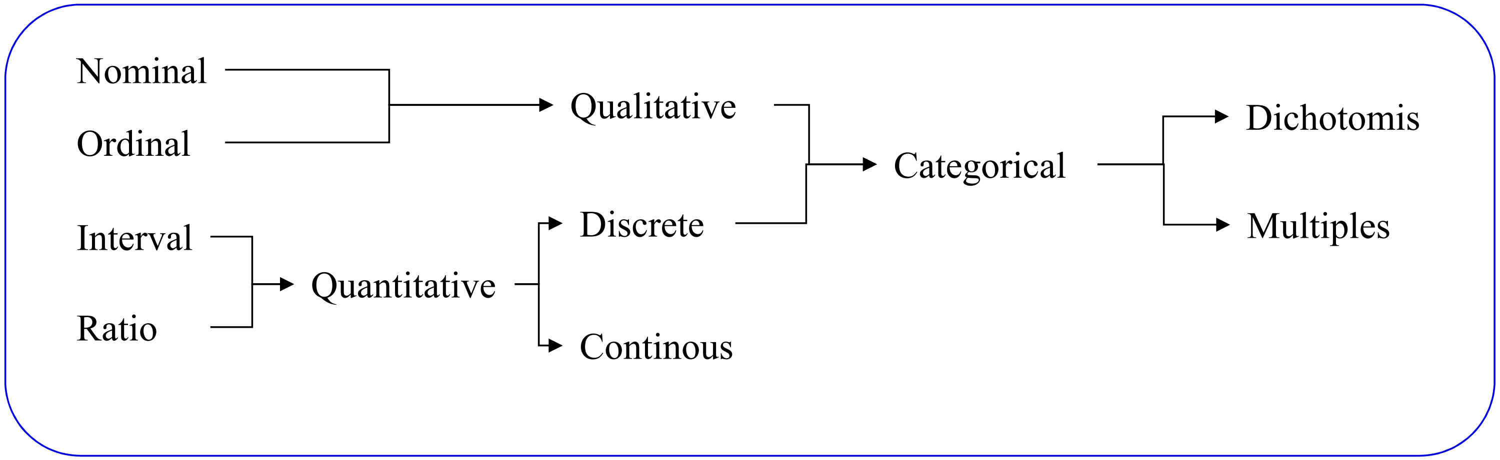
(Litschmannová, 2011)

MUNI SPORT

MaSiSS

Variable Types and Scaling **Levels of Measurement**

Levels of Measurement (Scale of Measure) are *Nominal, Ordinal, Interval, Ratio*



(Řezanková et al., 2017)

MUNI SPORT

MaSiSS

Variable Types and Scaling

Levels of Measurement

Levels of Measurement (Scale of Measure) are *Nominal, Ordinal, Interval, Ratio*

- **Nominal data**:
  - The number of female athletes in football association
  - Your political party affiliation
  - The state/region/city where you were born
  - The color of your hair
- **Ordinal data**:
  - Order of finish in race/contest/tournament
  - A school grades (A, B, C, D, F)
  - Ranking of chilli peppers (hot, hotter, hottest; not Scoville scale, SHU)
  - Student's year of study (freshman, Sophomore, junior, senior)
  - Cancer stage (I, II, III, IV)
- **Interval data**:
  - Intelligence Quotient scores
  - Dates on calendar
  - The heights of waves in the ocean
  - Shoe size
  - Longitudes on map/globe
- **Ratio data**:
  - Height
  - Pulse
  - length
  - Money in your bank/wallet/pocket
  - Monthly Income/expenses

MaSiSS

Misuse of statistics

- Target (2002): Can we determine which customers are pregnant?
  - *Even if they don't want us to wnow.*

- Andrew Pole:
  - yes + also they expectd due date is

= Target send right coupons at the right time

MUNI SPORT

MaSiSS

Misuse of statistics

- **<u>Elderly woman was robbed</u>** (1964):
  - She saw: *Blonde woman, ponytail*
  - a passing man saw: *yellow car driven by a black man (had beard and a mustache)*
  - Police catche: Janet and Malcom Collins

    They matched all the descriptions

Prove of guilt for Collins by mathematician

- Mathematician calculated the <u>probability</u> of just randomly selecting a couple that was inocent and also share all charakteristice
  - 1 in 12 million chance that Collins are innocent

MaSiSS

Misuse of statistics

- **Sally Clark**: *guilty of murdering* her two infant son‘s in 90s.
  - 1st son: died suddenly due to *unknown causes*
  - 2nd son: found dead 8 week after birth of suddent *unknown causes*
  - During the trial a pediatrician professor: chance of two <u>suddent unknown causes</u> is **1 in 73 million**

MaSiSS

Misuse of statistics

Elderly woman was robbed (1964)

- People v. Collins
  - Example of **Prosecutor‘s fallacy**
  - $P(A/B) \neq P(B/A)$

MUNI SPORT

MaSiSS

Misuse of statistics

Given:

- Behind curtain is an animal with 4 leg

What is the probability that it's a dog

- **1 in 1000**

Given:

- Behind curtain is a dog

What is the probability that it has 4 legs?

- **Almost 100%**

By switching the given and the question = change of probability

MUNI SPORT

MaSiSS

Misuse of statistics

Elderly woman was robbed (1964)

– People v. Collins
  – Example of Prosecutor‘s fallacy
  – $P(A/B) \neq P(B/A)$

Given:

  – Couple fit all description (blonde woman,…).

– < 1 in 12 million chance of innocence

WRONG

MaSiSS

Misuse of statistics

Given:
- Innocent couple
- < 1 in 12 million
- Probability of fiffing all descriptions

Right
- Example: random couple (of a mall) had very small chance of fitting the description

People v. Collins

Given:
- Couple fit all description (blonde woman,…).
- < 1 in 12 million chance of innocence

WRONG

(eg. Prosecutors fallacy)

MaSiSS

Misuse of statistics

Sally Clark - guilty of murdering her two infant son's

Bacterial test reveal *more specific information* that a simple *multiplication of two probabilities*

They were independent of each other (genetic and enviromental factor)

Lost childrens (due natural caused), accused of murdering, found guilty (due to a misuse of statistic), spent 3 years in prison, after release not able to recovery and died (alcohol poisoning)

MaSiSS

Misuse of statistics

If high school dropout rates go from 5% to 10%

5% increase

5% → 10%

100% increase

MUNI
SPORT

MaSiSS

Misuse of statistics

If dropout rates are 1 in a million people and then next year go to 2 in million people

- Headlines 1: Dropout rates go up by 100%
- Heallines 2: Dropout rates go up by .0001%

100% increase

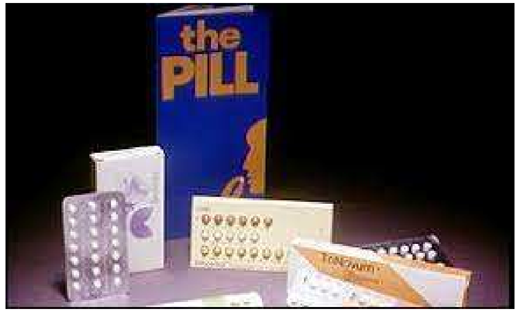
1 in 1 mil → 2 in 1 mil

– .0001%   + .001%   – .0002%

MUNI SPORT

MaSiSS

Misuse of statistics



UK committee on sefty of medicines *(1995)*

- *certain type of birth control pill increaased the risk of life-threating blood clots by 100%*

Older pill

1 in 7000

– ≈.014%

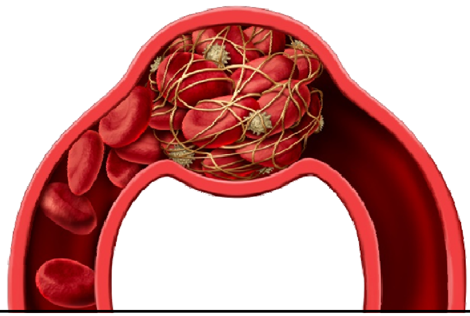
100% increase

+ .014%

New pill

2 in 7000

– ≈.028%



MaSiSS

Misuse of statistics

Correlation or Causation?

- Or both?

- If A *correlated* with B, it don‘t mean that A *cause* B

- Rotating turbines correlated with wind
  - but tubine don‘t cause wind, wind cause rotation of turbine

MUNI SPORT

MaSiSS

Misuse of statistics

Correlation or Causation?

- Or both?

- Violent show cause kids to be more violent?
  - it is posible but …
- More violent kids watch more violent TV shows
  - Also seens reasonable

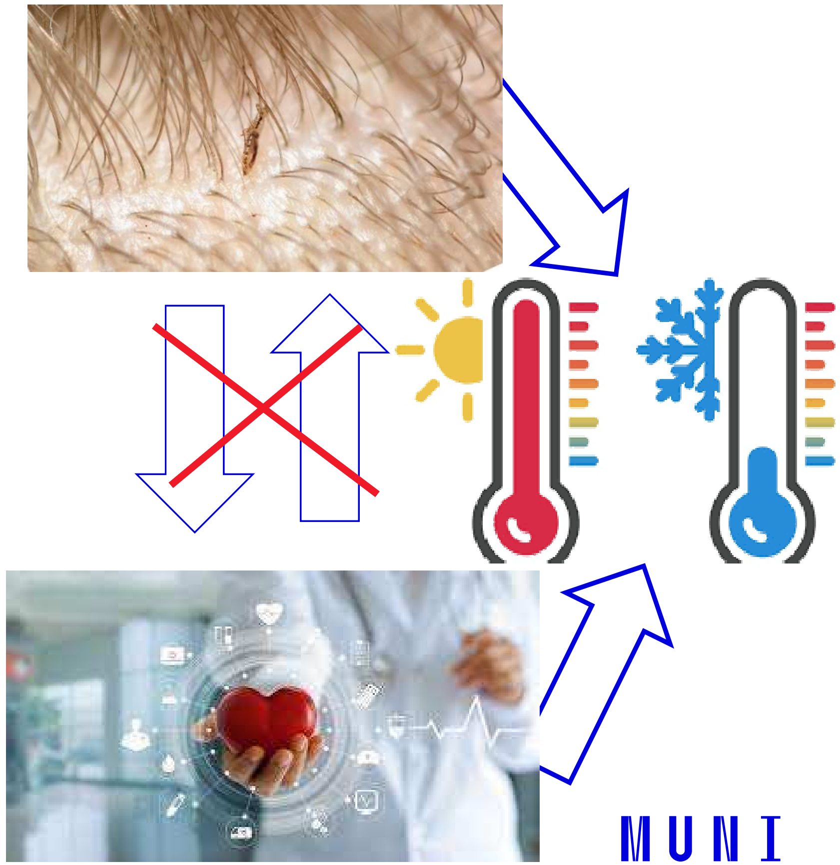
MaSiSS

Misuse of statistics

Correlation or Causation?

- Or both?

- Peole how had head lice were healthy and sick people were ralely even had head lice
  - Head lice *cause* bettet health
  - But head lice is very *sentive* to the body temperature



MUNI SPORT

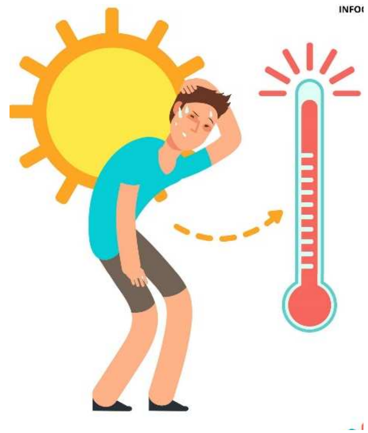
MaSiSS

Misuse of statistics

Correlation or Causation?

- Or both?

- Ice cream sales do not cause increase in heat strokes or other way, even though are correlated
  - Hot weather cause both (eg. *Third-Cause Fallacy*)



MaSiSS

Misuse of statistics

Correlation or Causation?

– Or both?

– CO_2 production increased along with obesity. Does one cause the other?

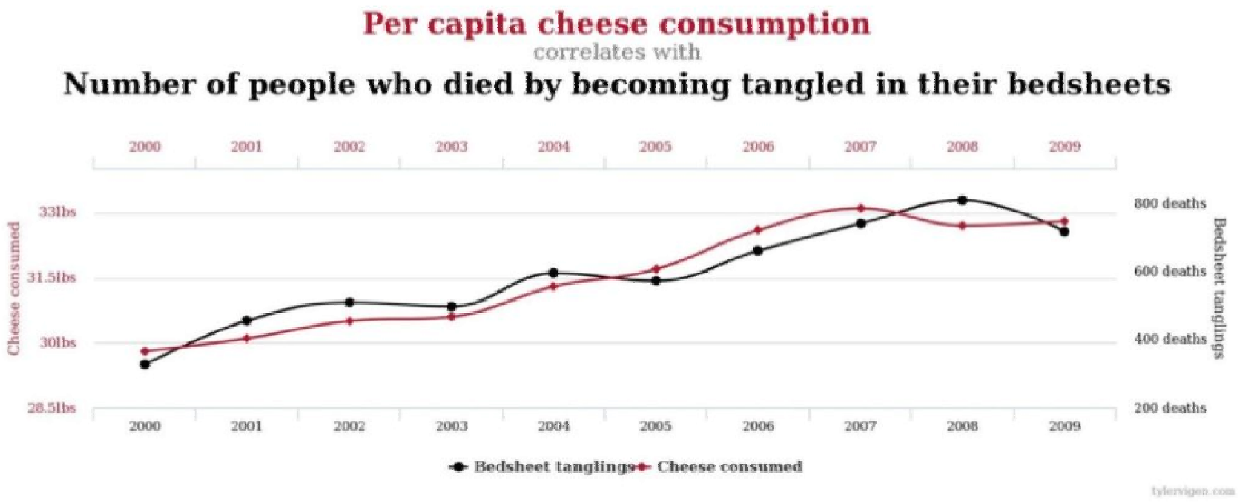
– *No, wealthy population eat more and produce more CO2*
  (eg. Third-Cause Fallacy)

MaSiSS

Misuse of statistics

Correlation or Causation?

- Or neither?



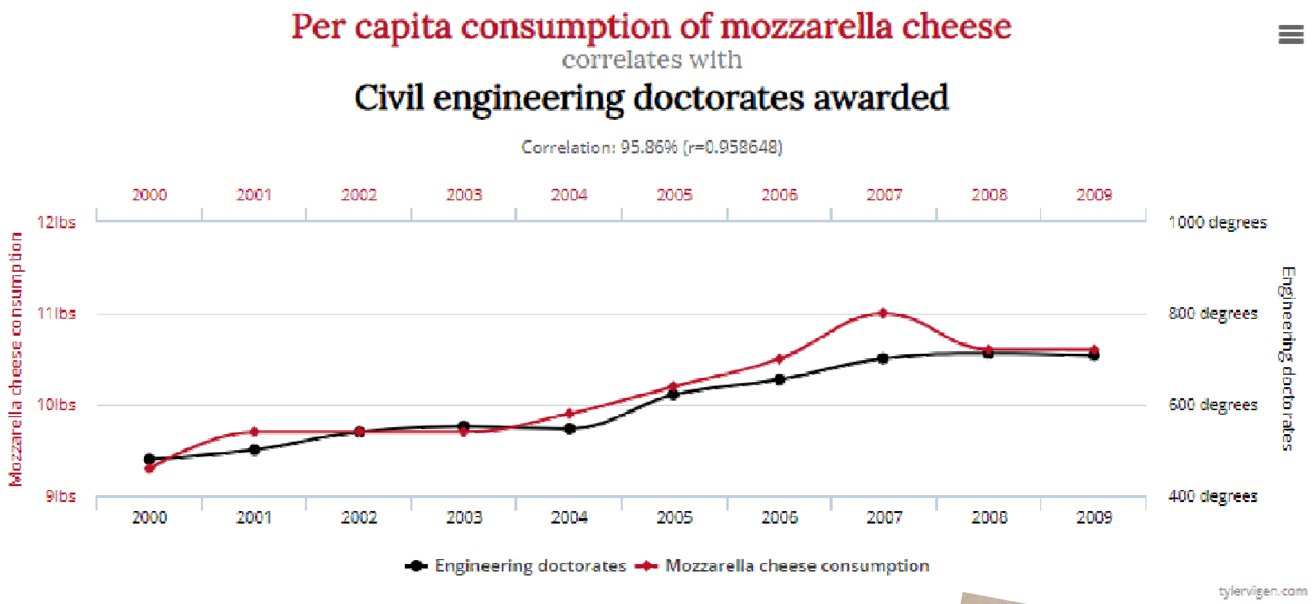
- Cheese consumption per person and the number of people who died as a result of tangling in the sheets
  (r = 0.947)

MaSiSS

Misuse of statistics

Correlation or Causation?

– Or neither?



– Consumption of mozzarella cheese per person with doctorates from engineering
  (r = 0.950)

Both examples are an example of a false correlation

MUNI
SPORT

MaSiSS

Descriptive Statistic

Presenting, organizing, simplifying, and summarizing data

1st step

Measures of Distribution or Frequency

Measures of Central Tendency

Measures of Dispersion or Variation

Measures of Position

Inferential Statistic

Drawing conclusions about a population base on data observed in a sample

2nd step

Comparison tests

Correlation tests

Regression tests

MaSiSS

Descriptive Statistic

1. Measures of Distribution or Frequency
2. Measures of Central Tendency
3. Measures of Dispersion or Variation
4. Measures of Position

MaSiSS

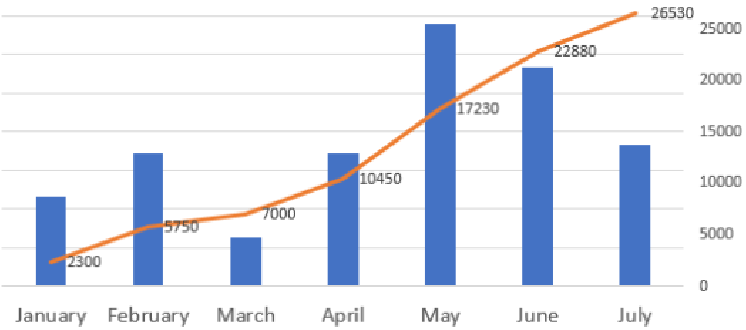
Descriptive Statistic

Measures of Distribution or Frequency

- a graph or data set organized to show the frequency of occurrence of each possible outcome of a repeatable event observed many times
- compare one part of the distribution to another part of the distribution
- *Count, Percent, Frequency*

My Classmate's Favorite Colors		
Color Choices	**Tally Marks**	**Frequency**
Red		4
Blue		7
Yellow		5
Orange		2

Weight (Kg)	Frequency	Cumulative Frequency
0 up to 20	2	2
20 up to 40	7	9
40 up to 60	12	21
60 up to 80	6	27
80 up to 100	3	30



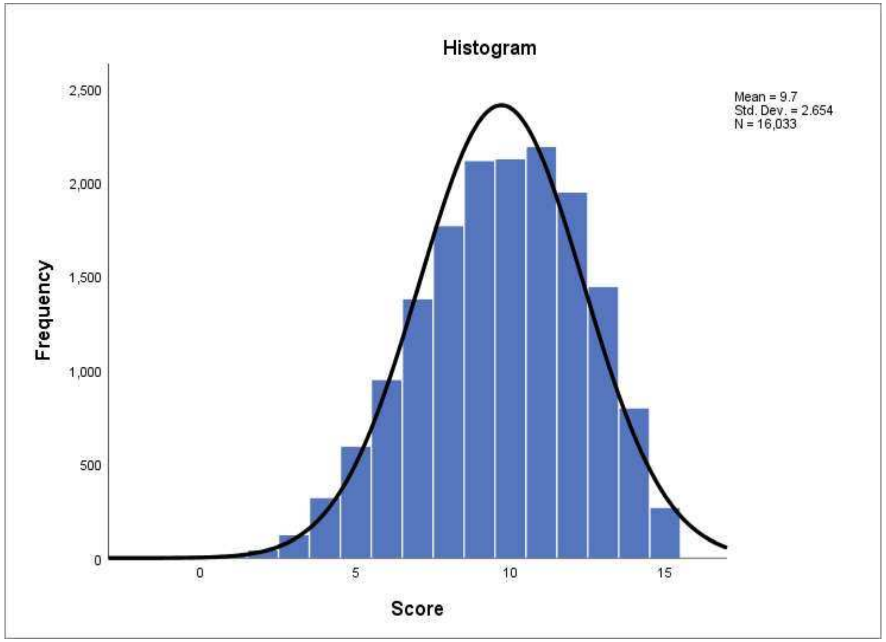
MUNI SPORT

MaSiSS

Descriptive Statistic

Measures of Distribution or Frequency

- a graph or data set organized to show the frequency of occurrence of each possible outcome of a repeatable event observed many times
- compare one part of the distribution to another part of the distribution
- *Count, Percent, Frequency*



MUNI
SPORT

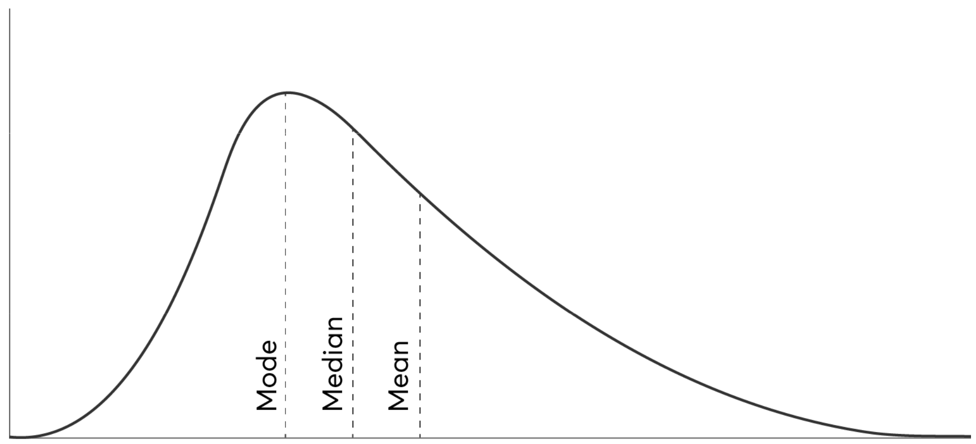
MaSiSS

Descriptive Statistic

Measures of Central Tendency

is defined as the number used to represent the center or middle of a set of data values

- *Mean, Median, Mode*

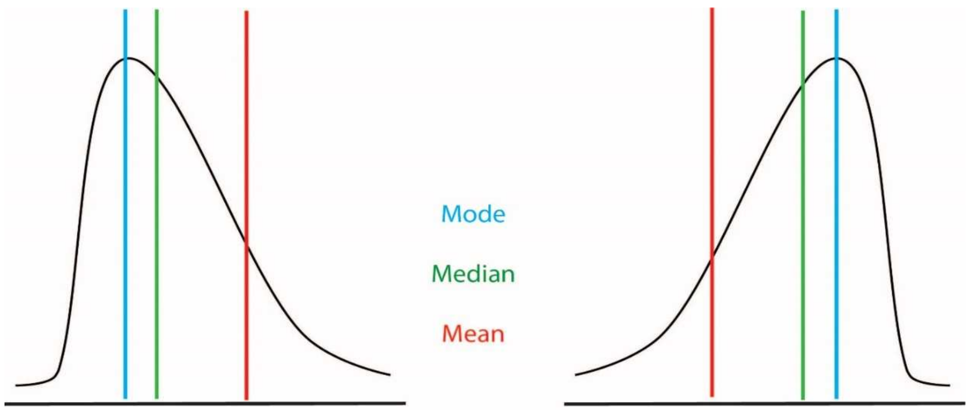


MUNI SPORT

MaSiSS

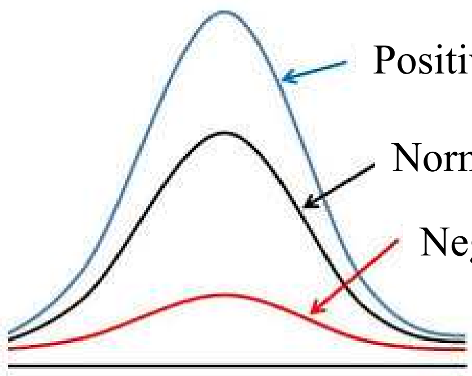
Descriptive Statistic

Measures of Central Tendency



Right or Positive **skew**

Left or Negative **skew**



MUNI SPORT

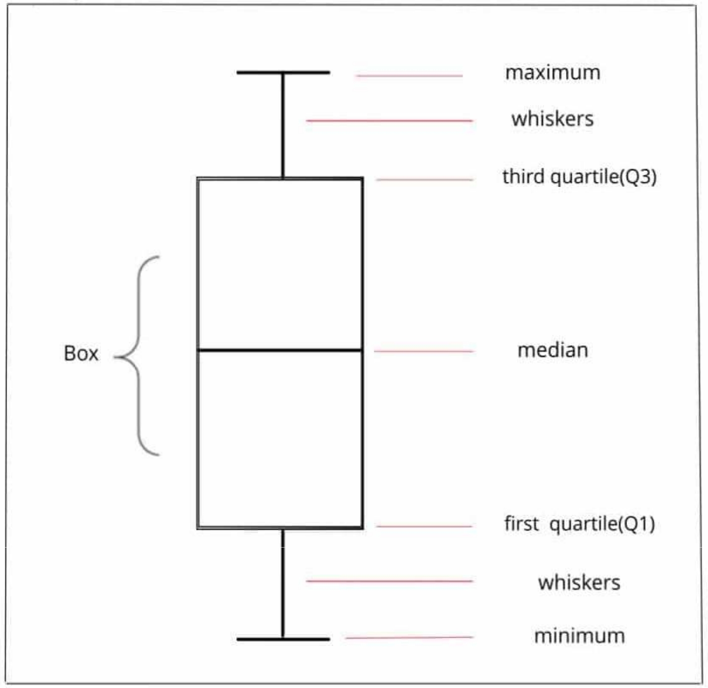
MaSiSS

Descriptive Statistic

Measures of Dispersion or Variation

how far apart data points lie from each other and from the center of a distribution (mean±SD)

- **Range, IQR, Standard Deviation, Variance**
  - *Range* (R) = $x_{Max} - x_{min}$
  - *Interquartile range (IQR)* = $Q_3 - Q_1$
  - *Standard deviation* (SD, s) is the average amount of variability in your dataset
    - s = sample standart diviation
    - σ = population standard deviation
  - *Variance* (s^2) is the average of squared deviations from the mean
    - To get variance, square the standard deviation (s)
    - s^2 = sample variance
    - σ^2 = population variance

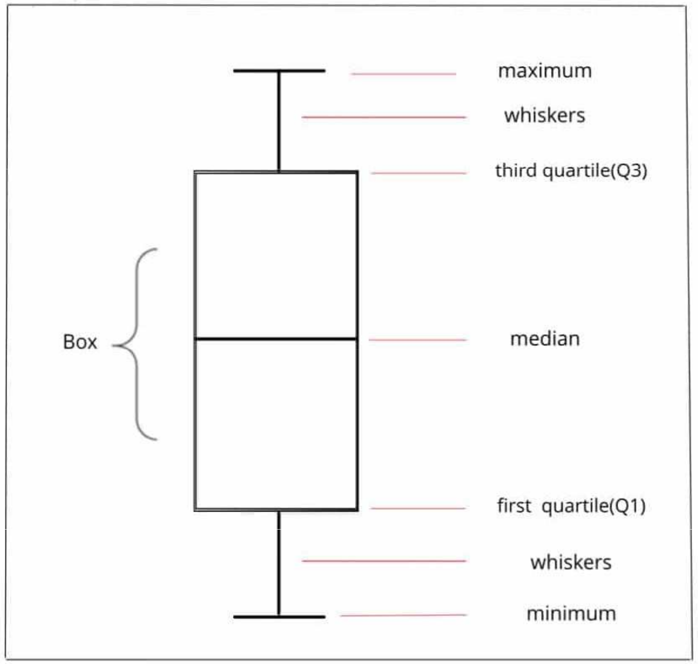


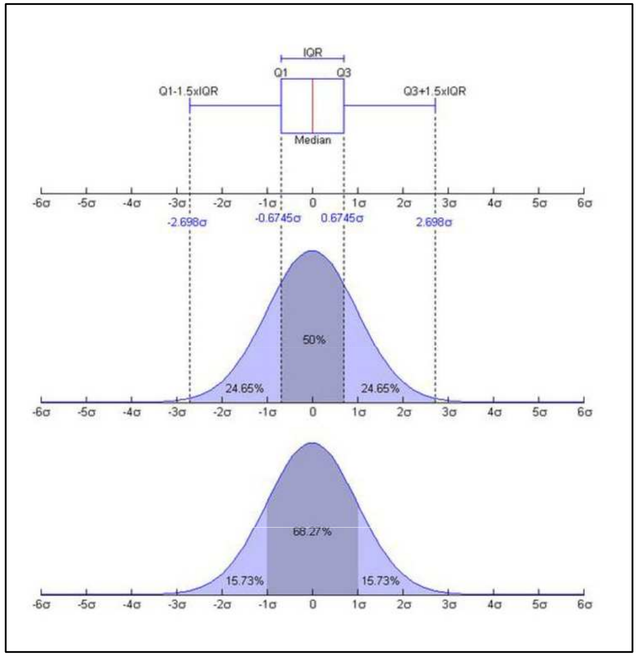
MUNI
SPORT

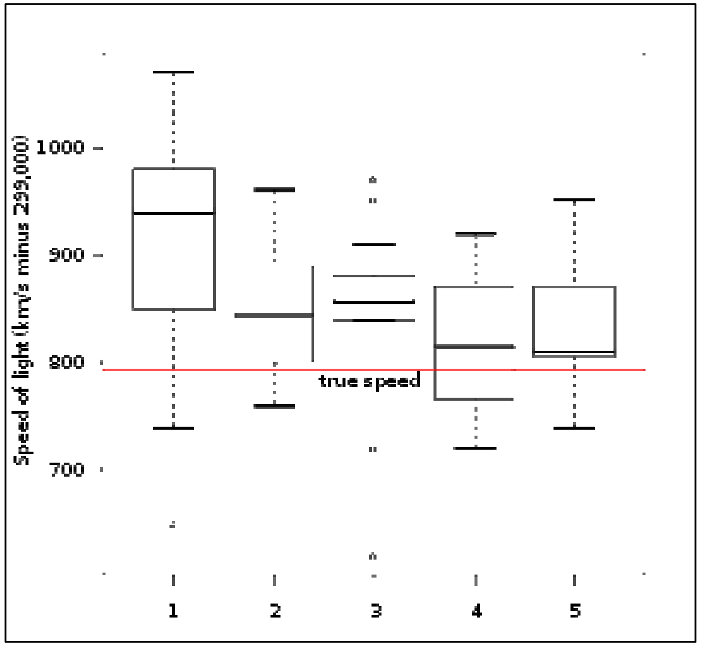
MaSiSS

Descriptive Statistic

Measures of Dispersion or Variation





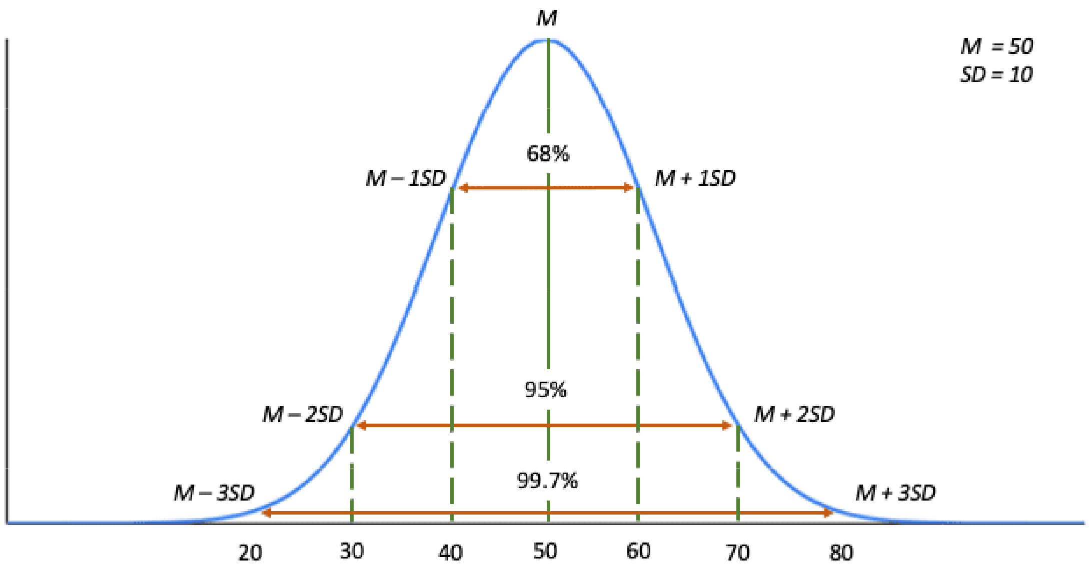


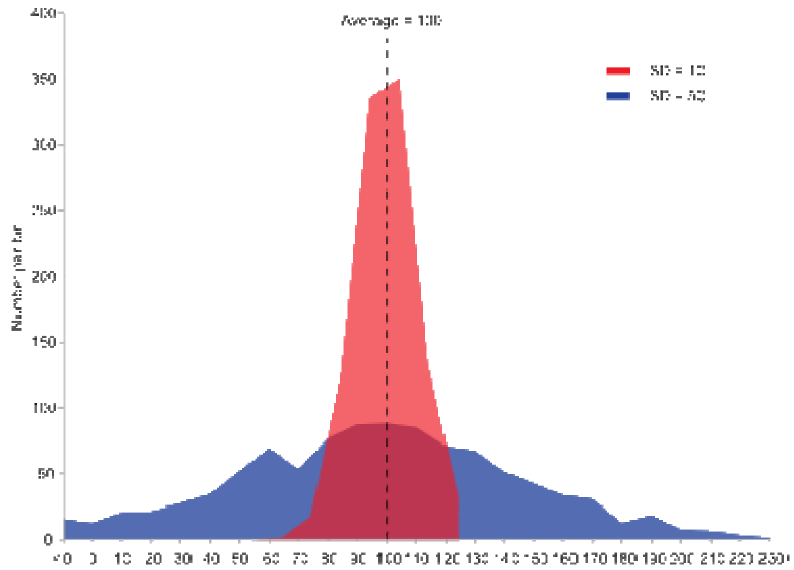
MUNI
SPORT

MaSiSS

Descriptive Statistic

Measures of Dispersion or Variation





MUNI SPORT

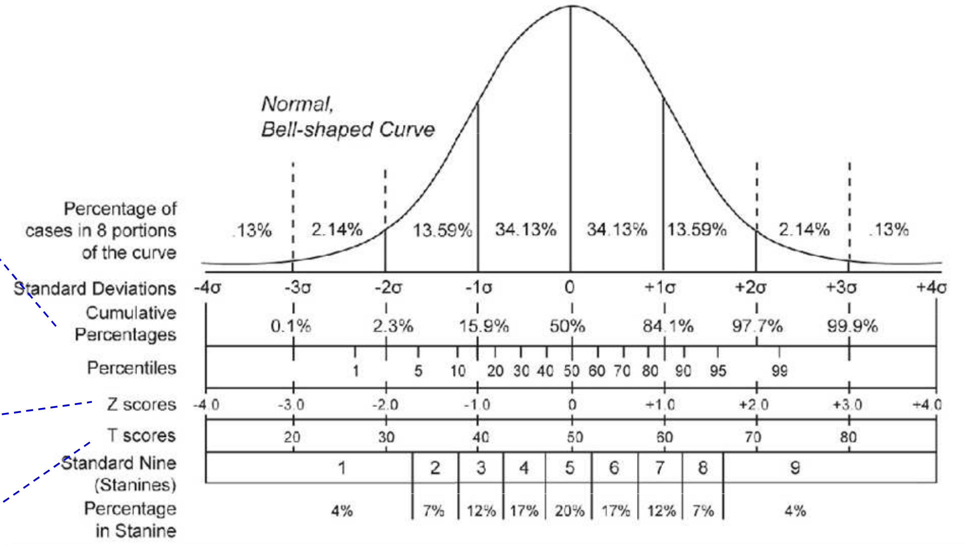
MaSiSS

Descriptive Statistic

2.3.4 Measures of Position

Determines the position of a single value in relation to other values in a sample or a population data set.

- **Quantiles (Percentile, Decile, Quartile), Outliers, standard scores**
  - *Percentile* (P_i)= divide a rank-ordered data set (from the smallest to the largest) into 100 equal parts
  - *Decile* - divide a rank-ordered data set into ten equal parts
  - *Quartiles* *(Q_i)* - divide a rank-ordered data set into four equal parts
    - $Q_1 = P_{25}$, $Q_2 = P_{50}$ = median
  - *Standard scores*: raw scores that, for ease of interpretation, are converted to a common scale of measurement, or z distribution
    - *z-Score* indicates how many standard deviations an element is from the mean
    - *t-score* enables you to take an individual score and transform it into a standardized form, which helps you to compare scores.



MaSiSS

Inferential or Analytical Statistic

- Inferential statistics takes data from a sample and makes inferences about the larger population from which the sample was drawn
- the goal of inferential statistics is to *draw conclusions* from a *sample* and *generalize* them to a *population*

MUNI SPORT

MaSiSS

Inferential statistic (p-value)

Ronald Fisher (1890–1962)

- The experiment provides a subject with 8 randomly ordered cups of tea
  - 4 prepared by first pouring the *tea*, then adding milk
  - 4 prepared by first pouring the *milk*, then adding the tea

Lady tasting tea

n = 8 total cups

k = 4 cup chosen

$$\binom{8}{4} = \frac{8!}{4!(8-4)!} = 70$$

MUNI SPORT

MaSiSS

Inferential statistic (*p*-value)

Lady tasting tea

$$n = 8 \text{ total cups}$$
$$k = 4 \text{ cup chosen}$$

Tea-Tasting Distribution Assuming the Null Hypothesis

Success count	Combinations of selection	Number of Combinations
0	oooo	1 × 1 = 1
1	ooox, ooxo, oxoo, xooo	4 × 4 = 16
2	ooxx, oxox, oxxo, xoxo, xxoo, xoox	6 × 6 = 36
3	oxxx, xoxx, xxox, xxxo	4 × 4 = 16
4	xxxx	1 × 1 = 1
	Total	70

MUNI SPORT

MaSiSS

Inferential statistic (*p*-value)

Lady tasting tea

How many cups must be correctly identified to concludes that subject can <u>truly</u> tell the difference?

Lady Ottoline Violet Anne Morrell (1873–1938)

English aristocrat and society hostess

MUNI SPORT

MaSiSS

Inferential statistic (p-value)

4 of 4 successe

- 1 / 70 = 0.0143 (1.14%)

3 of 4 success

- (16 + 1) / 70 = 0.243 (24.3%)

Fisher willing to *reject the null hypothesis* … acknowledging the lady's ability *at a 1.14% significance level*

Lady tasting tea

$n = 8$ total cups

$k = 4$ cup chosen

Tea-Tasting Distribution Assuming the Null Hypothesis

Success count	Combinations of selection	Number of Combinations
0	oooo	1 × 1 = 1
1	ooox, ooxo, oxoo, xooo	4 × 4 = 16
2	ooxx, oxox, oxxo, xoxo, xxoo, xoox	6 × 6 = 36
3	oxxx, xoxx, xxox, xxxo	4 × 4 = 16
4	xxxx	1 × 1 = 1
	Total	70

Fisher's Exact Test

MaSiSS

Inferential or Analytical Statistic

- Comparison tests
  - t-test, ANOVA, Mann-Whitney U test, Wilcoxon test, Kruskal-Wallis H test
- Correlation tests
  - Pearson’s r, Spearman’s rho, Chi square test
- Regression tests
  - Simple linear regression, Multiple linear regression, …

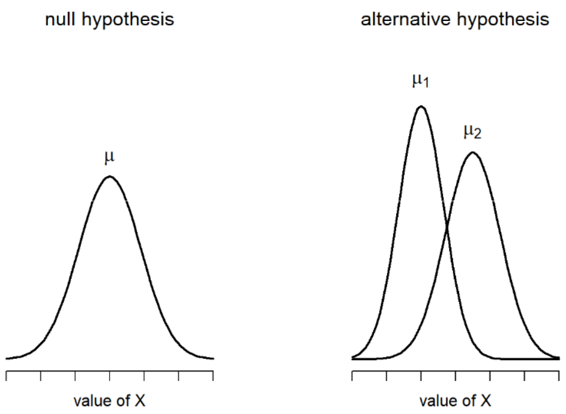
Hypothesis (or Predictions) Testing

MUNI SPORT

MaSiSS

Inferential or Analytical Statistic

- **<u>Comparasion tests</u>**
  - t-test, ANOVA, Mann-Whitney U test, Wilcoxon test, Kruskal-Wallis H test
- Correlation tests
  - Pearson's r, Spearman's rho, Chi square test
- Regression tests
  - Simple linear regression, Multiple linear regression, …



Hypothesis (or Predictions) Testing

For Choosing adequate Comparasion test You need to know:

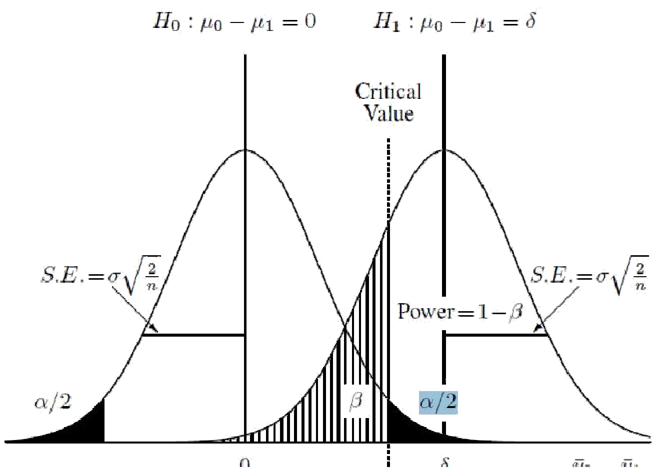
- Which test answer the research question (*Validity*)
- *Scales of measurement*
  - Nominal, Ordinal, Interval, Ratio data
  - Binary, Multiple, Discrete Continuous data
  - Binary, Nominal, Ordinal, Discrete, Continuous data
- *Parametric* or *nonparametric* data
  - using the normal distribution test (just for quantitative data)
- *Paired* (repeated measurements) or *independent* (unpaired) sample

MUNI SPORT

MaSiSS

Inferential or Analytical Statistic

- <u>**Comparasion tests**</u>
  - t-test, ANOVA, Mann-Whitney U test, Wilcoxon test, Kruskal-Wallis H test
- Correlation tests
  - Pearson's r, Spearman's rho, Chi square test
- Regression tests
  - Simple linear regression, Multiple linear regression, …



Hypothesis (or Predictions) Testing

Example of Comparison Tests

– Suitable tests for <u>*nominal*</u> data
  - o Odds ratio test (OR),
  - o Relative risk (RR)

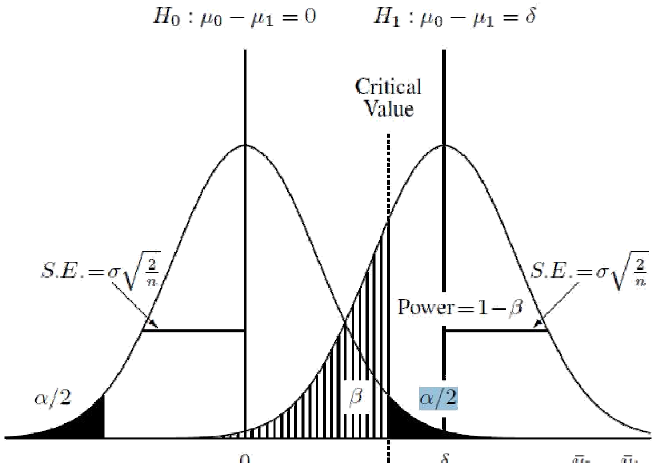
\- Suitable tests for <u>*ordinal*</u> data
  - o Chi-square test, χ^2
    - o Goodness of fit (observed vs. Expected)
    - o Homogeneity (separate subgroups)
    - o Independence (same population)

\- Suitable tests for <u>*interval*</u> and <u>*Ratio*</u> data (Scale data)
  - o *t*-test
    - o one sample
    - o two independent samples
    - o two paired samples
  - o ANOVA test
    - o for three or more sample
    - o also for repeated measurements

MaSiSS

Inferential or Analytical Statistic

- **<u>Comparasion tests</u>**
  - t-test, ANOVA, Mann-Whitney U test, Wilcoxon test, Kruskal-Wallis H test
- Correlation tests
  - Pearson's r, Spearman's rho, Chi square test
- Regression tests
  - Simple linear regression, Multiple linear regression, …



Hypothesis (or Predictions) Testing

William Sealy Gosset
(aka *Student*)

Example of Comparison Tests

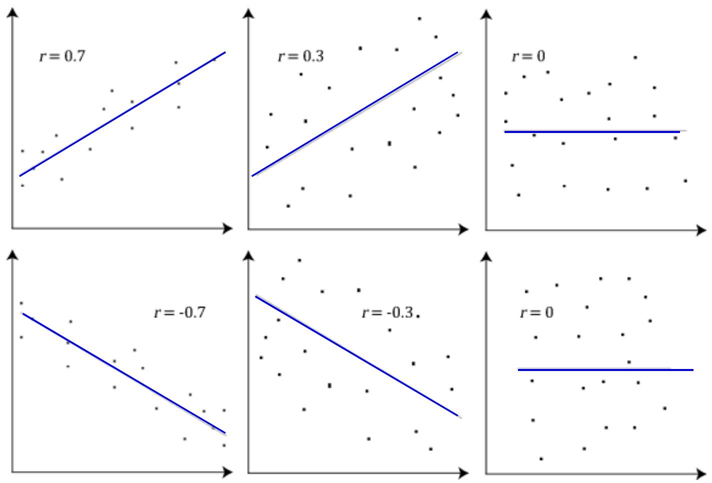
- Suitable tests for <u>*nominal*</u> data
  - Odds ratio test (OR),
  - Relative risk (RR)
- Suitable tests for <u>*ordinal*</u> data
  - Chi-square test, χ^2
    - Goodness of fit (observed vs. Expected)
    - Homogeneity (separate subgroups)
    - Independence (same population)
- Suitable tests for <u>*interval*</u> and <u>*Ratio*</u> data (Scale data)
  - *t*-test
    - one sample
    - two independent samples
    - two paired samples
  - ANOVA test
    - for three or more sample
    - also for repeated measurements

MUNI SPORT

MaSiSS

Inferential or Analytical Statistic

- Comparasion tests
  - t-test, ANOVA, Mann-Whitney U test, Wilcoxon test, Kruskal-Wallis H test
- **Correlation tests**
  - Pearson's r, Spearman's rho, Chi square test
- Regression tests
  - Simple linear regression, Multiple linear regression, …



Hypothesis (or Predictions) Testing

– Pearson's *r*
  – parametric, quantitative data
– Spearman's *rho*
  – nonparametric, quantitative which are handled like qualitative data
– Chi-square test (χ^2)
  – qualitative data
– Dimension reduction techniques
  – *Factor analysis* (use when you assume association if you want to understand the latent factors)
  – *Principal Component Analysis* (seeks to identify, to predict using the factors, PCA)

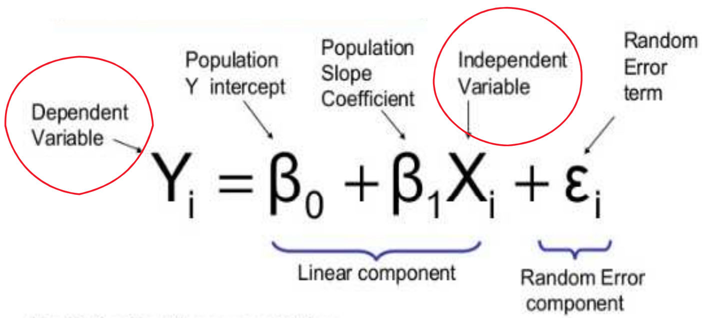
Size of Correlation	Interpretation
.90 to 1.00 (−.90 to −1.00)	Very high positive (negative) correlation
.70 to .90 (−.70 to −.90)	High positive (negative) correlation
.50 to .70 (−.50 to −.70)	Moderate positive (negative) correlation
.30 to .50 (−.30 to −.50)	Low positive (negative) correlation
.00 to .30 (.00 to −.30)	negligible correlation

MUNI SPORT

MaSiSS

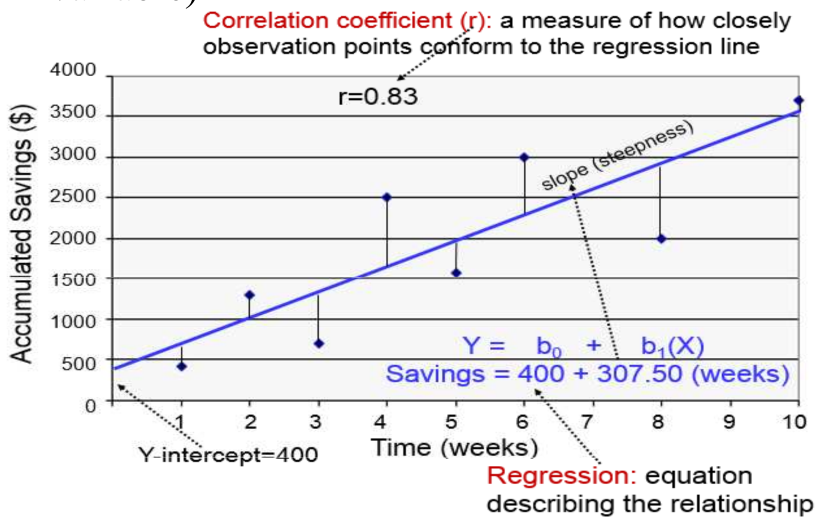
Inferential or Analytical Statistic

- Comparasion tests
  - t-test, ANOVA, Mann-Whitney U test, Wilcoxon test, Kruskal-Wallis H test
- Correlation tests
  - Pearson's r, Spearman's rho, Chi square test
- **Regression tests**
  - Simple linear regression, Multiple linear regression, …



Hypothesis (or Predictions) Testing

– **Simple linear regression** (1 metric IV; 1 metric DV)
– Multiple linear regression (2+ metric IV; metric DV)
– Logistic regression (1 any IV; 1 binary variable)
– Nominal regression (1 any IV; 1 nominal variable),
– Ordinal logistic regression (1 any IV; 1 ordinal variable)



MUNI SPORT

MaSiSS

Inferential or Analytical Statistic

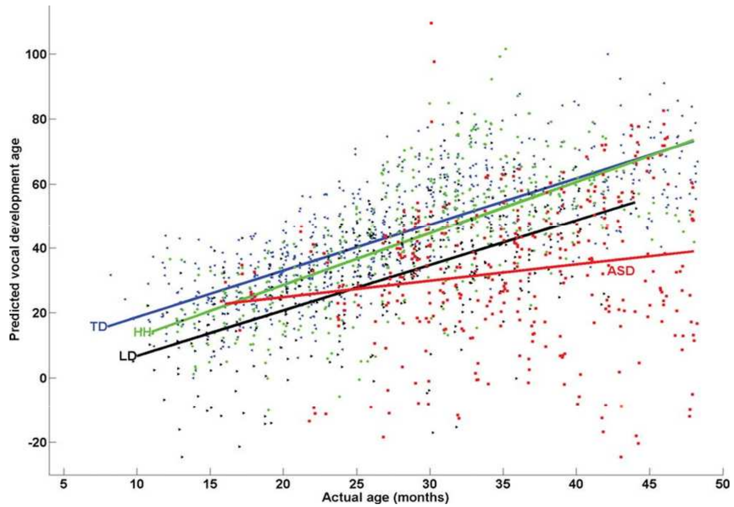
- Comparasion tests
  - t-test, ANOVA, Mann-Whitney U test, Wilcoxon test, Kruskal-Wallis H test
- Correlation tests
  - Pearson's r, Spearman's rho, Chi square test
- **<u>Regression tests</u>**
  - Simple linear regression, Multiple linear regression, …

Dependent variable (DV) Independent variables (IVs)

$$y = b_0 + b_1 {}^* x_1 + b_2 {}^* x_2 + \dots + b_n {}^* x_I$$

Hypothesis (or Predictions) Testing

- Simple linear regression (1 metric IV; 1 metric DV)
- **Multiple linear regression** (2+ metric IVs; metric DV)
- Logistic regression (1 any IV; 1 binary variable)
- Nominal regression (1 any IV; 1 nominal variable),
- Ordinal logistic regression (1 any IV; 1 ordinal variable)



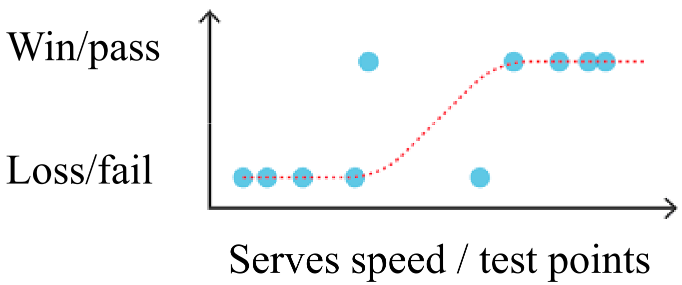
MaSiSS

Inferential or Analytical Statistic

- Comparasion tests
  - t-test, ANOVA, Mann-Whitney U test, Wilcoxon test, Kruskal-Wallis H test
- Correlation tests
  - Pearson's r, Spearman's rho, Chi square test
- **Regression tests**
  - Simple linear regression, Multiple linear regression, …

Hypothesis (or Predictions) Testing

- Simple linear regression (1 metric IV; 1 metric DV)
- Multiple linear regression (2+ metric IVs; metric DV)
- **Logistic regression** (1 any IV; 1 binary variable)
- Nominal regression (1 any IV; 1 nominal variable),
- Ordinal logistic regression (1 any IV; 1 ordinal variable)



MUNI SPORT

MaSiSS

Multiple comparisons problem

If for <u>1 test:</u> $p = 0.04$ ($\alpha < 0.05$)

If for <u>2 tests:</u> $p = 0.04$ and 0.02 ($\alpha < 0.05$) $\alpha \neq 0.05 = 0.025$

If for <u>4 tests:</u> $p = 0.04; 0.02; 0.01; 0.001$ ($\alpha < 0.05$) $\alpha \neq 0.05 = 0.0125$

α = 0.05

Adjusted α = 0.0041

MaSiSS

Multiple comparisons problem

Table 2. Mean and SD of analysed parameters in forward gait for 0.00% and 0.11% BrAC conditions in females and males.

	0.00% BrAC					0.11% BrAC					Dif. 0.00% and 0.11% BrAC	
	Females		Males		*p* (gender dif.)	Females		Males		p (gender dif.)	*p* (females)	*p* (males)
	Mean	SD	Mean	SD		Mean	SD	Mean	SD			
Foot rotation, °	3.24	4.06	7.87	5.84	**<0.001***	3.94	4.17	7.80	4.97	**<0.001***	0.484	**0.001***
Stride length, cm	128.40	12.40	135.07	13.42	**0.006***	134.87	13.84	138.47	14.75	0.167	0.572	0.401
Step width, cm	9.26	2.16	12.09	2.80	**<0.001***	9.54	2.47	11.96	3.56	**<0.001***	**0.025***	0.508
Stance phase, %	63.04	2.12	62.55	2.19	0.132	61.90	2.36	62.34	2.55	0.620	0.013*	0.308
Load response, %	12.77	1.78	12.80	2.13	0.615	12.42	2.02	12.32	2.23	0.482	0.806	0.496
Single limb support, %	37.39	2.17	37.49	2.34	0.807	37.49	1.94	37.67	2.58	0.566	0.498	0.140
Pre-Swing, %	12.89	1.71	12.24	1.88	0.069	12.08	1.71	12.35	2.27	0.684	0.295	**0.020***
Swing phase, %	36.96	2.12	37.45	2.19	0.132	38.10	2.36	37.66	2.55	0.620	**0.013***	0.308
Double stance phase, %	25.66	3.06	25.03	3.55	0.201	24.47	3.26	24.68	4.17	0.863	0.308	0.088
Stride time, sec	1.12	0.10	1.18	0.09	**0.002***	1.13	0.10	1.19	0.09	**<0.001***	0.057	0.134
Cadence, steps/min	107.58	9.45	102.20	7.53	**0.002***	106.91	8.83	101.55	6.96	**<0.001***	0.052	0.249
Velocity, km/h	4.17	0.65	4.14	0.48	0.885	4.34	0.67	4.22	0.49	0.391	0.521	0.197

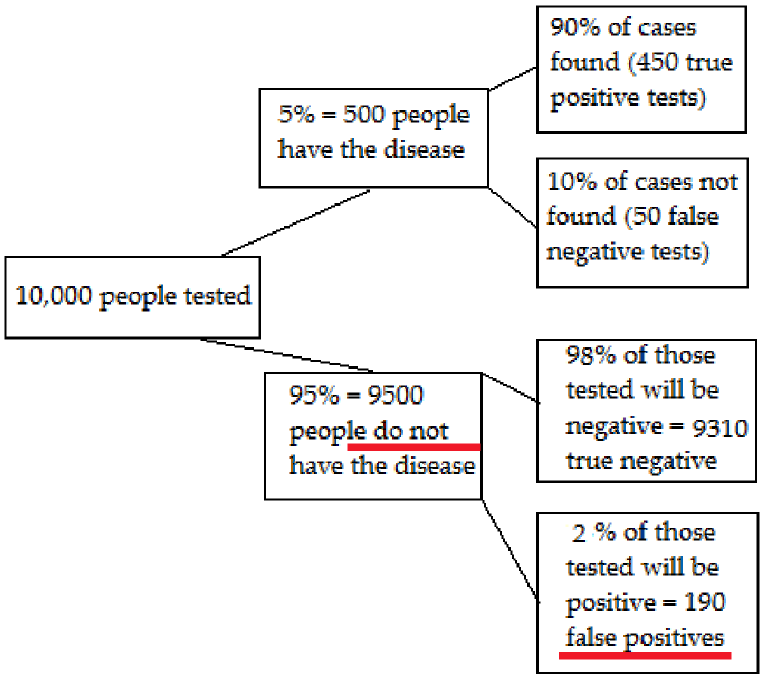
MUNI SPORT

MaSiSS

Multiple comparisons problem

- False Discovery Rate (FDR) – *expected proportion of type I error*
- Family-wise effer rate (FWER) – *probability of making at least one type I error*

If you **repeat a test** enough times, you will always get a number of **false positive**



MUNI SPORT

MaSiSS

Multiple comparisons problem

Benjamini-Hochberg ($p_k = (k / m) * Q$)

– Controls the FDR

Holm correction ($p_k = \alpha / (m + 1 - k)$)

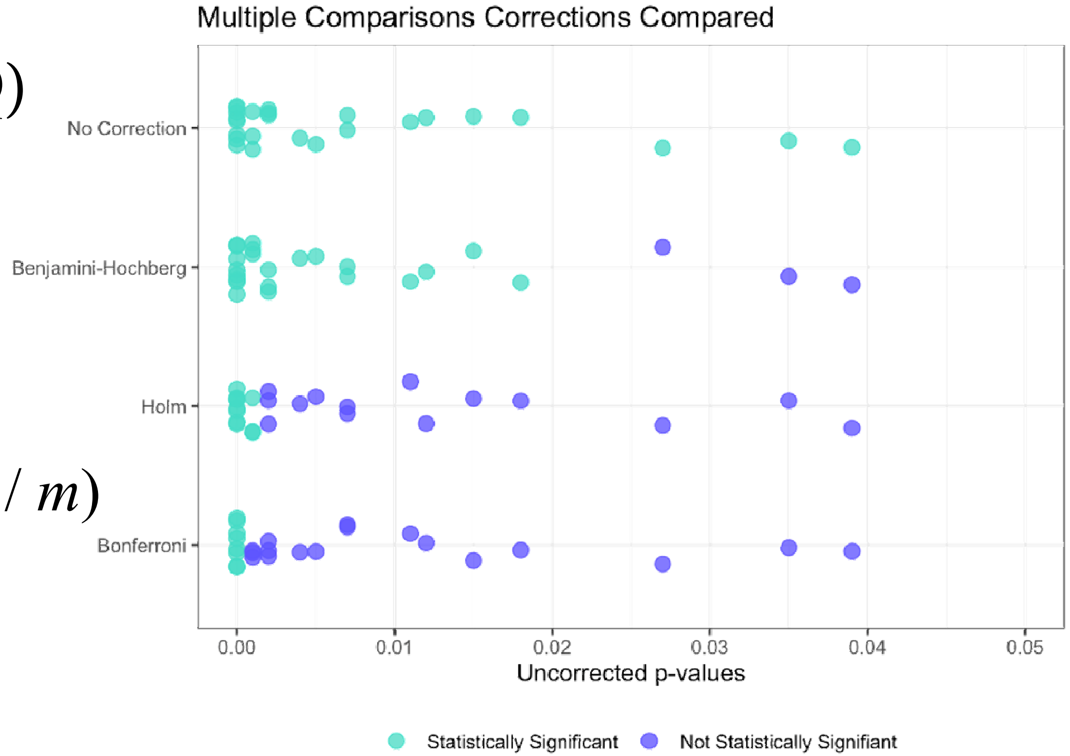
- Powerfull then Bonferroni
- Control the FWER

Bonferroni correction/adjustment ($p_k = \alpha / m$)

– Control the FWER

k = rank of ordered p-value; m = number of tests; Q = FDR (0.05-0.25)

How to correct rist of type I error



MUNI SPORT

MaSiSS

Multiple comparisons problem

When not to correct

If false negative are also importent for future research

- Example: you‘re researching a new AIDS vaccine. A high number of false positives may be hint to that you‘re on the right track

When the results are not statistically significant ($p > \alpha$)

For a single test

In the case of an exploratory study

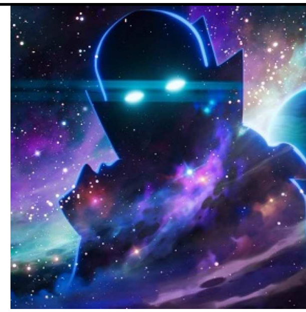
- we do not know the hypotheses in advance

(McDonald, 2014; Hendl & Remr, 2017)

MUNI SPORT

MaSiSS

Multiple comparisons problem



What if …?

Bonferroni correction ($p_k = \alpha / m$)

– Number of tests (m) = 6; $\alpha = 0.05$

– $p_k = 0.008$

– Number of tests (m) = 3; $\alpha = 0.05$

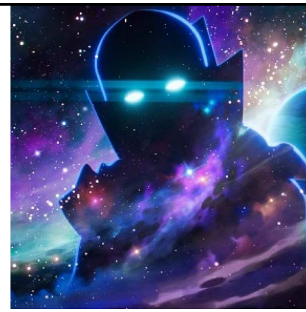
– $p_k = 0.017$

Class	Upper limb			Lower limb		
	Right	Left	*p*	Right	Left	*p*
1st	Mean±SD	Mean±SD	0.015	Mean±SD	Mean±SD	0.010
2nd	Mean±SD	Mean±SD	0.041	Mean±SD	Mean±SD	0.032
3rd	Mean ˙ SD	Mean ˙ SD	0.05	Mean ˙ SD	Mean ˙ SD	0.063

MaSiSS

Multiple comparisons problem

What if …?



Bonferroni correction ($p_k = \alpha / n$)

- Number of tests (n) = 6; $\alpha = 0.05$
- $p_k = 0.008$

- Number of tests (n) = 3; $\alpha = 0.05$
- $p_k = 0.017$

Class	**Upper limb**			**Lower limb**		
	Right	**Left**	***p***	**Right**	**Left**	***p***
Max GS [Nm]	Mean±SD	Mean±SD	0.015	Mean±SD	Mean±SD	0.010
Power GS [W]	Mean±SD	Mean±SD	0.041	Mean±SD	Mean±SD	0.032
Time to Max GS [s]	Mean ˙ SD	Mean ˙ SD	0.05	Mean ˙ SD	Mean ˙ SD	0.063

MUNI SPORT

MaSiSS

The Effect size (ES)

- Is a *quantitative measure* of *the magnitude* of the experimental effect
- Is considered an *essential complement* of the *statistical significance test*, because it allows to know the relevance of the difference and discerning between the statistical significance of a test and its practical importance.
- The effect size allows to make comparisons between the statistical significant differences from groups with a very *different number* of items, and studying groups from *different scientific works* (meta-analysis)
- Effect size hepl undestend *the magnitude of differences* found (statistical signifikance examines whether the findings are likely to be *due to chance*)

Most common effect size tests

- Cohen‘s *d* (Effect size index *d*)
  - Appropriate to campared two means (*t*-test)
  - Small (d = 0.2; 58%*), Medium (d = 0.5; 69%*), Large (d = 0.8; 79%)
- Pearson correlation (*r*)
- Confidence interval (CI)
  - is a range of values that you can be (95%) certain contains the true mean of the population.
  - there are variants of 90 % CI, 95% CI, 99% CI
  - *Cohen‘s d* [95% CI] = *d* [lower bound; higher bound] = 0.9 [0.12; 1.52]
- Odds ratio test (OR)
  - For 2x2 table, (AD) / (BC)
  - Example for OR (95% CI): 2.1 (2.0 to 2.2)
- Others Effect size tests
  - Epsilon-squared (ε^2), Cohen‘s ω (ω), Eta-squared (η^2), Cramer‘s V (V), Effect size Phi (φ), Relative risk or risk ratio (RR), Coefficient of determination (r^2), Common Language Effect Size (CLES), ...

*Procentige of control group below the mean of experimental group

MaSiSS

The Effect size (ES)

Test	Effect size	Small	Medium	Large
Difference between two means	d	0.20	0.50	0.80
Difference between many means	*f*	0.10	0.25	0.40
Chi-squared test	*w*	0.10	0.30	0.50
Pearson's correlation coefficient	*ρ*	0.10	0.30	0.50

Most common effect size tests

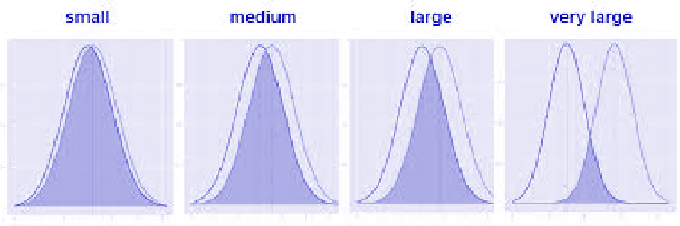
- Cohen's *d* (Effect size index *d*)
  - Appropriate to campared two means (*t*-test)
  - Small (d = 0.2; 58%*), Medium (d = 0.5; 69%*), Large (d = 0.8; 79%)
- Pearson correlation (r)
- Confidence interval (CI)
  - is a range of values that you can be (95%) certain contains the true mean of the population.
  - there are variants of 90 % CI, 95% CI, 99% CI
  - *Cohen's d* [95% CI] = *d* [lower bound; higher bound] = 0.9 [0.12; 1.52]
- Odds ratio test (OR)
  - For 2x2 table, (AD) / (BC)
  - Example for OR (95% CI): 2.1 (2.0 to 2.2)
- Others Effect size tests
  - Epsilon-squared (ε^2), Cohen's ω (ω), Eta-squared (η^2), Cramer's V (V), Effect size Phi (φ), Relative risk or risk ratio (RR), Coefficient of determination (r^2), Common Language Effect Size (CLES), ...

*Procentige of control group below the mean of experimental group

MUNI SPORT

MaSiSS

The Effect size (ES)



Relative size	Effect size	% of control group below the mean of experimental group
	0.0	50%
Small	0.2	58%
Medium	0.5	69%
Large	0.8	79%
	1.4	92%

Most common effect size tests

- Cohen's *d* (Effect size index *d*)
  - Appropriate to compared two means (*t*-test)
    - Small (d = 0.2; 58%*)
    - Medium (d = 0.5; 69%*)
    - Large (d = 0.8; 79%*)

*Procentige of control group below the mean of experimental group

MUNI SPORT

MaSiSS

The Effect size (ES)

Most common effect size tests

– Pearson correlation (*r*)
  – Summarises the strenght of the bivariate relationship
  – *r* correlation varies between -1 to +1
    (A perfect negative to a perfect positive correlation)

Variables	1	2	3	4	5	6	7	8	9	10	11	12	13
1. VATCI	1.00												
2. Sensation seek	.16**	1.00											
3. Aggression	.18**	.17**	1.00										
4. Risky driving	.11*	.39**	.20**	1.00									
5. Fighting	.20**	.23**	.65**	.32**	1.00								
6. Delinquency	.21**	.31**	.40**	.45**	.50**	1.00							
7. Risky sex	.02	.25**	.19**	.31**	.24**	.32**	1.00						
8. Drinking	.27**	.40**	.19**	.37**	.32**	.46**	.35**	1.00					
9. Smoking	–.02	.05	.03	–.02	.02	.05	–.01	.08	1.00				
10. Drugs	.20**	.38*	.19**	.32**	.28**	.43**	.38**	.45**	.08	1.00			
11. TV viewing	.28**	–.04	.18**	.02	.15**	.06	–.02	.06	–.04	.03	1.00		
12. Age	–.03	–.02	.03	.11*	.04	.08	.08	.10*	–.06	.04	.05	1.00	
13. Sex	–.12*	–.14**	–.26**	–.27**	–.39**	–.40**	–.06	–.28**	–.05	–.11*	–.02	–.09	1.00

*$p \le .05$, **$p \le .01$, ***$p \le .001$

Correlation matrix for all variables (Banerje et al. (2009)

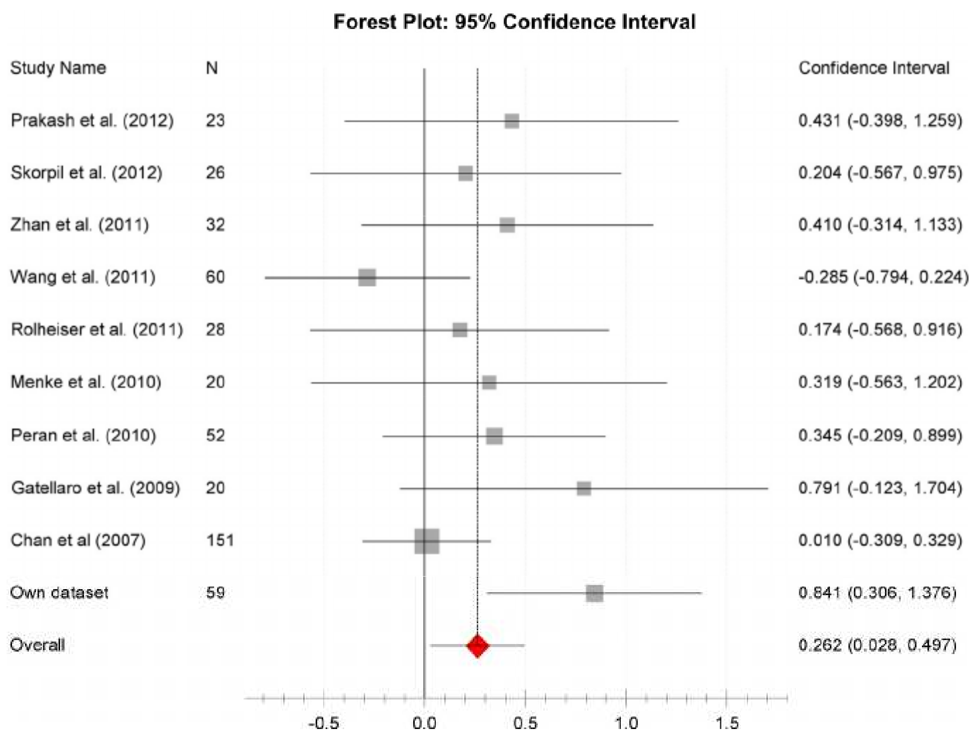
Range of r_s	Strength of Correlation
below 0.16	"very low"
0.16 - 0.29	"weak to low"
0 .30 - 0.49	"moderate to low"
0.50 - 0.69	"moderate"
0.70 - 0.89	"strong"
0.90 - 1.00	"very strong"

	Coefficient, *r*	
Strength of Association	Positive	Negative
Small	.1 to .3	-0.1 to -0.3
Medium	.3 to .5	-0.3 to -0.5
Large	.5 to 1.0	-0.5 to -1.0

MUNI SPORT

MaSiSS

The Effect size (ES)



Schwarz et al. (2013)

Most common effect size tests

– Confidence interval (CI)
  - is a range of values that you can be (95%) certain contains the true mean of the population.
  - there are variants of 90 % CI, 95% CI, 99% CI
  - *Cohen's d* [95% CI] = *d* [lower bound; higher bound] = 0.9 [0.12; 1.52]

MaSiSS

The Effect size (ES)

Exposure Status	Event Occurred	
	Yes	No
Exposed	a	b
Not Exposed	c	d

$$\text{Relative Risk} = \frac{a/(a+b)}{c/(c+d)}$$

$$\text{Odds Ratio} = \frac{a/b}{c/d} = \frac{ad}{cb}$$

Most common effect size tests

- Odds ratio test (OR)
  - For 2x2 table, (AD) / (BC)
  - Example for OR (95% CI): 2.1 (2.0 to 2.2)
- Relative Risk (RR

Interpreting Effect Size Results

- Cohen's "Rules-of-Thumb"
  - standardized mean difference effect size
    - small = 0.20
    - medium = 0.50
    - large = 0.80
  - correlation coefficient
    - small = 0.10
    - medium = 0.25
    - large = 0.40
  - odds-ratio
    - small = 1.50
    - medium = 2.50
    - large = 4.30

Practical Meta-Analysis -- D. B. Wilson 1

MUNI SPORT

MaSiSS

Probability

- Classic probability
- Frequentist probability
- Bayesian probability

MaSiSS

Probability

- **<u>Classic probability</u>**
  - Measure the likelihood (probability) of something happening.
- Frequentist probability
  - Null hypothesis signifikance testing
  - Probability of observed data under the assumption that the null hypothesis is true
- Bayesian probability
  - Probabilities of both hypothesis
  - Bayes theorem
  - Uses the idea of updating belief with new information

- Tossing a coin or dice, …

$$P(A) = f / N$$

  - P(A) = probability of event A
  - f = frequency (or number) of possible times the event could happen
  - N = the number of times the event could happen
- For example, the odds of rolling a 6 on (a fair) die are 1 out of 6 (1/6)

  = one possible outcome divided by the number of possible outcomes

MUNI SPORT

MaSiSS

Probability

- Classic probability
  - Measure the likelihood (probability) of something happening.
- **<u>Frequentist probability</u>**
  - Null hypothesis signifikance testing
  - Probability of observed data under the assumption that the null hypothesis is true
- Bayesian probability
  - Probabilities of both hypothesis
  - Bayes theorem
  - Uses the idea of updating belief with new information

$P(E|H_o)$

<u>Important</u>:

- The *p*-value is not the probability of a theory or hypothesis, but the probability of the observed data

MUNI SPORT

MaSiSS

Probability

- Classic probability
  - Measure the likelihood (probability) of something happening.
- Frequentist probability
  - Null hypothesis signifikance testing
- **Bayesian probability**
  - Probabilities of both hypothesis
  - Bayes theorem
  - Uses the idea of updating belief with new information

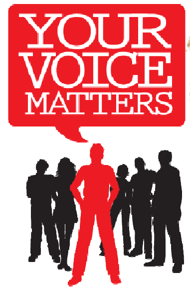
$$P(H|E)$$

MUNI SPORT

MaSiSS

Probability

Bayesian probability



– Treasure hunting

Tommy Thimpson (1988) – SS Central America with $700,000,000 worth of gold

– Scientists use Bayes how new data validetes or invalidates their models

– Programmers use to building artificial intelligence (kvantification of machines belief)

– How you view yourself,

  – your own opinions and what it takes for your mind to change (reframing your thought itself)

– Spam filter (looks at words)

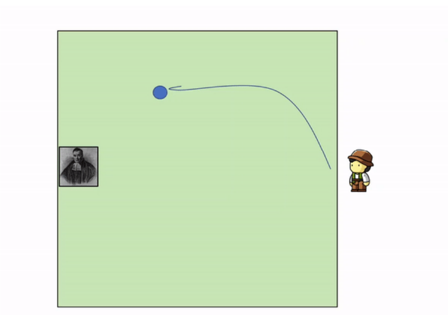
= probability that the email is spam, given that those words appear

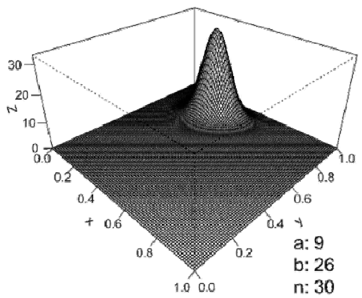
$$P(H|E) = \frac{P(H) * P(E|H)}{P(E)} = P(Spam|Word) = \frac{P(Spam) * P(Word|Spam)}{P(Word)}$$

MaSiSS

Probability

– Bayes never publish his treorem
– He submit it to the Royal Society
– Publish after he died (Richard Price)
  – Man coming first time out of a cave and saw the sun rise for the first time and ask himself: Is this one-off or does the sun always do this? An then every day after that, as the sun rose agan he yould get a little bit more confident, that the World works.





Bayesian probability

Origin thought experiment

– He was sitting back to a perfectly flat, perfectly squared table, then he ask assistant to throw a ball onto the table and he wanted to figure out where it was. So, he asked his assistant to throw on another ball and then tell him if it land to the left, or to the right, front, behind of the first ball. He would note that down and then ask for more and more balls to be thrown on the table. Through this method he yould keep updating his idea of where the first ball was. He will never be completely certain, but with each new piece of evidence, he would get more and more accurate

And that‘s how Bayes saw the world

$P(H|E) = \frac{P(H) * P(E|H)}{P(E)}$

MaSiSS

Probability

Bayesian probability

P(H|E) = ***Posterior*** = probability a hypothesis is true given some evidence

belief about the hypothesis after seeing the evidence

P(H) = ***Prior*** = probability a hypothesis is true (before any evidence)

hardest part of te equation (sometimes just guess)

$P(E|H)$ = ***likelihood*** = the probability that evidence given the hypothesis is true

probability of seeing the evidence if the hypothesis is true

P(E) = ***Marginalization*** = The probability evidence being true

probability of seeing the evidence

$$P(H|E) = \frac{P(H) * P(E|H)}{P(E)} = \frac{P(H) * P(E|H)}{P(H) * P(E|H) + P(\neg H) * P(E|\neg H)}$$

MaSiSS

Probability

Prior

likelihood

$$P(H|E) = \frac{P(H) * P(E|H)}{P(E)}$$

Posterior

Marginalization

Bayesian probability

- P(H|E) = *Posterior* = probability a hypothesis is true given some evidence
- P(H) = *Prior* = probability a hypothesis is true
- *P*(*E*|*H*) = *likelihood* = the probability that evidence given the hypothesis is true
- P(E) = *Marginalization* = The probability evidence being true

MUNI SPORT

MaSiSS

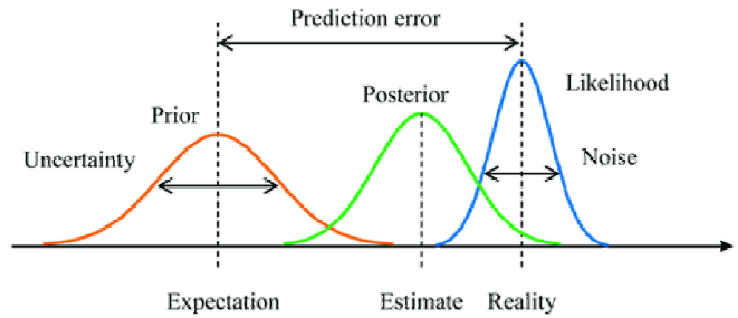
Probability

<u>Bayes‘ rules</u>

- You have hypothesis
- You‘ve observed some evidence
- You want know the probability that the your hypothesis holds given that the evidence is true

  = P(Hypothesis given Evidence)
  Given (“|“) = restricting view only to the possibilities where the evidence holds

Bayesian probability



$$P(H|E) = \frac{P(H)*P(E|H)}{P(E)} = \frac{P(H)*P(E|H)}{P(H)*P(E|H)+P(\neg H)*P(E|\neg H)}$$

MUNI SPORT

MaSiSS

Probability

Bayesian probability

$$\underbrace{\boxed{P(H|E)}}_{\textit{Posterior}} = \frac{\overbrace{\boxed{P(H)}}^{\textit{Prior}} \cdot \overbrace{\boxed{P(E|H)}}^{\textit{likelihood}}}{\underbrace{\boxed{P(E)}}_{\textit{Marginalization}}}$$

<u>Example</u>: you feel a little bit sick, without sympstoms, just not 100%. Doctors run a battery of tests and results are … you tested **positive** for a rare **disease** that affects about **0.1%** of population and it‘s a nasty disease. The **test correclty identify 99%** of people that have the disease and only **incorrecly identify 1%** of people who don‘t the disease.

- *What are the <u>chances</u> that you <u>actually have this disease?</u>*

99%???

MUNI SPORT

MaSiSS

Probability

Bayesian probability

Prior probability of having the disease

You would test + if you had the disease

$$\boxed{P(H|E)} = \frac{\boxed{P(H)} * \boxed{P(E|H)}}{\boxed{P(E)}}$$

Actually have the disease given you tested +

Probability of testing +

Prior

likelihood

$$P(H|E) = \frac{P(H) * P(E|H)}{P(E)}$$

Posterior

Marginalization

$$P(H|E) = \frac{P(H) * P(E|H)}{P(E)} = \frac{P(H) * P(E|H)}{P(H)*P(E|H)+P(\neg H)*P(E|\neg H)}$$

MaSiSS

Probability

Bayesian probability

Prior probability of having the disease

Actually have the disease

You would test + if you had the disease

given

$$P(H|E) = \frac{\boxed{0.001} * \boxed{0.99}}{\boxed{0.001*0.99+0.999*0.01}}$$ $\boxed{= 9\ \%}$

Actually having the disease after testing +

You tested +

Probability of testing +

Is it too Low?

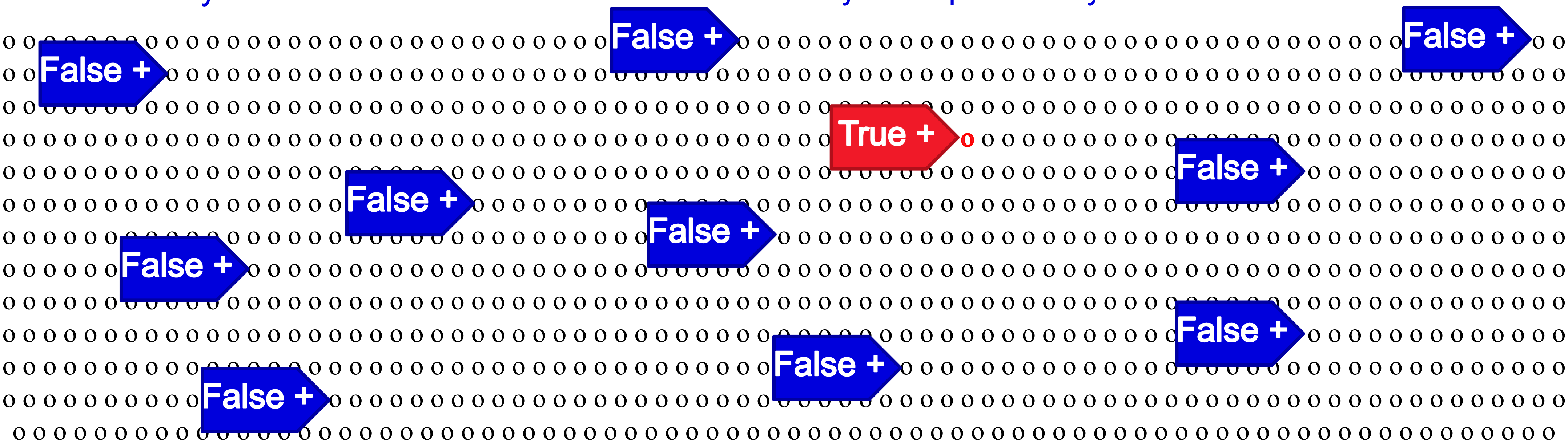
no, it is common sense applied to mathematics…

MUNI SPORT

MaSiSS

Probability

Bayesian probability



N = 1000; n = 1 (actually have the disease); n = 10 (1% of 999 people)

It just make sence

MaSiSS

Probability

Bayesian probability

o o o o o o o o o o o

1 in 11 people = 9%

N = 1000; n = 1 (actually have the disease); n = 10 (1% of 999 people)

MaSiSS

Probability

- Bayes‘ Theorem wasn‘t a formula intended to be used just once.
- Each tome gaining *new evidence* and *updating* your *probability*
  - That someting is true
  - you should update prior beliefs (update a belief based on evidence)

Bayesian probability

Back to example:

- You tested **positive** for a rare **disease**.
- You get another doctor opinion, get second test, but test aslo come back as **positive** …

What is the prabability that you actually have the disease?

MUNI SPORT

$$P(H|E) = \frac{P(H)*P(E|H)}{P(E)} = \frac{P(H)*P(E|H)}{P(H)*P(E|H)+P(\neg H)*P(E|\neg H)}$$

MaSiSS

Probability

Bayesian probability

Prior probability of having the disease

Actually have the disease

You would test + if you had the disease

Actually having the disease after testing +

given $$\boxed{P(H|E)} = \frac{\boxed{0.001} * \boxed{0.99}}{\boxed{0.001*0.99+0.999*0.01}} \boxed{= 9\ \%}$$

You tested +

Probability of testing +

<u>Posterion will be new prior</u>

MUNI SPORT

$$P(H|E) = \frac{P(H)*P(E|H)}{P(E)} = \frac{P(H)*P(E|H)}{P(H)*P(E|H)+P(\neg H)*P(E|\neg H)}$$

MaSiSS

Probability

Bayesian probability

Prior probability of having the disease

Actually have the disease

You would test + if you had the disease

Actually having the disease after testing + ***(2 times)***

given $$P(H|E) = \frac{0.09 * 0.99}{0.09*0.99+0.91*0.01} = 90.7\%$$

You tested +

Probability of testing +

New probability (after 2 positive tests)

MUNI SPORT

MaSiSS

Probability

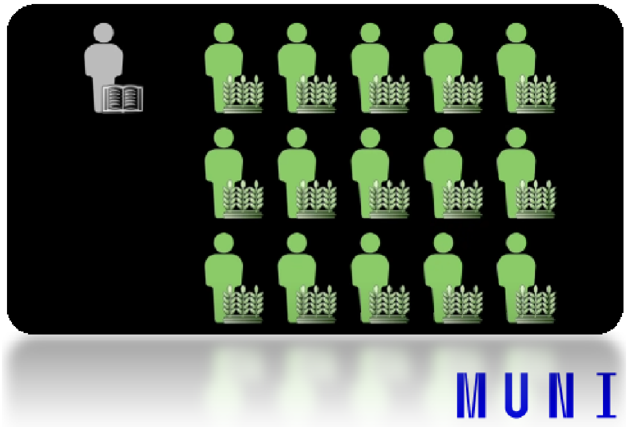
Bayesian probability

Kahneman and Tversky (2003), example 1

– Steve is very shy and withdrawn, invariably helpful but with very little interest in people or in the world of reality. A meek and tidy soul, he has a need for order and structure, and a passion for detail.

Is Steve a **librarian** or **farmer**?

Most comman anser: ***<u>Librarian</u>***
But people hold biased views about the personalities (of librarians or farmars)
= stereotypes
(almost) no one incorporate information about **ratio** of farmers to librarians in their judgments



MUNI SPORT

$$P(H|E) = \frac{P(H) * P(E|H)}{P(E)} = \frac{P(H) * P(E|H)}{P(H) * P(E|H) + P(\neg H) * P(E|\neg H)}$$

MaSiSS

Probability

!Prior! *likelihood*

$$P(H|E) = \frac{P(H) * P(E|H)}{P(E)}$$

Posterior *Marginalization*

Bayesian probability

P(H) = *Prior* = ratio of farmers to librarians in general population

The probability of H being true (this is knowledge)

P(E|H) = *likelihood* = proportion of librarians that fit this destriptions

The probality of E being true, given H is true

P(H|E) = *Posterior* = belief about the hypothesis after seeing the evidence

The probability of H being true, given E is true

P(E) = *Marginalization* = The probability evidence being true

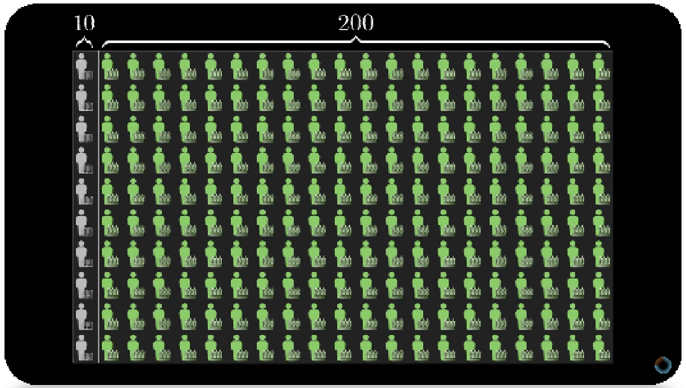
$$P(H|E) = \frac{P(H) * P(E|H)}{P(E)}$$

MaSiSS

Probability

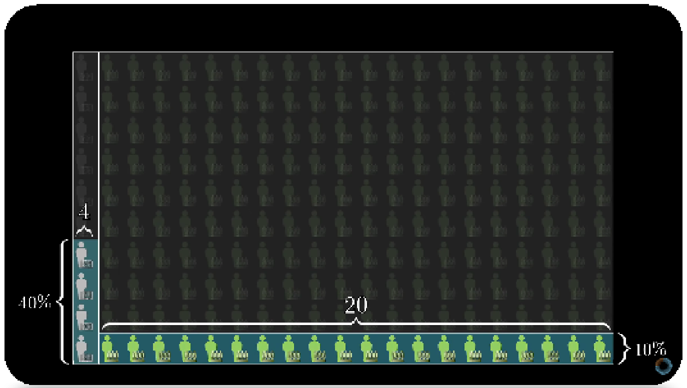
Kahneman and Tversky (2003)

- Farmars to libratians (20:1)



Population (10:200 = 1:20)

Bayesian probability



You expect from sample 4 librarians and 20 farmers to fit the descriptions

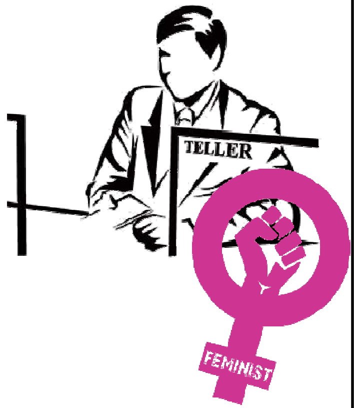
$$P(\text{Librarian given Descriptions}) = \frac{4}{4+20} = 16.7\%$$

MUNI SPORT

MaSiSS

Probability

Bayesian probability



Kahneman and Tversky (example 2)

- Linda is 31 yers old, single, outspoken, and very bright. She majored in philisiphy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations.

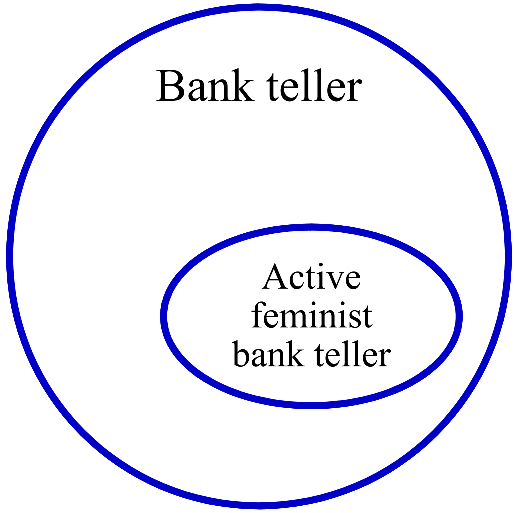
Is Linda more likely:

A) a bank teller.

B) a bank teller and is active in the feminist movement.

85% chose B !!!

even if it is a subsection



MaSiSS

Probability

Bayesian probability

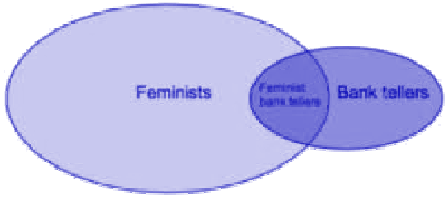
The same assignment (Linda)

– 100 people fit this description. How many are:

A) Bank teller? ___of 100
B) Bank teller and is active in the feminist movement? ___ of 100

100% people assigns a higher number to the first option than to the second

Results: people udestend
“40 out 100“ better “40%“ less than “0.4“ much less that abstractly referencing the idea (Steve, Linda)



MUNI SPORT

MaSiSS

Probability

Bayesian probability

Example:

- At the Campus you meet a guy name Tom.
  Alfter few minutes you notice that Tom is **shy**.
  - Is Tom more likely to be in IT program?
  - Is Tom more likely to be in a MNG program?

  - Shyness is more common in IT programs
    …but how many IT students are they relative to MNG students?

MUNI SPORT

MaSiSS

Probability

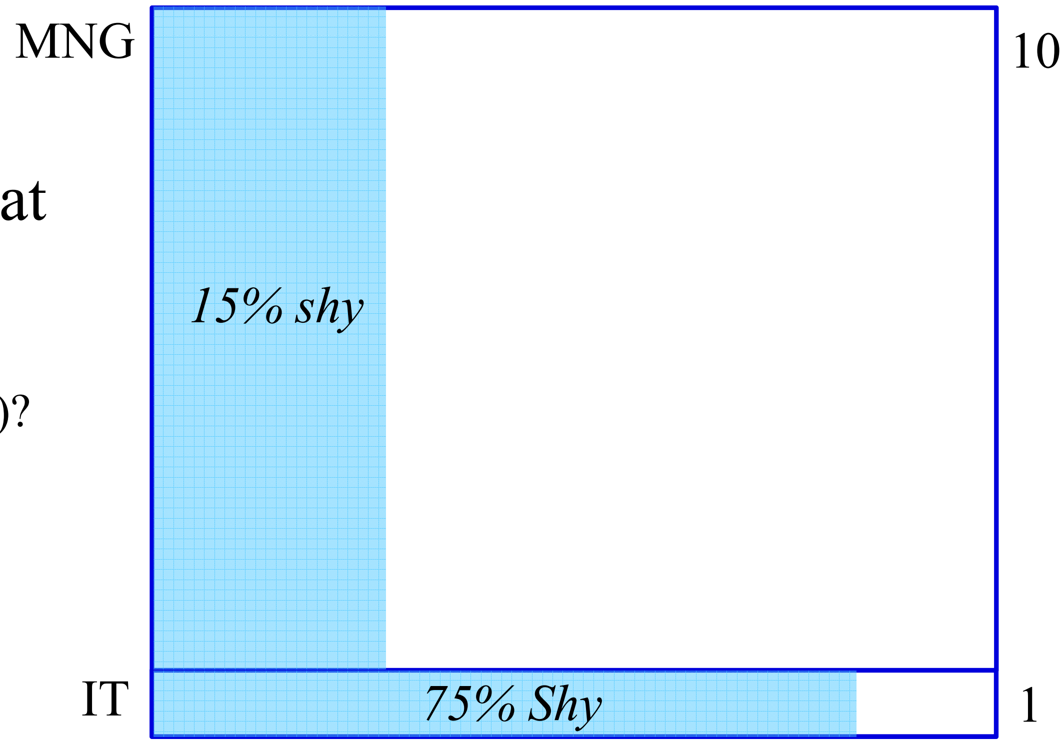
Example:

At the Campus you meet a guy name Tom. Alfter few minutes you notice that Tom is shy.

- Is Tom more likely to be in IT program (Bc.)?
- Is Tom more likely to be in a MNG program (Bc.)?

	IT : MNG
Prior odds ratio	1 : 10
Likelihood odds ratio	75 : 15
Posterio odds ratio	75 : 150 → $\boxed{1 : 2}$

Bayesian probability



MUNI SPORT

MaSiSS

Probability

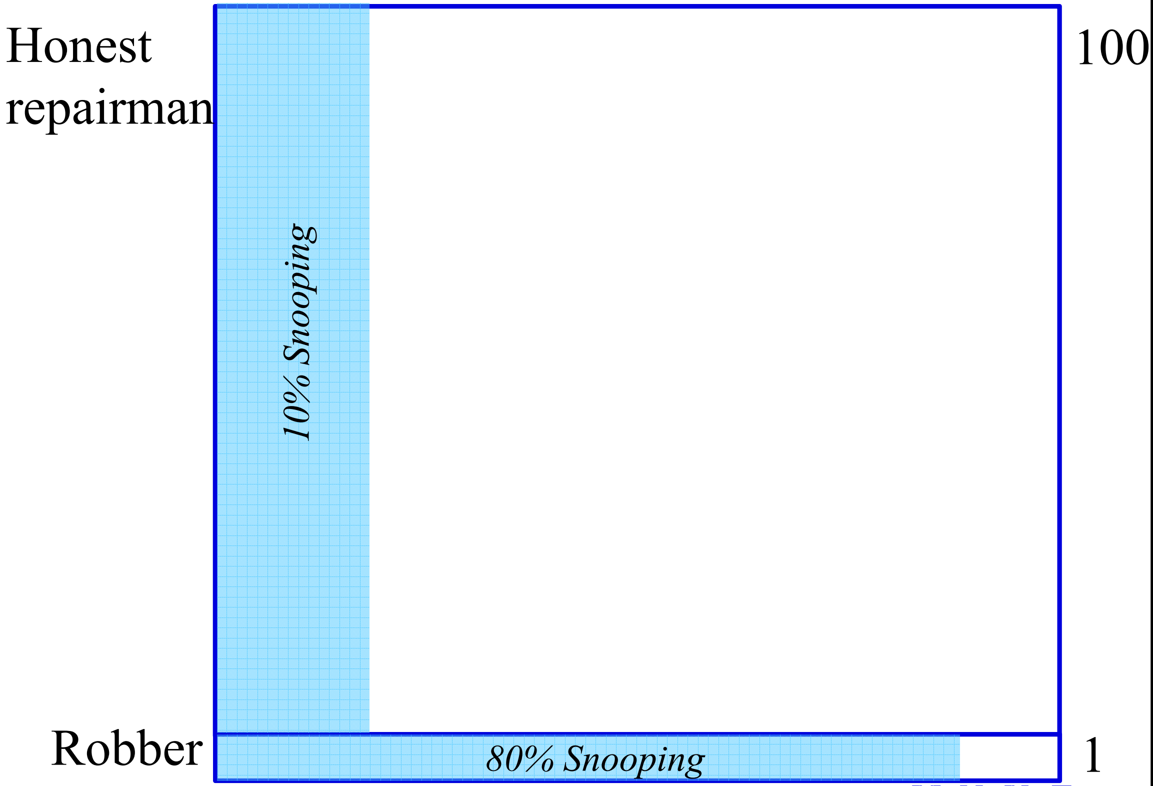
Example:

– The stove repairman looks suspiciously around the various rooms in the apartment
  – Is repairman more likely to be a robber?

	Robber : Repairman
Prior odds ratio	1 : 100
Likelihood odds ratio	80 : 1
Posterio odds ratio	8 : 100

→ $\boxed{8.0\%}$

Bayesian probability



MaSiSS

Probability

Bayesian probability

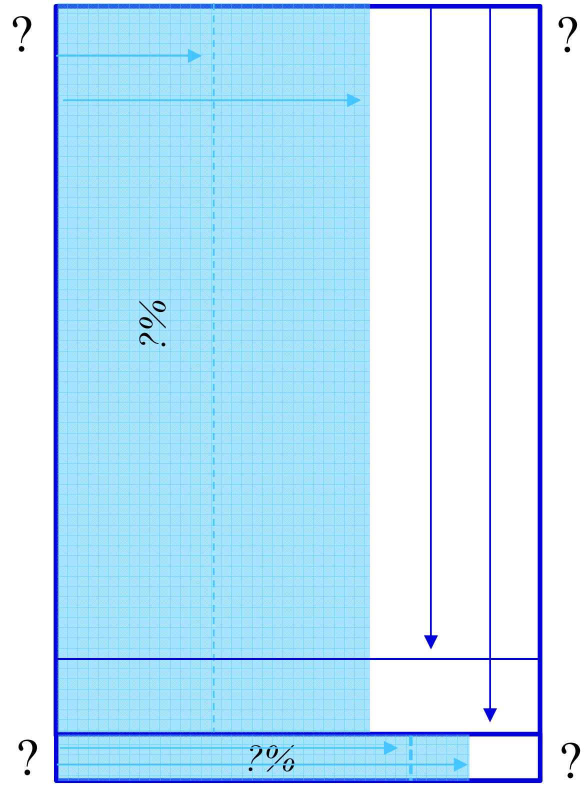
How much a colleague is jealous and complains?

Did I have more energy to the diet?

– Compering world in which the diet doesn‘t work to the world which it does

Somebody evaluate your work and thing it‘s great.

– How comman is to him that bad ideas are great?
– How much evidence is his approval that my work is great?



MUNI SPORT