MSc Applied Information and Data Science

Data Analytics for Data Scientists

Design of Experiments (DoE)

Lecture 03: Introduction to Design of Experiments (DoE)

2025

Prof. Dr. Jürg Schwarz

Program: 16:15 until 17:55

16:15	**Begin of the lesson**
	Lecture: Jürg Schwarz ◦ Why conduct trials / experiments? ◦ Variance as a basic concept ◦ Properties of measurement instruments ◦ Details and examples ◦ Preview of Lecture 04
	Tutorial: Students / Jürg Schwarz / Assistants ◦ Working on the exercise ◦ Support by Jürg Schwarz / Assistants
17:55	**End of the lesson**

Detailed table of contents → Slide 28

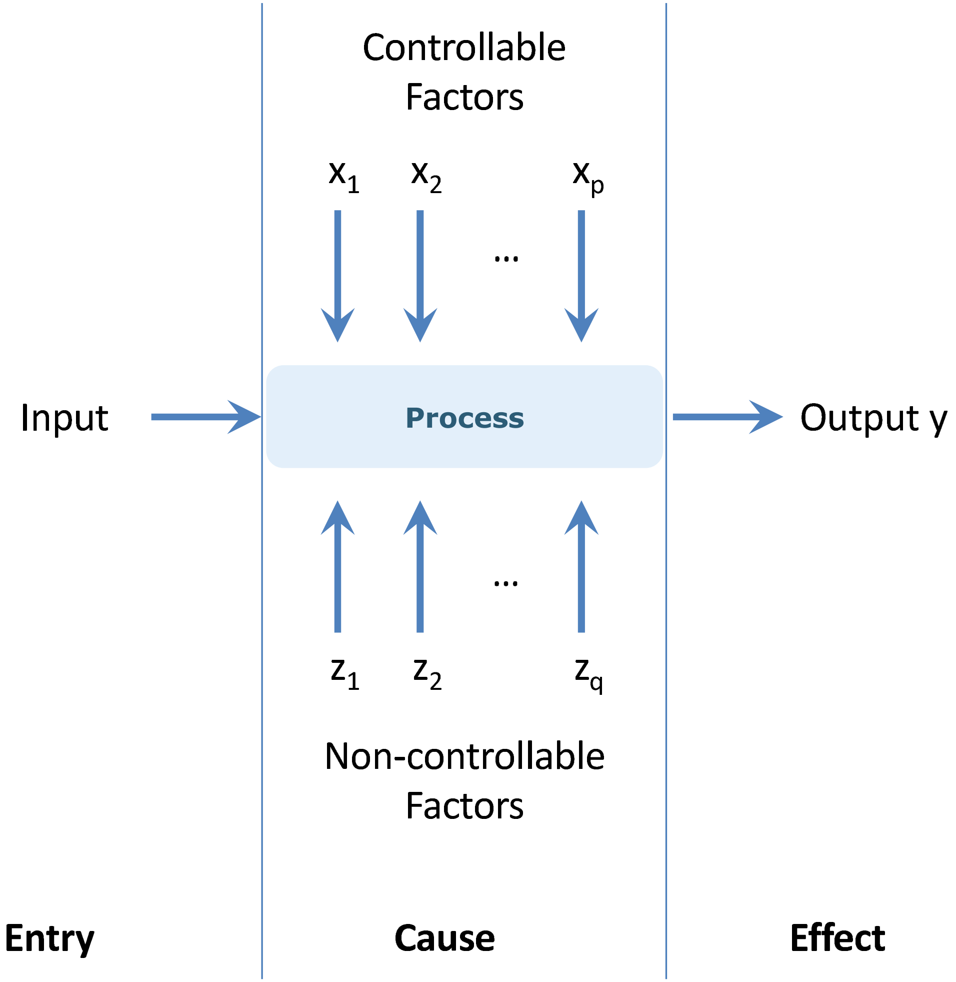
Why conduct trials / experiments?

Cause and effect (causality)

A **trial** / **experiment** is carried out to discover a cause-and-effect relationship in a process.

The lady tasting tea
A lady declares that by tasting a cup of tea made with milk she can discriminate whether the milk or the tea infusion was first added to the cup. We will consider the problem of designating an experiment by means of which this assertion can be tested.

Fisher in *The Design of Experiments* (1935)



A classic from R.A. Fisher

Factors affecting wheat yield

Agricultural experiment to study the effect of fertilizer on wheat yield.

Among other things, it was found that ammonium sulfate applied at four different concentrations, had a significant effect on wheat yield.

Table III.

Controllable factor
Fertilizer

Input
Wheat seeds

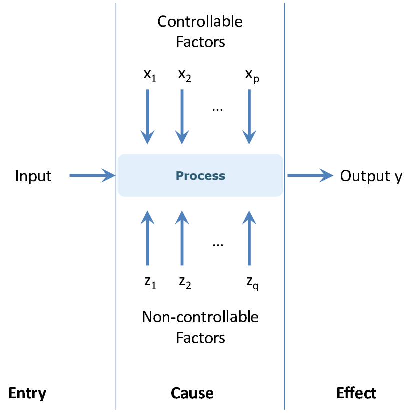
Plot	Mean yield (Bushels per acre)
5, no ammonia	14·18 ± ·44
6, single ammonia	22·58 ± ·71
7, double ,,	31·37 ± ·90
8, treble ,,	35·69 ± ·93

Output
Wheat yield

Non-controllable factors
Soil, Weather

Fisher (1921) / R.A. stands for Ronald Aylmer

Terms



Input

- Trial objects, test objects, test persons, probands, materials, …

Process

- Process in which controllable and non-controllable factors influence the input.

Output

- Input changed by the process, result of the test / experiment.
  Terms: End point, dependent variable (**DV**), target variable, target value, ...

Controllable factors

- Influencing factors whose strength can be adjusted within defined limits.
  Terms: Influencing variables, independent variables (**IV**), …

Non-controllable factors

- Influencing factors whose strength cannot be determined but that can be measured.
- Influencing factors whose strength cannot be determined and that cannot be measured.

Non-controllable factors

Non-controllable factors are also referred to as ...

- Nuisance variables in a general context
- Nuisance factors in the context of **blocking**
  Blocking = arranging of experimental units in groups (blocks)

Examples of influencing factors whose strength cannot be determined but that can be measured

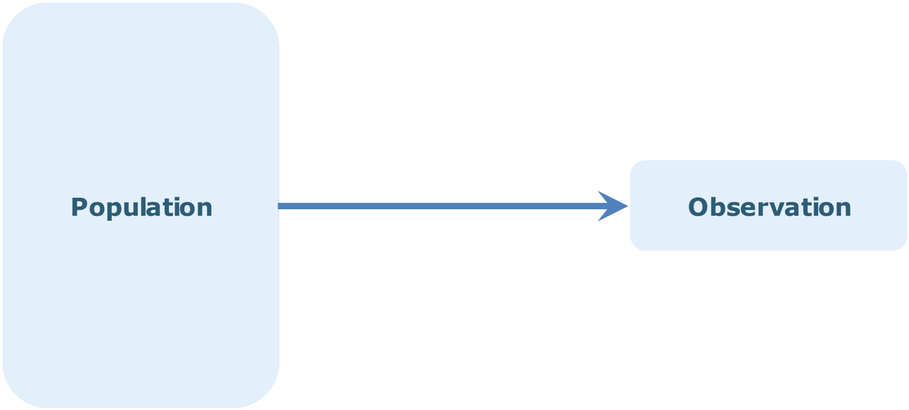
- General, using the example of agricultural science experiments
  Soil properties, weather effects, …
- With people as trial objects
  Body weight, socio-economic status, stress level, …

Influencing factors whose strength cannot be determined and that cannot be measured

- General
  Random fluctuations in output
  Random (quantum physical) fluctuations in measurement instruments

Causality in observational and experimental study designs

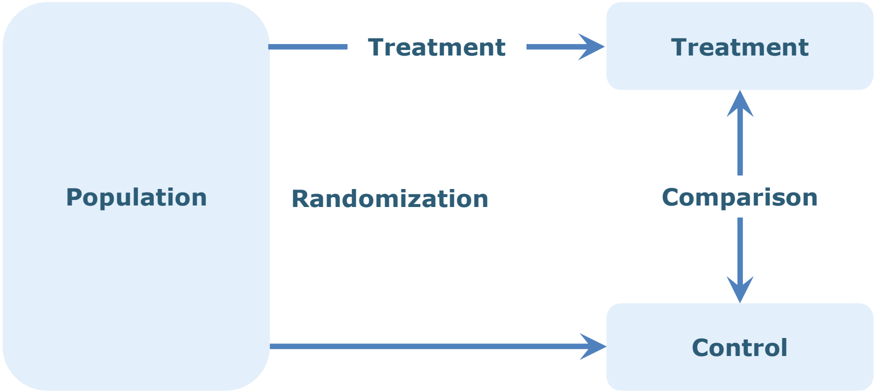
Observational study



Suitable for hypothesis **formation** (insufficient explanatory power for hypothesis testing)

Causality **cannot** be postulated

Experimental study



Suitable for hypothesis **testing**

Especially
Randomized Controlled Trial → **RCT**

Causality **can be** postulated

Examples of observational and experimental studies

	Observation	Experiment
Field	<u>Economic research</u> Indicators: GDP, Inflation, … <u>Social research</u> Indicators: Human development index, …	<u>Medical research</u> Field trial of Polio vaccination, … <u>Online market research</u> A/B testing of websites, …
Lab	<u>Physical research</u> Hubble Space Telescope, ... <u>Psychological research</u> Conditioning (Pavlov's dog), …	<u>Behavioral research</u> Stanford-Prison-Experiment, … <u>Market research</u> Eye tracking website optimization, …

Variance

Variance as a basic concept

Concept of variance in the context of statistics

In descriptive statistics, the empirical variance – also referred to as sample variance, or just variance – is a key figure of a sample.

The variance belongs to the dispersions and describes the mean square deviation of the individual measured values from the empirical mean.

The root of the empirical variance is the empirical standard deviation.
The empirical standard deviation is the most common measure of variation.

Concept of variance in the context of analysis of variance (ANOVA)

The analysis of variance was developed by **R. A. Fisher**.

ANOVA describes statistical procedures for analyzing data and testing structures.

Variances and test statistics are calculated to gain insight into the regularities in the data:
The variance of one or more dependent variables is explained by the impact of one or more influencing variables (factors).

Fisher (1918) / After Wikipedia

The term variance in the context of design of experiments

Primary variance

- Impact of (experimental) factors in an experiment on the change / variation of the output to be examined.

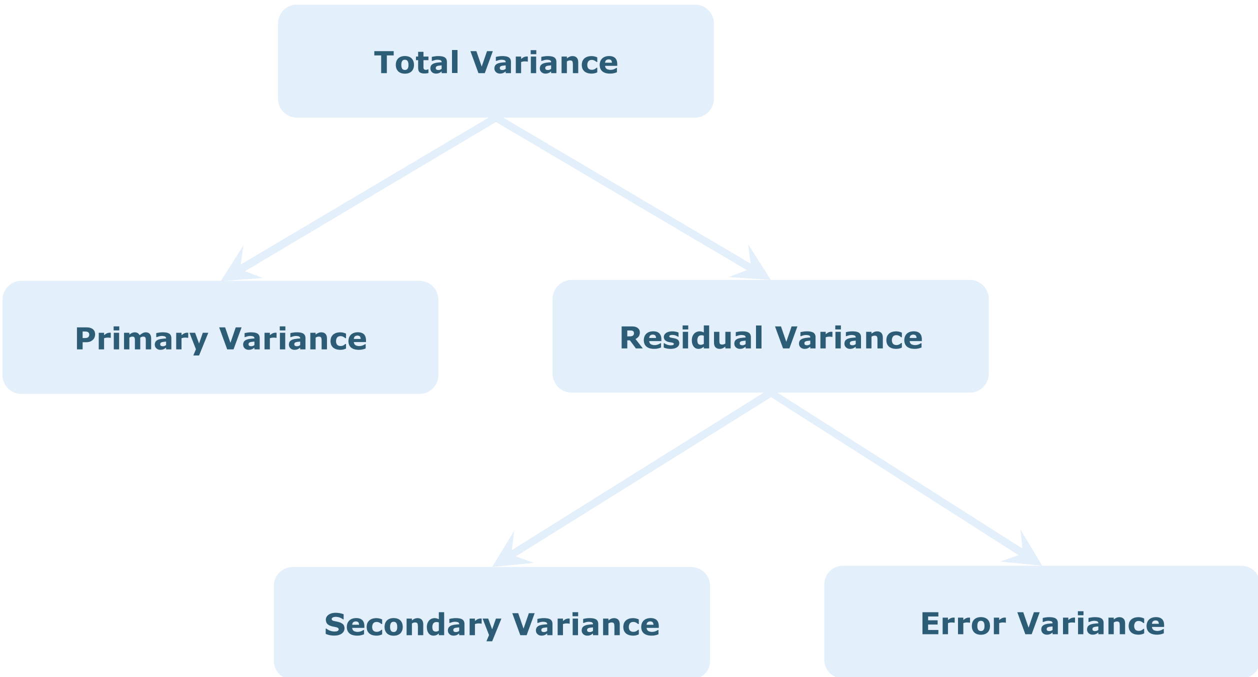
Secondary variance

- Variation of the output to be examined, caused by nuisance variables. Not in the focus of the study.

Error variance

- Variation caused by measurement errors and random processes.

Residual variance



Variation of the variance

Goal: The variance of the dependent variable **(DV)** (primary variance) should be attributed to the systematic variation of the independent variable **(IV)**.
The secondary variance should be controlled and the error variance minimized.

Primary Variance	**Maximize by ...**
Systematic change of the **DV**, which is **only** due to the change of the **IV**.	◦ Selection of extreme values in the **IV** ◦ Selection of optimal increments of **IV** ◦ Selection of many levels of **IV**
Secondary Variance	**Control by ...**
Systematic change of the **DV**, due to the effect of nuisance variables, but **not** to the change in **IV**.	◦ Elimination / Keeping constant ◦ Repetition / Randomization / Blocking ◦ Nuisance variable as covariate
Error Variance	**Minimize by ...**
Variation of **DV** is **not** due to the change in **IV** and **not** due to the influence of nuisance variables.	◦ Reliable measurement instruments ◦ Measurement repetition / Sample size ◦ Suitable analysis methods

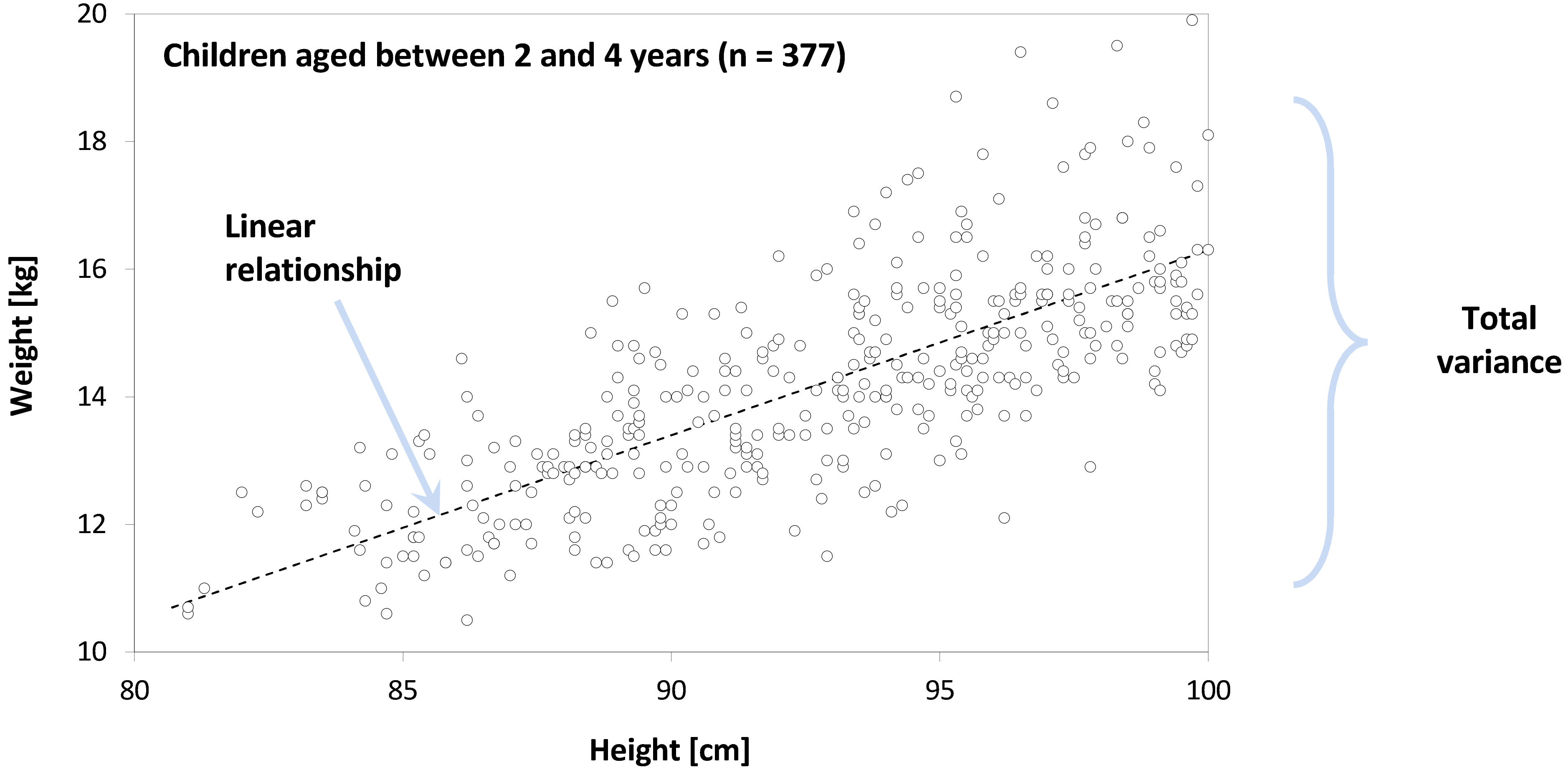
This overview is discussed in detail below.

Example – Body-growth in children

Relationship between height and weight

Sample of the *National Health and Nutrition Examination Survey* (approx. 9,500 persons)

Linear relationship: The bigger a child is, the heavier it is.

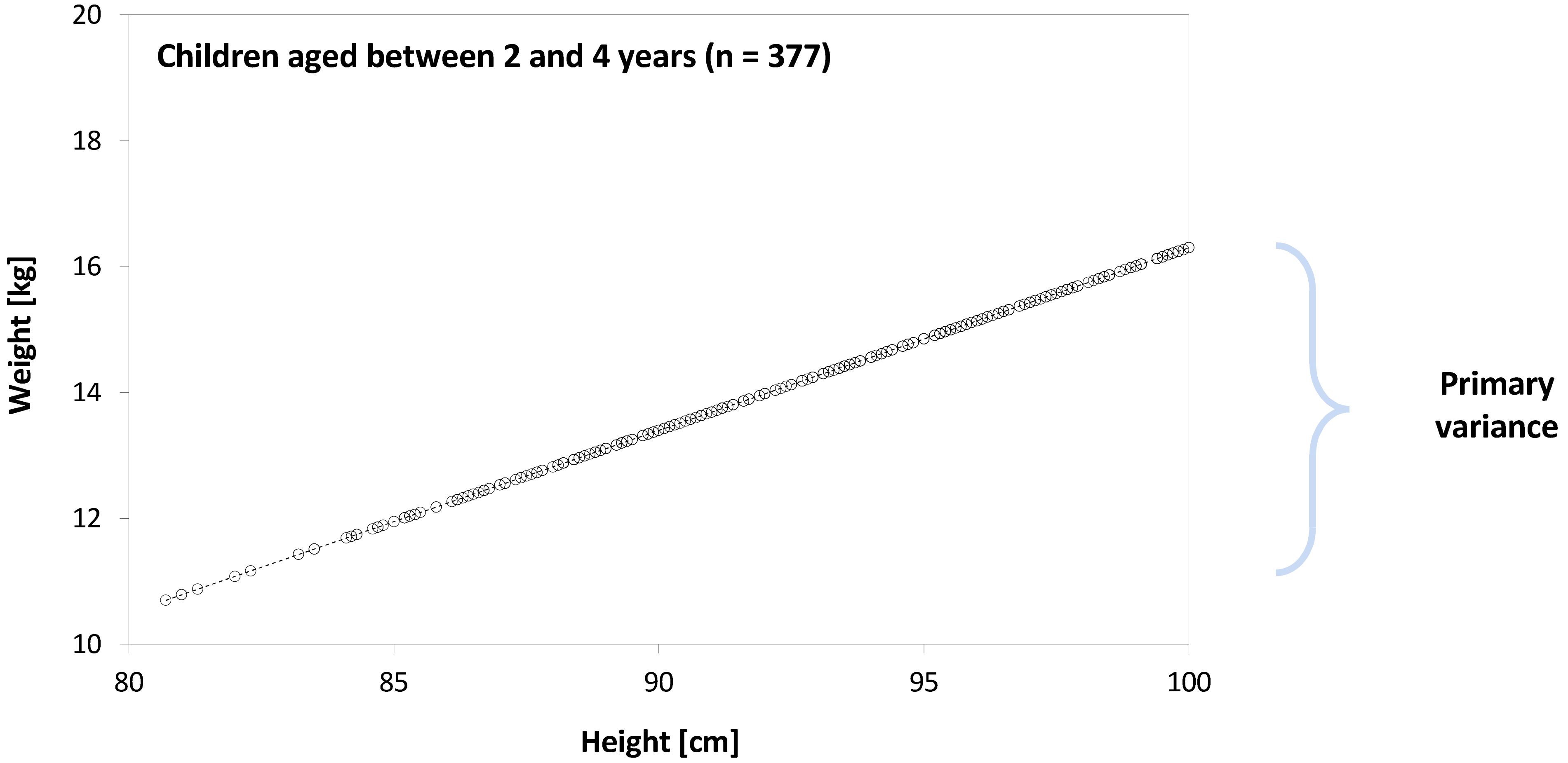


Primary variance

In theory, growth in children follows a biological law.

Body weight changes systematically when height changes.

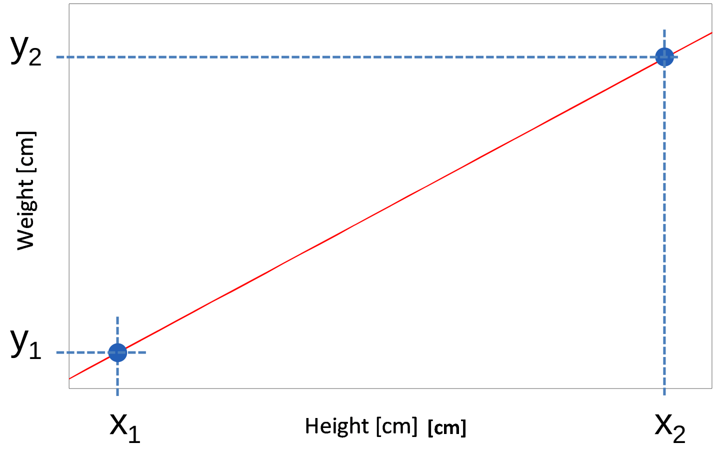
Linear relationship: The bigger a child is, the heavier it is.



Maximizing the primary variance

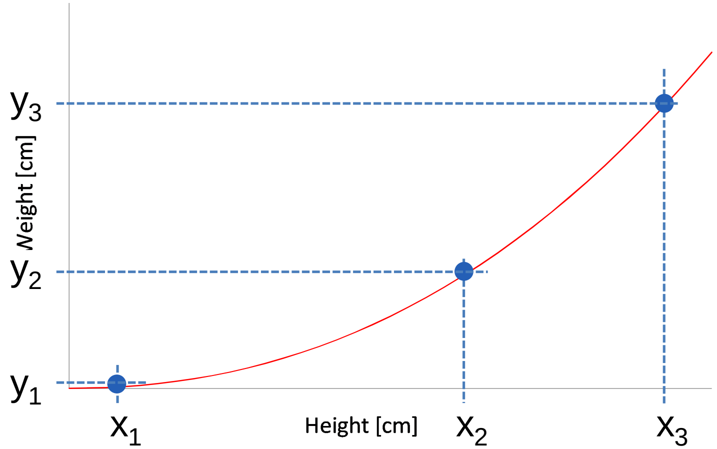
If the relationship is linear, the primary variance can be maximized by

- Selection of extreme values in the **IV**



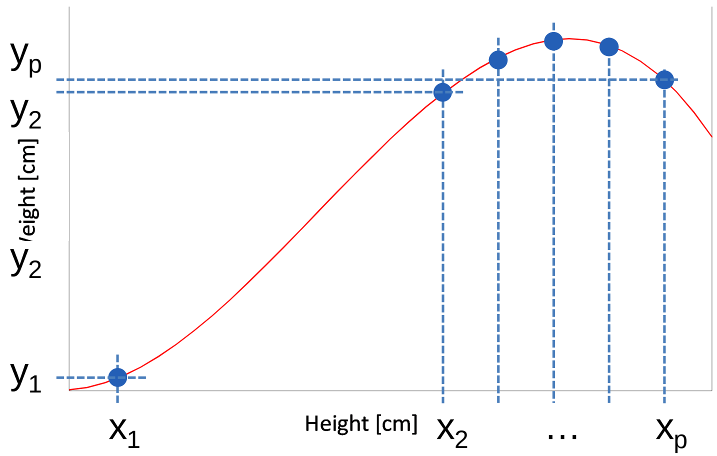
If the relationship were curvilinear, the primary variance can be maximized by

- Selection of optimal increments of **IV**



If the relationship were unknown,
the primary variance could be maximized by

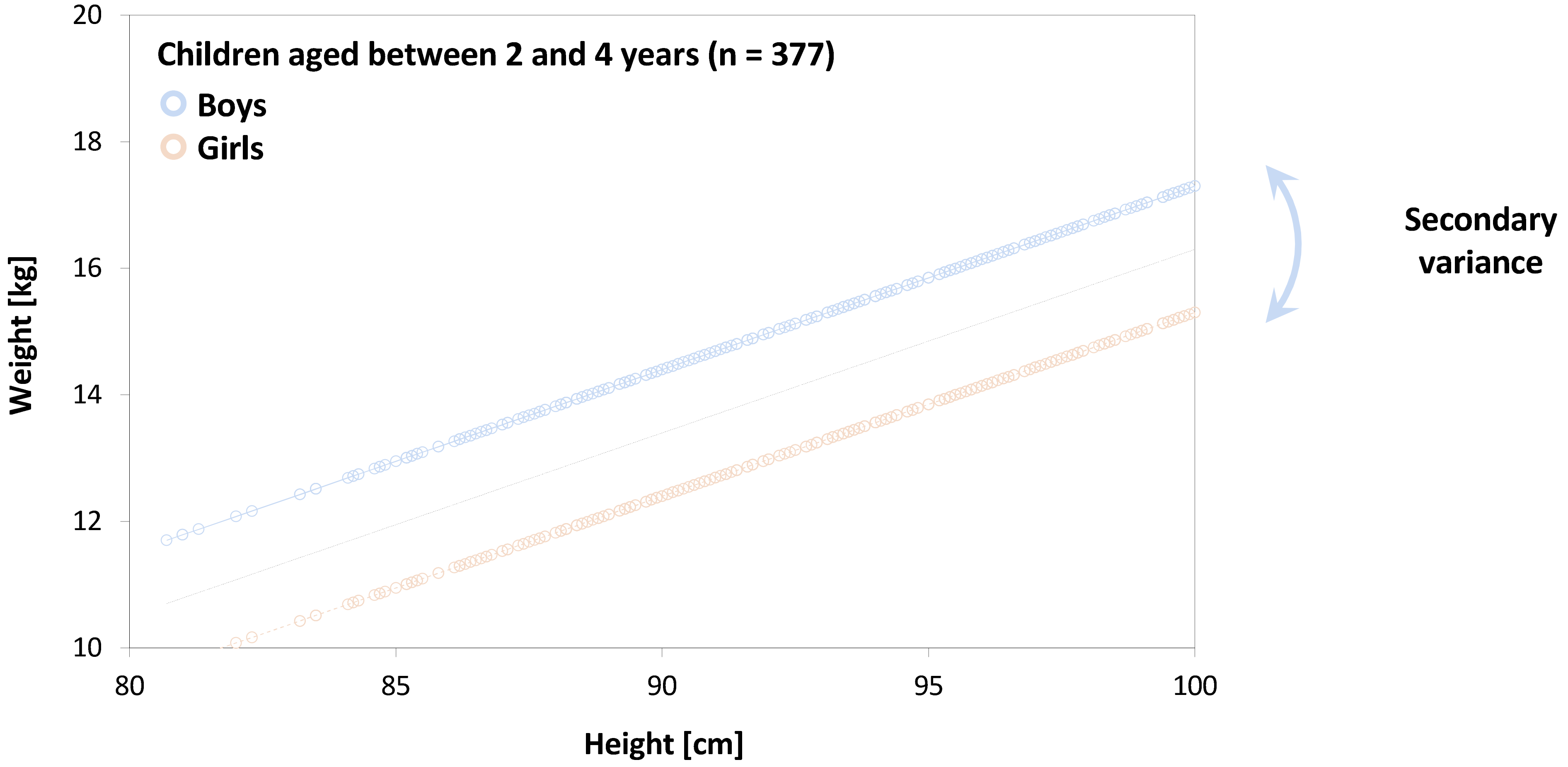
- Selection of many increments of **IV** and the smallest possible steps



Secondary variance

Boys and girls differ in the level of body weight.

The sample can be divided into two groups.



Control of the secondary variance

- Elimination / Keeping constant
- Repetition / Randomization / Blocking
- Nuisance variable as covariate

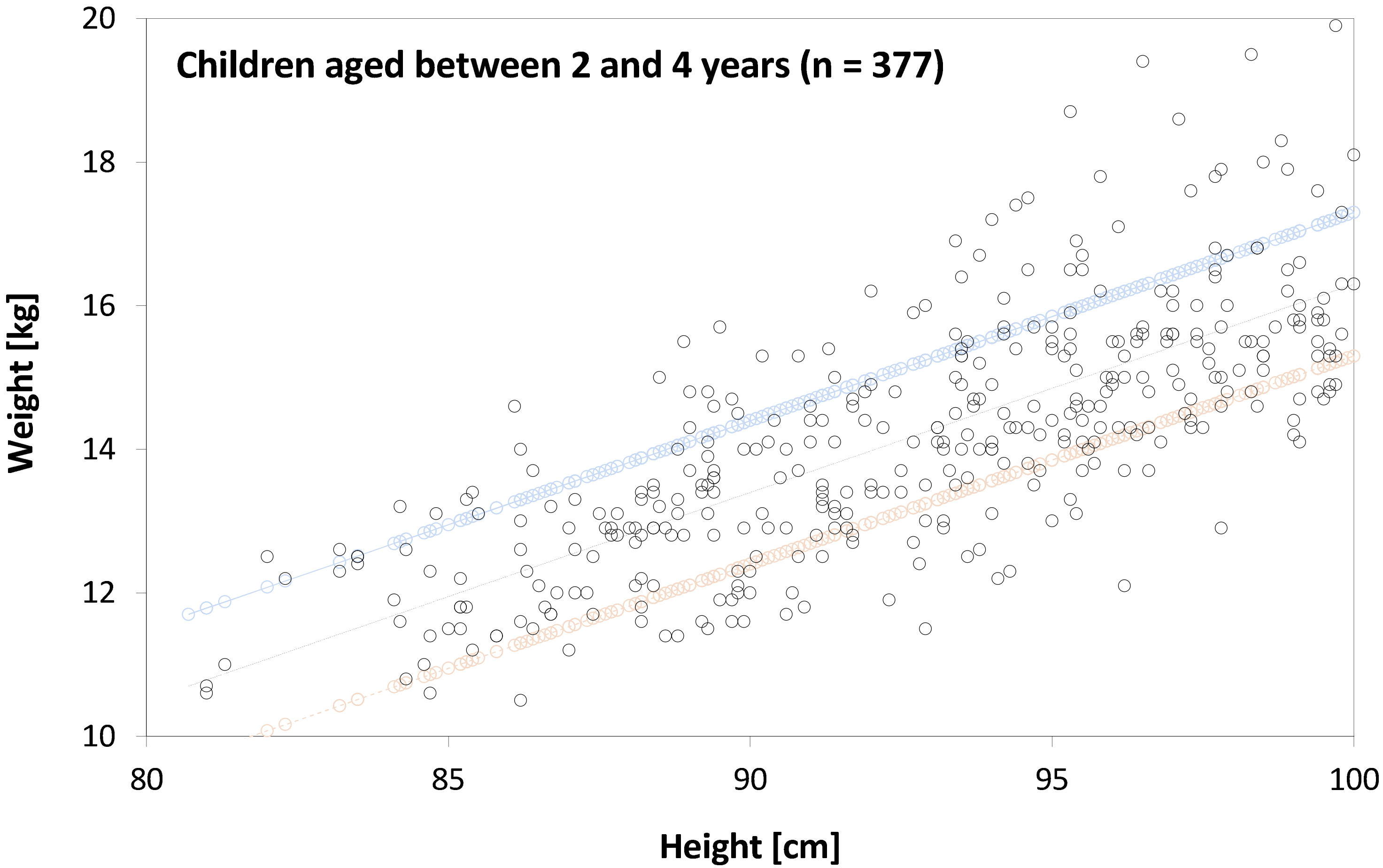
Methods for controlling secondary Variance (Individually or combined) …

- Keeping the experimental setup constant ↔ Particularly possible in laboratory experiments
- Repetition
  Several measurements are repeated on the same probands / trial objects.
- Randomization
  Trial objects are assigned randomly to *Treatment* and *Control* groups
  to eliminate systematic bias.
- Blocking
  Trial objects are grouped into homogeneous blocks based on one or more influential variables to reduce variability.
- Covariate adjustment
  Nuisance variables are included as covariates in the statistical model
  to account for their effects.

Error variance

Error variance can mostly be attributed to these two causes …

- Basically every measurement contains errors (by random processes and by measurement)
- Not only body size has an influence on body weight.
  There are many other unknown factors that can also influence weight.



Error variance
=
Random fluctuations that cannot be described by the primary variance or by the secondary variance

Minimizing the error variance

- ◦ Reliable measurement instruments
- ◦ Measurement repetition / Sample size
- ◦ Suitable analysis methods

Methods for minimizing error variance (Individually or combined) …

- Reliable measurement setup / measurement instruments
  - Standardization of the experimental conditions
  - Use of standardized equipment and procedures / staff training
  - Selection of high-quality measurement instruments
  - Calibration and verification of measurement instruments

- Measurement repetition / Sample size
  - Repeated measurements help average out random errors, improving reliability
  - Larger sample sizes reduce the impact of individual measurement errors

- Suitable analytical methods
  - Use of robust estimators to account for heterogeneous error variance

Properties of measurement instruments

Objectivity

Objectivity of an instrument is given when the results are **independent** of personnel and calculation methods.

Example
- Regardless of who reads a weighing scale, the result is the same.
- Negative example: Reading error due to different viewing angles.

Reliability

Reliability is the degree to which an instrument produces the same result **each time** under comparable conditions.

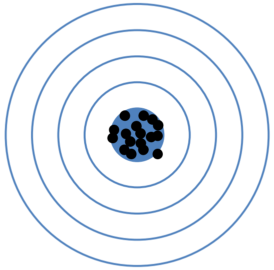
Example
- Weighing scale that always produces the same result at the same weight.
- Negative example: Use of a weighing scale that change over time.

Validity

Validity is the extent to which an instrument measures what was **intended**.

Example
- Weighing scale that leads to the measurement result of 75 kg at 75 kg.
- Negative example: Use of a weighing scale with insufficient accuracy class.

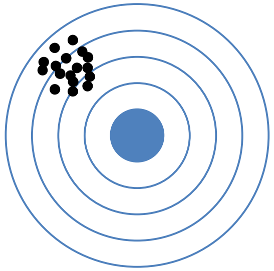
Hierarchy of validity and reliability (Analogy: Shooting at a target)



Valid and reliable

A valid instrument is always reliable.

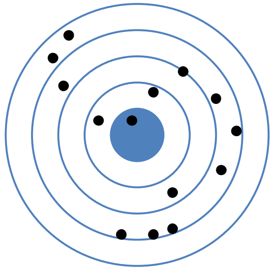
Validity is the most comprehensive criterion.



Reliable, but not valid

A reliable instrument is not necessarily valid.

(It does not necessarily measure what should be measured.)



Neither valid nor reliable

Note

This is the general concept of validity.

In Lecture 04 "Properties of DoE" we will look at the topic again in a different context.

In "Quality criteria of experiments" internal validity and external validity will be distinguished.

Further examples of how to control secondary variance

Control by … ◦ Repetition / Randomization / Blocking

Nuisance factors often cannot be completely eliminated or kept constant.

Blocking can help ensure equality of experimental conditions with respect to a nuisance factor.

General procedure

- Forming homogeneous blocks based on one or more nuisance factors.

  The block properties may be naturally occurring (e.g., age groups) or artificially determined (e.g., grouping by pre-test scores).

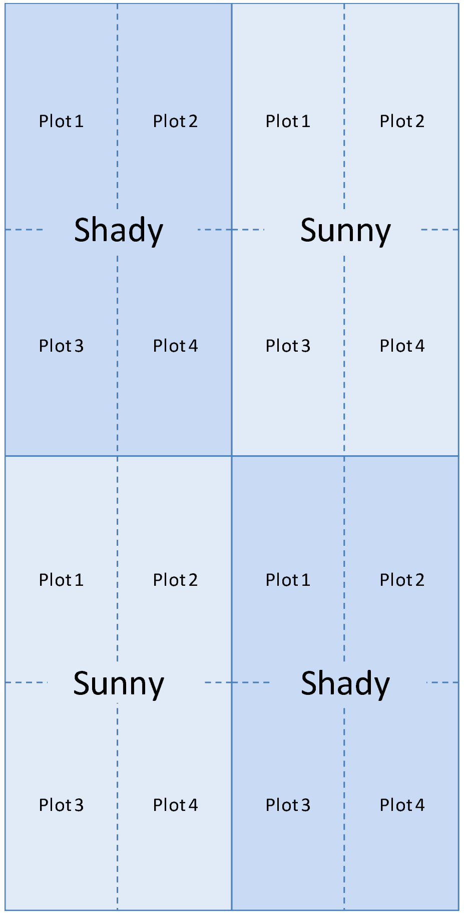
- Assigning experimental units to these homogeneous blocks.

- Randomly distributing the treatment conditions among the experimental units within each block to balance out systematic effects.

Example of blocking: Crop yield from wheat growing

Effect of two fertilizers (+/–) on the yield of two types of grain (+/–).

Problem: The areas of the field varies in terms of solar radiation → nuisance factor



→ **Block factor**
Shady = 1
Sunny = 2

Block	Plot	Fertilizer	Grain
1	1	+	+
1	2	–	–
1	3	+	–
1	4	–	+
2	1	+	+
2	2	–	–
2	3	+	–
2	4	–	+

In each block, 4 plots are created with all combinations of the two independent variables.

Control by means of … ◦ Nuisance variable as covariate

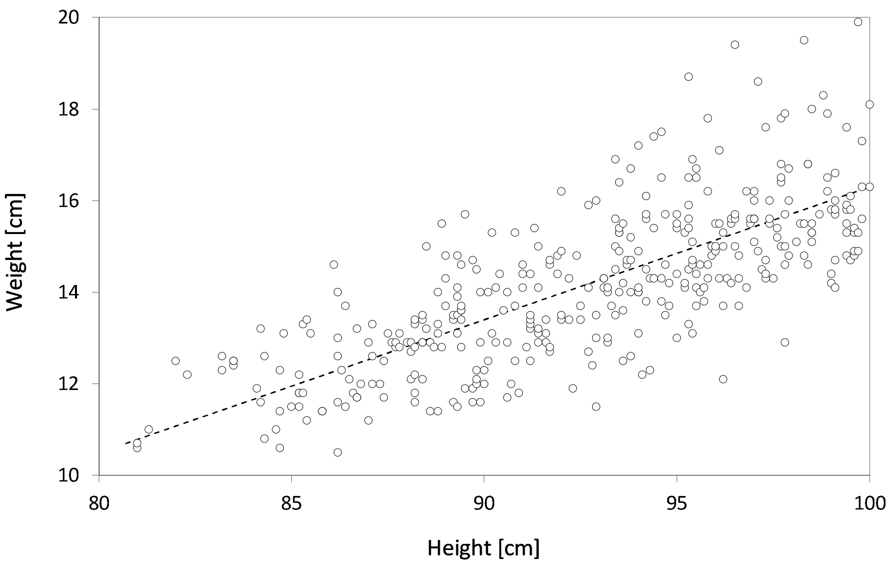
Conversion of a nuisance factor into an IV using the example of body growth in children

Primary variance is modeled using regression analysis.
The nuisance factor "sex" is included as an additional independent variable (control variable).

Model without sex as IV
Both sexes together

$y = \beta_0 + \beta_1 \cdot x$

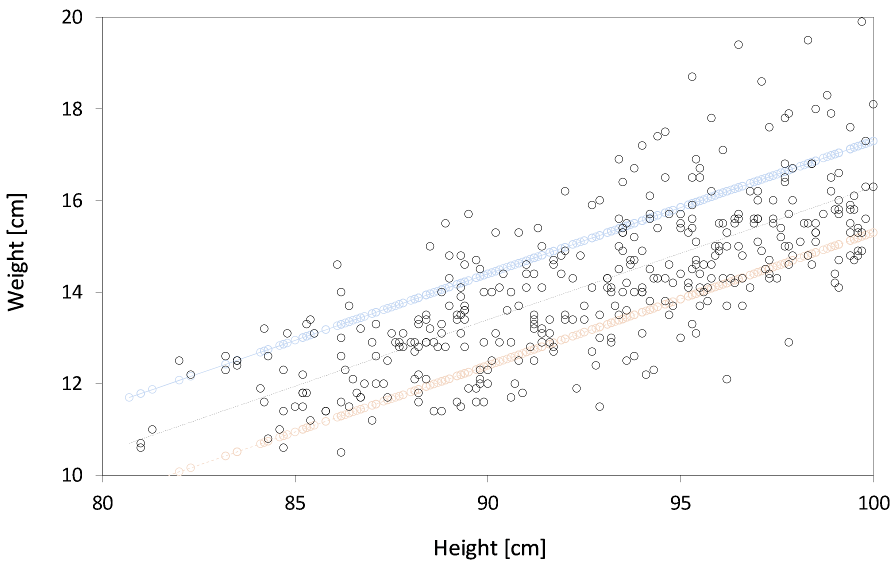
Model accuracy $R^2 = 0.557$



Model with sex as additional IV
Sexes separated

$y = \beta_0 + \beta_1 \cdot x + \beta_2 \cdot sex$

Model accuracy $R^2 = 0.601$



The relationship between height and body weight is estimated more precisely
↔ Secondary variance is controlled through the inclusion of sex as an additional IV.

Example of how to minimize error variance

- Reliable measurement instruments
- Measurement repetition / Sample size
- Suitable analysis methods

Ensuring measurement reliability

The *National Health and Nutrition Examination Survey* implements strict procedures to enhance measurement reliability by training medical staff and standardizing testing conditions.

Standardized Data Collection:

- Data were collected in the Mobile Examination Center (MEC) by trained health technicians.
- Rooms in all MECs were identical in layout and equipment to maintain consistency.

Quality Assurance and Control Measures:

- Equipment calibration was performed by health technicians and verified by supervisory staff.
- The manual provides detailed quality assurance and quality control protocols.

Consistent Measurement Procedures:

- Arm and leg measurements were taken on the right side of the body to ensure standardization.
- If a participant had an amputation, medical condition, etc. preventing measurements on the right side, the technician measured the left side instead.

wwwn.cdc.gov/Nchs/Nhanes

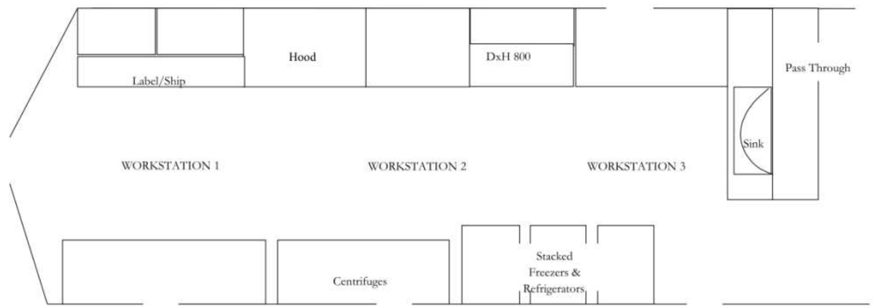
National Health and Nutrition Examination Survey (NHANES)

MEC Laboratory Procedures Manual

Ja

Table of Contents

Figure 1-1. Laboratory layout



719 pages (!)

Overview of "variation of variance"

Total Variance = **Desirable** + **Undesirable**

Primary Variance	**Seconary Variance**	**Error Variance**
Independent variable	**Nuisance variable**	**Measurement etc.**
Maximize	**Contol**	**Minimize**

Preview of Lecture 04

What has happened so far

The **question** "Why conduct trials / experiments?" has been clarified ↔ Causality

The **deepening** "Variance as a basic concept" lays the foundation.

Properties of measurement instruments have been covered.

What follows in Lecture 04

Design of experiments – an overview

- Placement of the experimental design into the phases of the research process
- Elements of design of experiments

Properties and functions of the design of experiments, in relation to …

- Goals
- Content
- …

Table of contents