SEE
Measurement as Process

Paul G. Plöger

Björn Kahl

Praveen Ramanujam

Outline

General remarks (on ALL of SEE)

1 Prelude

2 Foundations

- Basic Terms and definitions (2B)
- Errors (2E)
- Distributions (2D)

3 Summary

Labwork

General: Target

Learn how to design and conduct experiments decently

Learn how to present results decently (a.k.a. in a scientifically sound way)

	Precision	Recall	Time taken (s)
Instantanuous diagnosis			
Control run	-	-	1
Exogenous Intervention run	1	1	1
Progressive diagnosis			
Control run	-	-	1
Exogenous Intervention run	1	1	1

TABLE 6.1: Calculated precision and recall of exogenous intervention diagnosis in a simulated youbot. The calculation was done on a specific time interval when the exogenous interventions was occurring.

4,8376 +/- 0.1

	Assigned positive	Assigned negative
Actual positive	60	12
Actual negative	48	0

Table 5.2: Error matrix for PADI dataset

	Assigned positive	Assigned negative
Actual positive	103	4
Actual negative	13	0

Table 5.3: Error matrix for NUS dataset-I

General: Knowledge areas

- Theory of Measurement
  (“Messwerterfassung”)
- Mathematics
  - Probability theory
  - Statistics
- Physics
  - Instrumentation

General: Overview

PGP	30.09.13	lecture	Measurement as process
		Testat:	-none-
PR+PGP		Lab class	Run robots on sheet of paper, record distributions
you		Homework	Evaluate data, Do the fit
PR	07.10.13	lecture	Sample solution for 30.09.13
PGP+BK		Testat:	successfull Reprodution a normal distribution in 2D and proof that what you observed is actually a normal distribution, error analysis
PGP+BK		Lab class	read Zhang paper, explain it to newbies in Vision and projective geometry
you		Homework	finish undestanding of Zhang paper
PGP+BK	14.10.13	lecture	Intrinsic / extrinsic camera parameters, radial distortions, calibration
		Testat:	-none
PGP+BK		Lab class	run the CALTECH sw, do the calibration images, find intrisic / extrinsic params for your camera
you		Homework	Read Thrun / Probabal. Robotics / chapter 5.1, 5.2, 5.3
BK	21.10.13	lecture	Sampling of distributions and forward / backward model (Use of Prob. Models in mobile robotibs)
PR+BK		Testat:	correct intrinsic / extrinsic camera matrixes (compare on class level)
PR		Lab class	run the V,\omega experiemtent
		Homework	read 2 articles: On the repseentation and estimation of spatial uncertainty // Location estimation und uncertainty analysis for mobile robots
PGP	28.10.13	lecture	Covariance matrices and error propagation
		Testat:	-none
PR		Lab class	complete the experiment
you		Homework	Read Thrun / Probabal. Robotics / chapter 5.4, 5.5, 5.6
	04.11.13	lecture	-none
PGP+BK		Testat:	good v,\omega model für robot
PR		Lab class	do the ODOmetry Modion model // ticks_left ticks_right
you		Homework	Read Thrun / Probabal. Robotics / chapter 6

General: Overview

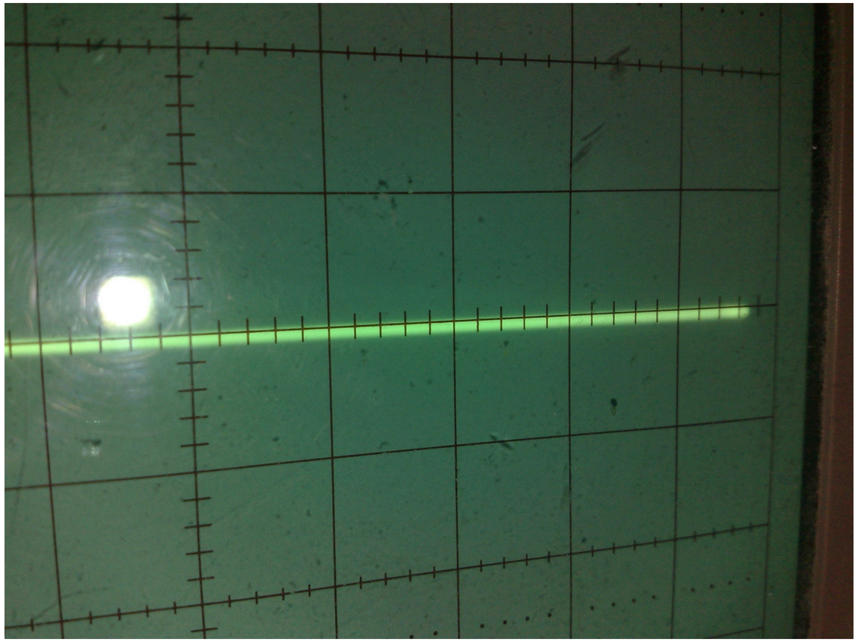
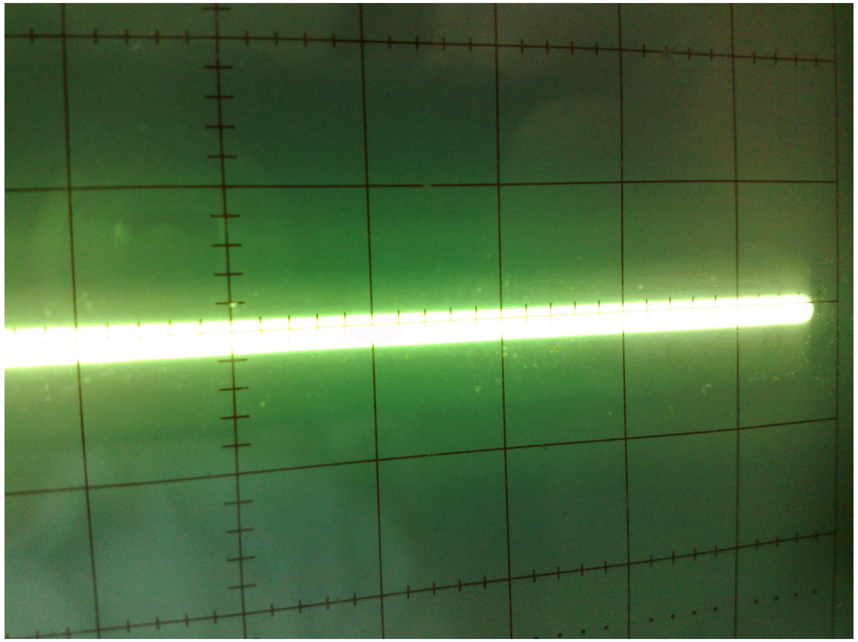
PGP	11.11.13	lecture	EM and Xpero Simulator, explanations
PGP+BK		Testat:	good ODO model
PR		Lab class	Build sonar radar on NXT robot, start experiemtens in cage
you		Homework	Refactor my matlab code in python (class project)
PGP	18.11.13	lecture	-none
PGP+BK		Testat:	-none
PR		Lab class	finish experiments on sensor model
you		Homework	read article: ROC Graphs: Notes and Practical Considerations for Researchers
BK	25.11.13	lecture	Confusion matrix and ROCs
PGP+BK		Testat:	good sensor model in XPEROSim
PR		Lab class	setup SVNs, learn hand-written digit classifier
you		Homework	read CT articles, hand in data piles from your R&D experiments
PGP	02.12.13	lecture	Data Visualization // based on Tufts book
PGP+BK		Testat:	good ROCS and explanation / comparisiosn: which classifier is best
PR		Lab class	get your own data from R&D, start visualize them
you		Homework	make nice diagrams
you	09.12.13	lecture	Explain your R&D data representation problem
PGP+BK		Testat:	on nice diagrams
		Lab class	TDB
you		Homework	TDB
you	16.12.13	lecture	Explain your R&D data representation problem
PGP+BK		Testat:	on nice diagrams
		Lab class	TDB
you		Homework	TDB

General: Overview

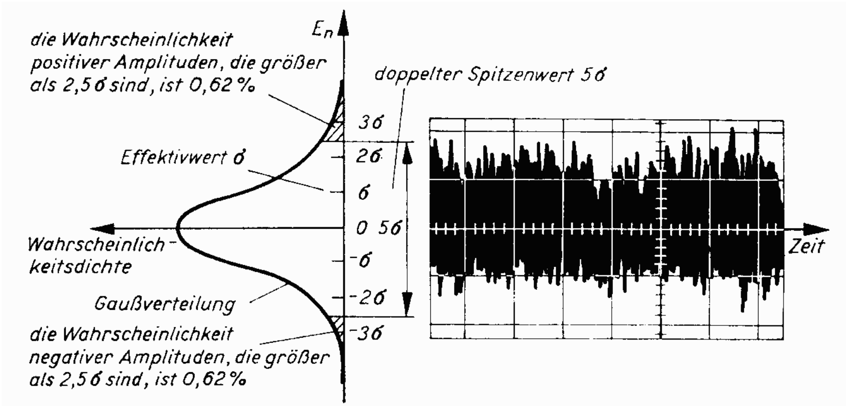
BK	06.01.14	lecture	you bot as a learning tool
		Testat:	-none
PR		Lab class	run youbot to get circle / square in air, record sensor traces
you		Homework	compare true ground to MEMS values and vision
BK	13.01.14	lecture	you bot as a learning tool
		Testat:	-none
PR		Lab class	run youbot to get circle / square in air, record sensor traces
you		Homework	compare true ground to MEMS values and vision
BK+PGP	20.01.14	lecture	resume / spare lecture
BK+PGP		Testat:	closing the gap between the sensor traces
PR		Lab class	none
you		Homework	none

- 75% must be solved and be attested (i.e. 6 out of 8)

1: Oscilloscope experiment

1: Questions



What is the true value („w"ahrer Wert x_w?)

Why can't we catch it?

If not catchable, how can we approximate it best?

How can we be sure that the measuring instrument did not add the noise? (called „burden")

What is maximum impression, is this instrument driven? How?

How to validate that the values follow a distribution?

How to formalize the process as such?

How to interpret the action in physical terms?

1: needed Sciences

Mathematics

- errors: absolute, relative, systematic, random, round off
- (Random variables)
- (Probability distributions)

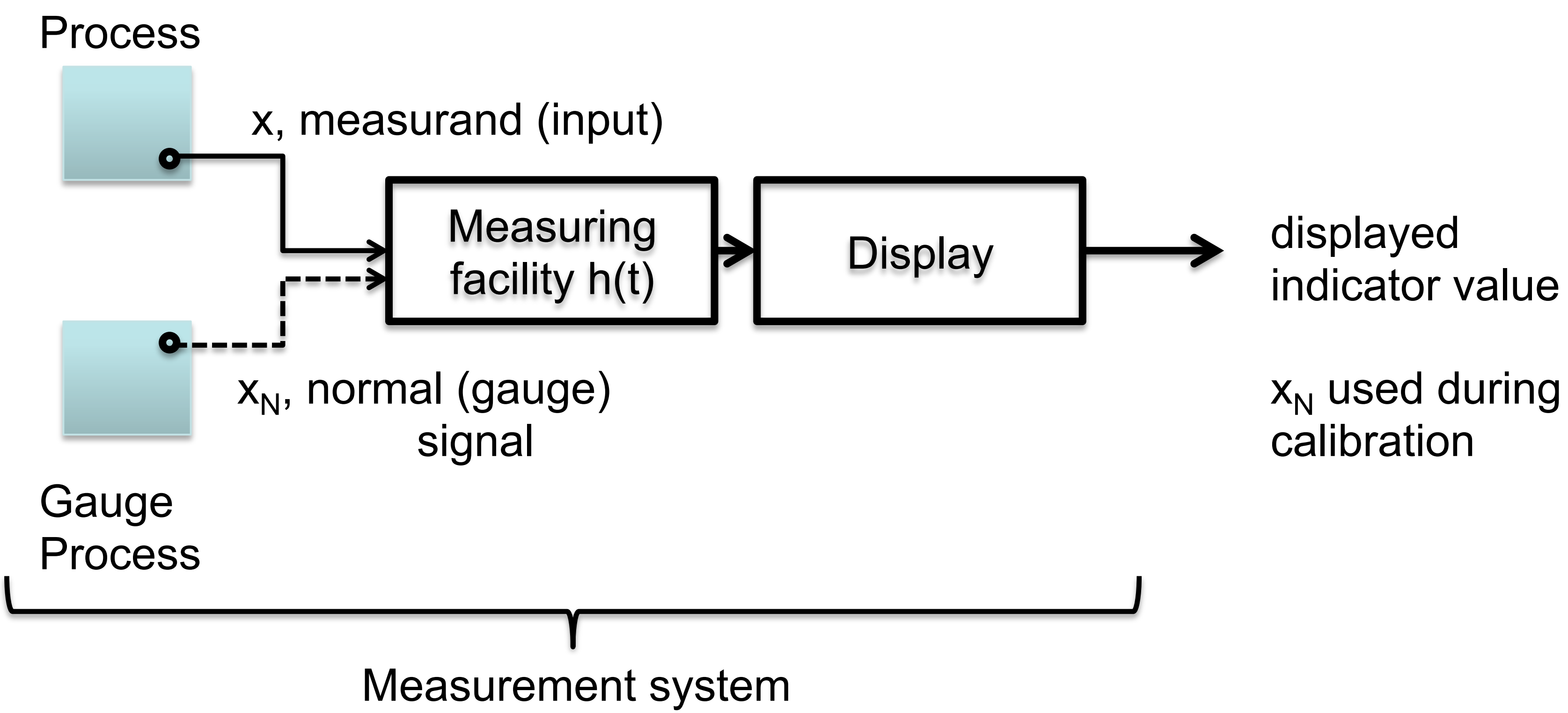
Statistics

- Empirical mean value
- (Maximum likelihood estimation)
- Statistical tests

Physics

- (Burden of instrument)
- (Equivalent circuit diagram)
- Measurement as signal flow graph

2B: System theory: Measuring System



x(t) == measurand
h(t) == transfer function of measuring facility (linear) =>
$x_a(t) = h(t) * x(t)$

2B: Formalization general terms

Measurand:
The measurand is **physical quantity, which is to be quantified** by the measurement (e.g. length, pressure, electrical resistance, etc.).

Measurement:
To „measure“ in the strict sense is, to **determine the size of a variable as a multiple** of a to be determined generally recognized unit size of the same physical dimension through experimental comparison with a solid measure of this unit.

To „measure“ in wider sense is, to determine experimentally on quantitative information of a particular property of a test object, which may be derived from one a or more measurable quantities of this test object. This property may be a size (for example a force or efficiency), or a mapping (for example a development over time). This property can also describe a system state, characterized via several variables which are recorded in parallel .

e.g. Count:
Counting is to **determine the number of elements** (e.g. piece count) or events (e.g. RPM). The count represents a special case of measuring.

e.g. Inspection:
Inspecting means to determine experimentally whether a **particular property of a test piece corresponds to the specified requirements**. If the check is carried out with the help of measurements, it can be considered a special form of measurement .

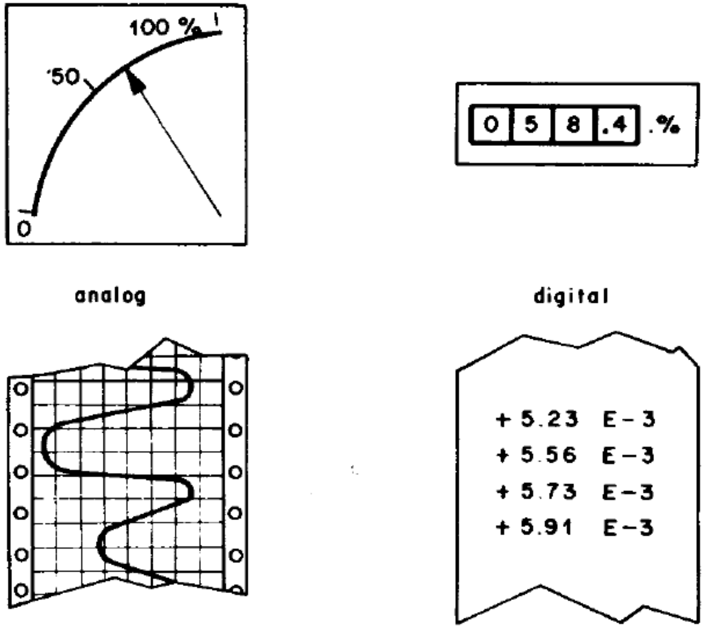
Display:
The **indicator shows the numerical value** of the to be measured variable. In analog displays one has to determine the **state of the indicator (arrow) on the scale by „reading“.** In digital displays it can be read directly as a number.

(Display) Range:
The display range of a measuring device is that range of **measured values, which can be read from the scale of the display (full scale, upper limit)**

Measurement range:
The range is part of the display area for which the **error remains within the guaranteed or prescribed limits of error.**

Suppression area:
The suppression area is that **area of measured values above which the display of the measuring device starts**.

2B: Example Displays



2B: Formalization general terms

Measured (Quantity) Value :
This is the **value of a specific variable. It will be determined from the displaying measuring instruments by reading this display and building the product of measured value and the unit of the measurand** (eg 3m). It can also be output in the form of a transferable measurement signal and may be sent for further automatic processing (storage, value processing, control, etc.).

Measurement result:
The measurement result is obtained **in general from several measured values by the help of a predetermined relationship**. In the simplest case, a single measurement is already the measurement result.

Device Under Test (DUT):
The DUT is **that part of a physical system, which carries the measurement variable** .

Measuring facility:
The measuring facility is **the entirety of device components used for the purpose of the measurement**. This includes sensor for detection of the measured quantity, amplifiers, computing devices and the output devices to display the observed value and possibly other components.

Measuring System:
The measuring system includes not only the measuring facility but also **those areas of the physical system which contains the DUT, which do affect the measurement process,** in particular the data acquisition .

Meter:
The meter is a either a part or the whole of the measuring facility.

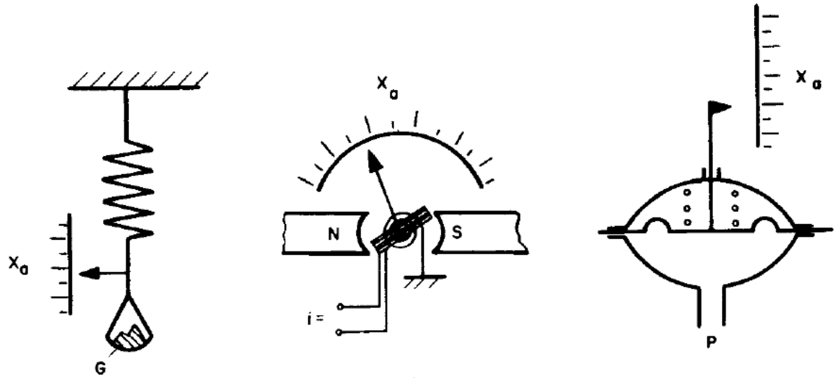
Measuring Principle:
The principle of measurement is **defined as the characteristic physical phenomenon which is used during the measurement** (for example, measurand: temperature, measurement principle: linear expansion, or thermoelectric effects, etc.).

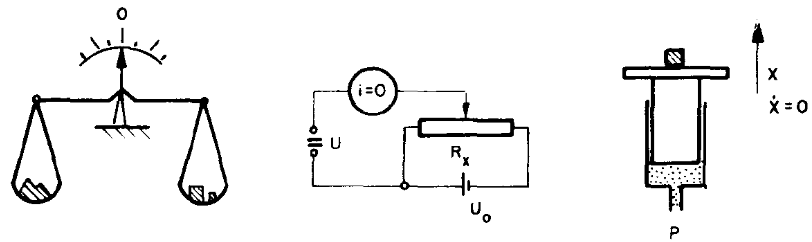
Measuring method:
Understand by this the kind of application or **implementation of the Measurement Principle**. In addition it may mean the function of the measuring device .

Sensitivity:
The Sensitivity is the **indicator pathway on the scale (in mm) per unit of the variable**. In light pointer devices, the vector length is defined as 1m. In digital instruments, the sensitivity is equal to the number of digits per unit of measured quantity. For non-linear meters the sensitivity is a function of the measured value or the display respectively.

2B:Example Methods

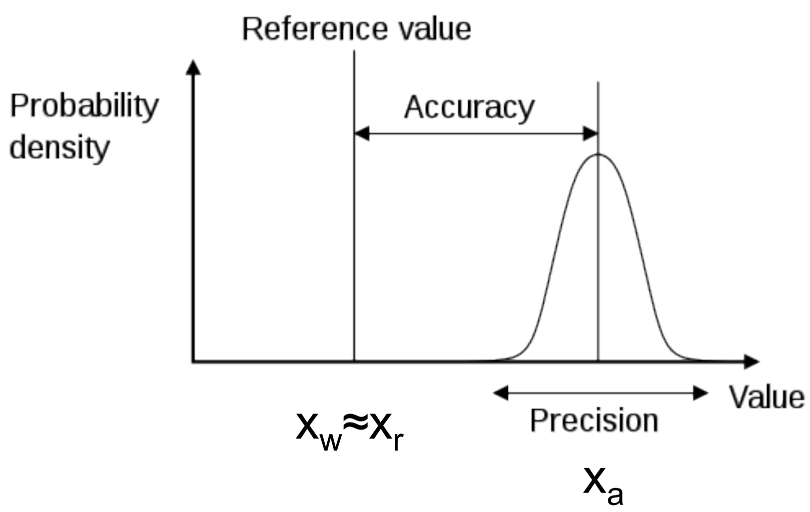
Methods:
- deflection of armatur
- compensation





2B: Accuracy and Precision

x_w “wahrer“ value (reference), i.g. unknown, e.g. distance too wall
x_r “right“ value, as close as we can get (assume: fault free measurement)
x_a “angezeigter“ value, as displayed



(Measurement) accuracy

closeness of agreement between a measured quantity value and a true quantity value of a measurand

(Measurement) precision

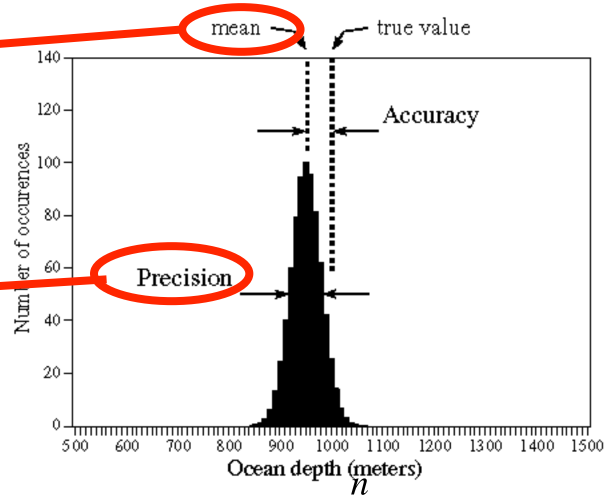
closeness of agreement between indications or measured quantity values obtained by replicate measurements on the same or similar objects under specified conditions

Precision is used to define measurement repeatability, intermediate measurement precision, and measurement reproducibility. Sometimes “measurement precision” is erroneously used to mean measurement accuracy.

2E: How to estimate x_w?

Do repeated empirical measurements => use samples
The empirical expectation is called “sample mean value”
similar: empirical variance S_x (“S”treuung or “S”catter):

$$\overline{x}_n = \frac{1}{n}\sum_{i=1}^{n} x_i$$



$$S_x^2 = \frac{1}{n-1}\sum_{i=1}^{n} (x_i - \overline{x}_n)^2$$

One can show: these are optimal estimators
(maximum likelihood estimators)
Furthermore for the expected value μ:

$$\mu = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} x_i$$

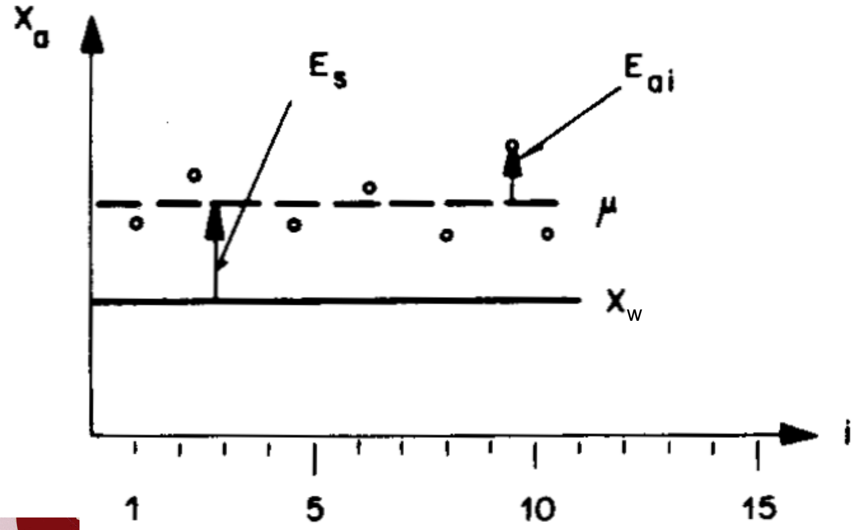
Precision depends of the meter (repeatability).
Accuracy depends on all components of the measurement facility.

Definition

2E: Systematic and Random Errors
Absolut error and relative Error

$$E_s = \mu - x_w$$

$$E_{a_i} = x_{a_i} - \mu$$



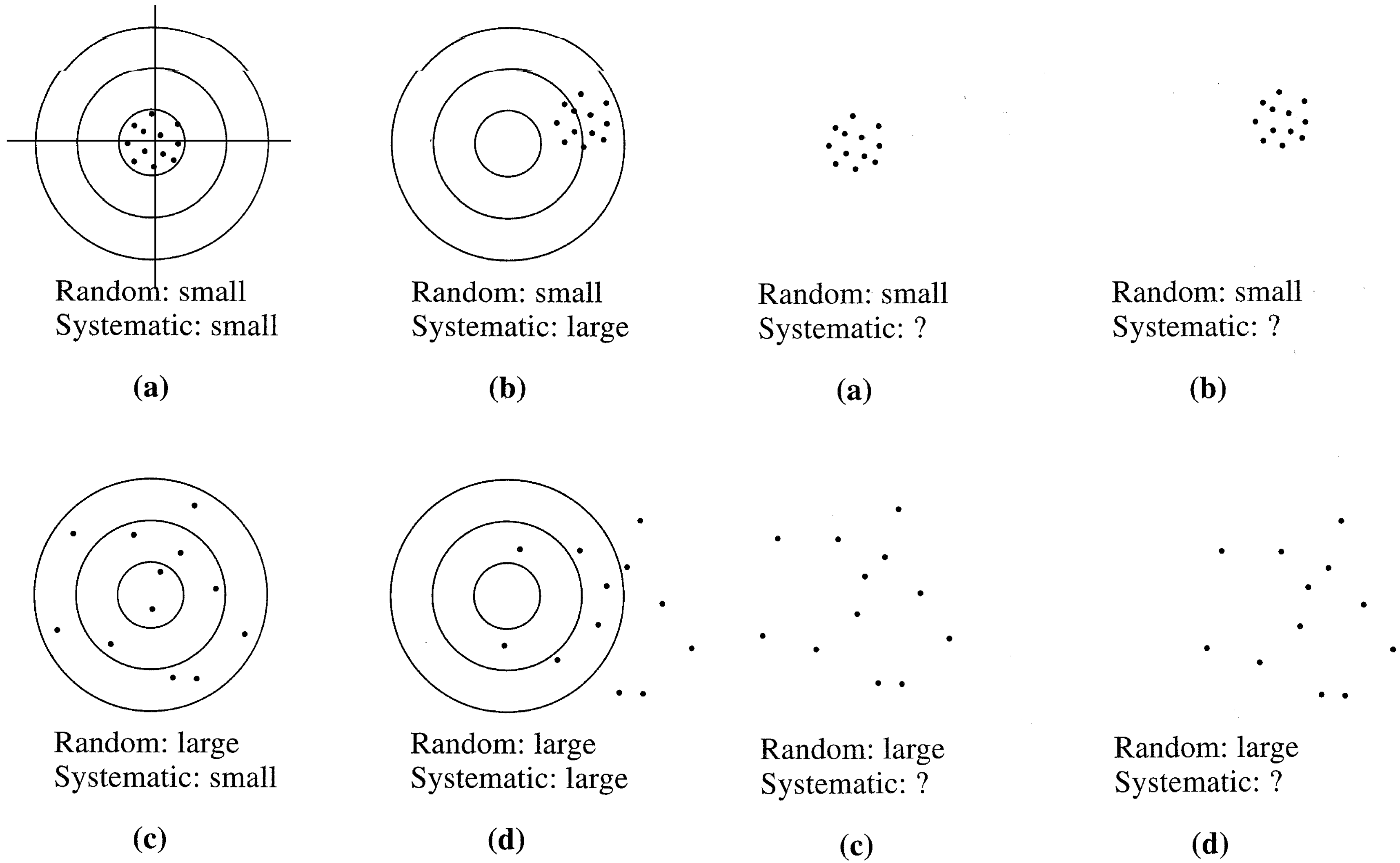
The repetitive measurement shows two types of errors:

a systematic error $\boldsymbol{E_s}$
and a random error $\boldsymbol{E_{ai}}$.

absolut error: $\Delta_{ai} := |\boldsymbol{E_{ai}}|$
relative error: $\rho_{ai} := \Delta_{ai} / \mu$

2E: Systematic vs. Random

Example



Target known => can judge systematic errors
target unknown => only random errors can be judged

shoots to a known / unknown target

2E: cure systematic errors by calibration

E_s can be repeated, every time it is the same thus it is a deterministic error.

- caused by factors that can (in theory) be modeled -> prediction
- e.g. **calibration** of a laser sensor
- e.g. distortion caused by the optic of a camera

E_{ai} is a random error -> non-deterministic

- no prediction possible
- however, they can be described probabilistically
- e.g. Hue, instability of camera, black level noise of camera .

Calibration:

An illustrative example is the calibration of a self-indicating weighing by putting of normalized weights. Taking into account systematic effects (previously determined by calibration measurement deviations of the weights, air pressure, temperature, buoyancy forces) and random influences the display of the scale is compared with the launched mass and then we estimate the uncertainty of this deviation.

A simple calibration result is: The weight displays at a load of 200 g, a deviation of +0.12 g, this result has an uncertainty of 0.20 g with a confidence interval of 95%.

(wikipedia on Calibration)

2D: $\bar{x}_n$ is Normally distributed

By central limit theorem:

If we repeat measurements with same instrument very often (I.I.D.) AND sum and build averages, then we are arbitrary close to a normal distribution

$$\lim_{n\to\infty} \overline{x}_n = \frac{X_1+\ldots+X_n}{n} \to \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

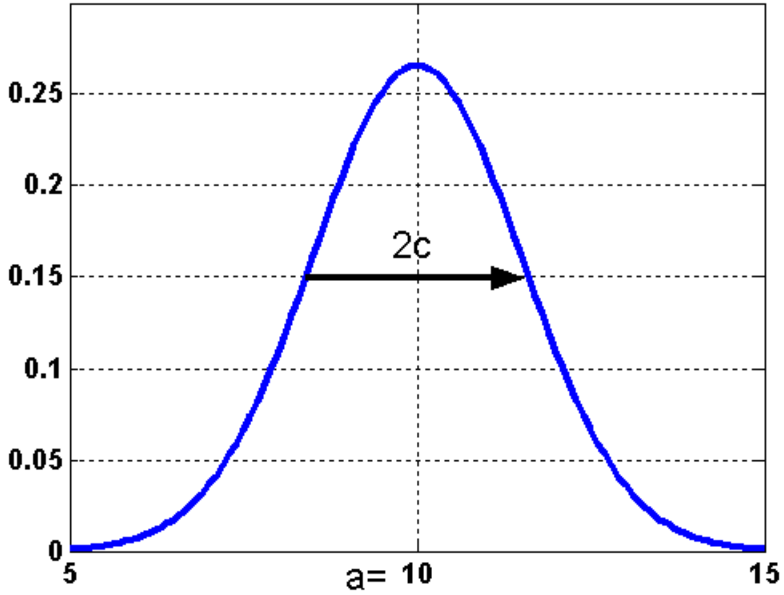
2D: Normal (Gaussian) distr.

$$f(t) = \frac{1}{c\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(t-a)^2}{c^2}}$$

abbreviated to

$$x \sim N(a, c^2)$$

$$F(x) = \int_{-\infty}^{x} f(t)\,dt$$



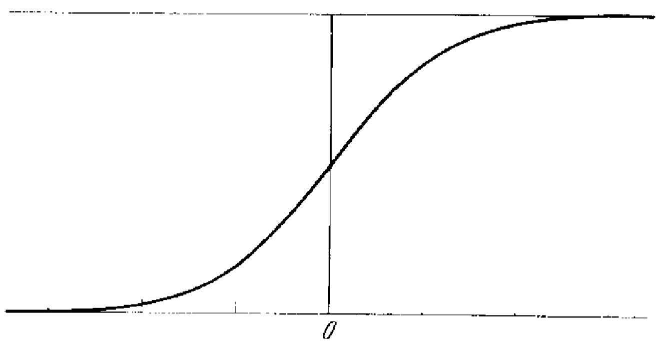


Fig. 2. Gaussian error function

Example

2D: Examples of 2 Gaussians

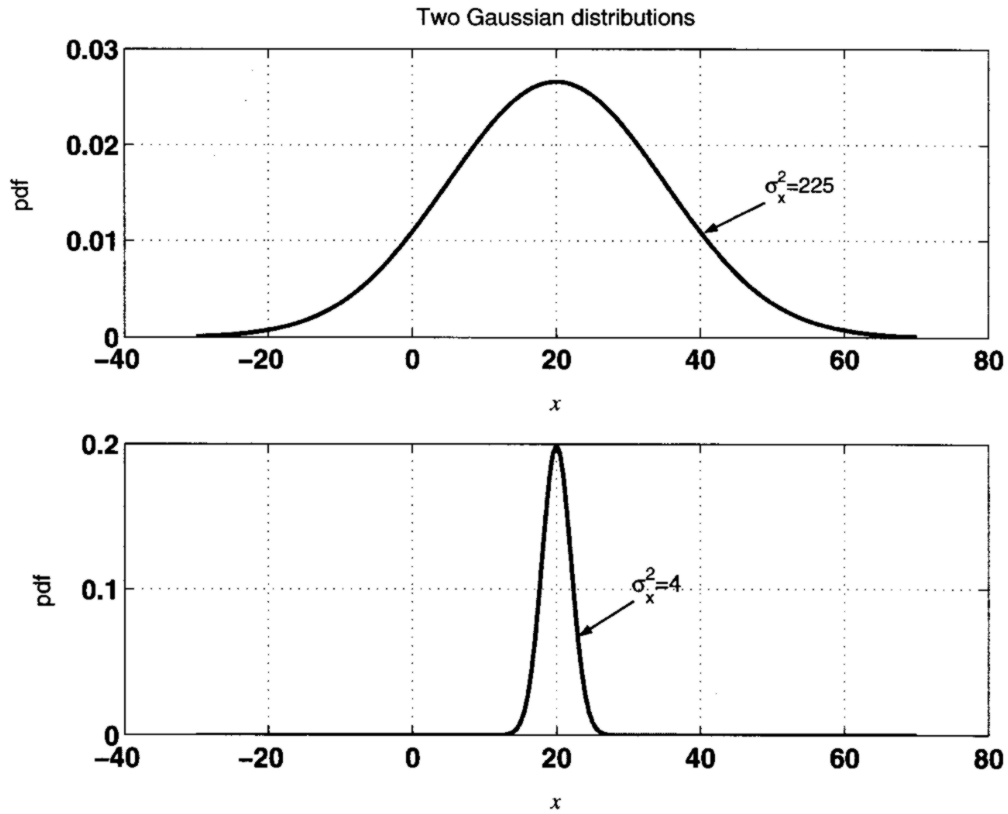
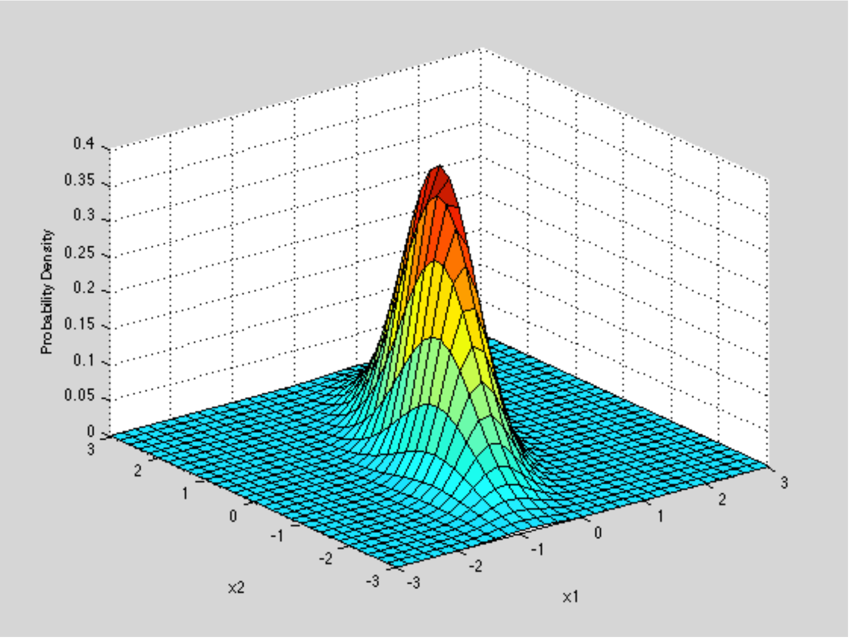


Figure 1.1. The figure shows the plots of the probability density functions of a Gaussian random variable $\boldsymbol{x}$ with mean $\bar{x} = 20$, variance $\sigma_x^2 = 225$ in the top plot, and variance $\sigma_x^2 = 4$ in the bottom plot.

Example

2D: higher dimensional N(μ,σ)

```
mu = [0 0];
Sigma = [.25 .3; .3 1];
x1 = -3:.2:3; x2 = -3:.2:3;
[X1,X2] = meshgrid(x1,x2);
F = mvnpdf([X1(:) X2(:)],mu,Sigma);
F = reshape(F,length(x2),length(x1));
surf(x1,x2,F);
caxis([min(F(:))-.5*range(F(:)),max(F(:))]);
axis([-3 3 -3 3 0 .4])
xlabel('x1'); ylabel('x2');
zlabel('Probability Density');
```



Example

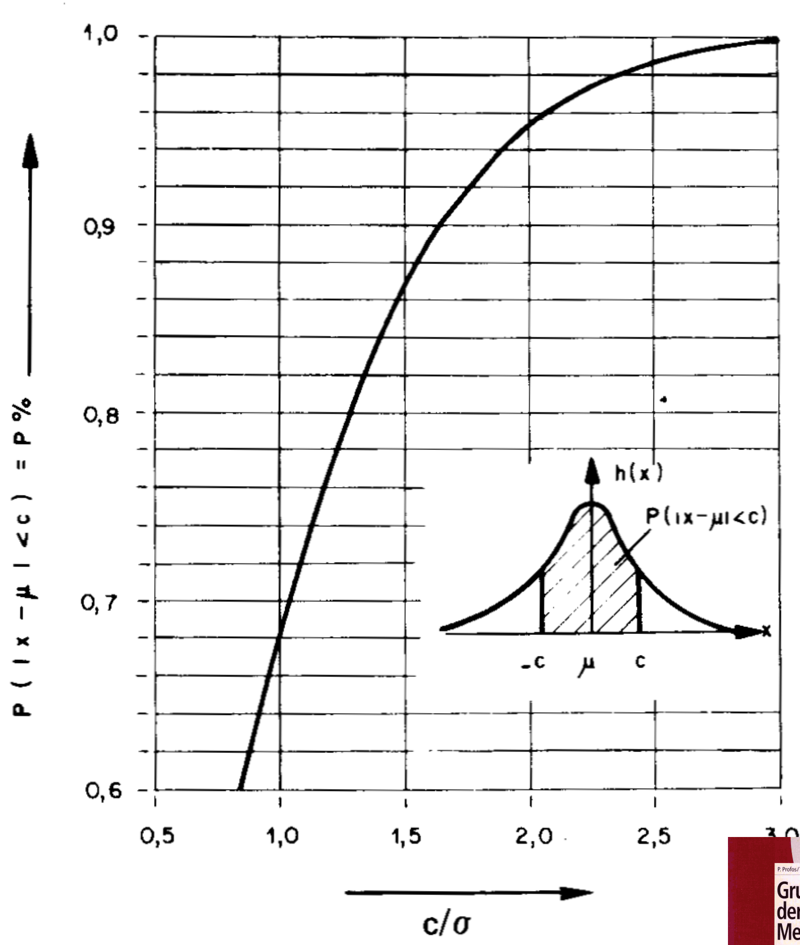
2D: If N(μ,σ) is known ...

Statistical test may be used, to see how likely (likelyhood bound c) a measurement value x may actually be.

Confidence level for

$|x-\mu| <= c$

versus multiples of σ.



Example

2D: testing for N(μ,σ)

Chi square test:

a) From n sample values (from an underlying totallity following some hidden distribution D) build estimators $\bar{x}_n$ and S_x !

b) Subdivide all n samples into K classes ($K >= 4$) s.t. in each class there are at least 5 samples (width of class may vary if necessary).

c) Get n_{ei}, i.e. the observed number of samples per class i.

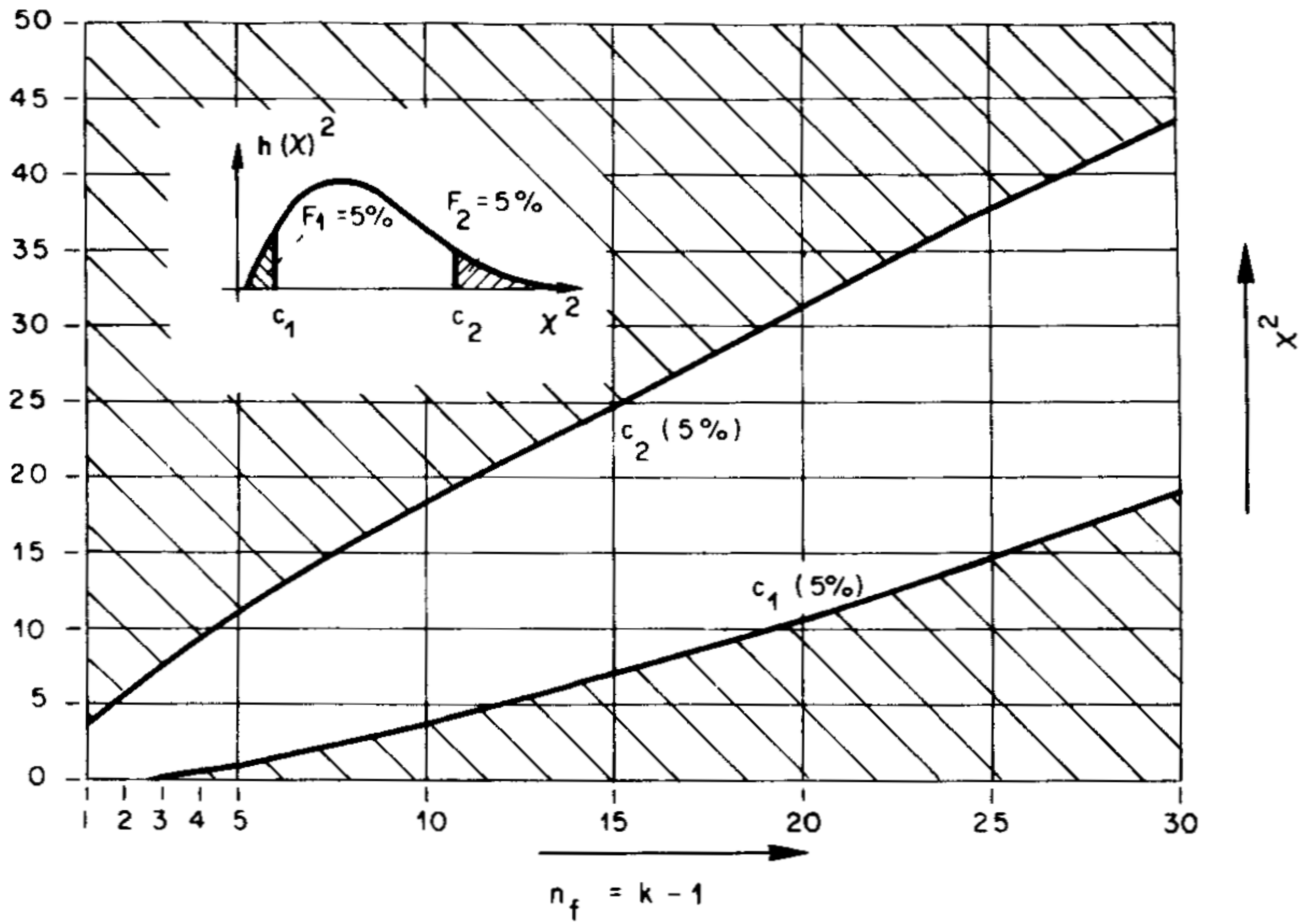
d) Build $N(\bar{x}_n, S_x)$ and P_i, i.e. the likelyhood that a sample lies in class i, then build $n_{oi} = P_i * n$, i.e. the number of to-be-expected samples in class i if the tatollity would be distributed according to $N(\bar{x}_n, S_x)$.

e) Build χ^2 and $n_f = K-1$.

$$\bar{x}_n = \frac{1}{n}\sum_{i=1}^{n} x_i \quad S_x = \sqrt{\frac{1}{n-1}\sum_i (x_i - \bar{x}_n)^2} \quad \chi^2 = \sum_{i=1}^{K} \frac{(n_{ei} - n_{oi})^2}{n_{oi}}$$

2D: testing for N(μ,σ)

If point (χ^2,n_f) lies in non-hashed area, there is no indication, that D would NOT be normal (up to a confidence level of 5% below and above)



3 Summary

- Measurements suffer from many effects (random and/or systematic).
- Systematic errors may be cured by calibration.
- Random errors are not totally random, there is distribution underneath (mostly: Gaussian, also in higher dimensions)
- Sampling and averaging helps to read „richtige“ values x_r (empirical mean and spread)
- Distributions may be checked versus model distributions using χ^2 test.