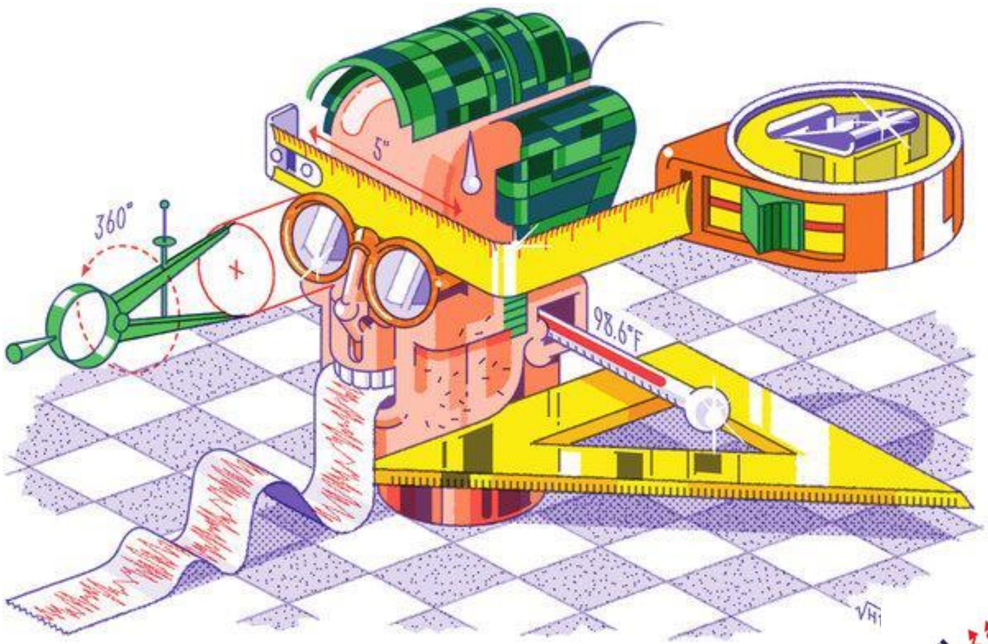
University of Zurich[UZH]

Department of Geography

RSL
measurements | products | policy

Basics of Uncertainty of Measurement

Simon Trim



SENSECO

EMPIR EURAMET
The EMPIR initiative is co-funded by the European Union's Horizon 2020 research and innovation programme and the EMPIR Participating States

Metrology for Earth Observation and Climate

Department of Geography

Basics of Uncertainty of Measurement

- Introduce the subject of measurement uncertainty and *why it matters*
- Understand when to use statistical analysis of a set of measurements and when to use other kinds of information about the measurement process in order to estimate measurement uncertainties
- Know the established rules for how to calculate an overall estimate of uncertainty from these individual pieces of information

University of Zurich[UZH]

Department of Geography

RSL measurements | products | policy

Basics of Uncertainty of Measurement

<u>End goal… measurement-function centred analysis of uncertainty and graphical representation in the form of an uncertainty analysis tree</u>

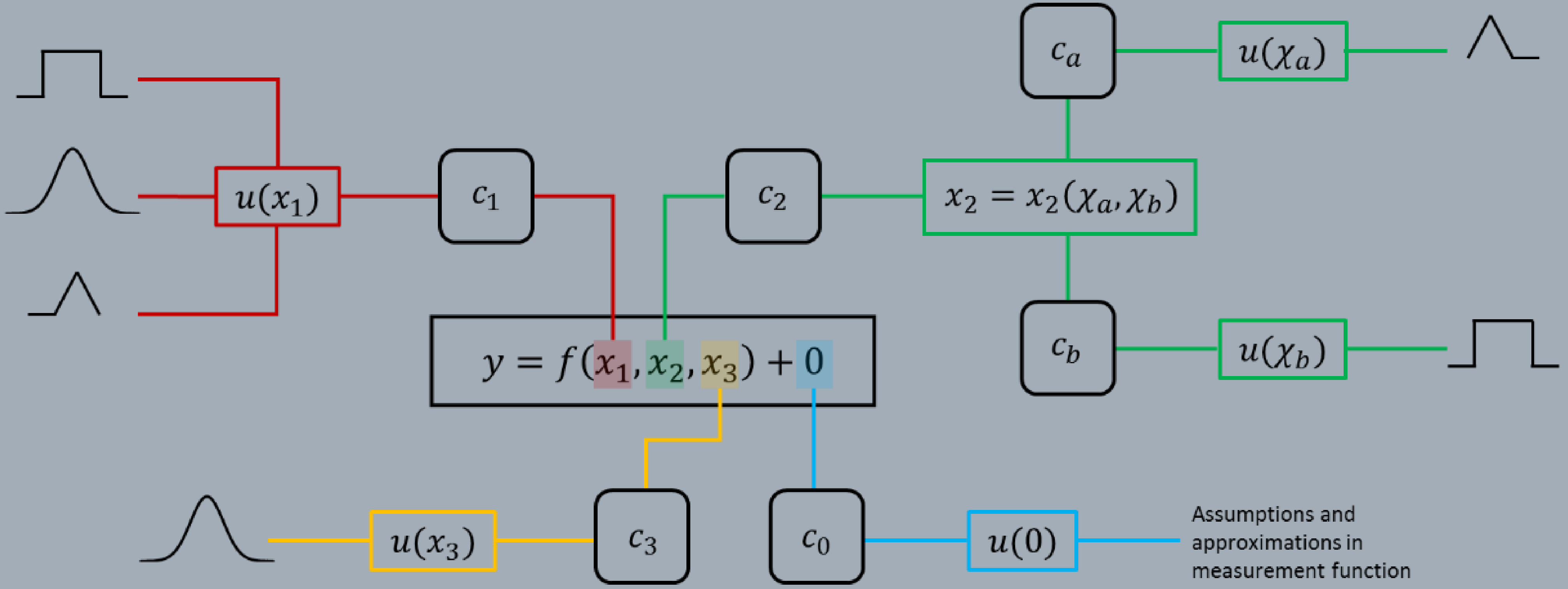


Table of Contents

1. What is a measurement (and what isn't)

1.1 What is a *measurement*?



➡ A measurement **tells us about a property of something**.

➡ A measurement **gives a number to that property**.



➡ Measurements are *always* made **using an instrument of some kind**.

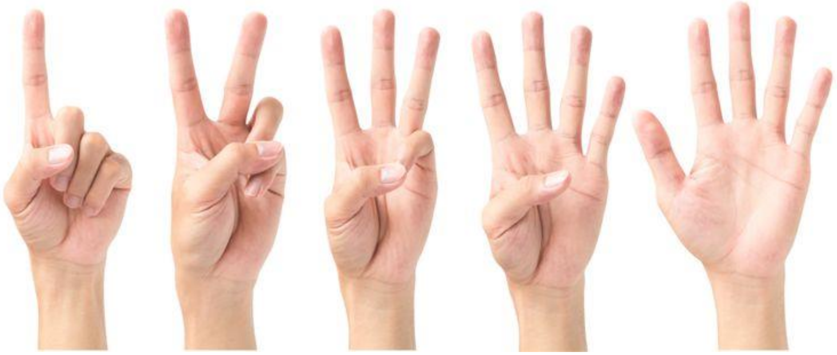
➡ The **result of a measurement** is generally in 2 parts: **a number and a unit of measurement**.

1.1 What is a *measurement*?

Example: “how long is it? … 81.8 metres.”

1.2 What is *NOT* a measurement?

- Comparing 2 pieces of string to see which is longer.



- Counting is not normally viewed as a measurement.

- **A test is not a measurement**: tests normally lead to a ‘yes/no’ answer or a ‘pass/fail’ result (however, measurements may be part of the process leading up to a test result).

2. Units of measurement and metrological traceability

2.1 About units of measurement…

Use the **International System of Units (SI)** which **defines seven units of measure as a basic set** from which all other SI units can be derived.

The **SI base units form a set of mutually independent dimensions**
as required by dimensional analysis commonly employed in science and technology

2.2 What is metrological traceability?

"The property of a measurement result whereby the result can be related to a reference through a documented unbroken chain of calibrations, each contributing to the measurement uncertainty."

JCGM 200, *International Vocabulary of Metrology. Basic and General Concepts and Associated Terms (VIM)*, 3rd ed., BIPM: Sèvres, France. http://www.bipm.org/utils/common/documents/jcgm/JCGM_200_2012.pdf accessed in September 2019.

2.3 Traceability? I love it when a plan comes together

Mars Climate Orbiter

- Total mission cost: **$327.6 million**

- Spacecraft lost due to ground-based computer software which produced output in non-SI units of pound-force seconds (lbf·s) instead of the SI units of newton-seconds (N·s) specified in the contract between NASA and Lockheed

- The discrepancy between calculated and measured position had been noticed earlier by at least two navigators, whose concerns were dismissed because they "did not follow the rules about filling out [the] form to document their concerns"[1].

[1]Oberg, James (December 1, 1999). "Why the Mars Probe went off course". *IEEE Spectrum.* IEEE.

3. The measurement process: from the sources of inaccuracy to the effects of inaccuracy

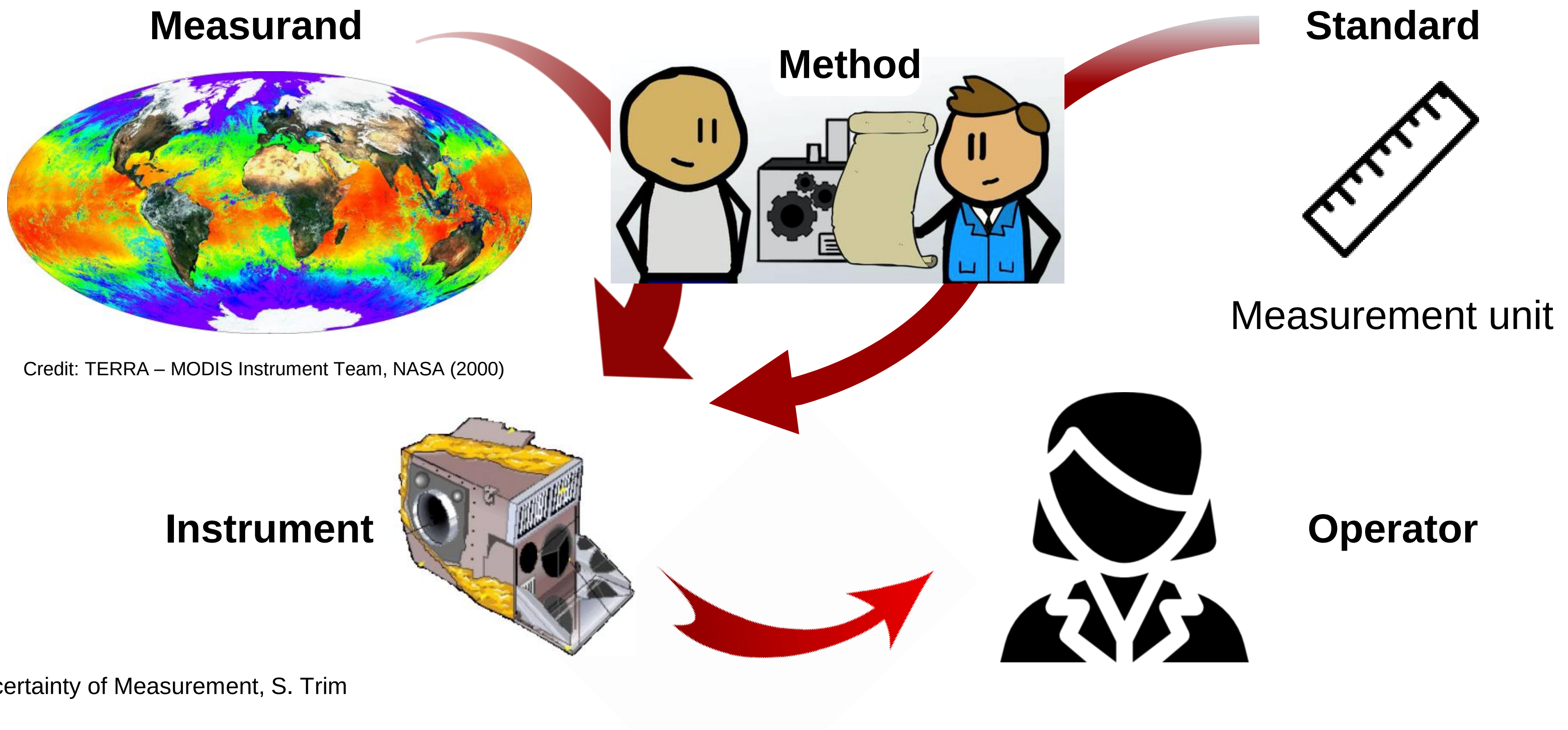
3.1 The measurement process

"The measurement of a quantity is generally defined as
the quantitative comparison of this same quantity with another one,
which is homogeneous with the measured one,
and is considered as the measurement unit."

Prof. Alessandro Ferrero, Electrical and Electronic Measurements, Polytechnic University of Milan

3.1 The measurement process

The measurement process can be described by the interaction of **5 agents**

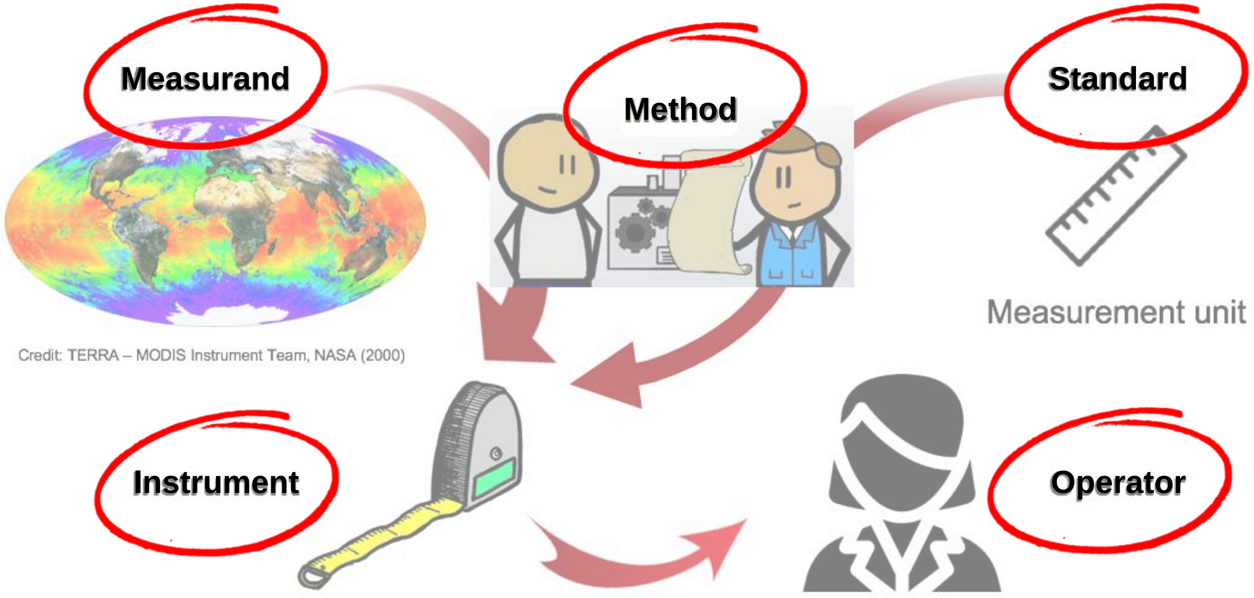


Credit: TERRA – MODIS Instrument Team, NASA (2000)

3.2 The measurement process: sources of inaccuracy

Can we get the true value of the measurand as the result of a measurement?

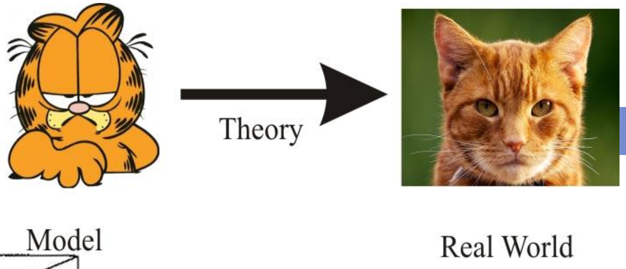
Hell no ➡ All agents taking part in the measurement process concur to make the measurement result different from the "true" value



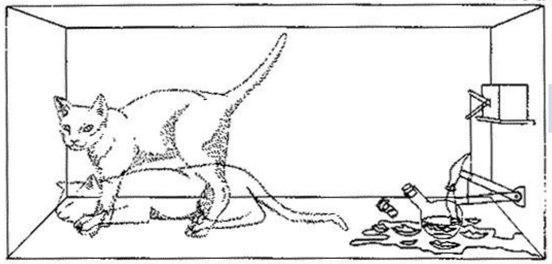
3.2.1 The measurement process: sources of inaccuracy

The measurand

- **Knowledge** of the measurand is very often **incomplete**



- **Mathematical model** of the measurand is **not totally correct**

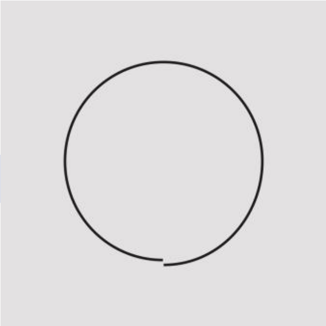


- **State** of the measurand is **not completely known**

- **Measurement process modifies** the measurand **state**

3.2.2 The measurement process: sources of inaccuracy

The standard



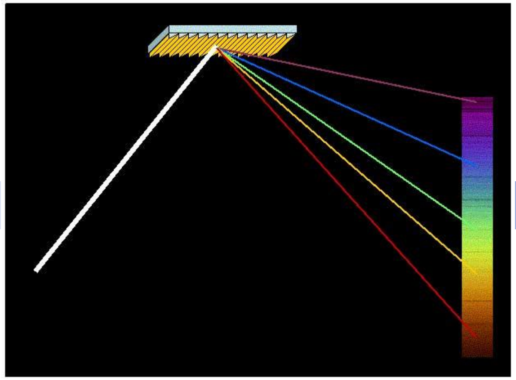
Not an ideal physical realisation of the measurement unit

Provides **only an approximation of the exact value of the measurement unit** ≈

"All exact science is dominated
by the idea of approximation."
Bertrand Russell

3.2.3 The measurement process: sources of inaccuracy

Credit: JPL, NASA



The method

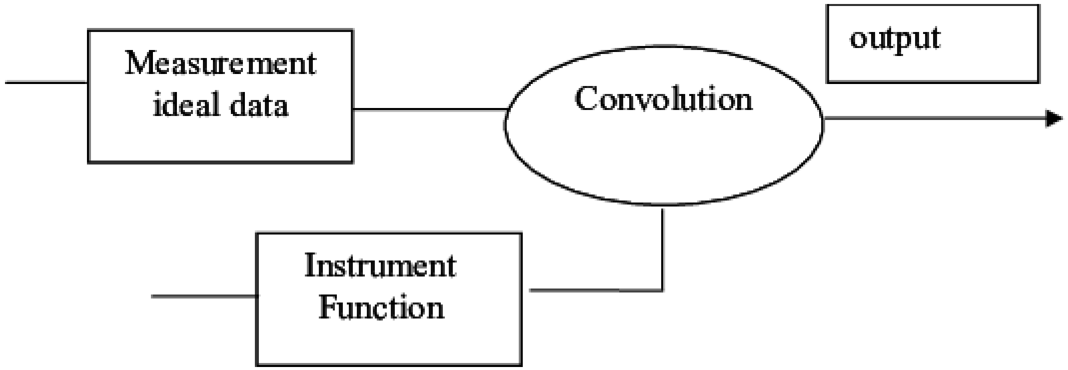
- The measurement method is **usually based on a physical phenomenon**
- It usually **does not take into account other phenomena that may interfere** with the considered one

3.2.4 The measurement process: sources of inaccuracy

The instrument

The **actual operating principle differs from the ideal one** due to:

- Non-ideal components
- Internally generated noise
- Sensitivity to environmental conditions
- Lack of calibration
- Ageing
- …



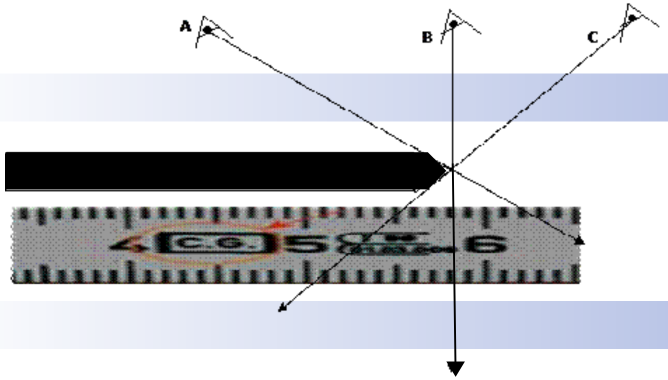
The ideal image is convolved by the instrument function to result in a non-ideal image

Teper, L., Dickstein, P., Ingman, D. (2006). ”The line-spread-function of an ultrasonic measurement system: theory and experimental characterization". *Third International Conference on Metrology*, IMS.

3.2.5 The measurement process: sources of inaccuracy

The operator

- **Measurement execution** must be “**just in time**”
(if measurement conditions are evolving…)



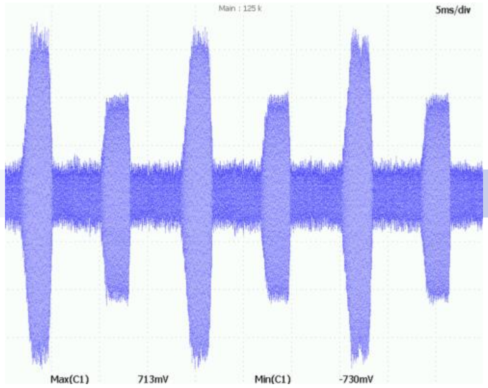
- **Evaluation of the pointer displacement** on a ruler
(mind the parallax errors)

- Correct **evaluation of the measurement results**

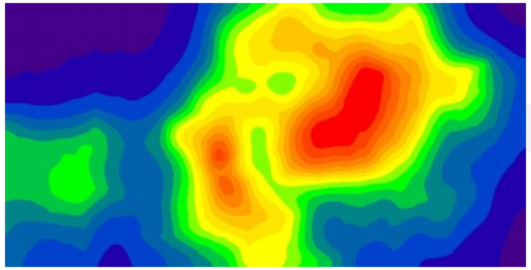
- Correct **post processing of the readings**

3.2.6 The measurement process: sources of inaccuracy

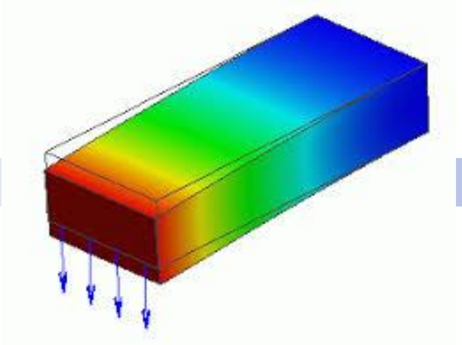
The rest of the world! (influence quantities)



- Electromagnetic interferences (EMI)



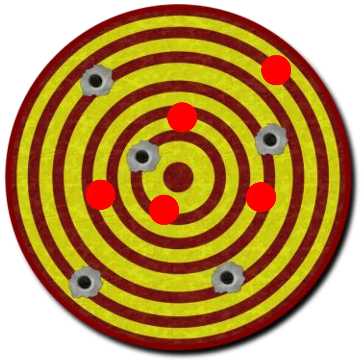
- Temperature variations



- Mechanical stress
- Other effects … all more or less impacting the final result depending on the kind of measurement!

3.3 The measurement process: effects of inaccuracy

If the measurement procedure is **repeated a number of times, the obtained measured values always differ from each other**, even if the measurement conditions are not changed



If the measurement procedure is **repeated by another operator, reproducing the same conditions** somewhere else with other instruments (albeit of the same type), **different results are obtained**

3.3.1 The measurement process: effects of inaccuracy

A first conclusion… and a new problem

- **Expressing a measurement result with a single number** (and a unit of measurement) **is meaningless**
- **We cannot use a single measured value in any comparison** with other measurement results obtained by means of the same measurement procedure
- **Is this acceptable?**

3.3.2 The measurement process: effects of inaccuracy

Can we still represent the result of a measurement with a single value and avoid the aforementioned drawbacks?

➡ **Yes, if we find a way to assess how good the measurement result is**

- **How well does the result of a measurement approximate the measurand value?**
- **Need to quantify how complete (or incomplete) the information retrieved from the measurand's measurement result truly is…**

4. Uncertainty of measurement (also: error theory versus uncertainty)

4.1 Error vs Uncertainty

Old way: the error theory

- **Axiom:** the **true value** x_t **of measurand** x is somehow **known**
- **If** x_m **is the measured value, an absolute error is defined as**:

$$\varepsilon = x_m - x_t$$

- **If** the **maximum error** ε_M is **considered**, an **interval is obtained about the measured value**:

$$[x_m - \varepsilon_M \,; x_m + \varepsilon_M]$$

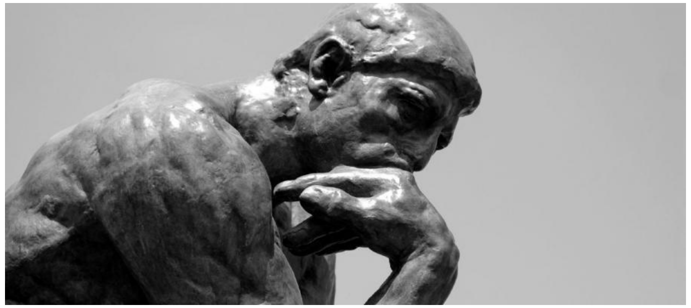
4.1 Error vs Uncertainty

Old way: the error theory

→ **Axiom:** the **true value** x_t **of measurand** x is somehow **known**

A philosophical question…

→ **If measurement results can provide only incomplete knowledge about the quantity subject to measurement, how can we ever know the true value of the measurand?**



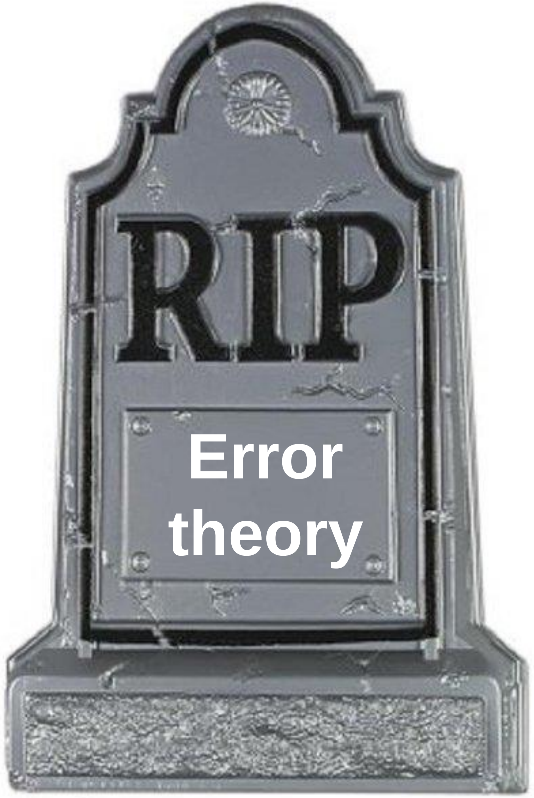
4.1 Error vs Uncertainty

Old way: the error theory

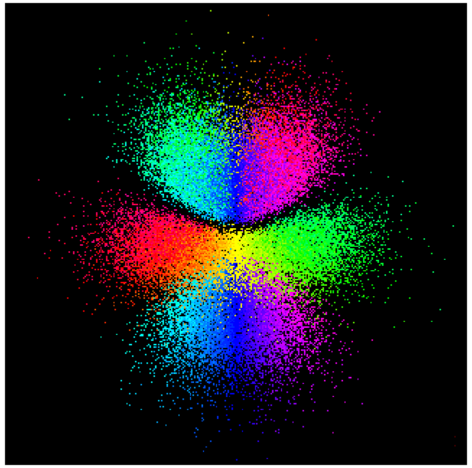
- Axiom: the true value x_t

The "true" value issue

- **The true value is unkno**
- **The true value is unkno**
- In some situations, **the tr**
  - The measurand can o



ty function



Credit: K. Sugiyama

4.1 Error vs Uncertainty

The uncertainty concept

- Introduced in the late 1980s

- **Uncertainty as a quantifiable attribute** of the measurement, able **to assess the quality** of the measurement process and result

- The concept comes from the awareness that:

“when all known or suspected components of error have been evaluated, and the appropriate corrections have been applied, there still remains an uncertainty about the correctness of the stated results, that is, a doubt about how well the result of the measurement represents the value of the quantity being measured”

Uncertainty of measurement, Part 3: Guide to the Expression of Uncertainty in Measurement (GUM:1995), ISO/IEC

4.2 Uncertainty of measurement

What is uncertainty of measurement?

Uncertainty of measurement is **the doubt that exists about the result of any measurement**.

Uncertainty meaning

A **parameter**, associated with the result of a measurement, that **characterises the dispersion of the values** that could ***reasonably*** be **attributed to the measurand**

The adverb ***reasonably*** is the key point in this definition:

- the **capability** of quantifying the dispersion in a reasonable way **depends on** the amount of **available information**

4.2 Uncertainty of measurement

Expressing uncertainty of measurement

Axiom

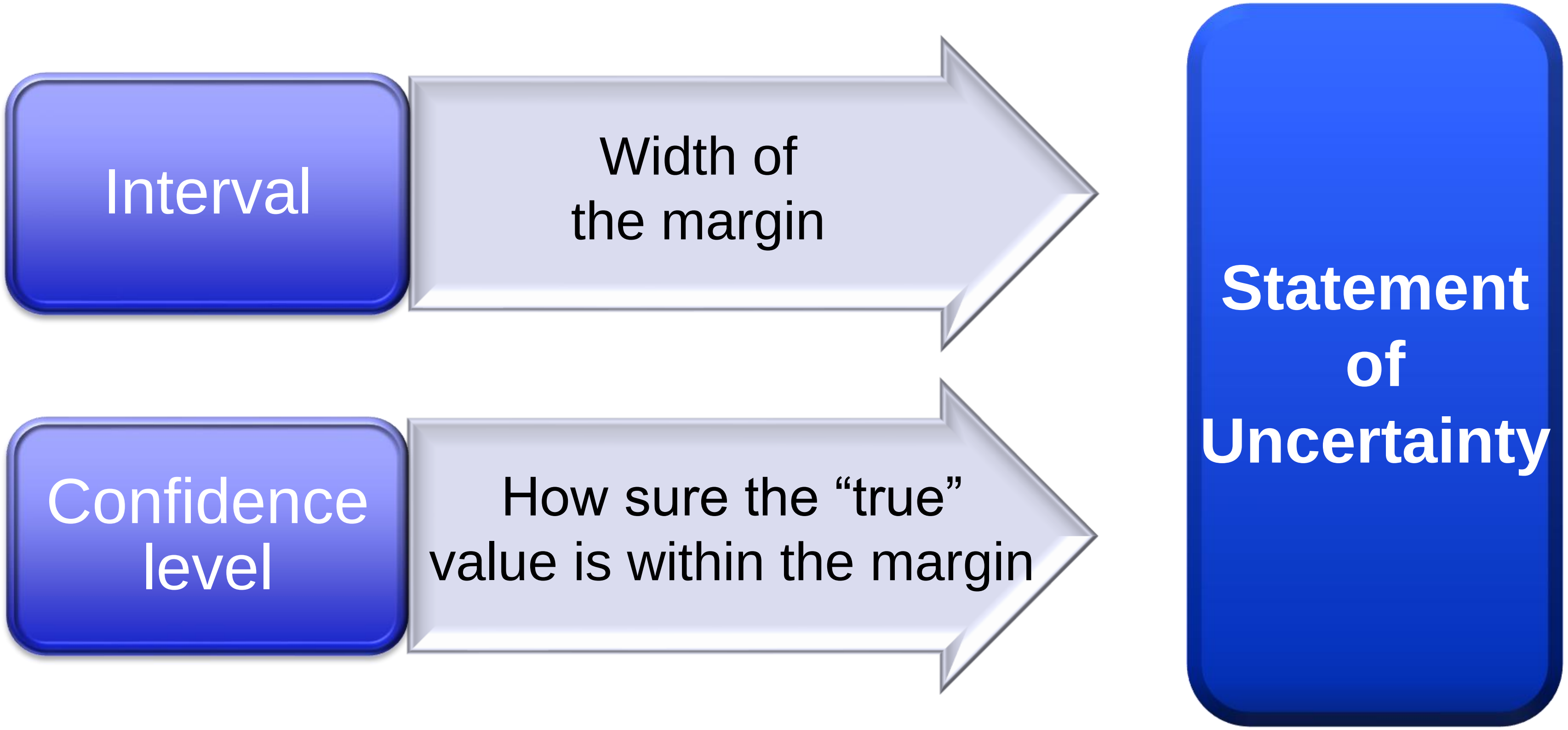
There is **always a margin of doubt** about any measurement

How big is the **margin**?

How bad is the **doubt**?

4.2 Uncertainty of measurement

Expressing uncertainty of measurement



4.2 Uncertainty of measurement

Expressing uncertainty of measurement

Example

Stick length: we find that it is 60 cm plus or minus 2 cm, at the 95 percent confidence level

Result: $$60\ cm\ \underbrace{\pm\, 2\ cm\text{, at a level of confidence of 95\%.}}_{\textbf{Uncertainty statement}}$$

4.3 Uncertainty of measurement: why is it important?

Calibration – where the uncertainty of measurement must be reported on the certificate

Test – where the uncertainty of measurement is needed to determine a pass or fail

Tolerance – where you need to know the uncertainty before you can decide whether the tolerance is met

4.4 What is not a measurement uncertainty?

- **Mistakes made by operators** are not measurement uncertainties – they don’t contribute to uncertainty

- **Tolerances** are not uncertainties: they are acceptance limits chosen for a process or a product

- **Specifications** are not uncertainties: they tell you what to expect from a product (including non-technical qualities of the item, such as its appearance)

- **Accuracy (or rather inaccuracy)**: a qualitative term (“a measurement was accurate/inaccurate”) – uncertainty is quantitative (when a $\pm$ figure is quoted, it may be called an uncertainty – not an accuracy)

4.5 'Error' in the age of uncertainty

- ~~**A measured value will differ from the "true" (haha!) value for several reasons**~~
- **A measured value will differ from the value of a particular quantity intended to be measured (*aka*, the measurand) for several reasons, *some of which we may know about***
  E.g.: the quantity realised for measurement is an approximation of the measurand
- **To account for known differences, we apply a *correction***
  E.g.: the measured value may be multiplied by a gain determined during the instrument's calibration
- **This correction will never be perfectly known** and there will also be other effects that cannot be corrected, so after correction **there will always be a residual, <u>unknown *error*</u>**
- This specific **<u>error</u>** is **<u>conceptualised as a draw from a probability distribution function</u>**, and the ***uncertainty*** associated with the measured value is ***a measure of that probability distribution function***
- **<u>Final (unknown) error</u>** on the measured value is **<u>drawn from the overall probability distribution</u>** described by the uncertainty associated with the measured value – this is **built up from the probability distributions associated with all the different sources of uncertainty**

4.5 'Error' in the age of uncertainty

~~The **final corrected result** is viewed as the **best estimate of the "true" value of the measurand**~~

The final corrected result is simply the best estimate of the value of the quantity intended to be measured (*i.e.*, the approximately defined measurand, itself a source of uncertainty)

<u>Reference reading:</u>

JCGM 100:2008 – GUM 1995 with minor corrections, *Annex D: "True" value, error, and uncertainty*, 1st ed. 2008.

5. Basic statistics on sets of numbers: from getting the best estimate to calculating an estimated standard deviation

5.1 Getting the best estimate

Taking the average of a number of readings

- An average gives you an estimate of the value of the measurand

- Average (arithmetic mean): $\bar{x} = \frac{1}{n} * \sum_{i=1}^{n} x_i$

- The more measurement results are used, the better the estimate.

5.2 Spread… quantified using standard deviation

- The **spread of values** tells us something about the uncertainty of a measurement

- **Standard deviation**: how different the individual readings typically are from the average of the set

- The “true” value for the the standard deviation can only be found from a very large (infinite) set of readings

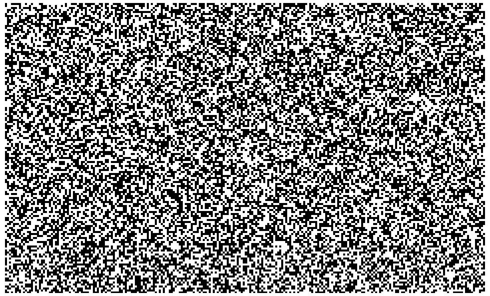
- **Estimated standard deviation** formula: $s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{(n-1)}}$

  Also known as **standard uncertainty**

6. The general kinds of uncertainty in any measurement: random or systematic, and distribution (the ‘shape’ of the errors)

6.1 Random or systematic

The effects that give rise to uncertainty in measurement can be either:

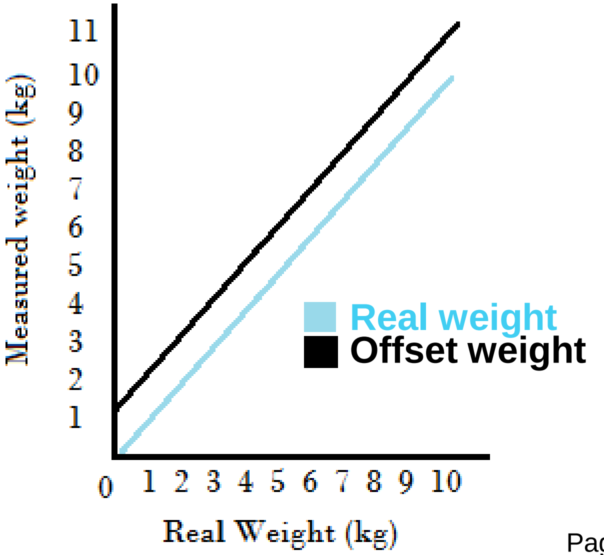


A randomly generated bitmap, made using RANDOM.ORG's True Random Number Generator (TRNG)

- **random** – where repeating the measurement gives a randomly different result

the more measurements are averaged, the better the estimate

- **systematic** – where the same influence affects the result for each of the repeated measurements



6.2 Distribution – the ‘shape’ of the errors

The spread of a set of values can take different forms

→ ***probability distributions***

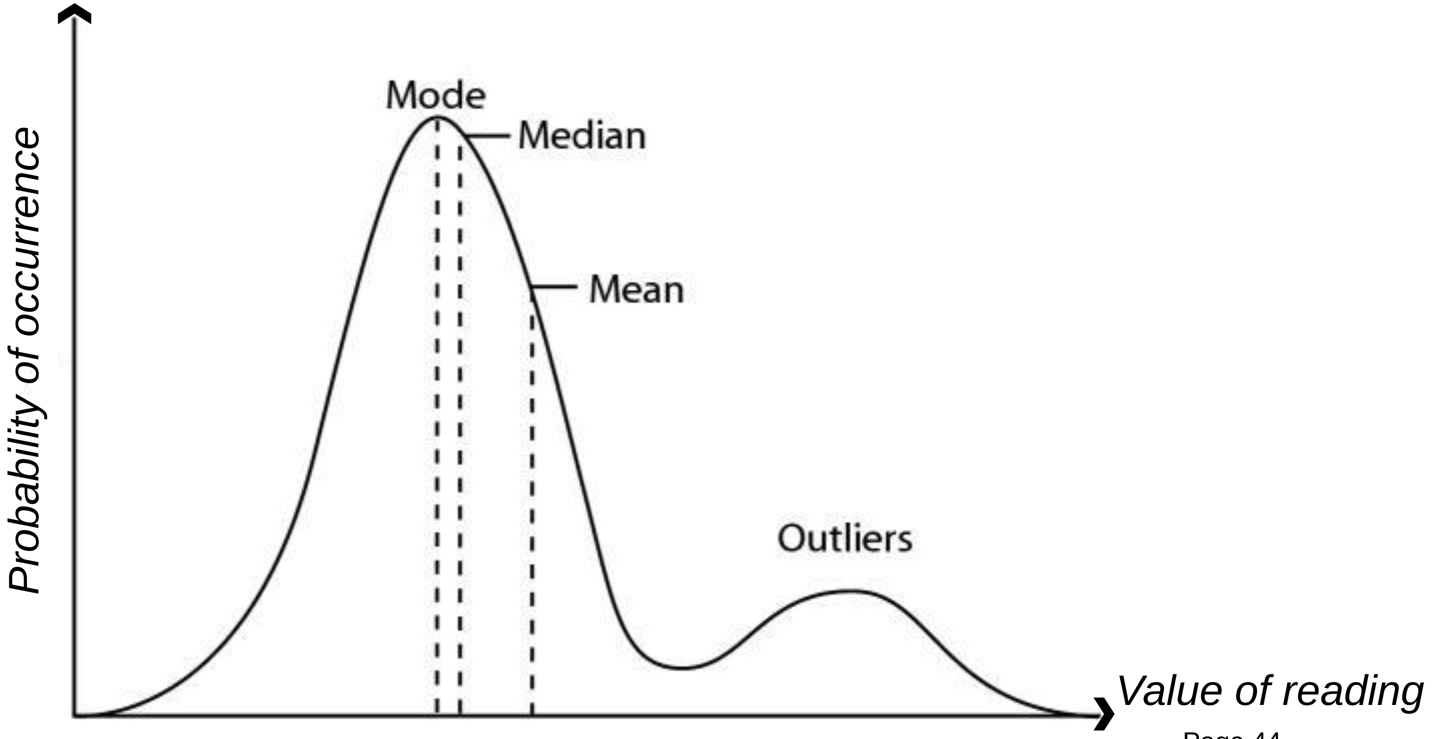
The different elements of a probability distribution:

- mode: the number that appears most frequently

- median: middle point between higher and lower half of the distribution

- mean: average

- outliers: unusual/extreme values, have the effect of skewing the data



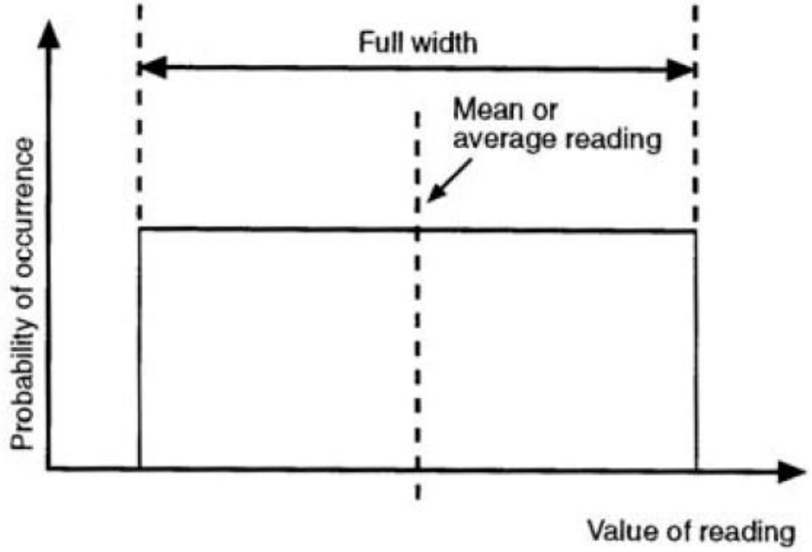
6.2 Distribution – the 'shape' of the errors

The uniform or rectangular distribution

When the measurements are quite evenly spread between the highest and the lowest values, a rectangular or uniform distribution is produced. Example: rain drops falling on a thin, straight telephone wire, would be as likely to fall on any one part as on another.

Probability density function (in terms of the mean μ and variance σ^2):

$$f(x) = \begin{cases} \dfrac{1}{2\sigma\sqrt{3}} & \text{for } -\sigma\sqrt{3} \leq x - \mu \leq \sigma\sqrt{3} \\ 0 & \text{otherwise} \end{cases}$$

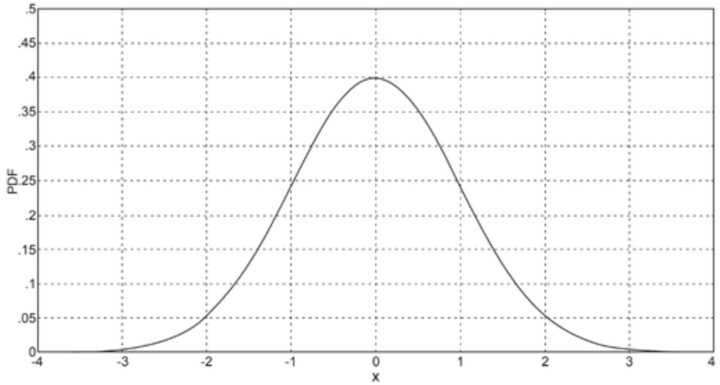


6.2 Distribution – the ‘shape’ of the errors

The normal (*aka* Gaussian) distribution

In a set of readings, sometimes the **values are more likely to fall near the average than further away** – typical of a *normal* or *Gaussian* distribution:

- physical quantities that are expected to be the sum of many independent processes (such as measurement errors) often have distributions that are *approximately* normal
- Standard formulation: $\boldsymbol{N(\mu, \sigma^2)}$ where
  - μ is the mean: controls the location of the peak of the distribution
  - σ^2 is the variance: a larger value for σ^2 results in a ‘fatter’ “bell curve” (distribution shape)
- Probability density function (PDF) is: $p(x) = \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{\frac{-(x-\mu)^2}{2\sigma^2}}$
- *Note that*:
  - 68.258% of the data is under the curve between $\mu \pm 1\sigma$
  - 95.45% of the data is under the curve between $\mu \pm 2\sigma$
  - 99.73% of the data is under the curve between $\mu \pm 3\sigma$



6.2 Distribution – the ‘shape’ of the errors

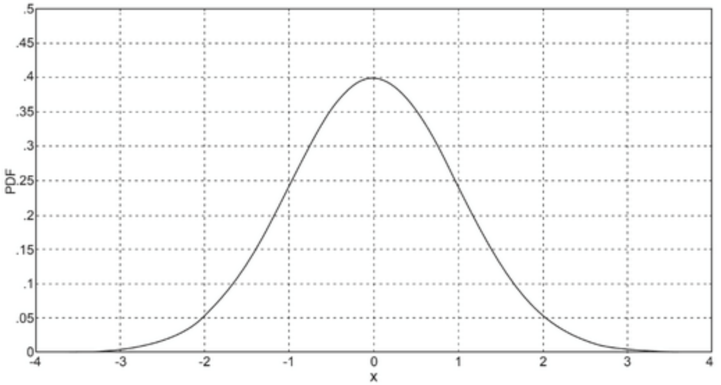
The normal (*aka* Gaussian) distribution

Out of all distributions with mean μ, variance σ^2, and support over all of $\mathbb{R}$, the normal distribution, $\boldsymbol{N(\mu, \sigma^2)}$, **maximises entropy**.

The class of normal distributions $\boldsymbol{N(\square, \sigma^2)}$ all maximize entropy subject to the constraints of a fixed variance σ^2 and support over $\mathbb{R}$.

About the notion of entropy:–

The **entropy of a distribution is a measure of how flat and smooth that distribution is**; the flatter and smoother a distribution, the higher its entropy. The flattest and smoothest distribution over $\mathbb{R}$ is the uniform distribution, and the flattest and smoothest distribution over $\mathbb{R}$ with a given finite mean and variance is the normal distribution with that mean and variance.



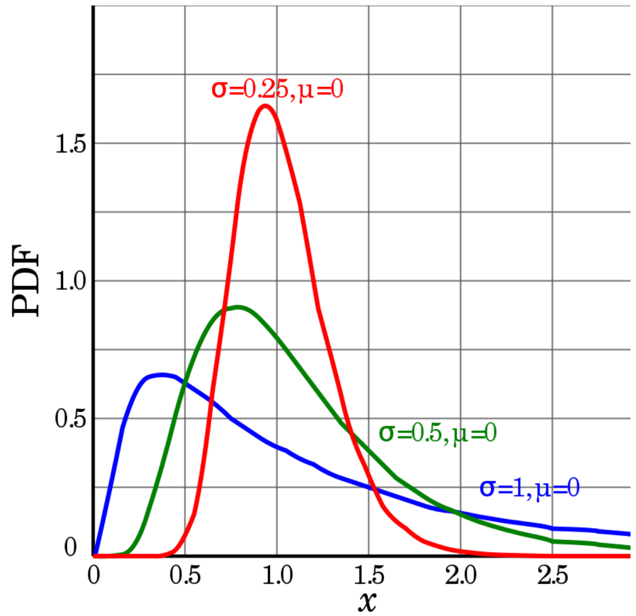
6.2 Distribution – the 'shape' of the errors

The log-normal distribution

A log-normal distribution is a continuous probability distribution of a random variable whose logarithm is normally distributed.
$\Rightarrow$ If random variable *X* is log-normally distributed, then *Y* = ln(*X*) has a normal distribution.
$\Rightarrow$ If *Y* has a normal distribution, then the exponential function of *Y*, *X* = exp(*Y*), has a log-normal distribution.

A random variable which is log-normally distributed takes only positive real values.

A log-normal process is the statistical realisation of the multiplicative product of many independent random variables, each of which is positive. This is justified by considering the central limit theorem in the log domain. The log-normal distribution is the maximum entropy probability distribution for a random variate *X* for which the mean and variance of ln(*X*) are specified.



7. How to calculate uncertainty of measurement: the 2 ways to estimate uncertainties, and the 8 main steps to evaluating uncertainty

7.1 The two ways to estimate uncertainties

Firstly: –

Identify sources of uncertainty in the measurement → Estimate the size of the uncertainty from each source → Combine the individual uncertainties to give an overall figure

Type A evaluations

- Uncertainty estimates using statistics (usually from repeated readings)

Type B evaluations

- Uncertainty estimates from any other information

There is a temptation to think of '**Type A**' as '**random**' and '**Type B**' as '**systematic**', but this is not necessarily true.

7.2 Eight main steps to evaluating uncertainty

1. Decide what you need to find out from your measurements. Decide what actual measurements and calculations are needed to produce the final result.

2. Carry out the measurements needed.

3. Estimate the uncertainty of each input quantity that feeds into the final result. Express all uncertainties in similar terms.

4. Decide whether the errors of the input quantities are independent of each other. If you think not, then some extra calculations or information are needed.

5. Calculate the result of your measurement (including any known corrections for things such as calibration).

6. Find the combined standard uncertainty from all the individual aspects.

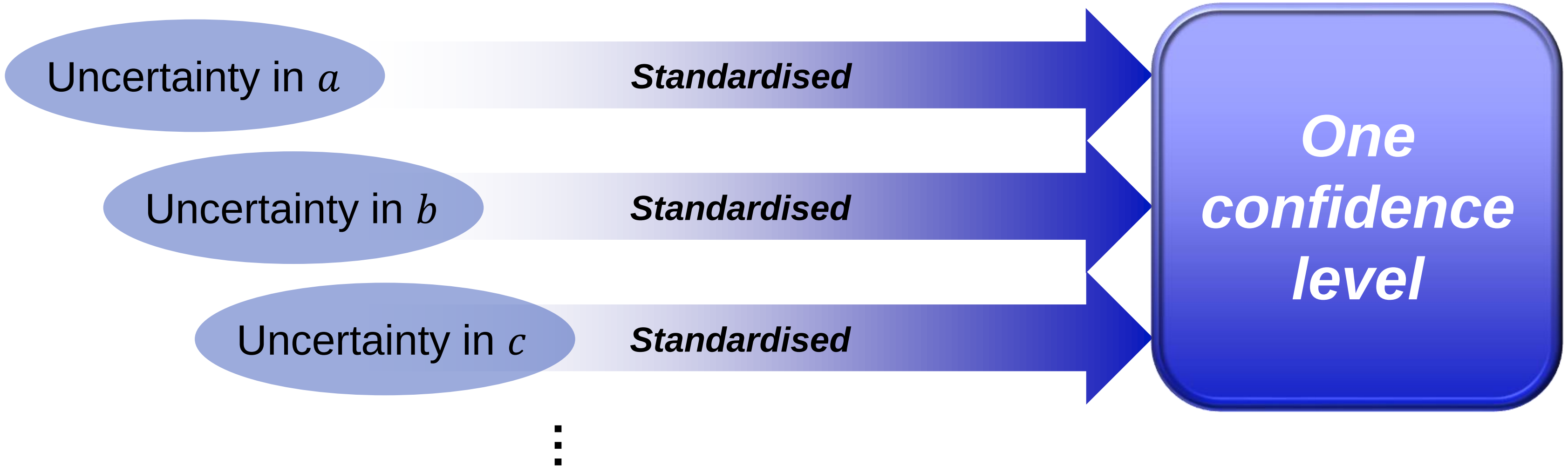
7. Express the uncertainty in terms of a coverage factor, together with a size of the uncertainty interval, and state a level of confidence.

8. Write down the measurement result and the uncertainty, and state how you got both of these.

8. Standard uncertainty for both Type A and Type B evaluations, combining standard uncertainties, correlation and coverage factor *k*

8.1 Standard uncertainty

All contributing uncertainties should be expressed at the same confidence level, by converting them into standard uncertainties.

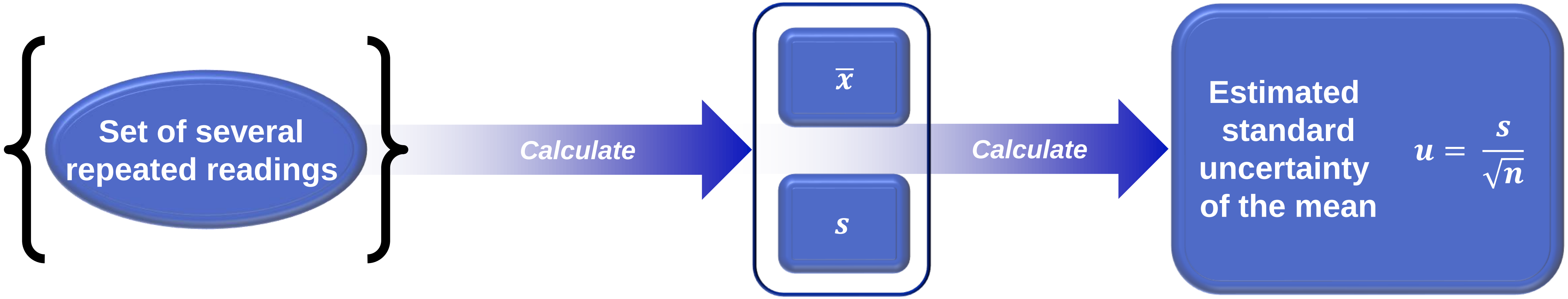


Standard uncertainty ≡ a **margin** whose size can be thought of as $\pm$ **one** (estimated) **standard deviation** (σ).

The standard uncertainty tells us about the *uncertainty of an average* (not just about the spread of values).
A standard uncertainty is usually shown by the symbol u, or $u(y)$ (*i.e.*, the standard uncertainty in y).

8.1 Standard uncertainty

Calculating standard uncertainty for a Type A evaluation



where n = number of measurements in the set.

*(The **standard uncertainty of the mean** has historically also been called the **standard deviation of the mean**, or the **standard error of the mean**.)*

8.1 Standard uncertainty

Calculating standard uncertainty for a Type B evaluation

Where the information is more scarce (in some Type B estimates, for example when the measurement procedure cannot be repeated), you might only be able to estimate the upper and lower limits of uncertainty (*i.e.*, interval *a priori* known, for instance by means of calibration results).

Uncertainty may be evaluated from assumed probability distributions based on experience or other information.

The standard uncertainty for a uniform distribution is found from:

$$\frac{a}{\sqrt{3}}$$

where a is the semi-range (or half-width) between the upper and lower limits.

N.B.: you can usually assume that uncertainties 'imported' from the calibration certificate for a measuring instrument are normally distributed.

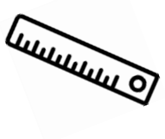
$$\bar{s} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}.$$

8.1 Standard uncertainty

Converting uncertainties from one unit of measurement to another

Uncertainty contributions must be in the same units before they are combined. As the saying goes, you cannot 'compare apples with oranges'.

<u>Example</u>: in making a measurement of length L, the measurement uncertainty will also eventually be stated in terms of length.



$\Rightarrow$ One source of uncertainty might be the variation in room temperature T. Although the source of this uncertainty is temperature, the effect is in terms of length, and it must be accounted for in units of length.

$\Rightarrow$ E.g.: for every degree rise in temperature, the material being measured expands in length by $0.1\ \%$.

$\Rightarrow$ For a piece of the material $100\ cm$ long, and $u(T)\ =\ \pm 2\ °\ C$, we have $\mathrm{u}(L) =\ \pm\ 0.2\ cm$.

8.2 Combining standard uncertainties

Individual standard uncertainties calculated by Type A or Type B evaluations can be combined validly by ‘summation in quadrature’ (also known as ‘root sum of the squares’).

The result of this is called the **combined standard uncertainty**, shown by u_c or $u_c(y)$.

Summation in quadrature is simplest where the result of a measurement is reached by addition or subtraction.

The more complicated cases are also covered here.

8.2 Combining standard uncertainties

Summation in quadrature for addition and subtraction

$$u_c = \sqrt{a^2 + b^2 + c^2 + \cdots etc}$$

Summation in quadrature for multiplication or division

$$\frac{u(A)}{A} = \sqrt{\left(\frac{u(L)}{L}\right)^2 + \left(\frac{u(W)}{W}\right)^2}$$

8.2 Combining standard uncertainties

Summation in quadrature for more complicated functions

⇒ Where a value is squared (e.g. Z^2) in the calculation of the final measurement result, then the relative uncertainty due to the squared component is in the form:

$$\frac{2u(Z)}{Z}$$

⇒ Where a square root (e.g. $\sqrt{Z}$) is part of the calculation of a result, then the relative uncertainty due to that component is in the form

$$\frac{u(Z)}{2Z}$$

8.2 Combining standard uncertainties

Summation in quadrature for more complicated functions

Some measurements are processed using formulae which use combinations of addition, subtraction, multiplication and division, etc.

For example, you might measure electrical resistance R and voltage V, and then calculate the resulting power P using the relationship:

$$P = \frac{V^2}{R}$$

In this case, the relative uncertainty u(P)/P in the value of power would be given by:

$$\frac{u(P)}{P} = \sqrt{\left(\frac{2u(V)}{V}\right)^2 + \left(\frac{u(R)}{R}\right)^2}$$

8.3 Correlation

Summation in quadrature for more complicated functions

The equations given here to calculate the combined standard uncertainty are only correct if the input standard uncertainties are not inter-related or correlated.

This means we usually need to question whether all the uncertainty contributions are independent.

Could a large error in one input cause a large error in another?

Could some outside influence, such as temperature, have a similar effect on several aspects of uncertainty at once - visibly or invisibly?

8.4 Coverage factor k

Having scaled the components of uncertainty consistently, to find the combined standard uncertainty, we may then want to re-scale the result. The combined standard uncertainty may be thought of as equivalent to ‘one standard deviation’, but we may wish to have an overall uncertainty stated at another level of confidence, e.g. 95 %.

This re-scaling can be done using a coverage factor, k. Multiplying the combined standard uncertainty, u_c, by a coverage factor gives a result which is called the expanded uncertainty, usually shown by the symbol U, *i.e.*,

$$U = k\,u_c$$

Some coverage factors (for a normal distribution) are:
k = 1 for a confidence level of approximately 68 %
k = 2.58 for a confidence level of 95 %
k = 3 for a confidence level of 99.7 %

Other, less common, shapes of distribution have different coverage factors.