Lecture "Experimental Physics I"

(Prof. Dr. R. Seidel)

Lecture 1

Introduction

- Origin and aim of physics (slides)
- Basic units (slides)
- Experimental errors

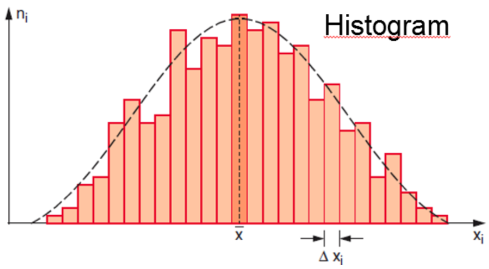
ERRORS OF EXPERIMENTAL MEASUREMENTS

1. Types of experimental errors

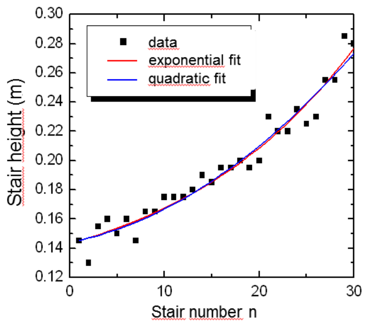
Starting point: Each measurement in physics has a certain imprecision, called the measurement error.

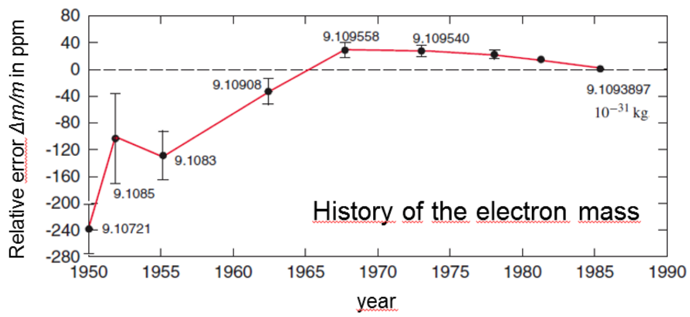
Experiments:

- Consider measuring the velocity of metal spheres after a fall through a glass tube (flight time through a photoelectric sensor at the end of the fall).
  One will see that the times vary significantly around a certain value. Most of the values are in vicinity of this value as can be seen by **binning the data in a histogram**.
- Error when reading an analog scale from different viewing angles.
- Error of the voltage of a battery when measured with three different voltmeters.



Other examples of data with errors are the scattering of the data for the stair height measurement in the lecture hall as well as the experimentally determined mass of the electron:





Question: What causes errors in experiments? Hereby we must distinguish between two general types of errors:

1) <u>Systematic errors:</u> errors that systematically and reproducibly lead to deviations from the real value. Often they are caused by the apparatus, e.g. wrong calibration of the measurement device, bad design of the measurement, external influence such as temperature drift. Often systematic errors are underestimated (see electron mass), since we must make assumption about the error sources. Alternatively, we could make further quantifications (comparison to better standards) to estimate systematic errors.

2) <u>Statistical errors:</u> statistical deviations of a measured quantity due to the random nature of the quantity or the measurement principle. They are typically non-reproducible and do not correlate to each other. Examples are random trajectories, vibrations/noise of the measurement apparatus, statistical deviations of the measured quantity itself

Definitions: When speaking about errors we define the following terms

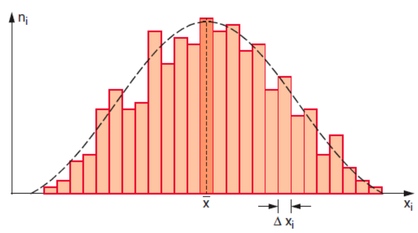
x_r	– real value (which we don't know!)
x_i	– measured value
$(x_i - x_r)$ or $\|x_i - x_r\|$	– absolute error
$\frac{x_i - x_r}{x_r}$ or $\frac{\|x_i - x_r\|}{x_r}$	– relative error

2. Quantifying statistical errors

Statistical errors: measured values with statistical errors have, by definition, **a randomness!** However, there are ways to deal and even exploit the randomness. When dealing with statistical errors, a **convenient trick to improve the measurement accuracy** is to **repeat the measurement**. This is explained in the following:

A) Mean value

Idea: The measured values x_i are scattered (distributed) around (unknown) real value x_r.



The **arithmetic mean** for a measurement of *n* values is then a **good approximation of x_r**

$$\bar{x} = \langle x \rangle = \frac{1}{n}\sum_{i=1}^{n} x_i$$

since one can show that it approaches the real value

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

B) Error of a single measurement:

To quantify the error of a single measurement with statistical errors one could try to calculate the mean of the absolute error:

$$\overline{(x_i - x_r)} = \frac{1}{n}\sum_i (x_i - x_r) \approx \frac{1}{n}\sum_i (x_i - \bar{x}) = \frac{1}{n}\sum_i x_i - \frac{1}{n}\sum_i \bar{x} = \bar{x} - \bar{x} = 0$$

This is not helpful since the mean (or x_r) lies in the middle of the distribution, where positive and negative deviations directly cancel each other out.

A better measure of the error of a single measurement is the (root)-mean square deviation, where all single error contributions are non-negative. It is called the **standard deviation**:

$$\sigma = \sqrt{\frac{1}{n}\sum_i (x_i - x_r)^2}\text{, in case we know } x_r.$$

$$\sigma \approx \sqrt{\frac{1}{n-1}\sum_i (x_i - \bar{x})^2}\text{ ,if we approximate } x_r \text{ using } \langle x \rangle$$

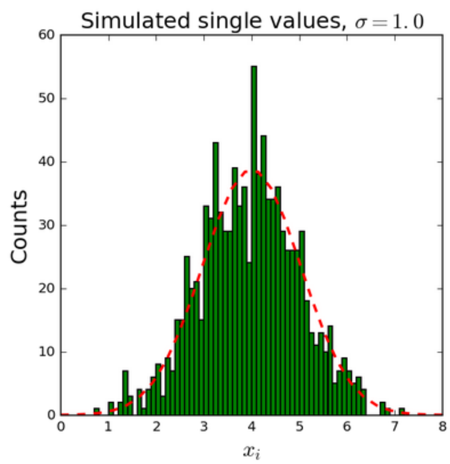
σ^2 is hereby **called the Variance**

The standard deviation is a measure of the **width of the distribution**.

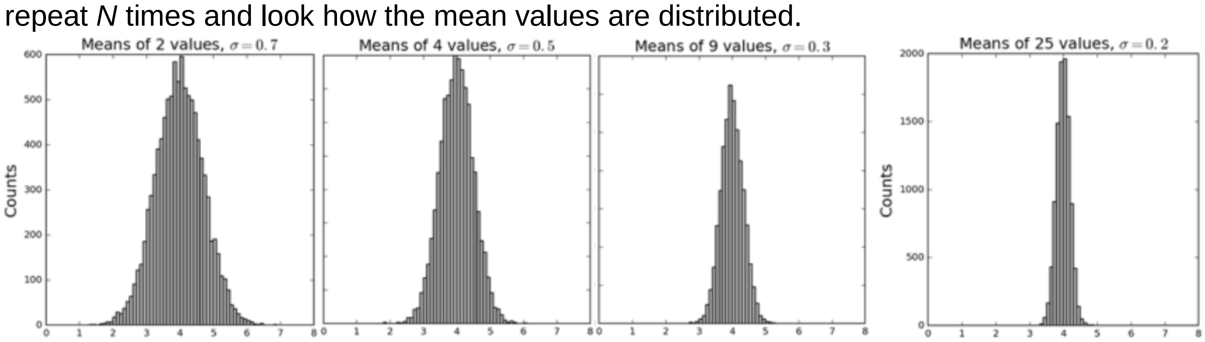
C) Error of the mean value (of multiple measurements)

Question: If we carry out our measurement with statistical error σ for *n* times, which residual error does the mean have from x_r?

For this we carry out a numerical experiment and draw random (normally distributed) measurement values on the computer being centered around a mean of 4 and having standard deviation of 1.



Now we draw *n* values from this distribution and calculate the mean. The mean calculation we repeat *N* times and look how the mean values are distributed.



We see that the width of these distributions is shrinking with increasing numbers of values used for the mean calculation, i.e. the mean approaches x_r more and more!

Now let us closely inspect the width, i.e. the standard deviation of these distributions. It can be easily from the plots that the **error of the mean decreases with** $1/\sqrt{N}$.

One can formally derive this found relationship between the error of the mean value σ_m and the error of a single measurement error σ:

$$\sigma_m = \frac{\sigma}{\sqrt{n}} = \sqrt{\frac{1}{n(n-1)}\sum_i (x_i - \bar{x})^2}$$

σ_m is called the **standard error of the mean.**

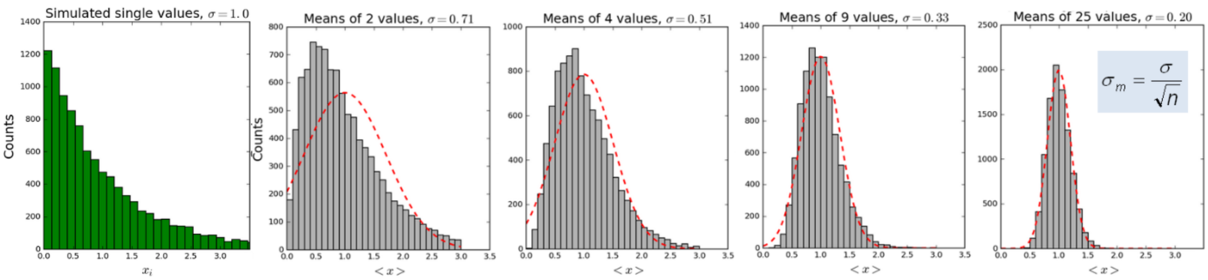
Thus, by repeating a measurement with statistical errors we can improve the accuracy of the final result!

D) Testing other distributions of single values

Let us see how general our result is by testing other distributions of single values.

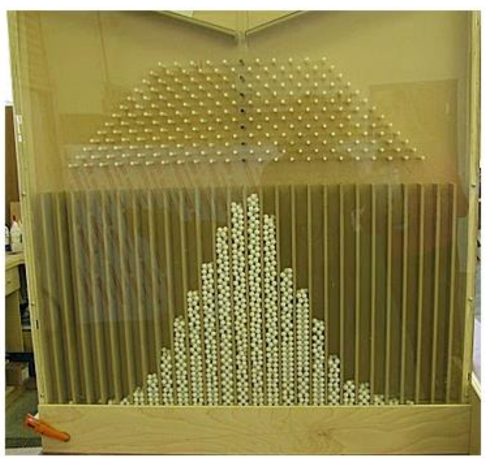
Exponential distribution of single values

We first test exponentially distributed values. The single value distribution has a mean of 1 and a standard deviation of 1.



The distribution of the means approaches again a bell-shaped distribution for increasing numbers of values for the mean! This bell-shaped distribution is called **Gaussian distribution** (red dashed line). As discussed later, one always approaches a Gaussian for the distribution of a sum (or mean) of many random variables, whatever distribution one takes for the single values.

Experiment with Galton board:



Spheres are inserted above the central pin in the upper row and can go at each pin that they encounter at the N rows with probability p to the left and with probability (1-p) to the right. This

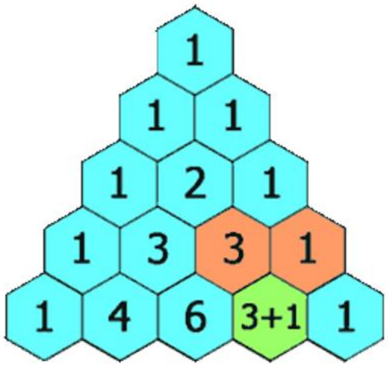
experiment can be interpreted as a **digital, binary distributed data** (either 0 or 1 step to the right per row, e.g as for the last digit of display alternates between two numbers).
The **board takes the sum of *N* binary result values** (from which one could calculate the mean be dividing by N). We see from the Galton board experiment that we get a **bell-shaped curve!** The obtained **distribution is called binomial distribution (see slide)** given by:

$$p(n - N/2) = \binom{N}{n} p^n (1-p)^{N-n} = \frac{N!}{n!\,(N-n)!} p^n (1-p)^{N-n}$$

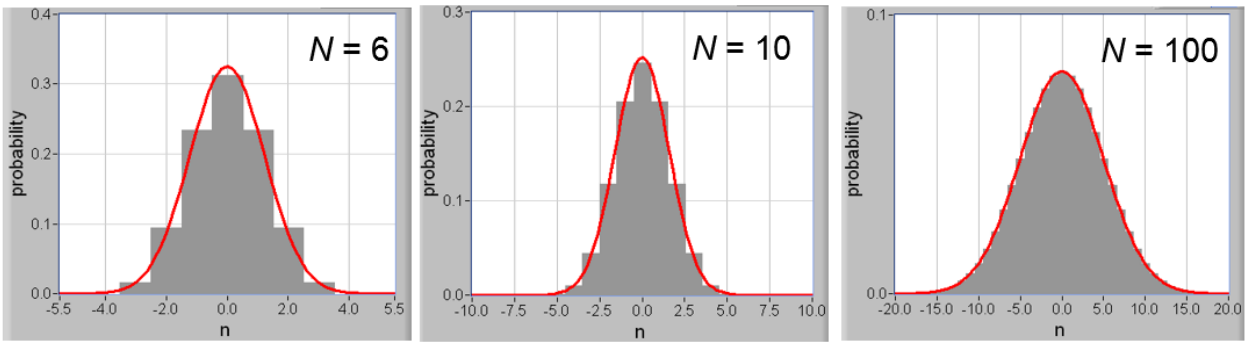
Where:

N is the number of downward steps,
n is the number of steps to the right
p is the probability for rightward step, (1-p) is the probability for leftward step

The probability of each path is given by the probability product in the equation above (equal for all paths if p = 0.5). However, there are different numbers of paths leading to one position. They are given by the **binomial coefficient** at the left side of the equation. The different path numbers can be illustrated using Pascal's triangle:



When plotting the binomial distribution for different *N* and *p* = 0.5 we get (gray bars):



When approaching large *N* also the binomial distribution approaches more & more a continuous distribution (see red line in plot above), which is **a Gaussian distribution!**

Experiment: Single photon counting with photons from a laser being collected by an avalanche photo diode. Here the number of photons collected per time interval also approaches a Gaussian (see slides)

E) Central limit theorem

Our finding that the sum of many statistically distributed data values will always look like a Gaussian, independent of the distribution of the individual values is remarkable and seems to be quite general. This finding is summarized in the **central limit theorem:**

"The sum of a large number of independent statistically distributed values approaches a Gaussian distribution"

The mean values of measurement values with statistical errors are thus always approximately Gaussian-distributed (since they were formed by such a sum):

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

A more formal statement of the Central limit theorem says the following:

Consider the sum of a large number *n* of independent statistically distributed (measurement) values then we get for the distribution of the sum:

- its expectation value is the sum of the mean values of the individual distributions
- its **variance is the sum of the individual variances**
- It is approximately Gaussian-distributed (even non-Gaussian single distributions)

From this one can easily derive the standard error of the mean. Let σ be the standard deviation of a single value, σ_s the standard deviation of the sum and σ_m the standard deviation of the mean (i.e. the standard error). Then we can write:

$$\sum_i x_i \longrightarrow \underset{\text{Variance of sum}}{\sigma_s^2 = N\sigma^2} \longrightarrow \underset{\text{Standard deviation of sum}}{\sigma_s = \sqrt{N}\sigma} \longrightarrow \underset{\text{Standard error of mean}}{\sigma_m = \frac{\sigma_s}{n} = \frac{\sigma}{\sqrt{N}}}$$

3. Confidence intervals

So far we defined standard deviation and standard error of the mean. What is however the **confidence that $\langle x\rangle - \sigma_m < x_r < \langle x\rangle + \sigma_m$?**

For this it is helpful to know that our mean **values are approximately Gaussian distributed**.

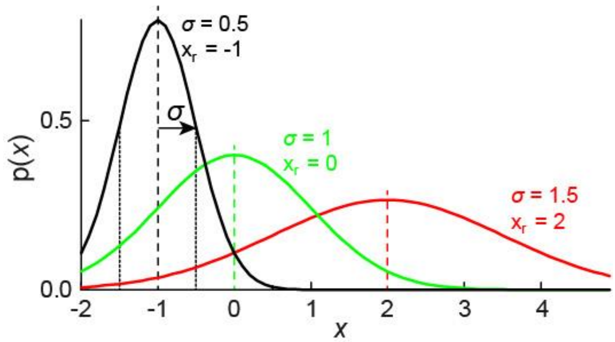
Watch out: Often it is said (also in textbooks), data that is subjected to statistical error has a bell-shaped distribution. This is often true, but generally speaking, wrong! However, for the mean values we can correctly rely on Gaussian statistics!

A) The Gaussian distribution (also Normal distribution or error function)

Let us first have a look at the Gaussian distribution. It is given by:

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-x_r)^2}{2\sigma^2}}$$

p(*x*) is a probability density, i.e. *p*(*x*)d*x* is the probability that *x* lies between x and $x + dx$



This provides the symmetric bell-shaped function that is centered around x_r.

Properties:

Normalization: $\int_{-\infty}^{\infty} p(x) \cdot dx = 1$

Mean of the distribution: $\int_{-\infty}^{\infty} x \cdot p(x) \cdot dx = x_r$

Variance of the distribution: $\int_{-\infty}^{\infty} (x - x_r)^2 p(x) \cdot dx = \sigma^2$

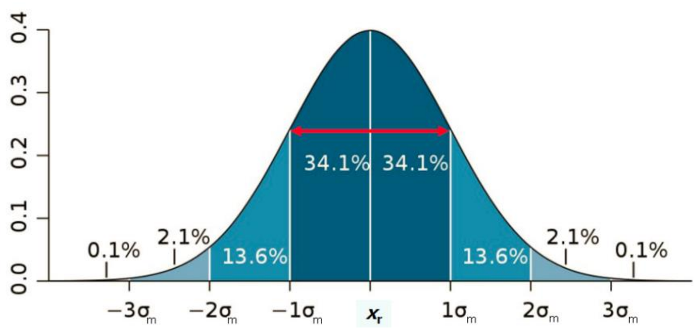
B) Confidence intervals

From the standard error of the mean we know: The obtained mean values <x> are Gaussian distributed around x_r with standard deviation $\sigma_m = \sigma/\sqrt{n}$

The probability, in this case called the confidence, that x_r lies within <x> ± Δ is thus given by integration:

$$P(\Delta) = \int_{x-\Delta}^{x+\Delta} p(x) \cdot dx$$

Carrying out a numeric integration over the Gaussian function provides the following:



$x_{r,estimated} = \langle x \rangle \pm \sigma_m$ with 68% confidence

$x_{r,estimated} = \langle x \rangle \pm 2\sigma_m$with 95% confidence

$x_{r,estimated} = \langle x \rangle \pm 3\sigma_m$with 99.7% confidence

This needs to be considered when providing statistical errors! A confidence interval of 2σ is provides already a very high probability that the obtained mean lies within the error boundaries. In particle physics 5σ are needed to make a certain measurement to be considered certain (e.g. about finding of the Higgs boson)

C) Significant figures

It is good practice to apply useful rounding of the obtained values according to their errors but also rounding of the errors themselves:

- Round values to the highest significant digit
- Provide errors with just a single significant digit
- Round errors always to the closest upper value to avoid decreasing them

Example: You obtain with the calculator for a mean and its standard error a value of :

123. 256 ± 3.7894

Sensible rounding will provide

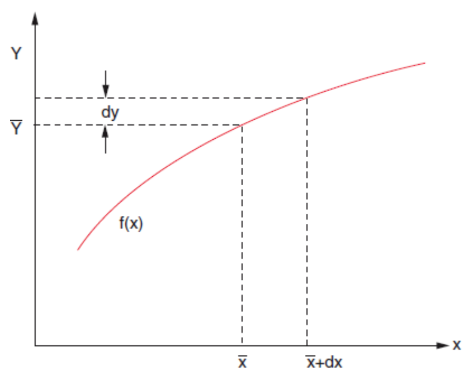
123 ± 4

in which the original confidence interval did not decrease.

D) Error propagation (not part of lecture)

Now we want to consider the error of a quantity that is calculated as a function y = f(x) of a measured value x, with known error. Consider that

$$dy = \frac{df(x)}{dx}$$



For a given deviation Δx_i from its mean <x>, we get for the corresponding deviation Δy_i from <y>:

$$\Delta y_i = y_i - \langle y \rangle \approx \left(\frac{df}{dx}\right)_{\langle x \rangle} (x_i - \langle x \rangle) = \left(\frac{df}{dx}\right)_{\langle x \rangle} \Delta x_i$$

Thus

$$\sigma_y = \sqrt{\frac{\sum(y_i - \langle y \rangle)^2}{n-1}} = \left(\frac{df}{dx}\right)_{\langle x \rangle} \sqrt{\frac{\sum(x_i - \langle x \rangle)^2}{n-1}} = \left(\frac{df}{dx}\right)_{\langle x \rangle} \sigma_x$$

Now let us ask, which is the error of a quantity that is calculated as a function of multiple measured parameters, i.e. of $f(x,y)$ (e.g. density being the ratio of mass and volume).

We start with $f(x,y) = x + y$. According to central limit theorem the variances add up in a sum:

$$\sigma_f^2 = \sigma_x^2 + \sigma_y^2$$

Now we combine both equations from above into the general expression:

$$\sigma_f = \sqrt{\sigma_x^2 \left(\frac{df}{dx}\right)^2_{\langle x \rangle} + \sigma_y^2 \left(\frac{df}{dy}\right)^2_{\langle y \rangle}}$$

Lecture 1: Experiments

1. PhyPhox: Centrifugal acceleration vs angular velocity measured with smartphone
2. Slide with step height in the lecture hall + fit with different functions
3. Spring pendulum
4. Time standards (pendulum, quartz clock, rubidium atomic clock)
5. Length standards (ruler, yard stick, caliper, laser-based distance measurement)
6. Mass measurement (weights, board scale (Roberval balance), digital balance)
7. Statistical error of a measurement: Velocity of a sphere after a fall through a glass tube
8. Systematic Error of a measurement: Views on the needle of a analog scale device from different directions (Parallaxenfehler), Voltage of battery measured with 3 different voltage meters
9. Galton's board
10. Single-photon counting: count statistics for small photonrates, transition of symmetric Poisson-distribution to Gaussian