
Data Reduction & Error Analysis

in brief

Anna Sinelnikova



UPPSALA UNIVERSITY
2017

Contents

1	Why Do We Calculate Errors?	3
2	Physical Quantity and its Measurement	3
3	Classification of errors	4
3.1	Classification of errors by nature of their occurrence	4
3.2	Classification of errors by reason of their occurrence	5
4	Quoting the Result and Scientific Notation	6
4.1	Rounding off	6
4.2	Significant digits	6
4.3	Quote the result	7
4.4	Scientific notation	7
5	Calculation of Errors in Direct Experiment	8
5.1	Average	8
5.2	Least Square Fit	8
5.3	Peaks	10
6	Calculation of Errors in Indirect Experiment	11
7	Example and Important Remark	11
7.1	Very Important Remark	12

Introduction

This booklet was written to give an idea about how errors propagate in physical measurements. For this purpose we will first discuss what a measurement of a physical quantity means and where errors come from. Then we will show how to write down the result in a proper way, using a scientific notation. We then discuss how to compute errors in different cases, in direct and indirect measurements. Finally, we give an example that explicitly shows how to use the methods.

The booklet can be useful for students who do laboratory work in physics. It might be used as a reference source for data reduction and propagation of errors.

1 Why Do We Calculate Errors?

When we solve problems, our aim can be to find some number. This number is the answer to a question like: “how many?”, “how long?”, “how much time?” etc. In physical problems the question can be: “what force?” , “what mass?” ... In all these examples we want to find *one exact number*. This number is the solution for the problem.

When we do an experiment we might also be interested in numbers that answer the same questions as above, but now instead of using a mathematical model we have to deal with real world. What is the difference?

We cannot measure anything with infinite precision in reality. This restriction is unavoidable. We can improve our devices and it can give us better and better precision, but we cannot reach this limit. So whatever we measure, we need to state how trustful the result is, i.e. how close it to the “true” value.

The quality of an experiment is defined by the value of the error. It is quite clear that if the uncertainty is larger than the found value itself, then the result of such experiment does not make any sense at all.

For example, one of the biggest discoveries of our century in physics was the first direct observation of gravitational waves by LIGO (Laser Interferometer Gravitational-Wave Observatory) on September 14, 2015¹. The experiments at LIGO started in 2002, more than 10 years ago. The theoretical predictions were made one century ago. The issue was the precision of the experiment. The errors (which were produced by noise) were larger than the expected value. It took one hundred years to invent and conduct the experiment where the result could be trusted, i.e. with small enough errors.

Error propagation is not an independent science. It uses different branches of statistics, depending on what experiment you have. That fact probably makes things even more complicated.

2 Physical Quantity and its Measurement

A physical quantity is a quantitative property of a phenomenon, body, or substance. Each physical quantity must be defined such that it unequivocally can be found experimentally, or theoretically, in any real situation. The value of a physical quantity can be expressed as a combination of a magnitude expressed by a number and a unit.

Measurement of a physical quantity is the determination or estimation of how many uniform physical quantities are contained in it. Physical quantities are uniform if they describe the same properties and differ only by value. Usually as uniform physical quantities we take

¹Nature News. doi:10.1038/nature.2016.19361. Retrieved 11 February 2016

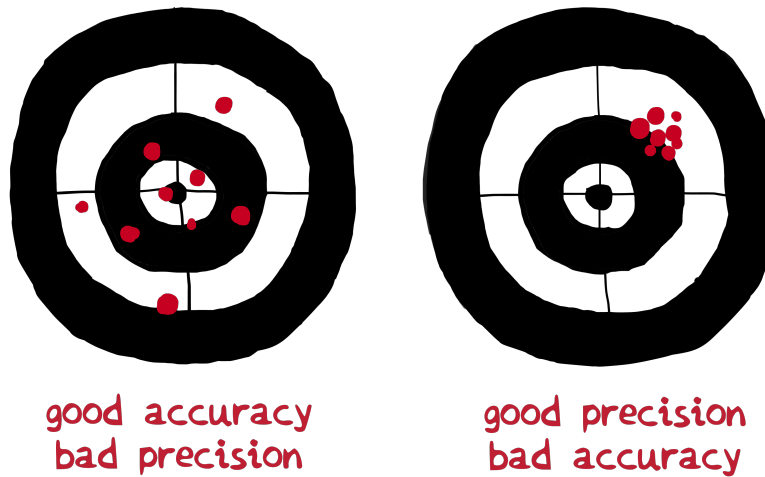


Figure 1: Illustration of the difference between precise and accurate data.

a measurement unit. Units of measurement are defined and adopted by convention or by law and form a system of units.

Example : Kilometre, metre, millimetre are units of length. Hence they are uniform. When you want to find the length of any object you have to find how many units of length are contained in the object. In the International System of Units (SI) the definition of a metre is: Length of the path travelled by light in a vacuum in $1/299\,792\,458$ of a second.

Measurement error (or Observational error) is the difference between a measured value of a quantity and its true value. The problem is that we do not know the true value, if we did there would be no reason to perform the experiment. Hence we need a way to systematically estimate errors both from data and from experimental conditions.

Errors have the same measurement units as the value itself!

3 Classification of errors

There are different ways of classifying errors. We will consider two of them here.

3.1 Classification of errors by nature of their occurrence

Random errors are errors which are caused by uncontrolled fluctuations in observations which appear randomly in each measurement. This type of errors can be estimated by repeating the same experiment several times and using statistical analysis. We can decrease this error by increasing the number of measurements. In this case we say that we increase the **precision** of the measurement value. Random errors vanish when the number of measurements is taken to infinity.

Systematic errors are constant or change according to some law during the experiment. The reason for this type of errors is errors in the tools we use in the experiment, both equipment and methods of measurement. This kind of errors can be decreased only by improving the tools themselves. We then say that we increase the **accuracy** of the measurement value (see fig. 1 for comparison accuracy and precision). There is no way to avoid systematic errors in any real experiment. The best way to reveal these errors is to measure the same quantity with completely different methods in completely different experiments.

3.2 Classification of errors by reason of their occurrence

Error of equipment is caused by fundamental imperfections of the set up. The reason can be the equipment itself (e.g a fast or slow stop-watch) or improper use of the equipment (e.g wrong calibration). This error is *systematic*.

Example : Bad graded ruler, slow or fast watch, bad calibration of the set up.

Roundoff error appears because of the result you get from any device (both analog or digital) has a finite number of significant digits. This error can be classified as a *random* error since if one performs a lot of measurements then it is equally likely that rounding off will increase and decrease the value. If there is no other source of random errors in the experiment then the device will however show the same result all the time and roundoff error becomes a *systematic* one.

Example part I: Let's assume you have quite strange scientific calculator with only 4 digits on the screen. If you now want to get value of π it shows you 3.141 (or 3.142) while π has infinite numbers of digits.

Error of calculations is similar to the roundoff error but now we are talking about rounding off during the calculations.

Example part II: Now you want to divide your π over 1 000. It should be 0.003141 but since you calculator shows only 4 digits you will see 0.003. You might think that it is not a big problem, but now try to divide π over 10 000. Instead of 0.0003141 you will get just 0!

Of course real calculators have more digits on the screen. And moreover they work with significant digits (see section 4.2) only. But as we said π , for example, has infinite number of significant digits so if you want to have π^2 you will face error of calculations again.

Representing numbers with finite number of significant digits is a big problem in computer science as well.

Error of method is caused by imperfections of the method we use in the experiment. In any physical problem, we have to use some model. It means that we have to neglect some phenomena in order to be able to solve the problem. These errors give *systematic* contributions.

Example : Not taking into account air resistance in experiments with falling bodies.

Subjective error is an error caused by the human who perform the experiment. In an experiment with a stop-watch, for example, all times will be overestimated because of the reaction time. Then we deal with *systematic* errors. Subjective errors can give a contribution to *random* errors as well. If one have to estimate half of a distance with the bare eye, for example, then the result can be biased in both sites.

Example : The reaction time (0.3 s) should be taken into account if you use manual stop-watch.

Mistakes (or illegitimate errors) are rough errors which can occur as a result of misreading the data, algebraic mistakes, damaged devices etc. These errors are usually apparent either as obviously incorrect data points or as results that are not reasonably close to the expected values. Mistakes can be corrected by carefully repeating the operations.

4 Quoting the Result and Scientific Notation

Uncertainty is the final error, after considering contributions from all errors above. Usually it is denoted by σ . It is computed as

$$\sigma_{tot} = \sqrt{\sigma_1^2 + \sigma_2^2 + \dots} \quad (1)$$

where $\sigma_1, \sigma_2, \dots$ are different types of errors which we assume to occur independently.

Discrepancy is the difference between two results. Either you compare two experimental results or the experimental result with the accepted value of the physical quantity.

We will now discuss how to write down physical quantities together with their uncertainties in a correct way (how to compute the errors we will discuss in the next chapters).

4.1 Rounding off

When some digits are dropped from a number, the last digit retained should be rounded off for the best accuracy. For example we want to round off 1.2345 to the third digit. Then the fraction after the third digit 0.45 will define the way how we should do this according the following rules:

1. If the fraction is greater than $\frac{1}{2}$, increment the last digit.
2. If the fraction is less than $\frac{1}{2}$, do not increment.
3. If the fraction equals $\frac{1}{2}$, increment the last digit only if it is odd.

The last rule reduces systematic error in rounding off numbers.

4.2 Significant digits

There are 4 rules for significant figures (let's consider number 12.34):

1. The leftmost nonzero digit is the most significant digit (1).
2. If there is no decimal point, the rightmost nonzero digit is the least significant digit (it is not our case)
3. If there is a decimal point, the rightmost digit is the least significant digit, even if it is a 0 (4).
4. All digits between the least and the most significant digit are counted as significant digits (2, 3).

So we have 4 significant digits in number 12.34. Here are other numbers with 4 significant digits:

1234; 1234000; 1.234; 0.001234; 0.01200; 123.0; 1.000

Note, that numbers **123** and **123.0** are mathematically equal but they have different numbers of significant digits!. 123.0 is more *precise* than 123.

4.3 Quote the result

Let's assume that you have already got some result in the experiment and calculated the uncertainty. We will show how the result can be quoted together with the error in the following table

Result	Answer	Comments
$L = 2.5492 \text{ m}$ $\sigma = 0.03 \text{ m}$	$L = (2.55 \pm 0.03) \text{ m}$	We have uncertainty in the third digit, i.e. in number "4" in our result. So the rest of the digits don't make sense. We should truncate the answer to those digits where we have our uncertainty.
$L = 77.323 \text{ m}$ $\sigma = 0.82 \text{ m}$	$L = (77.3 \pm 0.8) \text{ m}$	Here we have corrections to two digits in the result but the first significant digit in the error is larger than the second one. Then we can round off the result.
$L = 77.323 \text{ m}$ $\sigma = 0.28 \text{ m}$	$L = (77.32 \pm 0.28) \text{ m}$	The opposite situation: the first significant digit in the error is smaller than the second one. We can keep it then.
$T = 1322.7 \text{ s}$ $\sigma = 72.33 \text{ s}$	$T = (1320 \pm 70) \text{ s}$	Round off to the first significant digit in the uncertainty. We have to leave the zero, since $70 \neq 7$.

The last example reveals an interesting problem. We have only one significant digit in the error but we have to keep one more digit (it is not significant even) only because $70 \neq 7$. This leads us to the necessity to use *scientific notation* for writing down numbers.

4.4 Scientific notation

In scientific notation a number is written on form: m times ten raised to the power of n , that is

$$m \times 10^n,$$

where the coefficient m is any real number (called the significand or mantissa) and the exponent n is an integer (can be both positive and negative).

The table below might help to understand the idea.

Value	Scientific notation
1000	$1.0 \cdot 10^3$
0.001	$1.0 \cdot 10^{-3}$
237.83	$2.3783 \cdot 10^2$
$24.06 \cdot 10^8$	$2.4606 \cdot 10^9$
$0.003 \cdot 10^8$	$3.0 \cdot 10^5$
$0.045 \cdot 10^{-12}$	$4.5 \cdot 10^{-14}$

5 Calculation of Errors in Direct Experiment

A **direct measurement** is a measurement where you obtain the desired quantity itself and do not do any kind of calculations.

Example : measuring the length of a stick with ruler. If you instead would like to find half the length by dividing by 2 then it would be an indirect measurement. This case will be discussed in the next chapter.

You can find examples of calculation errors in chapter 7.

5.1 Average

In a lot of experiments you need to find only one value. In this case usually you do the same measurement under the same conditions several times in order to increase the precision of the result. Then the unknown quantity should be calculated as the average over all outcomes:

$$\langle x \rangle = \frac{1}{N} \sum_{i=1}^N x_i \quad (2)$$

The error of this quantity should be found as **standard deviation of average**:

$$S_{\langle x \rangle} = \sqrt{\frac{1}{N(N-1)} \sum_{i=1}^N (x_i - \langle x \rangle)^2} = \sqrt{\frac{1}{N}} S_x \quad (3)$$

where $S_x = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \langle x \rangle)^2}$ is a **standard deviation** of **one certain** measurement. It means that if you want to know the deviation of your third, for example, measurement from set of, say, 5 measurements then you use S_x with parameters $i = 3$, $N = 5$. The error of average, not of a single measurement, is given by equation 3.

Why it is so important not to mix up S_x and $S_{\langle x \rangle}$? In mathematical statistics it is shown than with increasing N , S_x tends to a constant value. Hence, $S_{\langle x \rangle}$ tends to 0:

$$\lim_{N \rightarrow \infty} S_x = \text{Const} \Rightarrow \lim_{N \rightarrow \infty} S_{\langle x \rangle} = 0 \quad (4)$$

It means that indeed, with an increasing number of measurements you will decrease random errors and increase the precision of the experiment. It can be reflected only with the correct formula 3.

5.2 Least Square Fit

Suppose that you have a spring and want to find the spring constant, having a set of blocks with different masses and a ruler. According to the Hooke's law:

$$|\vec{F}| = k|\Delta x| \quad (5)$$

In principle, you could make a set of measurements when you use only one block and measure the difference in the length of the spring. Then you would find the coefficient k as a result of indirect measurements (see section 6).

There is however a more efficient way to find k . Instead of using only one block you could use different ones with different masses. Hence you will have a table with weights F_i and the corresponding Δx_i . Theoretically all the data points should belong to the line (5), and the slope would be the desired spring constant k . But as we already know, the real experiment

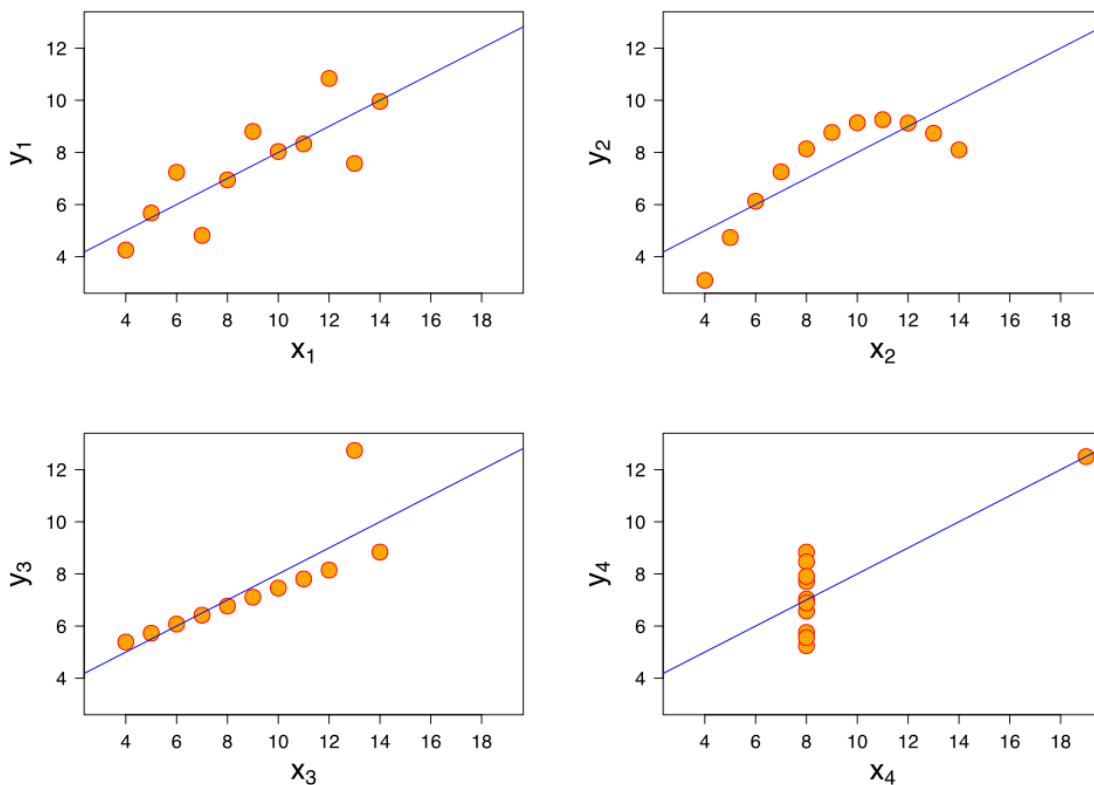


Figure 2: Illustration how the same linear regression line fits different data sets ⁴.

always has errors. In this case it means that the data points will *approximate* the line (5) but not strictly belong to it.

Least square fit is the most common way of fitting the data with any (not only linear) functions. Despite that the idea behind the least square method is simple, a proper description would be lengthy. For now just note, that least square fit is a part of *linear regression method*.

Most likely you will do the fitting of your data by using some software² such as GNUplot, Matlab, Mathematica etc. All these programs are able to do this for you. You can learn more by reading the description of particular function you are going to use. Unfortunately, a lot of programs don't present errors of fitting by default, so you have to call some extra functions.

Errors of fitting are really important, because in principle you can fit whatever data with whatever function, only errors³ will show you how good your fitting function is. The plot of a function together with the data is also useful, but it can only give subjective information, while quantitative estimation is objective information. Take a look at Figure 5.2. There you can see how different data are fitted with the same line.

Whatever software and function you decide to use **you must present the error of fitting together with the result**.

Let us come back to our example with spring constant. In that case you want to approximate the data with a line:

$$y = ax + b \quad (6)$$

where $y = F$ and $x = \Delta x$. Hopefully the coefficient b will be very small, then we can say that the Hooke's law works well in our case and the answer for stiffness is:

$$k = a \pm \sigma_a \text{ [units]} \quad (7)$$

²Scientific calculators are able to do linear fitting as well.

³ χ -square test is also important for estimation of fitting quality. But this topic lies beyond the scope of this booklet.

⁴Reference to the picture: Anscombe, Francis J. (1973) Graphs in statistical analysis. American Statistician, 27, 17–21

The coefficient b and its error are not the part of the result, however do not forget to check them as well. If b is not close to zero or the error is very big then it might be a signal that you made a mistake somewhere and have to fix it.

5.3 Peaks

In some experiments you want to measure quantities which are given to you as a peak in the measured result (e.g. wave length). Ideally you want to obtain a delta-function, where the height show the intensity of the signal and the position is the required quantity. Real equipment cannot give you this ideal precision. Instead of the desired delta-function you will have a Gaussian distribution ⁵:

$$P_G(\mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp \left\{ \frac{1}{2} \left(\frac{x - \mu}{\sigma} \right)^2 \right\} \quad (8)$$

μ is *mean- or expectation value* and this is the value of our desired quantity.

σ is a *standard deviation*, which give us the uncertainty of the measurement. You need then to fit your data with a Gaussian distribution and find μ and σ as the parameters of the fit. However, this way of doing it would be quite complicated, you would need some additional software which allows you to do this kind of fitting. In many cases (when you are not so interested in the precision of error) one can use this trick to find the standard deviation.

There is a value Γ which denotes the *full-width at the half maximum* or *half-width*.

$$P_G(\mu \pm \frac{1}{2}\Gamma, \sigma) = \frac{1}{2}P_G(\mu, \sigma) \quad (9)$$

It can be shown that Γ has the property,

$$\Gamma = 2.354 \sigma. \quad (10)$$

It means that in the first approximation **you can count the error as a half-width at the half maximum of the peak.**

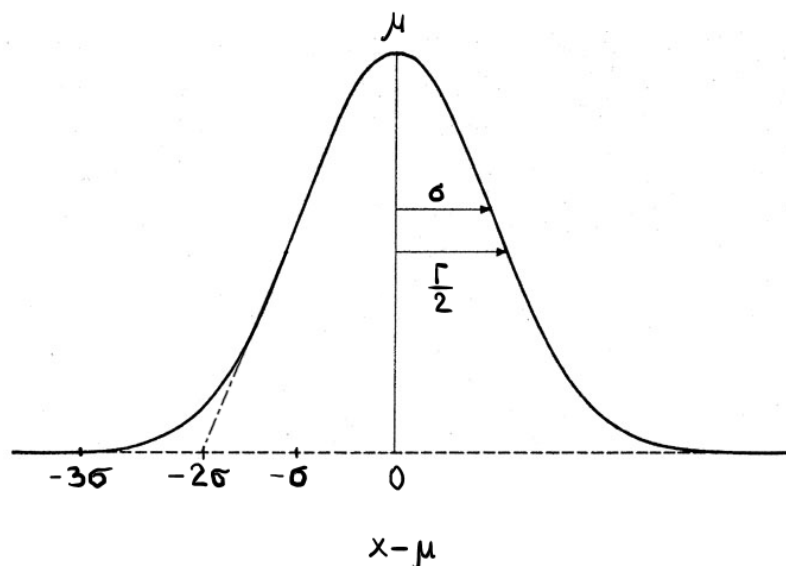


Figure 3: Gaussian probability distribution.

⁵Strictly speaking, it can be some other distribution as well, but in the overwhelming majority of cases it would be a Gaussian distribution.

6 Calculation of Errors in Indirect Experiment

If you do any kind of mathematical operations with a measured quantity then we say that the result was measured indirectly. Or that you did an **indirect measurement**.

Example : you measured two sequent time intervals. The total time which is just the sum of those intervals will be an indirect measurement.

In general, an indirect quantity u depends on direct quantities x, y, z (could be any number of quantities) as some function:

$$u = f(x, y, z) \quad (11)$$

The random error can then be calculated as a function of the direct quantities and their errors like this:

$$S_u = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 \cdot S_{\langle x \rangle}^2 + \left(\frac{\partial f}{\partial y}\right)^2 \cdot S_{\langle y \rangle}^2 + \left(\frac{\partial f}{\partial z}\right)^2 \cdot S_{\langle z \rangle}^2} \quad (12)$$

where $\frac{\partial f}{\partial i}$ is a partial derivative. For systematic errors the equation is the same but instead of random errors for direct measurements $S_{\langle i \rangle}$ one should take systematic ones $\sigma_{\langle i \rangle}$.

In the next chapter you can learn how to deal with this case by one simple example.

7 Example and Important Remark

Let's consider one simple example which covers a lot of what we have discussed above.

Kristofer wants to find the frequency of a swinging pendulum. He measures the time t for $n=50$ periods of the pendulum and then use the formula:

$$\nu = \frac{n}{t} \quad (13)$$

He repeats the measurement $N = 5$ times and puts his results into the table:

No.	t (s)
1	105.7
2	106,5
3	105.9
4	109.2
5	110.0

The average according to eq. 2 is:

$$\langle t \rangle = \frac{1}{N} \sum_{i=1}^N t_i = \frac{537.3}{5} = 107.46 \text{ s} \quad (14)$$

In order to calculate the standard deviation of the average (eq. 3), it is convenient to add an extra column to the table for $(t_i - \langle t \rangle)^2$:

No.	t (s)	$(t_i - \langle t \rangle)^2$ (s ²)
1	105.7	3.0976
2	106.5	0.9216
3	105.9	2.4336
4	109.2	3.0276
5	110.0	6.4516

The error is found as:

$$S_{\langle t \rangle} = \sqrt{\frac{1}{N(N-1)} \sum_{i=1}^N (t_i - \langle t \rangle)^2} = 0.8925245 \text{ s} \quad (15)$$

Note that Kristofer left more digits than needed, because he will use this number for further calculations. He will do a proper roundoff at the end.

The expectation value for the frequency is:

$$\langle \nu \rangle = \frac{n}{\langle t \rangle} = 0.465289 \frac{1}{\text{s}} \quad (16)$$

Random errors for indirect measurement (eq. 12) can be calculated as:

$$S_{\langle \nu \rangle} = \sqrt{\left(\frac{\partial \nu}{\partial t}\right)^2 \cdot S_{\langle t \rangle}^2} = \sqrt{\left(\frac{-n}{\langle t \rangle^2}\right)^2 \cdot S_{\langle t \rangle}^2} = 0.0038644 \quad (17)$$

Now Kristofer wants to take into account systematic errors. He assumes that equipment error can be neglected because it is much smaller than the reaction time: $\sigma_{react} = 0.3 \text{ s}$

Using the same formula, but for systematic errors, he gets:

$$\sigma_{\langle \nu \rangle} = \sqrt{\left(\frac{-n}{\langle t \rangle^2}\right)^2 \cdot \sigma_{\langle t \rangle}^2} = 0.00237 \quad (18)$$

Finally he finds the uncertainty using eq. 1:

$$\sigma_{\langle \nu \rangle}^{tot} = \sqrt{S_{\langle \nu \rangle}^2 + \sigma_{\langle \nu \rangle}^2} = 0.0045 \quad (19)$$

Then Kristofer presents the result of his experiment:

$$\nu = 0.465 \pm 0.005 \text{ s}^{-1} \quad (20)$$

7.1 Very Important Remark

The uncertainty in the example is quite small, only in the third digit. One can doubt if the real value of the frequency indeed lies in the interval $[0.460, 0.470]$? The answer is more likely *no* and *probably it should not!* From the first sight it seems we contradict our own definition of uncertainty from the beginning of the booklet but let us look closer.

As you remember, when we discussed random errors, we always referred to statistics. Everything here should be understood by means of probabilities.

First of all, we mentioned the law of large numbers when we talked about equation 3, for example. It means that the theory was developed for a number of measurements $N \rightarrow \infty$. In our example Kristofer has only $N = 5$ independent measurements, which is very far from infinity.

The second, even if N is large enough the probability that the true value lies in the interval $[\langle x \rangle - \sigma, \langle x \rangle + \sigma]$ is only about 70%! Even for this idealistic case, where $N \rightarrow \infty$ we can get 100% probability for the true value only in the interval $[\langle x \rangle - \infty, \langle x \rangle + \infty]$. However, it is possible to increase the probability up to 99%.

This subject lies beyond the scope of the booklet, but you can search for *confidence intervals*.

In practice in order to take into account both a limited number of measurements and a probability distribution, one can multiply the final uncertainty by some known coefficient. Search for *Student's coefficient* for example.

For the laboratory works you do not need these more advanced methods for error propagation but be sure that you understand why the true value can lie outside the interval $[\langle x \rangle - \sigma, \langle x \rangle + \sigma]$. And be ready to explain it.

Acknowledgements

I would like to thank Johann Schmidt and Tomas Löthman for productive discussions. As well as Lisa Freyhult and Peter Oppeneer for their corrections.

References

- [1] Митин И.В, Русаков В.С. *Анализ и обработка экспериментальных данных*. - М.:НЭВЦ ФИПТ. 1998. - 48с. (Russian). Available online in Russian: http://ffmgu.ru/images/4/4c/Mitin_rusakov.pdf
- [2] Philip R. Bevington, D. Keith Robinson. (2003). *Data reduction and error analysis for the physical sciences. Third edition*. New York, NY: McGraw-Hill.
- [3] Wikipedia - the free encyclopaedia. <http://www.wikipedia.org>.