Analysis of measured data and writing a protocol

1. Contents and principles of writing a protocol

The results of the measurements are recorded with a pen into a form handed out by the assistants. This form has to be signed by all students performing the given work, the assistant and the instructor after the work is finished. This form will be an inherent part of the submitted protocol.

The protocol has to contain:

a) Valid form (complete and signed) with measured data
b) Part with examples of calculations
c) Part with required graphs
d) Discussion of obtained data

b) The part with calculations will show one example of each calculation used in evaluating the measured data. Each example will contain the general expression, expression with substituted numbers, the result and the units. If you used specific mathematical software for doing these calculations, please refer to it in the protocol. You should obey rules for working with approximate numbers. The accuracy of results is given by the accuracy of the measured data which is not that high for technical measurements in a laboratory. That is why it is necessary to use a reasonable number of valid figures (rounding).

c) All graphs must contain a legend explaining the meaning of all plotted points and fitted lines. Both axes should have reasonable ranges so that the plots are clearly depicted in the graphs.

2. The use of statistical methods for the analysis of measured data

We often measure physical quantities (pressure, temperature, volumetric flow rate, etc.) that can affect the course of the observed or monitored processes or they can be the result of that processes. The measured set of data will then have to be processed in such a way that the set of data will characterize either a chosen invariable quantity or dependence between two or more quantities. Every measurement is biased with various errors that can be caused e. g. by the choice of the measurement method, sensitivity of the measurement device or its operators (experimentalists). The errors caused by the inaccuracies of measurement devices or chosen measuring methods are *systematic errors*. In this case all the measurements are biased with the same error. The size of systematic errors cannot be influenced by the students in the laboratory since the methodology of the measurement is given and described in a given instruction manual. The measuring instruments are also mounted on/to experimental apparatuses and the students are not allowed to unmount or calibrate them. *Large error* can sometimes occur during a measurement. This type of errors is usually caused by the negligence of the experimentalist or not keeping the given measurement methodology. Large errors are usually apparent in the final set of experimental data. If we obtain one or more results significantly different from the others, we will remove these results from the data set as being biased with large errors. If we are not sure whether the data is biased with large errors, we will use some of the statistical methods for exclusion of this biased data.

Most of the errors are random errors that are irregular and vary from experiment to experiment based on the immediate conditions of the experiment. These random errors consist of a large number of tiny errors that are hard to control (fluctuation of the measuring device, effect of the environment, etc...). By repeating the measurement of the same quantity, we will obtain data that are somewhat different due to the effect of random errors. We consider that the random errors have a normal (Gauss) probability distribution. The final result of such measurements is an estimate of the mean value and its error is a standard deviation. If the errors have normal distribution, the best estimate of the mean value is an arithmetic average.

Arithmetic average $\bar{x}$ calculated from *n* values is given as:

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Unbiased sample variance is defined as the sum over deviations of individual measurements from their arithmetic average raised to the second power and divided by the number of measurements minus 1:

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

Standard deviation is then a square root of the sample variance:

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

The final value of the measurement can be expressed in the form $\bar{x} \pm s_{\bar{x}}$ where $s_{\bar{x}}$ is the standard deviation of the arithmetic average defined as:

$$s_{\bar{x}} = \frac{s_x}{\sqrt{n}}$$

It can be proved that in case of random errors with normal distribution, the correct value lies within the interval $\bar{x} \pm s_{\bar{x}}$ with probability of only 68 %. The result given as the arithmetic average and its standard deviation also means that the mean value of the set of measured data can be off the given range with probability of 32 %. Such a high probability of a wrong estimate is often too large for practical applications. Intervals in which the correct value lies with much higher probability (e. g. 95 %) are often required. This interval is called confidence interval. For a mean value of the measured quantity, the $100 \cdot (1-\alpha)\%$ confidence interval is given by:

$$\left(\bar{x} - t_{n-1,\alpha} s_{\bar{x}},\ \ \bar{x} + t_{n-1,\alpha} s_{\bar{x}}\right)$$

where $\bar{x}$ is the average, $s_{\bar{x}}$ is the standard deviation of the average, $t_{n-1,\alpha}$ is the critical value of Student's distribution (given in table below), $100 \cdot (1-\alpha)\%$ is the confidence level expressed as a percentage, and $\nu = n-1$ is the number of degrees of freedom.

Critical values t_α of Student's distribution

Degrees of freedom	t_α				
ν	$\alpha = 10\ \%$	$\alpha = 5\ \%$	$\alpha = 2\ \%$	$\alpha = 1\ \%$	$\alpha = 0{,}1\ \%$
1	6,31	12,71	31,82	63,66	636,62
2	2,92	4,30	6,97	9,92	31,60
3	2,35	3,18	4,45	5,84	12,94
4	2,13	2,78	3,75	4,60	8,61
5	2,02	2,57	3,37	4,03	6,86
6	1,94	2,45	3,14	3,71	5,96
7	1,89	2,36	3,00	3,50	5,41
8	1,86	2,31	2,90	3,36	5,04
9	1,83	2,26	2,82	3,25	4,78
10	1,81	2,23	2,76	3,17	4,59
11	1,80	2,20	2,72	3,11	4,44
12	1,78	2,18	2,68	3,05	4,32
13	1,77	2,16	2,65	3,01	4,22
14	1,76	2,14	2,62	2,98	4,14
15	1,75	2,13	2,60	2,95	4,07
16	1,75	2,12	2,58	2,92	4,02
17	1,74	2,11	2,57	2,90	3,97
18	1,73	2,10	2,55	2,88	3,92
19	1,73	2,09	2,54	2,86	3,88
20	1,72	2,09	2,53	2,85	3,85
21	1,72	2,08	2,52	2,83	3,82
22	1,72	2,07	2,51	2,82	3,79
23	1,71	2,07	2,50	2,81	3,77
24	1,71	2,06	2,49	2,80	3,75
25	1,71	2,06	2,48	2,79	3,73
30	1,70	2,04	2,46	2,75	3,65
40	1,68	2,02	2,42	2,70	3,55
60	1,67	2,00	2,39	2,66	3,46
120	1,66	1,98	2,36	2,62	3,37
∞	1,64	1,96	2,33	2,58	3,29

Literature: Pechoč V., Vyhodnocování měření a početní metody v chemickém inženýrství, SNTL, Praha 1981