MINISTRY OF HEALTH OF UKRAINE
BOGOMOLETS NATIONAL MEDICAL UNIVERSITY

Department of Analytical, Physical and Colloid Chemistry

ELECTIVE COURSE

“Basics of Chemical Metrology”

Student notebook

Kyiv – 2022

Developers:

Yaroslava Pushkarova – PhD in Chemistry, Associate Professor, Associate Professor of the Department of Analytical, Physical and Colloid Chemistry;

Galina Zaitseva – PhD in Chemistry, Associate Professor, Head of the Department of Analytical, Physical and Colloid Chemistry.

“A measurement is a kind of language. It gives everyone something to agree upon so that effective communication and commerce can take place. Good measurements must be defined by their accuracy, precision and need. Of course, measurement capabilities should fit the need, but new measurement capabilities can often redefine those needs”.
David Layden, cited in Quality Progress, February 2001

This student notebook for elective course “Basics of Chemical Metrology” contains the necessary information for successful preparing for seminars and allows you to keep your written records.

The student notebook is for second year foreign students of Pharmaceutical Faculty.

Have fun,enjoy your works,and best of luck with it!

Contents

Topic 1. Rounding of the measurements results. Estimation of the presence of outliers in results. 4

Topic 2. Calculation of the basic metrological characteristics of analysis methods. Presentation of the results of quantitative analysis. 8

Topic 3. Comparison of two methods by reproducibility analysis. Comparison of the mean values of two samples. 11

Topic 4. Estimation of the repeatability of results. Estimation of the accuracy of measurements 15

Topic 5. Calculation and statistical evaluation of linear dependence parameters. Estimation of the limit of detection. 18

Appendixes 23

Topic 1. Rounding of the measurements results. Estimation of the presence of outliers in results.

Data may be considered to be one of the vital fluids of modern civilization. Data are used to make decisions, to support decisions already made, to provide reasons why certain events happen, and to make predictions on events to come.

The management of simple data sets is an important operation for many scientists and engineers. Data should be looked at, critically, for errors and consistency. Questions on whether all members of a set are peers or whether outliers are present need to be answered in most cases. Ways to combine data from various sources or that produced by the same source at different times is sometimes an important issue.

General aims of the seminar:

- ✓ to know rules for determining the number of significant figures;
- ✓ to know the scientific notation of very large or very small numbers;
- ✓ to know rules for rounding of data in a computation;
- ✓ to know what are outliers;
- ✓ to know the Rule of the Huge Error;
- ✓ to know the Dixon Test (Q-test);
- ✓ to know the Grubbs Test.

Key questions of the seminar:

1. Significant figures.
2. Rounding of data in a computation.
3. Outliers.
4. Procedures for a statistically supported decision with respect to rejection or retention of a suspected outlier.

Key formulas of the seminar:

Dixon test (the ***Q*-test**)

$$Q = \frac{|suspect\ value - nearest\ value|}{(largest\ value - smallest\ value)}$$

Grubbs Test

$$G = \frac{|suspect\ value - \ mean|}{standard\ deviation}$$

ID of course: 7412

https://likar.nmuofficial.com/md/course/view.php?id=7412

You can find on the platform for successful preparing for seminar:

– lectures;
– methodical instructions;
– recommended literature.

Questions and tasks for self-control:

1. Rules for determining the number of significant figures.
2. Rules for rounding of data in a computation (*addition*, *subtraction*, *multiplication*, *division*, *powers and roots*, *logarithms*).
3. Why is the question of outliers an important consideration? Answer from the point of view of their influence on a data set, the danger of false rejection, and the problem of false retention.
4. Describe the rule of the huge error.
5. Discuss the similarities and differences in the Dixon test and the Grubbs test for decisions on possible outlier data.

Example 1. How many significant figures?

1) 128.3; 2) 0.015; 3) 709;

4) 3×10^{-3}; 5) 5.71×10^{4} 6) 0.01300×10^{-2}

Your answer and explanation

Example 2. Calculate pH if molarity of H^+ is 4.3×10^{-11} mole/L. Give the answers with correct number of significant figures.

Your answer and explanation

Example 3. A set of data was taken to calibrate a measuring instrument. In the course of a preliminary manual plot of the data (eye estimate), it was noticed that one point appeared to be much farther from the line than the others, so it was ignored for the moment. The analyst was tempted to reject it because he feared (and rightly so) that its inclusion (if it were an outlier) would falsify his least squares fit. The average deviation of the plotted points from the line (observed – curve) was 0.012. The deviation of the suspected point from the line was 0.06. Use the Huge Error Rule for statistically supported decision with respect to rejection or retention of a suspected outlier.

Your answer and explanation

Example 4. Consider the following results. Use the Dixon test and Grubbs test for statistically supported decision with respect to rejection or retention of a suspected outlier.

9, 12, 12, 13, 13, 14, 15

Your answer and explanation

Topic 2. Calculation of the basic metrological characteristics of analysis methods. Presentation of the results of quantitative analysis.

This seminar reviews the procedures most commonly used to estimate the statistical parameters of measurement data upon which statistically based decisions about scientific and engineering data are made. A thorough understanding of these procedures will not only permit one to use them intelligently for data analysis but should also prove to be useful for better planning and execution of measurement programs.

The science of measurement is called metrology and it is fast becoming a recognized field in itself. Special branches of metrology include engineering metrology, physical metrology, chemical metrology, and biometrology. The distinguishing feature of all metrologists is their pursuit of excellence in measurement as a profession.

Metrologists do research to advance the science of measurement in various ways. They develop measurement systems, evaluate their performance, and validate their applicability to various special situations. Metrologists develop measurement plans that are cost effective, including ways to evaluate and assess data quality. Statistics play a major role in all aspects of metrology since metrologists must contend with and understand variability.Statistical techniques provide the basis for design of measurement programs including the number of samples, the calibration procedures and the frequency of their application, and the frequency of control sample measurement.

General aims of the seminar:

- ✓ to know examples of application of statistical methods in chemistry;
- ✓ to know difference between sample and population;
- ✓ to know formulas for calculation the basic metrological characteristics (mean, standard deviation, variance, coefficient of variation, relative standard deviation, standard deviation of the mean (standard error of mean), degrees of freedom, confidence interval (confidence limits)).

Key questions of the seminar:

1. Application of statistical methods in chemical experiments.
2. Basic metrological (statistical) characteristics.
3. Presenting data.

Key formulas of the seminar:

The mean, $\bar{x}$, of n measurements: $\bar{x} = \frac{\sum x_i}{n}$.

The standard deviation, s, of n measurements: $s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$.

Variance = the square of the standard deviation, s^2.

Coefficient of variation: CV = $\frac{s}{\bar{x}}$.

Relative standard deviation: RSD = $100 \times \frac{s}{\bar{x}}$.

Standard deviation of the mean or standard error of mean: $s_{\bar{x}} = \frac{s}{\sqrt{n}}$.

Degrees of freedom: $\nu = n - 1$.

Confidence limits of the mean: $\bar{x} \pm t_{n-1} \frac{s}{\sqrt{n}}$

ID of course: 7412

https://likar.nmuofficial.com/md/course/view.php?id=7412

You can find on the platform for successful preparing for seminar:

- lectures;
- methodical instructions;
- recommended literature.

Questions and tasks for self-control:

1. Give examples of the application of statistical methods in chemical experiments.
2. What is mean?
3. What is standard deviation?
4. What is variance?
5. What are coefficient of variation and relative standard deviation?
6. What is standard deviation of the mean (standard error of mean)?
7. What is degrees of freedom?
8. What is confidence interval (confidence limits)?

Example 1. When determining the Nickel content in the standard steel sample, the following results were obtained (%): 5.1; 5.5; 5.4; 5.8; 5.2. Calculate the basic metrological (statistical) characteristics.

Your answer and explanation

Example 2. The sodium ion content of a urine specimen was determined by using an ion-selective electrode. The following values were obtained: 102, 97, 99, 98, 101, 106 mM. What are the 95% and 99% confidence limits for the sodium ion concentration?

Your answer and explanation

Topic 3. Comparison of two methods by reproducibility analysis. Comparison of the mean values of two samples.

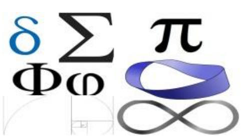
In order to decide whether the difference between the measured and standard amounts can be accounted for by random error, a statistical test known as a significance test can be employed. As its name implies, this approach tests whether the difference between the two results is significant, or whether it can be accounted for merely by random variations. Significance tests are widely used in the evaluation of experimental results. This seminar considers several tests which are particularly useful to analytical chemists.

General aims of the seminar:

- ✓ to know algorithm for comparison of standard deviations;
- ✓ to know algorithm for comparison of two experimental means if the standard deviations do not differ significantly;
- ✓ to know algorithm for comparison of two experimental means if the standard deviations differ significantly.

Key questions of the seminar:

1. F-test for the comparison of standard deviations.
2. Comparison of two experimental means.
2.1. Case I. The standard deviations do not differ significantly.
2.2. Case II. The standard deviations differ significantly.



Key formulas of the seminar:

F-test for the comparison of standard deviations:

$$F = \frac{s_{larger}^2}{s_{smaller}^2}$$

Comparison of two experimental means

Case I. The standard deviations do not differ significantly.

Step 1. Pool the *s* values to obtain s_p, a better estimate of the standard deviation.

$$s_p = \sqrt{\frac{s_1^2 \times df_1 + s_2^2 \times df_2 + \cdots + s_k^2 \times df_k}{df_1 + df_2 + \cdots + df_k}}$$

$$df = n_A + n_B - 2$$

Step 2. Calculate the variance of $\bar{x}_A$ and $\bar{x}_B$

$$V_A = \frac{s_p^2}{n_A}$$

$$V_B = \frac{s_p^2}{n_B}$$

Step 3. Choose the probability level (i.e., significance level) for the decision.

Step 4. Calculate the uncertainty of the difference of the two means at the probability level chosen $U_\Delta = t \times \sqrt{V_A + V_B}$ using a value for *t* corresponding to df = $n_A + n_B - 2$ and the probability level chosen.

Step 5. If Δ is greater than U_Δ, conclude that the means differ; if Δ is not greater than U_Δ, there is no reason to believe that the means differ

Comparison of two experimental means

Case II. The standard deviations differ significantly.

Step 1. Calculate the variances for the respective means

$$V_A = \frac{s_A^2}{n_A}$$

$$V_B = \frac{s_B^2}{n_B}$$

Step 2. Choose the probability level for the decision.

Step 3. Calculate the *effective number of degrees of freedom*, *f*, to be used for selecting *t* when calculating U_Δ. This is necessary because adding df_A and df_B is like adding apples and oranges!

$$f = \frac{(V_A + V_B)^2}{\frac{V_A^2}{n_A - 1} + \frac{V_B^2}{n_B - 1}}$$

Step 4. Calculate the uncertainly of $U_\Delta = t^* \times \sqrt{V_A + V_B}$ where t^* is based on f and the chosen probability level.

Step 5. If U_Δ is greater than Δ, there is no reason to believe that the means differ.

ID of course: 7412
https://likar.nmuofficial.com/md/course/view.php?id=7412
You can find on the platform for successful preparing for seminar:

- lectures;
- methodical instructions;
- recommended literature.

Questions and tasks for self-control:

1. What is the test for statistical significance?
2. Algorithm for comparison of standard deviations.
3. Algorithm for comparison of two experimental means if the standard deviations do not differ significantly;
4. Algorithm for comparison of two experimental means if the standard deviations differ significantly.

Example 1. Assume that an analyst used a certain method to determine the amount of lead in a water sample and estimated s = 2.5 µg/L based on 7 measurements. At a later date, she analyzed another water sample for its lead content, using the same method, and obtained s = 3.3 µg/L based on 10 measurements. Were both sets of measurements equally precise with 95% confidence?

Your answer and explanation

Example 2. Make calculations according to the case I. The two sets of data are the following:

$\bar{x}_A = 50$ mg/L, $s_A = 2$ mg/L, $n_A = 5$

$\bar{x}_B = 45$ mg/L, $s_B = 1.5$ mg/L, $n_B = 6$

$\Delta = 5$ mg/L

Your answer and explanation

Example 3. Make calculations according to the case II. The two sets of data are the following:

$\bar{x}_A = 50$ mg/L, $s_A = 2$ mg/L, $n_A = 5$

$\bar{x}_B = 45$ mg/L, $s_B = 1.5$ mg/L, $n_B = 6$

$\Delta = 5$ mg/L

Your answer and explanation

Topic 4. Estimation of the repeatability of results. Estimation of the accuracy of measurements

Since quantitative studies play a dominant role in any analytical laboratory, it must be accepted that the errors that occur in such studies are of supreme importance. Our guiding principle will be that no quantitative results are of any value unless they are accompanied by some estimate of the errors inherent in them. This principle naturally applies not only to analytical chemistry but to any field of study in which numerical experimental results are obtained.

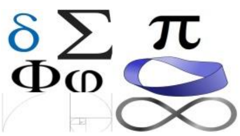
One of the most important properties of an analytical method is that it should be free from systematic error. This means that the value which it gives for the amount of the analyte should be the true value. This property of an analytical method may be tested by applying the method to a standard test portion containing a known amount of analyte.

General aims of the seminar:

- ✓ to know the difference between precision, bias and accuracy;
- ✓ to know the difference between repeatability and reproducibility;
- ✓ to know how estimate systematic error;
- ✓ to know how estimate repeatability;
- ✓ to know how compare an experimental mean with a known value.

Key questions of the seminar:

1. Precision, bias and accuracy.
2. Repeatability and reproducibility.
3. Comparison of an experimental mean with a known value.



Key formulas of the seminar:

Repeatability is evaluated by the value of the relative standard deviation.

Relative standard deviation (RSD) = $100 \times \frac{s}{\bar{x}}$.

In order to decide whether the difference between $\bar{x}$ and μ is significant, *t* is calculated:

$$t = \frac{(\bar{x}-\mu)\sqrt{n}}{s}.$$

ID of course: 7412
https://likar.nmuofficial.com/md/course/view.php?id=7412
You can find on the platform for successful preparing for seminar:
- lectures;
- methodical instructions;
- recommended literature.

Questions and tasks for self-control:

1. What is precision?
2. What is bias?
3. What is accuracy?
4. Explain the difference between precision, bias and accuracy.
5. What is repeatability?
6. What is reproducibility?
7. Explain the difference between repeatability and reproducibility.
8. How one can estimate systematic error in results?
9. How one can estimate repeatability?
10. How one compare an experimental mean with a known value?

Example 1. The absorbance scale of a spectrometer is tested at a particular wavelength with a standard solution which has an absorbance given as 0.470. Ten measurements of the absorbance with the spectrometer give = 0.461, and *s* = 0.003. Find the 95% confidence interval for the mean absorbance as measured by the spectrometer, and hence decide whether a systematic error is present. Make conclusion about accuracy of results.

Your answer and explanation

Example 2. In a new method for determining selenourea in water, the following values were obtained for tap water samples spiked with 50 ng / ml of selenourea: 50.4, 50.7, 49.1, 49.0, 51.1 ng / ml. Is there any evidence of systematic error?

Your answer and explanation

Example 3. The following values (expressed as percentages) give the antibody concentration in human blood serum for a sample of eight healthy adults. 1.13, 1.45, 1.35, 1.09. Estimate the repeatability of results using relative standard deviation.

Your answer and explanation

Topic 5. Calculation and statistical evaluation of linear dependence parameters. Estimation of the limit of detection.

The statistical procedures used with instrumental analysis methods must provide as always information on the precision and accuracy of the measurements. They must also reflect the technical advantages of such methods, especially their ability to cover a great range of concentrations (including very low concentrations), and to handle many samples rapidly.
One of the principal benefits of using instrumental methods of analysis is that they are capable of detecting and determining trace and ultra-trace quantities of analytes. These benefits have led to the appreciation of the importance of very low concentrations of many materials, for example in biological and environmental samples, and thus to the development of many further techniques in which lower limits of detection are a major criterion of successful application. It is therefore evident that statistical methods for assessing and comparing limits of detection are of importance.

General aims of the seminar:

- ✓ to know algorithm for calculation the product-moment correlation coefficient;
- ✓ to know algorithm for calculation the slope of regression line;
- ✓ to know algorithm for calculation the intercept of regression line;
- ✓ to know algorithm for calculation the errors in the slope and intercept of the regression line;
- ✓ to know algorithm for calculation the limit of detection.

Key questions of the seminar:

1. The product-moment correlation coefficient.
2. The line of regression of *y* on *x*.
3. Errors in the slope and intercept of the regression line.
4. Limit of detection.

Key formulas of the seminar:

Product-moment correlation coefficient:

$$r = \frac{\sum\{(x_i - \bar{x}) \times (y_i - \bar{y})\}}{\{[\sum(x_i - \bar{x})^2] \times [\sum(y_i - \bar{y})^2]\}^{1/2}}$$

Slope of least squares line:

$$b = \frac{\sum\{(x_i - \bar{x}) \times (y_i - \bar{y})\}}{\sum(x_i - \bar{x})^2}$$

Intercept of least squares line:

$$a = \bar{y} - b \times \bar{x}$$

Limit of detection can be defined as the analyte concentration giving a signal equal to the blank signal, y_B, plus three standard deviations of the blank, s_B:

$$Limit\ of\ detection = y_B + 3s_B.$$

ID of course: 7412

https://likar.nmuofficial.com/md/course/view.php?id=7412

You can find on the platform for successful preparing for seminar:

- lectures;
- methodical instructions;
- recommended literature.

Questions and tasks for self-control:

1. What is the product-moment correlation coefficient?
2. Calculation the product-moment correlation coefficient.
3. Test for a significant correlation.
4. Calculation the slope of regression line.
5. Calculation the intercept of regression line.

6. Calculation the errors in the slope of the regression line.
7. Calculation the errors in the intercept of the regression line.
8. What is the limit of detection?
9. Calculation the limit of detection.

*Example 1.*Standard aqueous solutions of fluorescein are examined in a fluorescence spectrometer, and yield the following fluorescence intensities (in arbitrary units):

Fluorescence intensities:	2.1	5.0	9.0	12.6	17.3	21.0	24.7
Concentration, pg ml^{-1}	0	2	4	6	8	10	12

Determine the correlation coefficient, *r*.

Your answer and explanation

Example 2. Calculate the slope and intercept of the regression line for the data given in the previous example 1.

Your answer and explanation

Example 3. Calculate the standard deviations and confidence limits of the slope and intercept of the regression line calculated in Example 2.

Your answer and explanation

Example 4. Estimate the limit of detection for the fluorescein determination studied in the previous examples.

Your answer and explanation

Appendixes

Table 1. Values for use in the Dixon Test for Outliers

Statistic	Number of Observations, n	Risk of False Rejection 0.5%	1%	5%	10%
	3	.994	.988	.941	.886
	4	.926	.889	.765	.679
τ_{10}	5	.821	.780	.642	.557
	6	.740	.698	.560	.482
	7	.680	.637	.507	.434
	8	.725	.683	.554	.479
τ_{11}	9	.677	.635	.512	.441
	10	.639	.597	.477	.409
	11	.713	.679	.576	.517
τ_{21}	12	.675	.642	.546	.490
	13	.649	.615	.521	.467
	14	.674	.641	.546	.492
	15	.647	.616	.525	.472
	16	.624	.595	.507	.454
τ_{22}	17	.605	.577	.490	.438
	18	.589	.561	.475	.424
	19	.575	.547	.462	.412
	20	.562	.535	.450	.401

Table 2. Values for Use in the Grubbs Test for Outliers

Number of Data Points	Risk of False Rejection 0.1%	0.5%	1%	5%	10%
	Critical Values				
3	1.155	1.155	1.155	1.153	1.148
4	1.496	1.496	1.492	1.463	1.425
5	1.780	1.764	1.749	1.672	1.602
6	2.011	1.973	1.944	1.822	1.729
7	2.201	2.139	2.097	1.938	1.828
8	2.358	2.274	2.221	2.032	1.909
9	2.492	2.387	2.323	2.110	1.977
10	2.606	2.482	2.410	2.176	2.036
15	2.997	2.806	2.705	2.409	2.247
20	3.230	3.001	2.884	2.557	2.385
25	3.389	3.135	3.009	2.663	2.486
50	3.789	3.483	3.336	2.956	2.768
100	4.084	3.754	3.600	3.207	3.017

Table 3. The t-distribution

Value of t for a confidence interval of Critical value of $\lvert t \rvert$ for P values of number of degrees of freedom	90% 0.10	95% 0.05	98% 0.02	99% 0.01
1	6.31	12.71	31.82	63.66
2	2.92	4.30	6.96	9.92
3	2.35	3.18	4.54	5.84
4	2.13	2.78	3.75	4.60
5	2.02	2.57	3.36	4.03
6	1.94	2.45	3.14	3.71
7	1.89	2.36	3.00	3.50
8	1.86	2.31	2.90	3.36
9	1.83	2.26	2.82	3.25
10	1.81	2.23	2.76	3.17
12	1.78	2.18	2.68	3.05
14	1.76	2.14	2.62	2.98
16	1.75	2.12	2.58	2.92
18	1.73	2.10	2.55	2.88
20	1.72	2.09	2.53	2.85
30	1.70	2.04	2.46	2.75
50	1.68	2.01	2.40	2.68
∞	1.64	1.96	2.33	2.58

Table 4. Critical values for the F-test, P=0.05, ν_1 is the number of freedom of the numerator, ν_2 or is the number of degrees of freedom of the denominator

ν_2	ν_1												
	1	2	3	4	5	6	7	8	9	10	12	15	20
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9	243.9	245.9	248.0
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45
3	10.13	9.552	9.277	9.117	9.013	8.941	8.887	8.845	8.812	8.786	8.745	8.703	8.660
4	7.709	6.944	6.591	6.388	6.256	6.163	6.094	6.041	5.999	5.964	5.912	5.858	5.803
5	6.608	5.786	5.409	5.192	5.050	4.950	4.876	4.818	4.772	4.735	4.678	4.619	4.558
6	5.987	5.143	4.757	4.534	4.387	4.284	4.207	4.147	4.099	4.060	4.000	3.938	3.874
7	5.591	4.737	4.347	4.120	3.972	3.866	3.787	3.726	3.677	3.637	3.575	3.511	3.445
8	5.318	4.459	4.066	3.838	3.687	3.581	3.500	3.438	3.388	3.347	3.284	3.218	3.150
9	5.117	4.256	3.863	3.633	3.482	3.374	3.293	3.230	3.179	3.137	3.073	3.006	2.936
10	4.965	4.103	3.708	3.478	3.326	3.217	3.135	3.072	3.020	2.978	2.913	2.845	2.774
11	4.844	3.982	3.587	3.357	3.204	3.095	3.012	2.948	2.896	2.854	2.788	2.719	2.646
12	4.747	3.885	3.490	3.259	3.106	2.996	2.913	2.849	2.796	2.753	2.687	2.617	2.544
13	4.667	3.806	3.411	3.179	3.025	2.915	2.832	2.767	2.714	2.671	2.604	2.533	2.459
14	4.600	3.739	3.344	3.112	2.958	2.848	2.764	2.699	2.646	2.602	2.534	2.463	2.388
15	4.543	3.682	3.287	3.056	2.901	2.790	2.707	2.641	2.588	2.544	2.475	2.403	2.328
16	4.494	3.634	3.239	3.007	2.852	2.741	2.657	2.591	2.538	2.494	2.425	2.352	2.276
17	4.451	3.592	3.197	2.965	2.810	2.699	2.614	2.548	2.494	2.450	2.381	2.308	2.230
18	4.414	3.555	3.160	2.928	2.773	2.661	2.577	2.510	2.456	2.412	2.342	2.269	2.191
19	4.381	3.522	3.127	2.895	2.740	2.628	2.544	2.477	2.423	2.378	2.308	2.234	2.155
20	4.351	3.493	3.098	2.866	2.711	2.599	2.514	2.447	2.393	2.348	2.278	2.203	2.124