UNIWERSYTET GDAŃSKI

Course title Statistics and chemometrics in chemical analytics / Statystyka I chemometria w analityce chemicznej	**ECTS code** 13.3.0865

Name of unit administrating study
Faculty of Chemistry

Studies

Field of study	Type	Form	
Chemistry	Bachelor	Full-time studies	

Teaching staff

Forms of classes, the realization and number of hours

A. Forms of classes, in accordance with the UG Rector's regulations
Lecture, auditorium classes, laboratory classes

B. The realization of activities
In-class learning

C. Number of hours
60 h (30 h lecture, 15 h auditorium classes, 15 h laboratory classes

ECTS credits 6

classes - 60 h
tutorial classes – 20 h
student's own work – 70 h

Total: 150 h - 6 ECTS

The academic cycle
First year, summer semester

Type of course
Obligatory

Language of instruction
Polish

Teaching methods

- Lecture with multimedia presentation
- In the computational laboratory students will conduct hands on exercises, based on the instructions prepared by the teacher
- During the auditorium classes students will solve tasks with the application of statistical test

Form and method of assessment and basic criteria for evaluation or examination requirements

A. Final evaluation, in accordance with the UG study regulations
Lectures – Course credit with a grade
Lab classes – Course credit with a grade
Auditorium classes – Course credit with a grade

B. Assessment methods
Lectures – final test
Laboratories – colloquia and written reports
Auditorium classes – colloquia

C. The basic criteria for evaluation or exam requirements

- All colloquia as well as written reports organized during lab and/or auditorium classes need to have a positive grade

Final test:

- Written part (obligatory): single choice test with 15 questions (1 point per question) plus three open questions (5 points per each) – max. 30 points in total. Positive grade if the number of points > 50%. For students having between 40% and 50% from the written part, oral part is obligatory. Students with the number of points < 40% do not pass the exam.
- Oral part (obligatory for students having between 40% and 50% from the written part and facultative for students with > 50%): discussion on three problems related to the topic, selected by the teacher
- Students are allowed passing the test twice (two attempts).
- All colloquia organized during lab and auditorium classes need to have a positive grade as well as written reports

Its obligatory to have a positive final grade from the lab exercises before passing the final test

Required courses and introductory requirements
Chemistry

Mathematics

Aims of education

- Achieving basic skills in statistic calculations by the students
- Presenting the applications of chemometrics in green chemistry to the students
- Achieving basic skills in chemometric methods by the students (performing basic analyses and interpreting the results)
- Familiarizing the students with the available chemometric software

Course contents

1. The course discusses the basic statistical theory that is frequently used in chemometric analysis: general population and statistical sample, characteristic of a single series of results, testing statistical hypothesis
2. Introduction to chemometrics: multivariate data, difference between statistics and chemometrics, review of the basic software
3. Methods of the initial data controlling, e.g.: problem of lacking data, outlying objects, variables transformation, distribution normalization, correlation, and covariation analysis
4. Methods of analyzing the internal structure of the data: similarity in the multivariable feature space, methods of similarity analysis, dimensionality reduction, hierarchical cluster analysis (HCA), principal component analysis (PCA). Examples of applying these methods in green chemistry and technology.
5. Modeling phenomena and processes with regression and classification techniques: linear regression (LR), multiple linear regression (MLR), principal component regression (PCR), partial least square regression (PLS), linear discriminant analysis (LSA), k-nearest neighbors classifier (k-NN); artificial neural networks (ANN) for solving classification and regression problems; methods of optimal variables selection (stepwise regression, genetic algorithms); validation of classification and regression models. Examples of applying these methods in green chemistry and technology.

Bibliography of literature

A. Literature required to pass the course

• A. Łomnicki: Wprowadzenie do statystyki dla przyrodników. Wydawnictwo Naukowe PWN, Warszawa 2003.
• J. Mazerski: Podstawy chemometrii. Gdańsk: Wydawnictwo Politechniki Gdańskiej, 2000
• P. Konieczka, J. Namieśnik i in.: Ocena i kontrola jakości wyników analitycznych. Centrum Doskonałości Analityki i Monitoringu Środowiskowego,
Gdańsk 2004.

A.1. Literatura wykorzystywana podczas zajęć

• Skrypt do ćwiczeń laboratoryjnych przygotowywany przez pracowników Pracowni Chemometrii Środowiska

B. Extracurricular readings

• J. B. Czermiński, A. Iwasiewicz i in.: „Metody statystyczne w doświadczalnictwie chemicznym", Wydawnictwo Naukowe PWN, Warszawa 1992 lub
wersja starsza tej książki zatytułowana „Metody statystyczne dla chemików".
• Praca zbiorowa pod redakcją H. Kassyk-Rokickiej: „Statystyka. Zbiór zadań". Polskie Wydawnictwo Ekonomiczne, Warszawa 1997.
• S. D. Brown, R. Tauler, B. Walczak (red): Comprehensive chemometrics: Chemical and biochemical data analysis. Amsterdam: Elsevier, 2009
• R. Kramer: Chemometric techniques for quantitative analysis. New York: Marcel Dekker, Inc, 2005
• D. Zuba, A Parczewski (red.): Chemometria w analityce: wybrane zagadnienia. Kraków: Wydawnictwo Instytutu Ekspertyz Sądowych, 2008
• JM. Dobosz: Wspomagana komputerowo statystyczna analiza danych. Warszawa: Akademicka Oficyna Wydawnicza EXIT, Warszawa 2004

Knowledge

At the end of the course every student:

- knows basic statistical parameters used to characterize a single series of result
- knows basis statistical distributions (e.g., t-Student)
- knows basis statistical tests, their applications and limitations
- knows basic classification of chemometric methods and provides examples of theirs applications in green chemistry
- knows basic software packages to be used for chemometric analyses
- explains theoretical background (algorithm) of the most important chemometric methods, including: HCA, PCA, LR/MLR, LDA, kNN

Skills

At the end of the course every student:

- calculates basis statistical parameters used to characterize a single series of results
- prepares the histogram for assessing the distribution of a single series of results

- uses statistical tests to solve various problems
- uses KNIME environment for chemometric analyses
- correctly prepares data for further analysis
- performs HCA and PCA analyses and correctly interprets the results
- develops regression and classification models, validates the models correctly and applies the models for predictions
- applies methods of chemometric optimization

Social competence

At the end of the course every student:

- is convinced that the use of a computer and chemometrics strengthens the potential of data analysis
- can critically evaluate experimental results and understand the necessity of their control
- understands the need of deeper learning of chemometric methods