Application of basic statistic and chemometric approaches in chemical studies

Prof. Dr. Stefan Tsakovski.

Faculty of Chemistry and Pharmacy, University of Sofia "St. Kl. Okhridski", Sofia, Bulgaria

Horarium: 30 academic hours (15 hours lectures + 15 hours seminars)

Abstract:

The goal of the course is to deepen participants' knowledge in statistical concepts used in chemical studies. In the first part of the course, connection between basic (univariate) statistics and chemometrics (multivariate statistics) will be explained. In the second part the most widely used chemometric methods like Cluster analysis, Principal component analysis, Partial least squares and some more advanced methods will be introduced. The application of variety of case studies and proper software packages during seminars will provide the students basic practical skills in multivariate data analysis.

At the end of the course, each participant will be able to:

- arrange data in appropriate data matrix for multivariate analysis;
- apply Cluster analysis, Principal component analysis, Partial least squares on new data and analyse the results;
- compare and contrast the methods for a given data analysis situation considering the benefits and the pitfalls of the methods;
- interpret chemometrics studies described in scientific literature and describe his/her own results in a scientific way.

Seminar Topics:

1. Statistics of repeated measurements and data quality. Presentation of results.
2. Parametric and non-parametric (robust) statistical tests.
3. Analysis of variation as a method for evaluation of experimental factors.
4. Software package STATISTICA 8.0 – basic functions and basic statistics
5. Chemometrics – multivariate data analysis. Multivariate data arrangement, Pretreatment of the data. Transformation of variables. Handling of missing data.

6. Clusters analysis. Similarity measures. Agglomerative hierarchical algorithms. Dendrograms. Clustering and classification.
7. Principal component analysis (PCA). Loading plots. Score plots. Choice of principal components (rank analysis), both numerically and graphically (scree plot). Data exploration.
8. Software package STATISTICA 8.0 – exploratory data analysis.
9. Data modeling aspects of PCA – Absolute principal component scores (APCS).
10. Partial least squares (PLS). Loadings, weights and score plots. Importance of "original" variables. Evaluation parameters for a regression model.
11. Classification applications based on PLS: PLS- discriminant analysis (PLS-DA)
12. PLS toolbox for MATLAB: basic functions, PLS and PLS-DA modules.
13. Advanced chemometric methods: Self organizing maps - principles and available software packages.
14. Advanced chemometric methods: Hasse diagram techniques – partial ordering theory basics and available software packages.
15. Implementation of "hybrid" chemometric approaches – combination of dimensionality reduction methods and partial ordering techniques.

Prof. Tsakovski will be available for consultation and discussion about the material presented during the lectures and key topics covered in the seminars. Consultation concerning proper selection of chemometric treatment of participants' data sets will be available.

Termin	Dzień tygodnia	Godzina	Miejsce
07.01.2019	**Poniedziałek**	**14.15 – 17.00**	**Minicentrum Konferencyjne (Luwr)**
08.01.2019	**Wtorek**	**12.15 – 15.00**	**Minicentrum Konferencyjne (Luwr)**
09.01.2019	**Środa**	**12.15 – 15.00**	**Minicentrum Konferencyjne (Luwr)**
10.01.2019	**Czwartek**	**12.15 – 15.00**	**Minicentrum Konferencyjne (Luwr)**
11.01.2019	**Piątek**	**12.15 – 15.00**	**Minicentrum Konferencyjne (Luwr)**