DI.109.2 HEP Data analysis in high energy physics: a practical guide to statistical methods II

1. Study program

1.1. University	University of Bucharest
1.2. Faculty	Faculty of Physics
1.3. Department	Department of Theoretical Physics, Mathematics, Optics, Plasma and Lasers
1.4. Field of study	Physics
1.5. Course of study	Master of Science
1.6. Study program	High Energy Physics (in English)
1.7. Study mode	Full-time study

2. Course unit

2.1. Course unit title	Data analysis in high energy physics: a practical guide to statistical methods I							
2.2. Teacher				Dr. Julien Maurer				
2.3. Tutorials/Practicals instructor(s)				Dr. Julien Maurer				
2.4. Year of study	I	2.5. Semester	2	2.6. Type of evaluation	E	2.7. Type of course unit	Content[1)]	**DS**
							Type[2)]	**DI**

[1)] fundamental (DF), specialized (DS); complementary (DC)
[2)] compulsory (DI), elective (DO), noncompulsory disciplines (DFC)

3. Total estimated time (hours/semester)

3.1. Hours per week in curriculum	**4**	distribution: Lecture	**2**	Practicals/Tutorials	**2**
3.2. Total hours per semester	**56**	Lecture	**28**	Practicals/Tutorials	**28**

Distribution of estimated time for study	**hours**
3.2.1. Learning by using one's own course notes, manuals, lecture notes, bibliography	**20**
3.2.2. Research in library, study of electronic resources, field research	**20**
3.2.3. Preparation for practicals/tutorials/projects/reports/homeworks	**52**
3.2.4. Preparation for exam	**4**
3.2.5. Other activities	**0**

3.3. Total hours of individual study	**96**
3.4. Total hours per semester	**150**
3.5. ECTS	**6**

4. Prerequisites (if necessary)

4.1. curriculum	Algebra, Programming languages
4.2. competences	Knowledge about: algebra, quantum mechanics, electrodynamics

5. Conditions/Infrastructure (if necessary)

5.1. for lecture	Video projector

5.2. for practicals/tutorials	

6. Specific competences acquired

Professional competences	• Identify and proper use of the main physical laws and principles in a given context: the use of the concepts of the standard model • Solving problems of physics under given conditions • Use of the physical principles and laws for solving theoretical or practical problems with qualified tutoring • Rigorous knowledge of quantum field theory, concepts, notions and problems in the area of theoretical particle physics and their interactions • Ability to use this knowledge in interpretation of experimental result and understand experiments at CERN; acquire the appropriate understanding of studied fundamental mechanisms
Transversal competences	• Efficient use of sources of information and communication resources and training assistance in a foreign language • Efficient and responsible implementation of professional tasks, with observance of the laws, ethics and deontology.

7. Course objectives

7.1 General objectives	To acquire practical experience analyzing and interpreting particle physics data with modern methods, including machine learning algorithms.
7.2 Specific objectives	Supplied datasets of collision events, either real or simulated, are analysed with the ROOT software and columnar analysis tools. Selected subsets of events are subjected to statistical interpretation with HistFactory-based software. More involved algorithms are presented with applications in particle physics event reconstruction and classification, including machine learning-based approaches. Theoretical aspects behind modelling uncertainties are investigated.

8. Contents

8.1. Lecture [chapters]	Teaching techniques	Observations/ hours
Overview of software tools for data analysis and hypothesis testing in high-energy physics; columnar analysis	Systematic exposition - lecture. Examples.	4 hours
Approximating empirical distributions: families of functions for fits; splines; interpolation; kernel density estimation; smoothing		2 hours
Classical regression techniques for event reconstruction: track and vertex fits; Hough transforms; Gaussian sum filtering; hadronic jet clustering and substructure;		6 hours
Classification and regression with machine learning: decision trees; neural networks; genetic algorithms		6 hours
Random sampling; Markov Chain Monte Carlo		4 hours
Theory uncertainties: factorization, scales; parton density functions; underlying event		4 hours
Unfolding		2 hours
Bibliography: 1. O. Behnke et al, *Data analysis in high energy physics: A practical guide to statistical methods*, Wiley-VCH, 2013.		

2. K. Hanagaki et al, *Experimental Techniques in Modern High-Energy Physics*, Springer Tokyo, 2021.
3. K. Cranmer et al, *Machine Learning*. In *The Review of Particle Physics*, edited by S. Navas et al., Phys. Rev. D 110, 030001 (2024).
4. G. Cowan, *Monte Carlo Techniques*. In *The Review of Particle Physics*, edited by S. Navas et al., Phys. Rev. D 110, 030001 (2024).

8.2. Tutorials/Practicals [main themes]	Teaching and learning techniques	Observations/hours
Hands-on introduction of software tools	Problem solving, computer-based	4 hours
Analysis of LHC open data; resonance searches		8 hours
Analysis of simulated data in more complex scenarios; statistical interpretation using Azimov datasets		8 hours
Training of multivariate classifiers with simulated data		8 hours

Bibliography:
1. R. Brun and F. Rademakers, *ROOT - An Object Oriented Data Analysis Framework*, Nucl. Inst. & Meth. in Phys. Res. A 389 (1997) 81-86.
2. K. Cranmer et al, *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*, CERN-OPEN-2012-016, 2012.
3. L. Heinrich et al, *pyhf: pure-Python implementation of HistFactory statistical models*, Journal of Open Source Software 6, 58 (2021) 2823.
4. CERN Open Data portal, https://opendata.cern.ch/
5. Inter-Experimental LHC Machine Learning Working Group, https://iml.web.cern.ch/homepage and https://github.com/iml-wg/HEP-ML-Resources#software

9. Compatibility of the course unit contents with the expectations of the representatives of epistemic communities, professional associations and employers (in the field of the study program)

This course unit develops some theoretical competences, which are fundamental for a Master student in the field of modern physics, corresponding to national and international standards. The contents is in line with the requirement of the main employers of research institutes and universities.

10. Assessment

Activity type	10.1. Assessment criteria	10.2. Assessment methods	10.3. Weight in final mark
10.4. Lecture	- coherence and clarity of exposition - correct use of equations/mathematical methods/physical models and theories - ability to indicate/analyse specific examples	Oral examination	40%
10.5.1. Tutorials/Practicals	- ability to use specific problem solving methods - ability to analyse the results	Homeworks/Lab reports	60%

10.6. Minimal requirements for passing the exam
Attendance of at least 50% for the lectures and at least 70% for the tutorials. **Correct solutions to the indicated subjects for obtaining the grade 5 (10 points scale) from all activities, part of the continuous evaluation.** **Correct solutions to the indicated subjects for obtaining the grade 5 (10 points scale) within the final exam.**

Date
4.10.2024

Teacher's name and signature

Dr. Julien Maurer, [redacted]

Practicals/Tutorials instructor(s) name(s) and signature(s)

Dr. Julien Maurer, [redacted]

Date of approval

Head of Department
Lect.dr. Roxana Zus
[redacted]