



Nazwa modułu:	Metody analizy danych doświadczalnych				
Rok akademicki:	2017/2018	Kod:	JFT-1-025-s	Punkty ECTS:	5
Wydział:	Fizyki i Informatyki Stosowanej				
Kierunek:	Fizyka Techniczna	Specjalność:	—		
Poziom studiów:	Studia I stopnia	Forma i tryb studiów:	Stacjonarne		
Język wykładowy:	Polski	Profil kształcenia:	Ogólnoakademicki (A)	Semestr:	0
Strona www:	<a href="http://home.agh.edu.pl/mariuszp/wfiis_madd/index.html">http://home.agh.edu.pl/mariuszp/wfiis_madd/index.html</a>				
Osoba odpowiedzialna:	prof. dr hab. inż. Przybycień Mariusz (mariusz.przybycien@agh.edu.pl)				
Osoby prowadzące:	dr inż. Adamczyk Leszek (Leszek.Adamczyk@agh.edu.pl) prof. dr hab. inż. Przybycień Mariusz (mariusz.przybycien@agh.edu.pl)				

### Krótką charakterystyka modułu

Przedmiot ten ma na celu zapoznanie studenta z metodami statystycznymi stosowanymi w analizie danych w fizyce cząstek elementarnych i fizyce jądrowej oraz innych działach fizyki eksperymentalnej.

### Opis efektów kształcenia dla modułu zajęć

Kod EKM	Student, który zaliczył moduł zajęć wie/umie/potrafi	Powiązania z EKK	Sposób weryfikacji efektów kształcenia (forma zaliczeń)
<b>Wiedza</b>			
M_W008	Student zna i rozumie wybrane zaawansowane metody analizy danych doświadczalnych.		Aktywność na zajęciach, Egzamin, Kolokwium, Wykonanie projektu
M_W013	Student zna podstawowe narzędzia programistyczne wykorzystywane w analizie danych.		Aktywność na zajęciach, Egzamin, Kolokwium, Wykonanie projektu
<b>Umiejętności</b>			
M_U011	Student potrafi przeprowadzić analizę statystyczną danych doświadczalnych z wykorzystaniem zaawansowanych metod i narzędzi.		Aktywność na zajęciach, Egzamin, Kolokwium, Wykonanie projektu
M_U016	Student potrafi posługiwać się wybranymi narzędziami programistycznymi wykorzystywanymi w zaawansowanej analizie danych.		Aktywność na zajęciach, Egzamin, Kolokwium, Wykonanie projektu

**Matryca efektów kształcenia w odniesieniu do form zajęć**

Kod EKM	Student, który zaliczył moduł zajęć wie/umie/potrafi	Forma zajęć										
		Wykład	Ćwiczenia audytoryjne	Ćwiczenia laboratoryjne	Ćwiczenia projektowe	Konwersatorium	Zajęcia seminaryjne	Zajęcia praktyczne	Zajęcia terenowe	Zajęcia warsztatowe	Inne	E-learning
Wiedza												
M_W008	Student zna i rozumie wybrane zaawansowane metody analizy danych doświadczalnych.	+	+	-	+	-	-	-	-	-	-	-
M_W013	Student zna podstawowe narzędzia programistyczne wykorzystywane w analizie danych.	+	+	-	+	-	-	-	-	-	-	-
Umiejętności												
M_U011	Student potrafi przeprowadzić analizę statystyczną danych doświadczalnych z wykorzystaniem zaawansowanych metod i narzędzi.	+	+	-	+	-	-	-	-	-	-	-
M_U016	Student potrafi posługiwać się wybranymi narzędziami programistycznymi wykorzystywanymi w zaawansowanej analizie danych.	+	+	-	+	-	-	-	-	-	-	-

**Treść modułu zajęć (program wykładów i pozostałych zajęć)****Wykład**Funkcje generujące i charakterystyczne oraz ich zastosowania (2h)

- Funkcja generująca rozkład prawdopodobieństwa.
- Funkcja generująca moment.
- Funkcja charakterystyczna.
- Losowe sumy zmiennych losowych.
- Procesy rozgałęzienia.

Statystyka porządku (2h)

- Statystyka porządku i jej zastosowania.

Proces Poissona i jego zastosowania (2h)

- Definicja procesu Poissona.
- Rozkłady warunkowe w procesach Poissona.
- Niezależne procesy Poissona.
- Złożony proces Poissona.

Bayesowska analiza statystyczna (4h)

- Prawdopodobieństwa logiczne, wnioskowanie bayesowskie.
- Estymacja parametrów, przedziały wiarygodności, testy hipotez statystycznych.

- Metoda największej wiarygodności a wnioskowanie bayesowskie.

#### Generacja rozkładów wielowymiarowych (2h)

- Generacja rozkładów kątowych we współrzędnych biegunowych i sferycznych.
- Generacja rozkładów z korelacjami.
- Generacja rozpadów dwu- i trzyciałowych.

#### Estymacja parametrów (2h)

- Metoda największej wiarygodności: rozszerzona oraz z wagami.
- Obszary ufności.
- Minimalizacja numeryczna: Minuit oraz RooFit.

#### Rozszerzona metoda najmniejszych kwadratów (MNK) (2h)

- Skalowana macierz kowariancji.
- MNK z korelacjami i więzami.

#### Unfolding (2h)

- Metoda bin-by-bin (akceptancja, wydajność, czystość).
- Macierz przejścia, regularyzacja (analiza danych pogrupowanych).
- Rozkład według wartości osobliwych, pakiet RooUnfold.
- Funkcja przejścia (analiza danych niegrupowanych).

#### Testowanie hipotez statystycznych (2h)

- Zaawansowane metody testowania hipotez.

#### Wielowymiarowa klasyfikacja przypadków (2h)

- Dyskryminacja Fishera (liniowa i nieliniowa).
- Sieci neuronowe, SVM, pakiet TMVA.

#### Oszacowanie niepewności systematycznych (2h)

- Źródła niepewności systematycznych.
- Szacowanie błędów systematycznych.

#### Bootstrap (2h)

- Zastosowania metody bootstrapu.

#### Funkcje korelacji (2h)

- Momenty faktorialne i kumulanty.
- Funkcje korelacji dla nieidentycznych cząstek.
- Dwu- i trój-cząstkowe funkcje korelacji.

#### Dyskusja przebiegu typowej analizy statystycznej od selekcji przypadków do prezentacji wyników (2h)

- Przedstawione zostaną kolejne kroki typowej analizy danych z dużego eksperymentu akceleratorowego na przykładzie eksperymentu ATLAS na akceleratorze LHC i STAR na akceleratorze RHIC.

### **Ćwiczenia audytoryjne**

#### Ćwiczenia rachunkowe tematyce zgodnej z programem wykładów.

Efekty kształcenia:

Student potrafi przeprowadzić analizę danych z eksperymentów fizyki wysokich energii wykorzystując zaawansowane narzędzia statystyczne (pakiet Root wraz z rozszerzeniami takimi jak RooFit, RooUnfold, TMVA).

Zajęcia prowadzone będą z wykorzystaniem rzeczywistych danych z eksperymentów takich jak ZEUS, ATLAS, STAR.

### **Ćwiczenia projektowe**

#### Ćwiczenia projektowe o tematyce zgodnej z programem wykładów.

### **Efekty kształcenia:**

Student potrafi przeprowadzić analizę danych z eksperymentów fizyki wysokich energii wykorzystując zaawansowane narzędzia statystyczne (pakiet Root wraz z rozszerzeniami takimi jak RooFit, RooUnfold, TMVA).

Zajęcia prowadzone będą z wykorzystaniem rzeczywistych danych z eksperymentów takich jak ZEUS, ATLAS, STAR.

### **Sposób obliczania oceny końcowej**

Oceny z ćwiczeń rachunkowych/projektowych obliczane są następująco: procent uzyskanych punktów przeliczany jest na ocenę zgodnie z Regulaminem Studiów AGH.

Ocena końcowa (OK) jest średnią arytmetyczną ocen z ćwiczeń rachunkowych i projektowych.

Student ma prawo do nieusprawiedliwionych nieobecności na 20% zajęć z ćwiczeń rachunkowych i projektu. Większa liczba nieobecności skutkuje brakiem zaliczenia bez możliwości pisania kolokwium poprawkowych.

### **Wymagania wstępne i dodatkowe**

Podstawowe wiadomości z rachunku prawdopodobieństwa i statystyki matematycznej.

Znajomość podstaw języka C++

### **Zalecana literatura i pomoce naukowe**

R. Nowak, Statystyka dla fizyków, PWN, 2002.

R. Nowak, Statystyka dla fizyków. Ćwiczenia, PWN, 2002.

P.D. Hoff, A First Course in Bayesian Statistical Methods, Springer, 2009.

P. Gregory, Bayesian Logical Data Analysis for the Physical Sciences, Cambridge, 2005.

W.L. Dunn, J.K. Shultis, Exploring Monte Carlo Methods, Elsevier, 2011.

O. Behnke et al., A Practical Guide to Data Analysis in high energy Physics, Wiley, 2012.

G. Bohm, G. Zech, Introduction to Statistics and Data Analysis, DESY, 2010.

G. Cowan, Statistical Data analysis, Oxford, 1998.

Strona domowa program Root: [root.cern.ch](http://root.cern.ch)

### **Publikacje naukowe osób prowadzących zajęcia związane z tematyką modułu**

[http://inspirehep.net/search?ln=pl&ln=pl&p=exactauthor%3A%22Przybycien%2C+Mariusz%22+or+exactauthor%3A%22Przybycien%2C+M.%22&action\\_search=Szukaj&sf=&so=d&rm=&rg=25&sc=0&of=hb](http://inspirehep.net/search?ln=pl&ln=pl&p=exactauthor%3A%22Przybycien%2C+Mariusz%22+or+exactauthor%3A%22Przybycien%2C+M.%22&action_search=Szukaj&sf=&so=d&rm=&rg=25&sc=0&of=hb)

### **Informacje dodatkowe**

Brak

**Nakład pracy studenta (bilans punktów ECTS)**

Forma aktywności studenta	Obciążenie studenta
Udział w wykładach	30 godz
Samodzielne studiowanie tematyki zajęć	30 godz
Udział w ćwiczeniach audytoryjnych	15 godz
Przygotowanie do zajęć	30 godz
Udział w ćwiczeniach projektowych	15 godz
Wykonanie projektu	30 godz
Sumaryczne obciążenie pracą studenta	150 godz
Punkty ECTS za moduł	5 ECTS