

Krakowska Interdyscyplinarna Szkoła Doktorska

Opis przedmiotu/ course description

Przedmiot/ Course :	All you've ever wanted to know about statistical analysis but never dared to ask?
Moduł kształcenia/ Training module:	Moduł ogólny
Okres realizacji/ Implementation period :	I rok, semestr letni
Język wykładowy/ Language:	język angielski/ English
Prowadzący/ Lecturer:	Dr hab. Paweł Brückman de Renstrom
Wymiar godzin przedmiotu/duration :	20 godzin
Forma prowadzenia zajęć/ Form of teaching :	Wykład (liczba godzin wykładu: 20)
Opis przedmiotu/ course content:	<p>Statistics in Data Analysis przeprowadzi studentów przez podstawowe zagadnienia statystycznej analizy danych eksperymentalnych. Wykład nie wymaga wcześniejszej znajomości rachunku prawdopodobieństwa i statystyki, Jedynym założeniem jest podstawowa znajomość analizy matematycznej oraz (w mniejszym stopniu) algebry liniowej. Wszystkie niezbędne pojęcia są definiowane na bieżąco. Najczęściej spotykane funkcje gęstości prawdopodobieństwa (pdf) są omówione, a następnie Centralne twierdzenie graniczne doprowadza do rozkładu Normalnego (Gaussowskiego) stanowiącego podstawę praktycznej analizy statystycznej.</p> <p>Metody Monte Carlo oraz liniowe i nieliniowe testowe zmienne losowe wraz z ich zastosowaniem do współczesnych technik analizy wielowymiarowej są po krótkce omówione. Specjalny nacisk położony jest na testy statystyczne i szacowanie parametrów modelu wraz z oszacowaniem błędu statystycznego. Metoda największej wiarygodności (ML) oraz metoda najmniejszych kwadratów (LS) są dokładnie omówione wraz z praktycznymi przykładami dopasowania do danych metodami ML i LS zarówno w wersji binowanej jak i niebinowanej. Osobny rozdział poświęcony jest znaczącości statystycznej obserwowanego sygnału lub stopnia wiarygodności jego wykluczenia, pojęciom kluczowym dla poszukiwań rzadkich lub nowych zjawisk. Wreszcie, zalety i ograniczenia metody unfolding są omówione w kontekście odzyskiwania pierwotnych rozkładów (lub parametrów modelu) z rozmytych i obciążonych danych eksperymentalnych.</p>

	<p>Wykład jest ilustrowany przykładami z Python Notebook z wykorzystaniem platformy Google Colaboratory.</p> <p><i>Statistics in Data Analysis guides students through the basic concepts of statistical analysis of experimental data. The lecture does not assume any prior knowledge of probability and statistics. The only assumed prerequisite is entry level calculus and (to lesser extent) linear algebra. All needed notions and fundamental concepts are introduced and defined as we go along. Most commonly encountered probability density functions (pdf) are discussed followed by the Central Limit Theorem and properties of the Normal (Gaussian) pdf - the cornerstone of practical statistical analysis.</i></p> <p><i>Monte Carlo methods are briefly introduced as well as linear and nonlinear test statistics leading to modern multivariate analysis techniques. Special emphasis is put on statistical tests and parameter estimation together with assessment of related uncertainties. The maximum likelihood (ML) and the least squares (LS) methods are discussed in detail followed by practical examples of ML and LS fits to experimental data in both binned and unbinned variants. A separate chapter is devoted to the statistical significance of the observed signal or the confidence level on its exclusion which is paramount for data analysis searching for rare or new phenomena. Finally, unfolding with its benefits and limitations is discussed in the context of extracting primary distributions (or model parameters) from smeared and biased experimental data.</i></p> <p><i>The lecture is illustrated by Python Notebook examples hosted on the Google Colaboratory platform.</i></p>
<p>Efekty uczenia się wg 8PRK zgodnie z Programem kształcenia KISD/ learning outcomes at level 8 of the PRK according to the KISD Training Program:</p>	<p>EU1, EU2, EU8, EU15</p>
<p>Forma weryfikacji efektów uczenia się/ Method of verification of learning outcomes:</p>	<p>Zaliczenie/ocena oparta będzie na uczestnictwie oraz łącznej punktacji zadań domowych. <i>The sign-off/grade will be based on attendance and summed notes from homework assignments.</i></p>
<p>Wymagania wobec uczestników/ Requirements for participants:</p>	<p>Podstawowa znajomość analizy matematycznej oraz (w mniejszym stopniu) algebry liniowej.</p>