



Centrum Studiów Zaawansowanych Politechniki Warszawskiej

Pl. Politechniki 1, 00-661 Warszawa, tel./fax +48 22 234 6003 (6002), www.csz.pw.edu.pl



Uczelniana Oferta Studiów Zaawansowanych	
SYLABUS 2020/2021	
Nazwa przedmiotu	Konstrukcja Modeli Statystycznych z Pakietem R (KMS) Building Statistical Models with R
Liczba punktów ECTS	Proponowana liczba punktów: 2 ECTS

Osoby prowadzące	Tytuł naukowy	Imię i nazwisko	Katedra / Instytut/ Centrum/ Inne
	Dr hab. inż., prof. PW	Anna Dembińska	Zakład Procesów Stochastycznych i Matematyki Finansowej, Wydział MiNI PW
Osoba odpowiedzialna za przedmiot	Dr hab. inż., prof. PW	Anna Dembińska	Zakład Procesów Stochastycznych i Matematyki Finansowej, Wydział MiNI PW

Semestr studiów	Semestr letni 2020/2021
Typ przedmiotu (możliwości wyboru) obowiązkowy O fakultatywny F	F
Wymagania wstępne	Podstawy z rachunku prawdopodobieństwa: pojęcie prawdopodobieństwa i jego własności, niezależność zdarzeń losowych, podstawowe rozkłady prawdopodobieństwa ciągłe i dyskretne, podstawowe wiadomości dotyczące zmiennych losowych jedno- i wielowymiarowych. Podstawy ze statystyki: znajomość podstawowych statystyk próbkowych i podstawowych metod graficznej prezentacji danych, znajomość pojęcia testu statystycznego.
Poziom przedmiotu Podstawowy P Średniozaawansowany Ś Zaawansowany Z	Ś
Charakter zajęć, liczba godzin w	1) L

semestrze, liczba godzin w tygodniu. 1) podać rodzaj prowadzonych zajęć dla danego przedmiotu: wykłady (W); ćwiczenia (Ć); laboratorium (L); projekt (P) 2) podać liczbę godzin w tygodniu np. W - 2; Ć - 2; L - 3; P - 0 3) podać liczbę godzin w semestrze np. W - 30; Ć - 30; L - 45; P - 0	2) L-2 3) L-15
Sugerowana liczba godzin pracy własnej	35 godzin obejmuje : 15 godzin przygotowywanie się słuchacza do zajęć, 20 godzin – analiza i rozwiązywanie projektów
Całkowita liczba godzin:	50 godzin
Aspekty międzynarodowe (jeśli są)	
Język wykładowy	Polski
Cel przedmiotu Opis zakładanych kompetencji i umiejętności, jakie student nabywa w wyniku zaliczenia przedmiotu. Uwaga: maksymalna objętość tekstu to 3 linie standardowej strony A4	Sprawność w rozwiązywaniu praktycznych problemów z wykorzystaniem metod statystycznych i pakietu R. Znajomość szerokiej gamy modeli statystycznych oraz umiejętność doboru stosownej procedury statystycznej i jej implementacji dla rzeczywistych danych.
Treść przedmiotu	
<ol style="list-style-type: none"> Ogólne zapoznanie z pakietem R: metody wczytywania danych, rodzaje zmiennych, operacje na zmiennych, generowanie danych. Wstępna analiza danych. Wyznaczanie prawdopodobieństw i kwantyli dla wybranych rozkładów. Generowanie liczb pseudolosowych z zadanego rozkładu i badanie własności rozkładów na tej podstawie. Analiza zgodności obserwowanych danych z zadanym rozkładem: metody graficzne i testy zgodności. Wykresy kwantylowe. Test zgodności chi-kwadrat Pearsona z prostą i złożoną hipotezą zerową. Test Kołmogorowa-Smirnowa. Regresja liniowa jednokrotna: dopasowywanie modelu do danych, sprawdzanie czy dopasowany model dobrze opisuje dane, obserwacje wpływowe i odstające, przekształcanie zmiennych, metoda ważonych najmniejszych kwadratów, prognozowanie przy użyciu modelu regresji. Regresja liniowa wielokrotna: przekształcanie zmiennych, współliniowość zmiennych objaśniających, dobór zmiennych objaśniających do modelu (częściowy test F, kryteria służące do wyboru najlepszego modelu, metoda dołączania, eliminacji i selekcji krokowej). 	
Spis zalecanych lektur	
LP.	Autor, Tytuł, Wydawnictwo,
1.	P. Dalgaard, „Introductory Statistics with R”, Springer, 2008
2.	J. Koronacki, J. Mielniczuk, „Statystyka dla studentów kierunków technicznych i przyrodniczych”, Wydawnictwa Naukowo-Techniczne, Warszawa, 2006
3.	P. Biecek, „Przewodnik po pakiecie R”, Oficyna Wydawnicza GIS, Wrocław, 2008
4.	J.J. Faraway „Practical Regression and ANOVA Using R”, http://www.maths.bath.ac.uk/~jif23/book/

Metody oceny (ocena, egz. pisemny, egz. ustny, projekt)	Praca podczas zajęć – rozwiązywanie projektów.
--	--

Uwagi dodatkowe	Zajęcia odbędą się, jeżeli zapisze się co najmniej 20 osób. Wystawiane jedynie oceny za zaliczenie przedmiotu.
------------------------	---

Tabela 1. Efekty kształcenia

Numer (symbol)	Efekty kształcenia słuchacza, który zaliczył przedmiot, potrafi	Sposób weryfikacji osiągnięcia efektu
WIEDZA		
KMS_W1	Ma uporządkowaną wiedzę z zakresu wstępnej analizy danych.	Projekt
KMS_W2	Zna testy statystyczne i metody graficzne służące do sprawdzania zgodności rozkładu.	Projekt
KMS_W3	Zna i wie jak dopasować do danych model regresji liniowej.	Projekt
UMIEJĘTNOŚCI		
KMS_U1	Potrafi dobrać odpowiednią procedurę statystyczną do rozwiązania konkretnego, praktycznego problemu.	Projekt
KMS_U2	Mając rzeczywiste dane, potrafi zastosować do nich wybraną procedurę statystyczną a następnie zweryfikować adekwatność modelu skonstruowanego w oparciu o wybraną procedurę.	Projekt
KMS_U3	Przeprowadzenie formalnych testów podczas statystycznej analizy danych umie poprzedzić wstępną analizą wykorzystującą metody graficzne.	Projekt
KMS_U4	Umie sprawnie posługiwać się pakietem R podczas przeprowadzania analizy danych.	Projekt
KOMPETENCJE		
KMS_K1	Rozumie konieczność diagnozowania modelu, użytego do opisu rzeczywistych danych, i wprowadzania w nim stosownych modyfikacji.	Obserwacja na zajęciach
KMS_K2	Rozumie, że automatyczne użycie procedur statystycznych, bez wnikięcia w istotę i charakter danych, grozi wyciągnięciem błędnych wniosków.	Obserwacja na zajęciach