

Sylabus przedmiotu

Przedmiot:	Analiza regresji
Kierunek:	Matematyka (specjalności nienauczycielskie), II stopień [4 sem], stacjonarny, ogólnoakademicki, rozpoczęty w: 2013
Specjalność:	statystyczna analiza danych
Rok/Semestr:	II/3
Liczba godzin:	15,0
Nauczyciel:	Bator Artur, dr
Forma zajęć:	wykład
Rodzaj zaliczenia:	egzamin
Punkty ECTS:	5,0
Godzinowe ekwiwalenty punktów ECTS (łącznie liczba godzin w semestrze):	30,0 Godziny kontaktowe z prowadzącym zajęcia realizowane w formie konsultacji 45,0 Godziny kontaktowe z prowadzącym zajęcia realizowane w formie zajęć dydaktycznych 22,5 Przygotowanie się studenta do zajęć dydaktycznych 22,5 Przygotowanie się studenta do zaliczeń i/lub egzaminów 30,0 Studiowanie przez studenta literatury przedmiotu
Poziom trudności:	nie dotyczy
Metody dydaktyczne:	<ul style="list-style-type: none"> • objaśnienie lub wyjaśnienie • wykład informacyjny
Zakres tematów:	<p>1. Analiza korelacji. Współczynnik korelacji liniowej Pearsona i jego statystyczna istotność. Współczynnik korelacji rang Spearmana i jego statystyczna istotność. Współczynnik korelacji rang τ Kendalla i jego statystyczna istotność oraz współczynnik korelacji rang τ Kendalla z rangami związanymi.</p> <p>2. Analiza regresji liniowej prostej. Estymacja i weryfikacja parametrów strukturalnych modelu. Założenia modelu oraz ich weryfikacja. Miary dopasowania modelu do danych empirycznych: współczynnik determinacji, współczynnik zbieżności.</p> <p>3. Regresja krzywoliniowa. Modele linearyzowalne (typu wykładniczego, potęgowego, logarytmicznego, wielomianowego, hiperbolicznego itp.) oraz modele nieliniowe. Składnia estymowanej funkcji. Funkcja straty. Algorytmy minimalizujące funkcję straty: quasi-Newtona, sympleksu, Hooke'a-Jeevesa przemieszania układu, Rosenbrocka poszukiwania układu. Ocena dobroci dopasowania modelu do danych empirycznych. Diagnostyka modelu w oparciu o analizę reszt.</p> <p>4. Macierz korelacji. Współczynnik korelacji zupełnej (całkowitej), współczynnik korelacji cząstkowej, współczynnik korelacji semicząstkowej, współczynnik korelacji wielorakiej i jego statystyczna istotność</p> <p>5. Klasyczny model regresji wielorakiej (wielokrotnej, wielowymiarowej). Zmodyfikowane zagadnienie regresji wielokrotnej. Założenia modelu. Estymacja i weryfikacja parametrów strukturalnych w modelu liniowym regresji wielokrotnej. Ocena dopasowania modelu do danych: odchylenie standardowe reszt, współczynnik zmienności resztowej, współczynnik determinacji wielorakiej, skorygowany współczynnik determinacji wielorakiej, współczynnik zbieżności. Zmienne jakościowe objaśniające w analizie regresji</p> <p>6. Analiza reszt regresyjnych. Test Durbina-Watsona. Standaryzowane wartości resztowe. Odległość Mahalanobisa, odległość Cooka. Pojęcie tolerancji. Czynniki inflacji wariancji CIW (VIF). Regresja grzbietowa</p> <p>7. Metody selekcji zmiennych objaśniających: regresja krokowa postępująca, regresja krokowa wsteczna, metoda Hellwiga, metoda najlepszych podzbiorów. Kryteria selekcji modelu. Predykcja na podstawie modelu regresji wielorakiej.</p> <p>8. Model regresji logistycznej. Estymacja współczynników regresji logistycznej i weryfikacja ich statystycznej istotności (odchylenie (deviance) D, test Walda, test ilorazu wiarygodności LR). Pojęcie szansy oraz ilorazu szans. Interpretacja modelu regresji logistycznej. Założenia modelu. Ocena dopasowania modelu do danych empirycznych: współczynnik McFaddena, współczynnik Cragga-Uhlera, współczynnik Nagelkerke'a, współczynnik Coxa-Snella, miara Akaike'a (AIC), miara BIC. Test Hosmera - Lemeshowa. Analiza reszt modelu. Zmienne objaśniające jakościowe. Miary oparte na macierzy klasyfikacji (Czułość, Specyficzność, PPV, NPV, PV, ACC, LR). Konstrukcja krzywej ROC.</p> <p>9. Model regresji probitowej. Estymacja i weryfikacja współczynników modelu regresji probitowej. Test dobroci dopasowania modelu do danych empirycznych - χ^2. Model probitowy i model logitowy różnice, podobieństwa, interpretacja parametrów. Predykcja na podstawie modeli probitowych i logitowych.</p>
Forma oceniania:	<ul style="list-style-type: none"> • egzamin pisemny

Literatura:	<ol style="list-style-type: none"> 1. D. T. Larose, Metody i modele eksploatacji danych, PWN, Warszawa 2008 2. A. Stanisław, Przystępny kurs statystyki z zastosowaniem Statistica PL na przykładach z medycyny, tom I-III, StatSoft, Kraków, 2007. 3. W. Kryszczyński, J. Bartos, W. Dyczka, K. Królikowska, M. Wasilewski, Rachunek prawdopodobieństwa i statystyka matematyczna w zadaniach. Część II. Statystyka matematyczna, PWN, Warszawa 2007 4. M. Dobosz, Wspomagana komputerowo statystyczna analiza wyników badań, Akademicka Oficyna Wydawnicza, Warszawa 2004. 5. J. Greń, Statystyka matematyczna. Modele i zadania, PWN, Warszawa, 1984. 6. N.R. Draper, H. Smith, Analiza regresji stosowana, PWN, Warszawa, 1973. 7. M. Gruszczyński, Modele i prognozy zmiennych jakościowych w finansach i bankowości, Oficyna Wydawnicza SGH, Warszawa, 2002. 8. D. Hosmer, S. Lemeshow, Applied Logistic Regression - second edition, John Wiley & Sons, Inc, 2000. 9. J.M. Hilbe, Logistic regression models, Chapman & Hall/CRC Press 2009.
Dodatkowe informacje:	Dodatkowe informacje znajdują się na stronie Instytutu Matematyki