

# **Książka streszczeń**

---

## **Książka streszczeń:**

Data wydania Wrzesień 2007, ISBN 83-89585-15-4

Copyright © 2007 pielaszek research

## **Informacje redakcyjne**

---

### **Recenzenci**

- Eugeniusz Gatnar
- Krzysztof Jajuga
- Andrzej Sokołowski
- Marek Walesiak
- Dorota Witkowska

### **Redakcja techniczna**

- Krzysztof Kompa
- Roman Pielaszek

### **Projekt okładki**

- Krzysztof Kompa

Streszczenia opublikowane zostały w postaci dostarczonej przez PT Autorów, bez ingerencji redakcyjnej w materię opracowań.

Publikacja nie może być powielana i rozpowszechniana w jakikolwiek sposób, bez pisemnej zgody posiadacza praw autorskich

---

Revision: 105.0.34, 2007-09-10 19:45 GMT

---

---

## Spis treści

XVI Konferencja Naukowa SKAD .....	1
Komitet Naukowy .....	2
Komitet Organizacyjny .....	2
Program Prezentacji SKAD 2007 .....	3
Streszczenia wystąpień i plakatów .....	7
Sesja plenarna .....	9
Sesja A .....	13
Sesja B .....	21
Sesja C .....	31
Sesja plakatowa .....	37
Sesja warsztatowa .....	43
Lista uczestników .....	47
Indeks .....	53

---

---

---

# XVI Konferencja Naukowa SKAD



Szanowni Państwo,

SKAD - Konferencja Naukowa Sekcji Analizy i Klasyfikacji Danych Polskiego Towarzystwa Statystycznego to inicjatywa o ugruntowanej pozycji, szacowna i wielce zasłużona. Rokrocznie organizowana przez inny ośrodek akademicki jest okazją do przedstawienia osiągnięć i wymiany doświadczeń z zakresu teoretycznych i aplikacyjnych zagadnień klasyfikacji i analizy danych. Stanowi coroczne forum podsumowań stanu wiedzy, prezentacji i promocji dokonań nowatorskich. Służy wreszcie wytyczaniu kierunków dalszych prac i badań. Warto zauważyć, że publikowane materiały pokonferencyjne w postaci Zeszytów Naukowych Akademii Ekonomicznej we Wrocławiu pod nazwą *Taksonomia*, zajmują znaczące miejsce wśród publikacji z tego zakresu.

Tegoroczną XVI już Konferencję Naukową SKAD PTS, a zarazem XXI Konferencję Taksonomiczną „Klasyfikacja i Analiza Danych – Teoria i Zastosowania” organizuje Szkoła Główna Gospodarstwa Wiejskiego w Warszawie, we współpracy z Sekcją Klasyfikacji i Analizy Danych Polskiego Towarzystwa Statystycznego oraz Akademią Ekonomiczną we Wrocławiu. Patronat honorowy nad Konferencją roztoczyli Jego Magnificencja Rektor SGGW prof. dr hab. Tomasz Borecki i Dziekan Wydziału Ekonomiczno-Rolniczego SGGW prof. dr hab. Bogdan Klepacki.

Miejsce SKAD'07 wybraliśmy nieprzypadkowo – spotykamy się w dniach od 19 do 21 września 2007r. w Krynicy-Zdroju, w sercu Beskidów, w mieście o bogatych tradycjach uzdrowskich, na trwałe wpisane w historię malarstwa naiwnego, w mieście Festiwalu Kiepurowskiego i Europejskiego Forum Gospodarczego. Krynica-Zdrój jest miastem przyjaznym i gościnnym, którego senna atmosfera przeszłości spleta się ze śmiałymi projektami na przyszłość. Naszemu zjazdowi patronuje Burmistrz Krynicy-Zdroju.

Formuła tegorocznej Konferencji została w stosunku do poprzednich nieco poszerzona. Po pierwsze – wszystkie wystąpienia pogrupowane jak zwykle w sesje plenarne, grupy sesji tematycznych (A, B, C) i sesję plakatową, uzupełniono o wydzieloną sesję warsztatową.

Po drugie - wprowadzono całkowicie z informatyzowaną, „bezpapierową” obsługę Konferencji na każdym etapie jej przygotowania: od rejestracji uczestników, poprzez zgłaszanie streszczeń do recenzji, opiniowanie zgłoszonych opracowań, rozsyłanie powiadomień i zaproszeń do publikacji materiałów w postaci tego opracowania i w Internecie włącznie. Po trzecie – uruchomiono portal internetowy Konferencji SKAD'07 z podstawowymi informacjami o przygotowaniach; tu także obok wiadomości konferencyjnych i książki streszczeń opublikowane będą wszystkie dopuszczone materiały pokonferencyjne. Mamy nadzieję, że te działania spotkają się z przy-

chylną oceną uczestników Konferencji, przyspieszą wymianę informacji i będą zaczątkiem „nowej tradycji” SKAD.

Staraliśmy się aby niniejsze opracowanie w wyczerpujący sposób dokumentowało organizację Konferencji, odzwierciedlało jej program oraz zawierało streszczenia wszystkich wystąpień.

Streszczenia pogrupowane zostały w układzie sesji. W pierwszej kolejności prezentowane są streszczenia wystąpień w sesjach plenarnych, następnie w sesjach tematycznych w kolejności leksykograficznej oznaczeń sesji, następnie streszczenia plakatów i na koniec – streszczenia wystąpień w sesji warsztatowej. Tematyka wystąpień w zakresie teorii obejmuje:

- *taksonomię,*
- *analizę dyskryminacyjną,*
- *metody porządkowania liniowego,*
- *metody statystycznej analizy wielowymiarowej,*
- *metody analizy zmiennych ciągłych,*
- *metody analizy zmiennych dyskretnych,*
- *metody analizy danych symbolicznych,*
- *metody graficzne,*

zastosowania natomiast dotyczą:

- *analiz danych finansowych,*
- *analiz danych marketingowych,*
- *analiz danych przestrzennych,*
- *aplikacji w innych dziedzinach nauki jak: medycyna, psychologia, archeologia, itd. oraz*
- *komputerowych implementacji metod statystycznych.*

Opracowanie kończy alfabetyczny spis uczestników SKAD'07 z afiliacjami i adresami korespondencji elektronicznej oraz indeks alfabetyczny autorów wystąpień.

Konferencja w istniejącym kształcie programowym nie byłaby możliwa bez wsparcia PT Sponsorów: SAS Institute Sp. z o.o., GPW w Warszawie SA, Sarnowski & Wiśniewski Spółka Audytorska Sp. z o.o.

**Dorota Witkowska, prof. zw. dr hab.**

Przewodnicząca Komitetu Organizacyjnego

## Program Ramowy

### **18-IX-2007**

15:00 - 22:00 Rejestracja

18:30 Kolacja

### **19-IX-2007**

7:45 - 22:00 Rejestracja

8:00 - 8:45 Śniadanie

9:00 - 9:15 Otwarcie XVI Konferencji Naukowej SKAD

9:15 - 10:15 Sesja Plenarna I

10:15 - 10:30 Przerwa

10:30 - 11:45 Sesja Plenarna I (c.d.)

11:45 - 12:00 Przerwa

12:00 - 13:00 Sesja Plakatowa

13:00 - 14:00 Obiad

15:00 - 16:00 Sesje: A1, B1, C1

16:00 - 16:15 Przerwa

16:15 - 17:15 Sesje: A2, B2, C2

17:15 - 18:00 Zebranie Sekcji Klasyfikacji i Analizy Danych

19:30 Kolacja integracyjna

### **20-IX-2007**

8:00 - 8:45 Śniadanie

9:00 - 10:00 Sesja Plenarna II

10:00 - 10:15 Przerwa

10:15 - 11:30 Sesje: A3, B3

11:30 - 11:45 Przerwa

11:45 - 12:30 Sesje: B4, C4

10:15 - 12:30 Sesja Warsztatowa

12:30 - 13:30 Obiad

14:30 Wycieczka

19:30 Uroczysta Kolacja

### **21-IX-2007**

8:00 - 8:45 Śniadanie

9:00 - 10:00 Sesje: A5, B5, C5

10:00 - 10:15 Przerwa

10:15 - 11:45 Sesja Plenarna III

11:45 - 12:00 Zamknięcie XVI Konferencji Naukowej SKAD

12:00 - 13:00 Obiad

## Komitet Naukowy

### **Przewodniczący**

**Prof. zw. dr hab. Krzysztof Jajuga**

Akademia Ekonomiczna we Wrocławiu  
Wydział Zarządzania i Informatyki  
Katedra Inwestycji Finansowych i Zarządzania Ryzykiem

### **Prof. AE dr hab. Andrzej Sokołowski**

Akademia Ekonomiczna w Krakowie  
Wydział Zarządzania, Katedra Statystyki

### **Prof. AE dr hab. Eugeniusz Gatnar**

Akademia Ekonomiczna w Katowicach  
Wydział Zarządzania, Katedra Statystyki

### **Prof. zw. dr hab. Marek Walesiak**

Akademia Ekonomiczna we Wrocławiu  
Wydział Gospodarki Regionalnej i Turystyki w Jeleniej Górze  
Katedra Ekonometrii i Informatyki

### **Prof. zw. dr hab. Dorota Witkowska**

Szkoła Główna Gospodarstwa Wiejskiego w Warszawie  
Wydział Ekonomiczno-Rolniczy  
Katedra Ekonometrii i Statystyki

## Komitet Organizacyjny

### **Przewodniczący:**

**prof. zw. dr hab. Dorota Witkowska**

Szkoła Główna Gospodarstwa Wiejskiego w Warszawie  
Wydział Ekonomiczno-Rolniczy  
Katedra Ekonometrii i Statystyki

### **Sekretarz naukowy:**

dr inż. Krzysztof Kompa  
e-mail: kkompa@morssggw.waw.pl

### **Członkowie:**

prof. SGGW dr hab. Zbigniew Binderman,  
prof. dr hab. Bolesław Borkowski  
dr Aleksandra Matuszewska  
dr inż. Iwona Staniec

### **Sekretariat konferencji:**

dr inż. Grzegorz Koszela  
e-mail: skad2007@morssggw.waw.pl

---

# Program Prezentacji SKAD 2007

## 19-IX-2007

9:00 - 9:15      Otwarcie XVI Konferencji Naukowej SKAD

### 9:15 - 10:15      Sesja Plenarna I:

1. Reinhold Decker: Growing Neural Gas-like Networks and their Application to Data Analysis and Clustering in Marketing

10:15 - 10:30      Przerwa

### 10:30 - 12:00      Sesja Plenarna I (c.d.)

1. Krzysztof Jajuga: Parametry statystyczne mające u podstaw analizę połączeń
2. Marek Walesiak: Procedura analizy skupień z wykorzystaniem programu komputerowego clusterSim i środowiska R
3. Łukasz J. Kociuba: Automatyzacja prognozowania dużej ilości szeregów czasowych - podejście lean.

### 12:00 - 13:00      Sesja Plakatu

1. Grażyna Dehnel: Estymacja typu Winsora w badaniach podmiotów gospodarczych
2. Stanisław K. Jabłonowski: Klasyfikacja danych rolniczych w oparciu o metody analizy skupień stosowane w genetyce
3. Marek Karwanski, Wiesław Szczesny: Analiza ryzyka szacowania rezerw IBNR w zakładzie ubezpieczeniowym - modele stochastyczne dla danych zagregowanych i indywidualnych.
4. Jadwiga H. Kostrzewska: Zastosowanie wybranych modeli tobitowych do opisu tygodniowej liczby godzin pracy
5. Tadeusz Kufel, Marcin Błażejowski: Ekonometryczne modelowanie asymetrii reakcji klienta na zmiany cen
6. Małgorzata L. Łuniewska: Wykorzystanie fundamentalnych baz danych w analizie portfelowej
7. Anna Szmít: Analiza zapotrzebowania na kwalifikacje kadry inżynierskiej i menedżerskiej w regionie łódzkim na podstawie wyników badania ankietowego
8. Mirosława Sztemberg-Lewandowska: Analiza czynnikowa z wykorzystaniem środowiska R
9. Krzysztof Szwarz: Determinanty głębokości ubóstwa gospodarstw domowych
10. Paweł Ulman: Awersja do nierówności ekonomicznych w społeczeństwie polskim
11. Mirosław Wójciak, Aleksandra Wójcicka-Krenz: Zdolności dyskryminacyjne wskaźników finansowych w ocenie kondycji finansowej podmiotów gospodarczych
12. Aleksandra Wójcicka-Krenz: Wrażliwość modelu MKMV ze względu na metodę estymacji wzrostu wartości aktywów

15:00 - 16:00      Sesje A1, B1, C1

### Sesja A1: Metody Klasyfikacji i Analizy Danych I

1. Barbara Pawełek: Wybór parametrów normalizacyjnych i miary odległości a hierarchiczne grupowanie aglomeracyjne obiektów
2. Ewa M. Witek: Metoda taksonomii oparta na modelach mieszanych
3. Andrzej Bąk: Metody symulacyjne w badaniach preferencji z wykorzystaniem programu R
4. Marcin Szymkowiak, Jan Paradysz: Taksonometryczne podstawy kalibracji w statystyce małych obszarów

### Sesja B1: Zastosowania w Finansach I

1. Katarzyna Kuziak: Modelowanie rozkładu strat w pomiarze ryzyka operacyjnego
2. Mariola Chrzanowska, Dorota Witkowska: Analiza wrażliwości rozwiązań wyznaczonych dla różnej struktury próby kredytobiorców: analiza dyskryminacyjna i drzewa klasyfikacyjne.
3. Krzysztof Kompa: Badanie efektywności ekonomicznej portfeli losowych generowanych dla kompozycji wybranych indeksów GPW w latach 2000-2006
4. Krzysztof L. Piontek: Zastosowanie wielowymiarowych modeli GARCH do szacowania współczynnika zabezpieczenia dla kontraktów futures na WIG20

### Sesja C1: Różne zastosowania

1. Małgorzata Machowska-Szewczyk: Porównanie wyników klasyfikacji nauczycieli akademickich pod względem wpływu osobowości na popularność wykładów ze statystyki
2. Magdalena C. Mojsiewicz, Katarzyna Wawrzyniak, Barbara Bątko: Wykorzystanie analizy czynnikowej w segmentacji behawioralnej na przykładzie skali semiometrycznej
3. Paweł Lula: Eksploracyjna analiza polskojęzycznych ofert zatrudnienia specjalistów z zakresu informatyki
4. Mariusz Grabowski: Wykorzystanie metod eksploracyjnej analizy tekstu do identyfikacji kluczowych zagadnień zawartych w dużych zbiorach publikacji naukowych

16:00 – 16:15      Przerwa

16:15 – 17:15      Sesje A2, B2, C2

### Sesja A2: Metody Klasyfikacji i Analizy Danych II

1. Kamila Najman: Analiza porównawcza struktur hierarchicznych skupień uzyskanych w oparciu o hybrydowe metody grupowania
2. Iwona Kasprzyk: Wykorzystanie drzew klasyfikacyjnych w analizie klas ukrytych
3. Artur P. Zaborski: Identyfikacja preferencji z wykorzystaniem modeli PREFMAP
4. Krzysztof Najman: Symulacyjna analiza wpływu wyboru metody podziału obiektów na jakość uzyskanej klasyfikacji w algorytmach k-średnich

### Sesja B2: Zastosowania Międzynarodowe

1. Elżbieta Sobczak: Wielowymiarowa analiza korespondencji jako narzędzie segmentacji międzynarodowej
2. Małgorzata Markowska, Tadeusz Borys: Obszary wykorzystania miar syntetycznych i subsyntetycznych w kontekście ich własności formalnych

- 
3. Dominik A. Rozkrut: Propozycja metody klasyfikacji krajów z wykorzystaniem metod korekcji sezonowej

### Sesja C2: Zastosowania w Medycynie

1. Iwona E. Laskowska, Karolina E. Lewandowska: Badanie efektywności służby zdrowia z wykorzystaniem metod wielowymiarowej analizy porównawczej
2. Małgorzata A. Misztal: Zagregowane modele dyskryminacyjne i regresyjne w prognozowaniu czasu pobytu na OIOM pacjentów z chorobą wieńcową
3. Joanna Banaś, Małgorzata Machowska-Szewczyk: Badanie poziomu świadomości praw pacjenta w środowisku medycznym

17:15 - 18:00 *Zebranie Sekcji Klasyfikacji i Analizy Danych*

## **20-IX-2007**

### 9:00 - 10:00: Sesja Plenarna II

1. Jose Maria Montero Lorenzo, Beatriz Larraz Iribas: Estimating Housing Prices: Kriging and Cokriging as an Alternative
2. Esteban Alfaro Cortés, Matías Gámez Martínez, Noelia García Rubio: Multiclass corporate failure prediction by aggregation classification trees

10:00 - 10.15 Przerwa

10.15 - 11.30 Sesje A3, B3

### Sesja A3: Metody Klasyfikacji i Analizy Danych III

1. Małgorzata Gliwa: Wykorzystanie pewnego kryterium przy budowie drzewa klasyfikacyjnego dla zbioru obiektów symbolicznych
2. Dorota E. Rozmus, :Wykorzystanie podejścia zagregowanego w taksonomii
3. Aleksandra Łuczak, Feliks Wysocki: Wykorzystanie rozmytej metody TOPSIS opartej na zbiorze alpha przekrojów do porządkowania liniowego obiektów
4. Joanna Trzęsiok: Ocena zasadności łączenia wybranych nieparametrycznych modeli regresji
5. Michał Trzęsiok: Walidacja wyników klasyfikacji otrzymanych taksonomiczną metodą wektorów nośnych

### Sesja B3: Zastosowania w analizach przestrzennych

1. Barbara Dańska-Borsiak: Różnicowanie poziomu rozwoju gospodarczego województw w Polsce a wielkość migracji międzywojewódzkich. Zastosowanie wybranych procedur taksonomicznych
2. Monika Rozkrut: Typologia regionów Polski ze względu na dysproporcje rozwojowe
3. Elżbieta Wiszniewska: Taksonomiczna analiza porównawcza poziomu zrównoważonego rozwoju województw w Polsce
4. Stanisław W. Matusik: Wielowymiarowa analiza porównawcza poziomu rozwoju gmin woj. małopolskiego w latach 1996-2005
5. Iwona A.Foryś: Wykorzystanie metod porządkowania liniowego do oceny polskiego rynku nieruchomości.

11:30 - 11:45 Przerwa

11:45 - 12:30 Sesje B4, C4

### Sesja B4: Zastosowania w Finansach II

1. Katarzyna Wawrzyniak, Barbara Batóg, Magdalena C. Mojsiewicz: Porównanie wyników segmentacji na różnych rynkach ubezpieczeń gospodarstw domowych
2. Paweł Rokita: Porównanie koncepcji zależności ekstremalnej i warunkowo zmiennej kowariancji w modelowaniu zależności między indeksami wybranych rynków akcji
3. Marcin Salamaga: Wykorzystanie wybranych metod analizy wielowymiarowej do klasyfikacji funduszy inwestycyjnych

### Sesja C4: Analiza Rynku Pracy

1. Aleksandra Witkowska, Marek Witkowski: Zmienna syntetyczna z medianą w analizie lokalnego rynku pracy
2. Joanna M. Landmesser: Aktywność ekonomiczna ludności: klasyfikacja osób za pomocą wielomianowych modeli logitowych oraz jej związek z modelami hazardu dla czasów trwania.
3. Iwona Staniec, Marek Martin: Analiza taksonomiczna ofert pracy pojawiających się w dodatku "Praca" w regionie Łódzkim w 2006 roku

### Równolegle: 10:15 - 12:30 Sesja Warsztatowa

1. Aleksandra Matuszewska, Dorota Witkowska: Modelowanie i prognozowanie kursu euro/dolar: modele autoregresyjne z rozkładami opóźnień i sztuczne sieci neuronowe
2. Mirosław Wójciak, Aleksandra Wójcicka-Krenz: Analiza porównawcza nowych metod oceny ryzyka kredytowego
3. Dorota Żebrowska-Suchodolska, Dorota Witkowska: Badanie efektywności GPW na przykładzie wybranych indeksów: test autokorelacji
4. Agnieszka Sompolska-Rzechuła, Beata Szczecińska: Statystyczna analiza efektywności gospodarowania spółek notowanych na Gieldzie Papierów Wartościowych w Warszawie
5. Arkadiusz J. Orłowski, Urszula Grzybowska, Piotr Jałowicki: Selekcja i klasyfikacja spółek notowanych na Warszawskiej Gieldzie Papierów Wartościowych przy pomocy pakietu SAS
6. Artur Mikulec: Zastosowanie wybranych metod rangowania obiektów do oceny stopnia zróżnicowania starości demograficznej ludności w krajach UE i EFTA
7. Danuta J. Rozpędowska-Matraszek, Barbara Dańska-Borsiak: Zastosowanie wybranych metod statystycznych i taksonomicznych do oceny sytuacji na rynku pracy ze szczególnym uwzględnieniem kształtowania się wynagrodzeń
8. Danuta J. Rozpędowska-Matraszek: Zastosowanie metod statystycznych i ekonometrycznych do analiz popytu i podaży pracy w sektorze ochrony zdrowia w Polsce
9. Piotr Jałowicki, Ewa Jałowicka, Marek Karwański: Porównanie prognoz spożycia papierosów w Polsce w latach 2006-2010 na podstawie modeli wyrównywania wykładniczego oraz autoregresyjnych
10. Włodzimierz Kołodenny, Tadeusz Żdanowicz: Modelowanie uzysku energii elektrycznej z modułów fotowoltaicznych na podstawie długoterminowych badań w warunkach naturalnych
11. Kesra Nermend: Koncepcja systemu informatycznego do analizy danych regionalnych



---

## **21-IX-2007**

9:00 - 10:00     Sesje A5, B5, C5

### **Sesja A5: Metody Klasyfikacji i Analizy Danych IV**

1. Małgorzata Markowska, Danuta Strahl: Klasyfikacja dynamiczno-przestrzenna europejskiej przestrzeni regionalnej ze względu na poziom gospodarki opartej na wiedzy
2. Marcin Pełka, Andrzej Dudek: SymScal: Metoda skalowania wielowymiarowego obiektów symbolicznych
3. Jerzy M. Korzeniewski: Badanie efektywności zmodyfikowanych metod aglomeracyjnych na zbiorach danych ze świata realnego
4. Mariusz Kubus: Zastosowanie metody boosting w indukcji reguł

### **Sesja B5: Zastosowania w Finansach III**

1. Daniel Papła: Przykład zastosowania funkcji powiązań Farliego-Gumbela-Morgensterna w analizie związków zależności między wybranymi akcjami notowanymi na GPW w Warszawie
2. Jacek Batóg: Klasyfikacja regionów według stopnia specjalizacji i koncentracji sektorowej oraz zróżnicowania dochodowego
3. Paweł Żuraw, Piotr Tarka: Model analizy conjoint w ocenie oferty ubezpieczeń od ryzyka do kredytu hipotecznego
4. Tomasz Ząbkowski: Zastosowanie sztucznych sieci neuronowych do oceny ryzyka kredytowego klienta w telefonii komórkowej

### **Sesja C5: Analiza gospodarstw domowych**

1. Hanna Dudek: Wielowymiarowe aspekty ubóstwa - statystyczna analiza zasobów dóbr trwałego użytku w gospodarstwach domowych rolników
2. Joanna Kisielińska: Klasyfikacja gospodarstw rolniczych w oparciu o dane w postaci szeregów czasowych
3. Izabela Kurzawa, Feliks Wysocki: Wykorzystanie analizy koszykowej do identyfikacji zachowań konsumpcyjnych gospodarstw domowych w Polsce
4. Piotr Łukasiewicz, Grzegorz Koszela, Arkadiusz J. Orłowski: Klasyfikacja rozkładów dochodów gospodarstw domowych

10:00 – 10:15     Przerwa

### **10:15 – 11:45     Sesja Plenarna III**

1. Waldemar S. Tarczyński: Analiza i ocena potencjału społeczno-gospodarczego województw w Polsce.
2. Mirosława Gazińska: Z badań nad rozwojem ludności wiejskiej Pomorza Pruskiego w latach 1762-1772. Analiza taksonomiczna
3. Eugeniusz Gatnar: Wpływ metody łączenia modeli na wielkość błędu klasyfikacji w podejściu wielomodelowym

**11:45 - 12:00     Zamknięcie XVI Konferencji Naukowej SKAD**



---

# **Streszczenia wystąpień i plakatów**

---

---

---

# Sesja plenarna

## Program

Środa, 19 Wrzesień

### Sesja Plenarna I

Środa rano, 19 Wrzesień, 9:15

9:15

Invited oral

#### Growing Neural Gas-like Networks and their Application to Data Analysis and Clustering in Marketing

Reinhold Decker

Universitaet Bielefeld, Fakultae fuer Wirtschaftswissenschaften,  
Universitaetstr. 25, Bielefeld 33615, Germany

*e-mail: RDecker@wiwi.uni-bielefeld.de*

The talk will present new developments in the field of neural gas-like networks and will show how this comparatively young and powerful class of neural networks can be applied successfully to data analysis and clustering. Due to their flexible structure (and unlike SOM, which are fixed to predefined lattices) these algorithms are able to preserve the topological structure of the data to be analyzed (i.e. the data determines the topology of the neural network). This proves helpful in clustering for market segmentation, among others.

The performance of the relevant algorithms in pattern recognition and clustering is demonstrated using synthetic as well as real marketing data. The applications discussed include market segmentation, market basket analysis, and missing value imputation.

**Keywords:** Clustering, data analysis, marketing, neural gas network

### Przerwa

Środa rano, 19 Wrzesień, 10:15

### Sesja Plenarna I (c.d.)

Środa rano, 19 Wrzesień, 10:30

10:30

Oral

#### Parametry statystyczne mające u podstaw analizę połączeń

Krzysztof Jajuga

Akademia Ekonomiczna (AEWROCLAW), Komandorska 118/120,  
Wrocław 53-345, Poland

*e-mail: krzysztof.jajuga@ae.wroc.pl*

Referat przedstawia koncepcję podstawowych parametrów statystycznych dla rozkładu wielowymiarowego, w przypadku, gdy do opisu tego rozkładu wykorzystana jest analiza połączeń (*copula analysis*). Podejście to wiąże dystrybuantę rozkładu wielowymiarowego

z dystrybuantami rozkładów brzegowych poprzez funkcję połączenia.

W pierwszej kolejności w referacie przedstawione są teoretyczne podstawy analizy połączeń. Właściwa część referatu zawiera omówienie głównych parametrów statystycznych dla rozkładu wielowymiarowego w podziale na: parametry położenia, parametry rozrzutu i parametry zależności. Przedstawione są również praktyczne interpretacje tych parametrów, głównie w odniesieniu do finansów.

11:00

Oral

#### Procedura analizy skupień z wykorzystaniem programu komputerowego clusterSim i środowiska R

Marek Walesiak

Akademia Ekonomiczna (AEWROCLAW), Komandorska 118/120,  
Wrocław 53-345, Poland

*e-mail: marek.walesiak@ae.jgora.pl*

W referacie zaprezentowano typową procedurę analizy skupień (por. np. Milligan [1996], s. 342-343; Walesiak [2005]) obejmującą wybór obiektów i zmiennych, wybór formuły normalizacji wartości zmiennych, wybór miary odległości, wybór metody klasyfikacji, ustalenie liczby klas, ocena wyników klasyfikacji, opis (interpretacja) i profilowanie klas. Następnie scharakteryzowano funkcje pakietów clusterSim, stats, cluster i ade4 środowiska R służących realizacji poszczególnych jej etapów. Ponadto przedstawiono przykładowe składnie poleceń (procedury) z wykorzystaniem analizowanych funkcji. Procedury te ułatwiają potencjalnemu użytkownikowi realizację wielu zagadnień klasyfikacyjnych niedostępnych w podstawowych pakietach statystycznych (np. SPSS, Statistica, S-Plus, SAS).

Zakres tematyczny: Zastosowania – aplikacje komputerowe metod statystycznych

#### Literatura

MILLIGAN, G.W. (1996), *Clustering validation: results and implications for applied analyses*, In: P. Arabie, L.J. Hubert, G. de Soete (Eds.), *Clustering and classification*. World Scientific, Singapore, 341-375.

WALESIAK M. (2005), *Rekomendacje w zakresie strategii postępowania w procesie klasyfikacji zbioru obiektów*, W: Zeliaś A. (red.), „Przestrzenno-czasowe modelowanie i prognozowanie zjawisk gospodarczych”, Wydawnictwo AE, Kraków, 185-203.

11:30

Oral

#### Automatyzacja prognozowania dużej ilości szeregów czasowych – podejście lean.

Łukasz J. Kociuba

SAS Institute Sp. z o.o., Gdańska27/31, Warszawa 01-633, Poland

*e-mail: Lukasz.Kociuba@spl.sas.com*

Podejście lean zastosowane w prognozowaniu, łańcuchu dostaw, księgowości czy jakimkolwiek innym procesie biznesowym ma na celu identyfikację i eliminację nieefektywności procesu. Nie zawsze można kontrolować dokładność wyznaczanych prognoz, nie zawsze

można wyznaczyć prognozy o żądanej dokładności. Niemniej podanie kontroli samego procesu prognozowania umożliwia kontrolowanie nakładów przeznaczanych na wyznaczenie prognoz. Artykuł przedstawia aktualne metody automatyzacji prognozowania dużej ilości szeregów czasowych oraz zastosowanie podejścia znanego z procesów produkcyjnych – lean manufacturing do wyznaczania prognoz.

### Zakończenie obrad plenarnych

Środa popołudniu, 19 Wrzesień, 12:00

## Czwartek, 20 Wrzesień

### Sesja Plenarna II

Czwartek rano, 20 Wrzesień, 9:00

9:00

Invited oral

#### Estimating Housing Prices: Kriging and Cokriging as an Alternative

Jose Maria Montero Lorenzo, Beatriz Larraz Iribas

*University of Castilla-La Mancha, Faculty of Economics and Business Studies of Albacete (UCLM), Plaza de la Universidad, Albacete 02071, Spain*

*e-mail: Jose.M.Lorenzo@uclm.es*

Nowadays, the housing prices are one of the burning issues in some of the current main economies in the world (such is the case of the United States of America, the United Kingdom, Spain, among others, as well as in Poland). This is the reason of the increasing literature about methodologies for estimating them, because the valuation of the real estate properties is crucial for fiscal issues, it influences the consumption and the rate of economic growth, etc. Model-driven orientation has been traditionally used for estimating housing prices (hedonic models, spatial autorregressive models, mixed regressive spatial autorregressive models, models with spatial dependence in the error term, spatial Durbin models, etc.). But there is another alternative approach, the data-driven approach, which includes, among others, shape functions, methods based on inverse distance weighting, kriging and cokriging, the last two ones having shown to provide excellent results due to the fact that they take into account the structure of the spatial dependence existing in the data. Having said that, the paper shows how kriging and cokriging work, and provides their equations in several situations (constant and unknown mean, existence of trend, homotopic, partial and complete heterotopic cases, etc.). It includes three different applications in the real estate market of Toledo (Spain). Out of the real estate valuation context, the univariate and multivariate kriging methods are used to estimate electoral results in the region of Valencia (Spain) and to analyze the neighbourhood crime in Columbus (Ohio, USA).

JEL classification: C21, C31, R31, P25.

Keywords: kriging, cokriging, variogram, spatial dependence, real estate valuation.

9:30

Invited oral

#### Multiclass corporate failure prediction by aggregation classification trees.

Esteban A. Cortes, Matias Gamez Martinez, Noelia Garcia Rubio

*Castilla-La Mancha University., Plaza de la Universidad, 1, Albacete 02-071, Spain*

*e-mail: 1Esteban.Alfaro@uclm.es*

The goal of this study is to compare the behaviour of three ensemble methods (AdaBoost, Bagging and Random Forest) in a financial application. Since the sixties several classification techniques have been used for corporate failure prediction. In last decades, the interest in this task has widely increased owing to the existence of data bases with available information.

Classification trees are a powerful alternative to the more traditional statistical models. This model has the advantage of being able to detect non-linear relationships and showing a good performance in presence of qualitative information as it happens in corporate failure prediction problems. As a result, they are widely used as base classifiers for ensemble methods.

AdaBoost constructs its base classifiers in sequence, updating a distribution over the training examples to create each base classifier. Bagging combines the individual classifiers built in bootstrap replicates of the training set. Random Forest is a combination of tree predictors such that each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest.

In this paper, we compare the prediction accuracy of these techniques for the corporate failure prediction task.

**Key Words:** Bankruptcy Prediction, AdaBoost, Bagging, Random Forest.

### Zakończenie obrad plenarnych

Czwartek rano, 20 Wrzesień, 10:00

## Piątek, 21 Wrzesień

### Sesja Plenarna III

Piątek rano, 21 Wrzesień, 10:15

10:15

Oral

#### Analiza i ocena potencjału społeczno-gospodarczego województw w Polsce.

Wldemar S. Tarczyński, Małgorzata L. Łuniewska

*Uniwersytet Szczeciński, Jedności Narodowej 22, Szczecin 70-453, Poland*

*e-mail: wtarc@uoo.univ.szczecin.pl*

Potencjał społeczno-gospodarczy województwa jest bardzo ważną kategorią ekonomiczną. Zarówno rozwój jak i poziom społeczno-gospodarczy mają istotne znaczenie chociażby w procesie dystrybu-

cji środków europejskich przeznaczanych na rozwój danego regionu. Jest to również bardzo ważna kwestia z punktu widzenia atrakcyjności i stymulowania rozwoju przedsiębiorczości. Oznacza to, możliwość wyodrębnienia regionów atrakcyjnych pod tym względem. Znajomość tego problemu może wspomagać rozwiązywanie wielu istotnych kwestii ekonomicznych, społecznych i gospodarczych (np. walkę z bezrobociem).

Badaniu poddano 16 województw Polsce, scharakteryzowanych wybranymi zmiennymi ekonomiczno-financeowymi. Do analizy i oceny potencjału społeczno-gospodarczego wykorzystano wybrane metody wielowymiarowej analizy porównawczej. Na tej podstawie dokonano klasyfikacji województw oraz oceniono ich potencjał i atrakcyjność z punktu widzenia przyjętego celu badania. Podjęto również próbę oceny wpływu i znaczenia zmiennych diagnostycznych na wynik badania. Dodatkowo analizie i ocenie poddano sektory, występujące w województwach. W badaniu wykorzystano podstawowe elementy statystyki opisowej oraz wielowymiarowej analizy porównawczej. Badaniem objęto lata 2005-2006.

10:45

Oral

### Z badań nad rozwojem ludności wiejskiej Pomorza Pruskiego w latach 1762-1772. Analiza taksonomiczna

Mirosława Gazińska, Radosław Gaziński

*Uniwersytet Szczeciński, Jedności Narodowej 22, Szczecin 70-453, Poland*

*e-mail: mgazi@univ.szczecin.pl*

W Archiwum Państwowym w Szczecinie w zespole Rękopisy i Spiszczyzny zachowało się 11 ksiąg finansowych z lat 1762-1772, odnoszących się do Pomorza Pruskiego. Posłużyły one do przeprowadzenia wszechstronnego badania dotyczącego zjawisk demograficznych i społeczno-ekonomicznych w 42 domenach państwowych na terenach wiejskich Pomorza Pruskiego. Analiza objęła między innymi dynamikę rozwoju ludności wiejskiej (na Pomorzu Pruskim oraz w poszczególnych domenach państwowych), zmiany w strukturze ludności według płci oraz funkcjonalnych grup wieku, a także wybrane procesy w ruchu naturalnym. Badaniem objęło także domeny państwowe z uwzględnieniem: liczby wsi, liczby folwarków, liczby domów, liczby inwentarza żywego.

Ponadto przeprowadzono analizę taksonomiczną, różnicującą pomorskie domeny wiejskie ze względu na: strukturę mieszkańców, udział ludności najemnej w ogólnej liczbie ludności, a także tempo odbudowy demograficznej wsi domenalnej po zakończeniu wojny siedmioletniej.

Podstawowe metody stosowane w badaniach to metody statystyczne zastosowane do opisu materiału statystycznego oraz taksonomiczne metody grupowania i klasyfikacji, a także taksonomiczny miernik rozwoju.

11:15

Oral

### Wpływ metody łączenia modeli na wielkość błędu klasyfikacji w podejściu wielomodelowym

Eugeniusz Gatnar

*Akademia Ekonomiczna im. K. Adamieckiego w Katowicach, Katowice 40-287, Poland*

*e-mail: egatnar@ae.katowice.pl*

W nieparametrycznej analizie dyskryminacyjnej (Gatnar, 2001) coraz większe znaczenie odgrywa podejście wielomodelowe (agregacja modeli) polegające na łączeniu  $M$  modeli składowych  $D_1(\mathbf{x}), \dots, D_M(\mathbf{x})$  w jeden model globalny za pomocą pewnej funkcji  $F$ :

$$D^*(\mathbf{x}) = F\{D_1(\mathbf{x}), D_2(\mathbf{x}), \dots, D_M(\mathbf{x})\}$$

gdzie  $D_m(\mathbf{x})$  jest wynikiem predykcji  $m$ -tego modelu składowego.

W literaturze można znaleźć kilkanaście różnych propozycji postaci funkcji  $F$ , począwszy od najprostszych jak np. uśrednianie lub głosowanie (Gatnar, 2005), aż do bardzo wyszukanych, wykorzystujących całki rozmyte lub szablony decyzyjne.

W referacie zostaną przedstawione wyniki analizy porównawczej dla różnych postaci funkcji  $F$  łączącej modele składowe. Zostanie także dokonana ocena ich własności, wpływających na wielkość błędu klasyfikacji modelu zagregowanego.

#### Literatura

GATNAR E. (2001): Nieparametryczna metoda dyskryminacji i regresji, PWN, Warszawa.

GATNAR E. (2005): „Dobór zmiennych do zagregowanych modeli dyskryminacyjnych”, *Taksonomia* 12, *Prace Naukowe Akademii Ekonomicznej we Wrocławiu*, nr 1076, str. 79-86.

### Zakończenie Obrad SKAD 2007

Piątek rano, 21 Wrzesień, 11:45

### Obiad

Piątek popołudniu, 21 Wrzesień, 12:00





# Sesja A

Metody Analizy i Klasyfikacji Danych

## Program

Środa, 19 Wrzesień

### Sesja A1

Metody Klasyfikacji i Analizy Danych I  
Środa popołudniu, 19 Wrzesień, 15:00

15:00 Oral

#### Wybór parametrów normalizacyjnych i miary odległości a hierarchiczne grupowanie aglomeracyjne obiektów

Barbara Pawełek

Akademia Ekonomiczna w Krakowie, Katedra Statystyki (AE), Rakowicka 27, Kraków 31-510, Poland

e-mail: [Barbara.Pawelek@ae.krakow.pl](mailto:Barbara.Pawelek@ae.krakow.pl)

Głównym celem proponowanego referatu jest podzielenie się wynikami teoretycznych rozważań dotyczących wpływu wyboru parametrów normalizacyjnych i miary odległości na wyniki grupowania obiektów z wykorzystaniem hierarchicznych metod aglomeracyjnych. Metody te można opisać centralną procedurą aglomeracyjną. Przedstawiony zostanie zatem schemat postępowania przedstawiający działanie tej procedury, w którym można zauważyć zasadniczą rolę relacji między dwiema wielowymiarowymi odległościami przy podejmowaniu decyzji o połączeniu, w danym powtórzeniu, dwóch skupień. To spostrzeżenie pozwoli nawiązać w rozważaniach do prowadzonych przez autorkę badań nad wpływem wyboru formuły normalizacyjnej i miary odległości na relację między wielowymiarowymi odległościami.

Zgodnie z ideą omawianych metod, w każdym powtórzeniu etapów centralnej procedury aglomeracyjnej, w celu m.in. prześledzenia „losów” badanych obiektów, rejestruje się, które skupienia zostały połączone, ich skład oraz, jaka była wielowymiarowa odległość między nimi. W badaniach należy liczyć się z możliwością otrzymania różnych wniosków w zależności od podjętych decyzji, co do wyboru: wstępnego uporządkowania liniowego obiektów, formuły normalizacyjnej, miary odległości oraz hierarchicznej metody aglomeracyjnej.

W wystąpieniu zaprezentowany będzie przykład pokazujący powyższe możliwości/zagrożenia. Analizie poddanych będzie: 5 wybranych wariantów parametru skalującego, 3 wybrane odległości, tzn. miara miejska, Euklidesa i Czebyszewa oraz 7 wybranych kombinatorycznych hierarchicznych metod aglomeracyjnych.

W świetle wniosków z zaprezentowanych rozważań, w przypadku analizowania „ścieżki połączeń” dla wybranego obiektu lub wyboru jednego spośród uzyskanych podziałów, jako najlepiej reprezentującego grupy podobnych obiektów, mogą zrodzić się wątpliwości, co do uniwersalności, wiarygodności, dopuszczalności, czy też trafno-

ści wniosków wyciąganych z tego typu analiz. Autorka uważa, że reakcją na te wątpliwości powinien być świadomy wybór takich statystycznych narzędzi, jak formuły normalizacyjne, miary odległości, czy też taksonomiczne metody. Śledzenie zmian w połączeniach skupień, w zależności od decyzji podjętych na wcześniejszych etapach badania porównawczego i poszukiwanie odpowiedzi na pytania: dlaczego otrzymuje się różne przyporządkowania; dlaczego dotyczą one tylko wybranych obiektów; czy przyczyny tych odmienności są do zaniedbania, czy też są ważną informacją o kształtowaniu się badanego złożonego zjawiska ekonomicznego w rozważanej zbiorowości; mogą być pomocne w uzyskiwaniu bardziej wiarygodnych wniosków.

15:15 Oral

#### Metoda taksonomii oparta na modelach mieszanych

Ewa M. Witek

Akademia Ekonomiczna im. K. Adamieckiego w Katowicach, Katowice 40-287, Poland

e-mail: [eewwaaww@interia.pl](mailto:eewwaaww@interia.pl)

W ostatnim czasie coraz większą popularnością cieszy się modelowe podejście w taksonomii (*model-based clustering*). Podejście to zakłada, że obserwacje pochodzą z pewnego skończonego zbioru prostych rozkładów prawdopodobieństwa o nieznanymi parametrach. W praktyce najczęściej wykorzystywany jest rozkład normalny. W modelu mieszanym każda zmienna, której prawdopodobieństwa nie znamy przypisana jest odpowiedniej klasie. Problem wyboru najlepszej z metod klasyfikacji i optymalnej liczby klas sprowadza się do wyboru modelu statystycznego o najlepszej jakości dopasowania.

W metodzie taksonomii opartej na modelach mieszanych, wykorzystuje się połączenie hierarchicznej metody aglomeracyjnej opartej na modelach (*model-based hierarchical clustering*), algorytmu EM oraz statystyki BIC. Hierarchiczne metody aglomeracyjne oparte na modelach dokonują podziału zbioru obiektów na klasy. Podział ten ma na celu zainicjowanie algorytmu EM, który daje lepsze wyniki wtedy, gdy znamy wartości startowe, zaś kryterium informacyjne BIC pozwala na wybór właściwej liczby klas.

W części empirycznej referatu metoda taksonomii oparta na modelach mieszanych zostanie porównana z powszechnie stosowanymi metodami taksonomii tj. metodą hierarchiczną oraz metodą k-średnich. Badania przeprowadzone zostaną za pomocą pakietów Mclust i Mclust02 programu komputerowego R.

Literatura:

1. Biernacki, Ch., Celeux, G., Govaert, G., Langrognet, F. (2006), *Model-based cluster and discriminant analysis with the MIXMOD software*, “Computational Statistics and Data Analysis”, 51, 587-600
2. Fraley, C., and Raftery, A.E. (2002), *Model-Based Clustering, Discriminant Analysis, and Density Estimation*, “Journal of the American Statistical Association”, 97, 611-631
3. Fraley, C., and Raftery, A.E. (1998), *How many clusters? Which Clustering Method? Answers via Model-Based Cluster Analysis*, “The Computer Journal”, 41, 577-588.

15:30

Oral

### Metody symulacyjne w badaniach preferencji z wykorzystaniem programu R

Andrzej Bąk

Akademia Ekonomiczna (AE), Nowowiejska3, Jelenia Góra 58-500, Poland

e-mail: andrzej.bak@ae.jgora.pl

Metody symulacyjne mogą być stosowane w połączeniu z metodami dekompozycyjnymi w badaniu następujących grup zagadnień o charakterze teoretycznym i aplikacyjnym mieszczących się w kontekście badań preferencji konsumentów (zob. Bąk 2004):

- symulacyjne badanie własności formalnych i aplikacyjnych poszczególnych procedur dekompozycyjnych (*conjoint analysis* i wyborów dyskretnych),
- symulacyjne generowanie danych wykorzystywanych w badaniach formalnych i aplikacyjnych własności algorytmów dekompozycyjnych,
- symulacyjne modelowanie rynku na podstawie danych zgromadzonych metodą *conjoint analysis*,
- wykorzystanie metod symulacyjnych w prognozowaniu preferencji konsumentów na podstawie danych zgromadzonych metodą *conjoint analysis*.

Program R nie oferuje pakietu bezpośrednio wspierającego badania preferencji za pomocą metod dekompozycyjnych (*conjoint analysis*, metody wyborów dyskretnych). Procedury obliczeniowe zawarte w różnych pakietach mogą być jednak wykorzystane w badaniach preferencji oraz analizach symulacyjnych z zastosowaniem metod dekompozycyjnych.

W referacie zostaną zaprezentowane dwie grupy zagadnień:

- pakiety i procedury obliczeniowych dostępne w programie R, które mogą znaleźć zastosowanie w badaniach preferencji za pomocą metod dekompozycyjnych,
- możliwości wykorzystania metod symulacyjnych w analizie danych o preferencjach z wykorzystaniem pakietów i procedur programu R.

#### Literatura

Bąk A. [2004], *Dekompozycyjne metody pomiaru preferencji w badaniach marketingowych*, Prace Naukowe nr 1013, Seria: *Monografie i Opracowania nr 157*, Wydawnictwo AE we Wrocławiu, Wrocław.

Everitt B.S. [2004], *An R and S-PLUS Companion to Multivariate Analysis*. Springer-Verlag, London.

Everitt B.S., Dunn G. [2001], *Applied Multivariate Data Analysis*. Arnold, London.

15:45

Oral

### Taksonometryczne podstawy kalibracji w statystyce małych obszarów

Jan Paradysz, Marcin Szymkowiak

Akademia Ekonomiczna w Poznaniu, al. Niepodległości 10, Poznań 60-967, Poland

e-mail: m.szymkowiak@ae.poznan.pl

Statystyka małych obszarów (SMO) obejmuje metody estymacji z wykorzystaniem wszelkich istniejących zasobów informacji (różnego rodzaju rejestrów administracyjnych, spisów, wykazów itd.) w przypadku, gdy liczebność próby jest niewystarczająca dla uzyskania wiarygodnych bezpośrednich ocen szacowanych parametrów.

W statystyce małych obszarów podstawowym pojęciem w klasycznym ujęciu jest „pożyczanie siły” (strength borrowing). Pożyczanie siły odbywa się bądź z dużego obszaru, bądź z obserwacji trendów w szeregach dynamicznych. W swoim czasie zaproponowaliśmy pożyczanie siły z poziomu obszarów typowych. W niniejszym opracowaniu proponujemy wykorzystanie tego podejścia w kalibracji danych dla małego obszaru. Jest to kontynuacja problemu zaprezentowanego na poprzedniej konferencji SKAD w 2006r. w Wiśle. Jak wykazano w ub. roku, kalibracja jako remedium na duże odsetki braków odpowiedzi w budżetach gospodarstw domowych sprawdza się na poziomie stosunkowo „dużych” małych obszarów, jakimi są województwa. Tutaj wschodzimy na poziom powiatów, a to wymaga pogłębienia analizy, między innymi o metody taksonometryczne.

Małym obszarem będą powiaty województwa wielkopolskiego charakteryzujące się dużym odsetkiem osób bezrobotnych (powyżej 20%) w 2002r. Tymi powiatami były: chodzieski, czarnkowsko - trzecieński, gnieźnieński, jarociński, kolski, koniński, międzychodzki, obornicki, ostrowski, pilski, pleszewski, słupecki, wągrowiecki, wrzesiński, złotowski, oraz miasta na prawach powiatu: Kalisz i Konin. Wykorzystując metody taksonomiczne, przeprowadzono klasyfikację wszystkich powiatów ze względu na wybrane cechy społeczno – ekonomiczne skorelowane z rynkiem pracy (współczynnik aktywności zawodowej, wskaźnik zatrudnienia i stopa bezrobocia). W dalszym ciągu, z wykorzystaniem wybranych estymatorów kalibracyjnych - w oparciu o informacje „zapożyczone” z małych obszarów - dokonano estymacji parametrów niektórych cech charakteryzujących powiaty, znajdujące się w tej samej klasie co rejony, z których „pożyczono siłę”.

Paradysz J., „Statystyka małych obszarów” w Zeliaś A., (red.) „Przestrzenno – czasowe modelowanie i prognozowanie zjawisk gospodarczych”, Kraków 1999r, materiały z XX Ogólnopolskiego Seminarium Naukowego zorganizowanego przez Zakład Teorii Prognoz Katedry Statystyki AE w Krakowie.

#### Przerwa

Środa popołudniu, 19 Wrzesień, 16:00

#### Sesja A2

*Metody Klasyfikacji i Analizy Danych II*  
Środa popołudniu, 19 Wrzesień, 16:15

16:15

Oral

### **Analiza porównawcza struktur hierarchicznych skupień uzyskanych w oparciu o hybrydowe metody grupowania**

Kamila Najman

Uniwersytet Gdański Wydział Zarządzania (WZR), Armii Krajowej 101, Sopot 81-824, Poland

e-mail: [kmn@panda.bg.univ.gda.pl](mailto:kmn@panda.bg.univ.gda.pl)

Jednym z celów analizy wielowymiarowych zbiorów danych jest wyróżnienie homogenicznych grup obiektów. Grupy te skupiają obiekty, które z punktu widzenia badanych cech są do siebie podobne. Wyróżnienie takich grup stanowi istotę analizy skupień i jest zadaniem dobrze znanym. Innym aspektem analizy danych wielowymiarowych jest ustalenie hierarchii obiektów. Hierarchia ta wskazuje, które obiekty i w jaki sposób są do siebie podobne. W badaniach empirycznych pojawia się coraz większa potrzeba połączenia obu tych aspektów badań zbiorów wielowymiarowych. Gwałtownie rozwijająca się informatyzacja życia ekonomicznego powoduje między innymi znaczny rozwój wielkości baz danych. Nierzadko bazy te zawierają dziesiątki tysięcy obiektów charakteryzowanych kilkudziesięcioma zmiennymi. Jedną z konsekwencji analizy takich zbiorów danych jest to, że ustalenie hierarchii obiektów staje się trudne do przeprowadzenia a uzyskana struktura jest często tak nieczytelna, że niemożliwa do praktycznego zastosowania. Jednocześnie badane zbiory danych charakteryzują się dużą liczbą skupień przekraczającą czasami liczbę 50. Obie powyższe konsekwencje ilościowych zmian w badanych zbiorach danych powodują, że konsumenci wyników badań statystycznych są dziś zainteresowani nie tylko strukturą hierarchiczną obiektów i wyodrębnieniem skupień, ale także strukturą hierarchiczną samych skupień.

Celem prezentowanych badań jest prezentacja i opis własności dwustopniowych metod grupowania obiektów w rozpoznawaniu hierarchicznej struktury skupień. Do wyodrębnienia skupień zostaną zastosowane metody optymalizacyjne: k-średnich, w-średnich k-medoids i sieć neuronowa typu SOM. Hierarchiczna struktura skupień wyodrębnionych przy pomocy powyższych metod będzie następnie analizowana klasycznymi metodami aglomeracyjnymi.

Badania zostaną oparte na danych symulacyjnych, co ułatwi ocenę własności zastosowanych metodologii i pozwoli na wyprowadzenie ogólniejszych wniosków dotyczących wziętych pod uwagę klas problemów klasyfikacyjnych.

Wnioski z badań powinny pomóc w lepszym zrozumieniu informacji zawartych w wielowymiarowych zbiorach danych i w skuteczniejszym stosowaniu tej wiedzy w praktyce.

Literatura:

1. Deboeck G., Kohonen T. (1998). Visual Explorations in Finance, with Self-Organizing Maps. Springer. Finance.
2. Hand D.J. (1981). Discrimination and Classification, John Wiley & Sons Ltd.
3. Kohonen T. (1997). Self-Organizing Maps. Springer – Verlag Berlin-Heidelberg
4. Lampinen J., Oja E. (1992). Clustering properties of hierarchical self-organizing maps. J. Math Imag. Vis. Vol. 2, No. 2-3.
5. Migdał Najman K., Najman K. (2001). Zastosowanie sieci neuro-

nowej typu SOM do wyboru najatrakcyjniejszych spółek na WGPW. Prace Naukowe Akademii Ekonomicznej we Wrocławiu nr 952.

6. Migdał Najman K., Najman K. (2003). Próba zastosowania sieci neuronowej typu SOM w badaniu przestrzennego zróżnicowania powiatów w Polsce. Wiadomości Statystyczne 4/03.
7. Mitra S., Acharya T. (2003). Data mining Multimedia, Soft Computing and Bioinformatics. A John Wiley & Sons, Inc. Publication.
8. Rencher A.C. (2002). Methods of Multivariate Analysis. Wiley-Interscience. A John Wiley & Sons, Inc. Publication.
9. Späth H. (1982). Cluster Analysis Algorithms for data reduction and classification of objects. John Wiley & Sons, Inc.
10. Vesanto J. (1997). Data mining techniques based on the Self-Organizing Maps. Helsinki University of Technology. Department of Engineering Physics and Mathematics. ESPOO. Finland.
11. Vesanto J., Alhoniemi E. (2000). Clustering of the Self-Organizing Map. IEEE Transactions on Neural Networks. Vol. 11, No. 3.

16:30

Oral

### **Wykorzystanie drzew klasyfikacyjnych w analizie klas ukrytych**

Iwona Kasprzyk

Akademia Ekonomiczna im. K. Adamieckiego w Katowicach, Katowice 40-287, Poland

e-mail: [i\\_kasprzyk@interia.pl](mailto:i_kasprzyk@interia.pl)

W badaniach ekonomicznych często spotyka się dane mierzone na skali nominalnej oraz porządkowej. Jedną z metod pozwalających na analizę tego typu danych jest analiza klas ukrytych (ang. *latent lass model*), która została zaproponowana przez Lazarsfelda (1968).

Drzewa są nieparametryczną metodą, które zostały wprowadzone przez Breimana i in. (1984). W zależności od tego czy zmienna zależna jest zmienną metryczną czy niemetryczną będziemy rozważać w pierwszym przypadku *drzewa regresyjne*, natomiast w drugim - *drzewa klasyfikacyjne*.

Celem referatu będzie zaprezentowanie, w jaki sposób drzewa klasyfikacyjne mogą usprawnić analizę klas ukrytych. Pokazane zostanie, jak tego typu drzewa można wykorzystywać do ograniczenia liczby zmiennych wprowadzanych do analizy klas ukrytych. W końcowym etapie referatu zostanie również pokazana graficzna prezentacja danych za pomocą tzw. wykresu współrzędnych barycentrycznych, który to pozwoli na przedstawienie zależności między kategoriami zmiennych zawartych w tablicy kontyngencji.

Literatura:

- Breiman L., Friedman J.H., Olshen R.A., Stone C.J. (1984), *Classification and Regression Tree*, Wadsworth, Belmont, CA.
- Heijden, P.G.M. van der, Gilula, Z., & Ark, L.A. van der (1999), An extended study into the relationship between correspondence analysis and latent class analysis. *Sociological Methodology*, nr 29, s.147-186.
- Lazarsfeld, P.F., Henry N.W.(1968), *Latent structure analysis*, Boston: Houghton Mill.
- Vermunt, J.K (1997), *Log-linear models for event histories*. Thousand Oaks: Sage

16:45

Oral

### Identyfikacja preferencji z wykorzystaniem modeli PREFMAP

Artur P. Zaborski

Akademia Ekonomiczna (AE), Nowowiejska3, Jelenia Góra  
58-500, Poland

e-mail: [artur.zaborski@ae.jgora.pl](mailto:artur.zaborski@ae.jgora.pl)

Znajomość postaw i preferencji jest niezbędna przy badaniu zachowań konsumentów na rynku oraz przy podejmowania decyzji rynkowych. Konsument postrzega każdy produkt jako zespół cech (punkt w przestrzeni wielowymiarowej) o różnych zdolnościach dostarczenia pożądanego korzyści i zaspokajania potrzeb. W analizie zmierzającej do określenia korzyści jakich oczekują konsumenci, istotnym zadaniem jest więc określenie atrybutów cenionych przez konsumentów, ocena znaczenia poszczególnych atrybutów, a także ocena poszczególnych atrybutów marki w porównaniu do marek konkurencyjnych.

Wśród metod, których wyniki są prezentowane w formie graficznej, w analizie preferencji szczególne znaczenie ma grupa modeli map preferencji PREFMAP (*PREFerence MAPping*). Idea konstrukcji map preferencji opiera się na analizie *unfolding* i obejmuje model wektorowy oraz modele punktu idealnego. Dzięki nim na jednej mapie percepcyjnej możliwe jest przedstawienie konfiguracji punktów reprezentujących zarówno respondentów jak i badane obiekty lub obiekty oraz opisujące je cechy. Modele te, oprócz ukazania preferencji konsumentów wobec produktów, pozwalają na wykrycie związku między zmiennymi a badanymi obiektami, określenie preferowanego przez respondentów zbioru cech, które powinny posiadać obiekty, a w przypadku modeli z wagami, również na wskazanie indywidualnych różnic w ocenach preferencji poszczególnych respondentów.

Celem referatu jest przedstawienie podstawowych cech modeli PREFMAP oraz możliwości ich wykorzystania na gruncie badań marketingowych. Szczególną uwagę zwrócono na zasady interpretacji wyników oraz problemy, które mogą występować przy interpretacji wyników dla modeli z wagami.

Zakres tematyczny: Analiza danych marketingowych

#### Literatura

1. Borg I., Groenen P., *Modern multidimensional scaling. Theory and applications. Second Edition*, Springer-Verlag, New York, 2005.
2. Coxon A.P.M., *The Users Guide to Multidimensional Scaling*, Heinemann Educational Books, London, 1982.
3. Zaborski A., *Metryczna i niemetryczna wielowymiarowa analiza unfolding*. W: *Ekonometria 11*, Prace Naukowe Akademii Ekonomicznej we Wrocławiu nr 981, 2003, s. 50-58.

17:00

Oral

### Symulacyjna analiza wpływu wyboru metody podziału obiektów na jakość uzyskanej klasyfikacji w algorytmach k-średnich

Krzysztof Najman

Uniwersytet Gdański Wydział Zarządzania (WZR), Armii Krajowej 101, Sopot 81-824, Poland

e-mail: [knajman@wzr.pl](mailto:knajman@wzr.pl)

Zasadniczym celem analizy skupień jest taki podział obiektów, który zapewnia maksymalną homogeniczność obiektów w wyróżnionych skupieniach i maksymalną heterogeniczność między skupieniami. Metody realizujące tą strategię nazywamy ogólnie metodami podziałowymi. Jedną z klasycznych i najczęściej stosowanych metod podziałowych jest metoda k-średnich. Jakość klasyfikacji obiektów uzyskana dzięki metodzie k-średnich zależy od wielu czynników. Do ważniejszych z nich należy zaliczyć: 1) wstępny (prototypowy) podział obiektów na skupienia, 2) kolejność obiektów w zbiorze danych, 3) metodę przenoszenia obiektów między skupieniami a także 4) przyjęte kryterium optymalności podziału obiektów. Każdy z wymienionych czynników wskazuje na problem z podziałem zbioru obiektów i w ramach każdego z nich znane jest kilka rozwiązań tego problemu. Z tego powodu nie istnieje jedna metoda k-średnich. Należy ta wskazuje raczej na pewien sposób postępowania prowadzący do podziału obiektów niż konkretną jego realizację. Niestety w praktyce przyjęło się rozumienie tej metody jedynie dla algorytmu MacQueen'a, który jest przecież tylko jedną z możliwych realizacji metody k-średnich. Wyrazem tego stanu jest fakt, że w kluczowych pakietach statystycznych znajduje się jedynie ten algorytm, co wydaje się zbyt dużym uproszczeniem.

Celem prezentowanych badań jest przegląd i krytyczna ocena różnych implementacji algorytmu k-średnich, uwzględniających powyższe cztery kluczowe czynniki wpływające na jakość uzyskanego grupowania. Analiza będzie oparta na danych symulacyjnych i empirycznych o znanej strukturze grupowej. Uzyskane struktury grupowe będą oceniane wskaźnikami jakości grupowania: *Silhouette*, *Davies'a-Bouldin'a*, *Calinski'ego-Harabasz* i *Hubert'a-Levin'a* a także porównywane ze wzorcem w oparciu o indeksy *Rand'a*, *Jaccard'a* i *Fowlkes'a-Mallows'a*. Uzyskane wyniki powinny pomóc w lepszym zrozumieniu i właściwym stosowaniu algorytmów k-średnich.

Literatura:

1. Fisher D., Xu L., Zard N. (1992). Ordering effects in clustering. Proceedings of the Ninth International Conference on Machine Learning. Morgan Kaufman, San Mateo, USA, s. 163-168
2. Gordon A.D. (1999). Classification. 2<sup>nd</sup> Edition, Chapman & Hall, London
3. MacQueen J. (1967). Some methods for classification and analysis of multivariate observations, in: L.M. LeCam, J. Neyman (Eds.), Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, vol. 1-Statistics, UCLA Press, Berkeley, USA, s. 281-297
4. Peña J.M., Lozano J.A., Larrañaga P. (1999). An empirical comparison of four initialization methods for the k-means algorithm.

- Pattern Recognition Lett. 20, s. 1027-1040
- Selim S.Z., Ismail M.A. (1984). K-means type algorithm: generalized convergence theorem and characterization of local optimality, IEEE Trans. Pattern Anal. Mach. Intell. 6, s. 81-87
  - Tarsitano A. (2003). A computational study of several relocation methods for k-means algorithms. Pattern Recognition 36, s. 2955-2966

**Zakończenie obrad w sesji A2**

Środa popołudniu, 19 Wrzesień, 17:15

**Czwartek, 20 Wrzesień****Sesja A3***Metody Klasyfikacji i Analizy Danych III*

Czwartek rano, 20 Wrzesień, 10:15

10:15

Oral

**Wykorzystanie pewnego kryterium przy budowie drzewa klasyfikacyjnego dla zbioru obiektów symbolicznych**Małgorzata Gliwa*Akademia Ekonomiczna im. K. Adamieckiego w Katowicach, Katowice 40-287, Poland**e-mail: wojcikm8@wp.pl*

Proces rekurencyjnego podziału zbioru obiektów symbolicznych na podzbiory aż do uzyskania ich jednorodności ze względu na przynależność do  $K$  klas to tworzenie drzewa klasyfikacyjnego. Efektywność tworzenia drzew zależy jednak od sposobu podziału zbioru obiektów symbolicznych w węzłach drzewa. W każdym kroku algorytmu należy wskazać tę cechę, która daje dwa najbardziej różniące się podzbiory. M. Vrac i E. Diday zaproponowali nowe kryterium, w którym postulują, aby wyboru tego dokonywać optymalizując równocześnie kryterium homogeniczności ( $H(W)$ ,  $W$ - węzeł drzewa) i kryterium zróżnicowania ( $D(W)$ ). Takie podejście wykorzystuje jednocześnie uczenie z nauczycielem i bez nauczyciela. Otrzymane w wyniku podziału klasy obiektów symbolicznych można opisywać, tj. wskazywać cechy charakterystyczne dla danej klasy.

Kontynuując rozważania z poprzednich Konferencji SKAD dotyczące analizy danych symbolicznych, w referacie przedstawione zostanie kryterium pozwalające na optymalny wybór cechy w oparciu, o którą nastąpi podział zbioru obiektów symbolicznych. Polega ona maksymalizacji funkcji  $\Delta Q(W)$ , gdzie:

$$\max \Delta Q(W) = \max [\alpha H(W) + \beta D(W)], \alpha + \beta = 1.$$

Funkcja  $D(W)$  najczęściej wyrażana jest przez wskaźnik zróżnicowania Ginięgo, natomiast omówione zostaną różne postaci funkcji  $H(W)$  w zależności od rodzaju zmiennych, którymi opisany jest obiekt symboliczny. Jako podsumowanie zaprezentowany zostanie przykład obliczeniowy.

**Literatura:**

- Bock H.H., Diday E., *Analysis of Symbolic Data*, Springer-Verlag, Berlin 2000.

- Breiman L., Friedman J.H., Olshen R. A., Stone C. J., *Classification and regression trees*, Wadsworth, Belmont 1984.
- Limam M. M., Diday E., Winsberg S., *Symbolic Class Description with Interval Data*, Journal of Symbolic Data Analysis, t.1, 2004.
- Quinlan J. R., *Induction of Decision Trees*, Machine Learning, t. 1, 1986.
- Vrac M., Diday E., Winsberg S., Limam M. M., *Symbolic class description*, [w:] Jajuga K. (red.), *Classification, Clustering and Data Analysis*, Springer-Verlag, Heidelberg 2002.

10:30

Oral

**Wykorzystanie podejścia zagregowanego w taksonomii**Dorota E. Rozmus*Akademia Ekonomiczna im. K. Adamieckiego w Katowicach, Katowice 40-287, Poland**e-mail: drozmus@ae.katowice.pl*

Podejście wielomodelowe (agregacja modeli) jest z powodzeniem stosowane w dyskryminacji celem poprawy dokładności predykcji. Ostatnie badania pokazały jednak, że agregacja może być także wykorzystana w taksonomii celem uzyskania poprawniejszej i bardziej stabilnej klasyfikacji obiektów (Fred, Jain, 2005; Kuncheva et al., 2006). Co więcej, podejście takie zmniejsza ryzyko uzyskania błędnej klasyfikacji poprzez niewłaściwy wybór algorytmu. Ma to o tyle znaczenie, że w analizie dyskryminacyjnej poprzez pomiar błędu klasyfikacji istnieje możliwość oceny wybranej metody, natomiast możliwości takiej już nie ma w taksonomii.

Celem referatu jest pokazanie możliwości zastosowania jednej z najpopularniejszych metod agregacji, jaką jest metoda losowania bootstrapowego (ang. *bagging*, Breiman, 1996) w taksonomii. Pokazane zostaną także wyniki badań empirycznych mających na celu porównanie zgodności wyników klasyfikacji z istniejącą strukturą klas na podstawie klasycznych i zagregowanych algorytmów taksonomicznych. Obliczenia wykonane zostaną z zastosowaniem programu R.

**Literatura:**

- Breiman L., Bagging predictors, *Machine Learning*, **26**(2), 123-140, 1996.
- Fred N. L., Jain A. K., Combining multiple clusterings using evidence accumulation, *IEEE Transactions on PAMI*, **27**(6): 835-850, 2005.
- Kuncheva L. I., Hadjitodorov S. T., Todorova L. P., Experimental comparison of cluster ensemble methods, *Proc FUSION 2006*, Florence, Italy, 2006.

10:45

Oral

**Wykorzystanie rozmytej metody TOPSIS opartej na zbiorze alpha przekrojów do porządkowania liniowego obiektów**Aleksandra Łuczak, Feliks Wysocki*Akademia Rolnicza (AR), ul. Wojska Polskiego 28, Poznań 60-637, Poland**e-mail: luczak@au.poznan.pl*

TOPSIS (*Technique for Order Preference by Similarity to an Ideal Solution*) jest metodą tworzenia cechy syntetycznej. Jej ogólna idea polega na wyznaczeniu odległości każdego obiektu wielocechowego od wzorca i antywzorca rozwoju oraz zagregowaniu ocen w każdym ocenianym obiekcie.

Celem pracy jest przedstawienie możliwości wykorzystania wybranych metod TOPSIS do porządkowania liniowego obiektów dla różnych typów danych: rzeczywistych, przedziałowych i rozmytych. Rozważane są metody TOPSIS:

- klasyczna (Hwang, Yoon 1981, Chen, Hwang 1992), w przypadku której dane wejściowe są dokładnie znanymi wartościami rzeczywistymi cech dla badanych obiektów,

- interwałowa (Jahanshahloo, Lotfi, Izadikhah 2006 a) – wartości cech opisujących obiekty są liczbami przedziałowymi, dla których początek przedziału określa minimalną, a koniec – maksymalną wartość cechy,

- rozmyta (Jahanshahloo, Lotfi, Izadikhah 2006 b, Kahraman, Büyüközkan, Ates 2007) – wartości cech nie są precyzyjnie określone lub są wyrażone za pomocą poziomów zmiennej lingwistycznej, którym odpowiadają trójkątne liczby rozmyte, reprezentowane przez trzy oceny: pesymistyczną, najbardziej prawdopodobną i optymistyczną.

Proponowane podejścia zostały zilustrowane przykładem dotyczącym analizy przestrzennego zróżnicowania poziomu życia ludności województwa wielkopolskiego według powiatów.

Literatura:

1. Chen S.J., Hwang C.L. (1992): *Fuzzy Multiple Attribute Decision Making: Methods and Applications*, Springer-Verlag, Berlin, 1992.
2. Hwang C.L., Yoon K. (1981): *Multiple Attribute Decision Making: Methods and Applications*, Springer-Verlag, New York.
3. Jahanshahloo, G.R., Lotfi, F. H., Izadikhah, M. (2006 a): An algorithmic method to extend TOPSIS for decision-making problems with interval data, *Applied Mathematics and Computation* 175 (2), s. 1375-1384.
4. Jahanshahloo, G.R., Lotfi, F. H., Izadikhah, M. (2006 b): Extension of the TOPSIS method for decision-making problems with fuzzy data. *Applied Mathematics and Computation Volume*. 181 (2), s. 1544-1551.
5. Kahraman C., Büyüközkan G.; Ates N. Y. (2007): A two phase multi-attribute decision-making approach for new product introduction. *Information Sciences* 177 (7) s. 1567-1582.

11:00

Oral

### Ocena zasadności łączenia wybranych nieparametrycznych modeli regresji

Joanna Trzęsiok

*Akademia Ekonomiczna im. K. Adamieckiego w Katowicach, Katowice 40-287, Poland*

*e-mail: trzesiok@ae.katowice.pl*

W rozwoju nieparametrycznych metod regresji zauważalne jest coraz częstsze wykorzystywanie podejścia wielomodelowego. W celu

uzyskania większej dokładności predykcji, do prognozowania stosowane są wskazania wielu modeli skonstruowanych na losowo dobranych podpróbach, których wyniki predykcji łączone są według określonej formuły – najczęściej przez uśrednianie.

Jak wykazują badania symulacyjne i wyniki teoretyczne, predykcja modelu zagregowanego jest na ogół dokładniejsza niż każdego z pojedynczych modeli składowych. Nie jest tak jednak zawsze. Warunkiem skuteczności zastosowania podejścia wielomodelowego jest to, by wyniki predykcji modeli składowych charakteryzowały się relatywnie silnym zróżnicowaniem.

W referacie przedstawione zostaną wyniki analizy zróżnicowania wybranych nieparametrycznych modeli regresji, co będzie pierwszym etapem w zweryfikowaniu hipotezy o zasadności agregacji tych modeli. Zaprezentowane zostanie również porównanie wyników predykcji pojedynczych oraz zagregowanych modeli.

Analiza zostanie przeprowadzona przy pomocy pakietu statystycznego R, na zbiorach standardowo wykorzystywanych do badania własności różnych metod wielowymiarowej analizy statystycznej.

**Zakres tematyczny:** metody statystycznej analizy wielowymiarowej.

**Literatura:**

1. Hastie T., Tibshirani R., Friedman J. H., *The Elements of Statistical Learning*, Springer – Verlag New York, 2001.
2. Joanna Trzęsiok, *Metoda rzutowania w budowie modelu regresyjnego, „Postępy ekonometrii”* – praca zbiorowa pod redakcją A. S. Barczaka, Wydawnictwo Akademii Ekonomicznej w Katowicach, 2004, s. 121 – 130.
3. Joanna Trzęsiok, *Adaptacyjne metody jądrowe w regresji*, W: Jajuga K., Walesiak M. (red.) „Klasyfikacja i analiza danych – teoria i zastosowania, Taksonomia 12”, Prace Naukowe Akademii Ekonomicznej we Wrocławiu (nr 1076), Wrocław 2005, s. 493 – 500.

11:15

Oral

### Walidacja wyników klasyfikacji otrzymanych taksonomiczną metodą wektorów nośnych

Michał Trzęsiok

*Akademia Ekonomiczna im. K. Adamieckiego w Katowicach, Katowice 40-287, Poland*

*e-mail: trzesiok@ae.katowice.pl*

W roku 2001 Ben-Hur, Horn, Siegelmann i Vapnik zaproponowali nową metodę taksonomiczną – SVC (*ang. Support Vector Clustering*) – wykorzystującą podejście analogiczne jak w dyskryminacyjnej metodzie wektorów nośnych (SVM). Metoda SVC jest metodą nieparametryczną, tzn. nie wymaga nakładania a priori założeń dotyczących liczby i kształtu skupień. Jednocześnie jest metodą bardzo elastyczną – pozwala na generowanie skupień o nieliniowych, bardzo nieregularnych kształtach.

Zdolność metody SVC do tak elastycznego dopasowywania się do analizowanego zbioru danych niesie ze sobą ryzyko, że metoda nie będzie miała własności stabilności klas, rozumianej jako generowanie podobnej struktury skupień dla różnych, losowo wybieranych próbek ze zbioru danych.

Celem referatu jest przedstawienie wyników analizy stabilności i zgodności klasyfikacji otrzymanych metodą SVC oraz innymi metodami taksonomicznymi. Do pomiaru zgodności wykorzystano indeks Randa. Obliczenia zrealizowano w pakiecie statystycznym R na zbiorach danych standardowo wykorzystywanych do badania własności metod wielowymiarowej analizy statystycznej.

**Zakres tematyczny:** taksonomia.

#### Literatura

Ben-Hur A., Horn D., Siegelman H. T., Vapnik V. (2001), *Support Vector Clustering*, Journal of Machine Learning Research, 2, str. 125–137.

Gatnar E., Walesiak M. [red.] (2004), *Metody statystycznej analizy wielowymiarowej w badaniach marketingowych*, Wydawnictwo Akademii Ekonomicznej we Wrocławiu, Wrocław, str. 316–350.

Schölkopf B., Smola A.J. (2002), *Learning with Kernels. Support Vector Machines, Regularization, Optimization, and Beyond*, MIT Press, Cambridge.

#### Przerwa

Czwartek rano, 20 Wrzesień, 11:30

#### Zakończenie obrad w sesji A3

Czwartek popołudniu, 20 Wrzesień, 12:30

## Piątek, 21 Wrzesień

#### Sesja A5

*Metody Klasyfikacji i Analizy Danych IV*  
Piątek rano, 21 Wrzesień, 9:00

9:00 Oral

#### Klasyfikacja dynamiczno-przestrzenna europejskiej przestrzeni regionalnej ze względu na poziom gospodarki opartej na wiedzy

Danuta Strahl, [Małgorzata Markowska](mailto:Malgorzata.Markowska@ae.jgora.pl)

*Akademia Ekonomiczna (AE), Nowowiejska3, Jelenia Góra 58-500, Poland*

*e-mail: mmarkowska@ae.jgora.pl*

Referat przedstawia klasyfikację dynamiczno-przestrzenną europejskiej przestrzeni regionalnej szczebla NUTS 2, ze względu na poziom gospodarki opartej na wiedzy (GOW). Ocena GOW obejmuje dwa sektory gospodarki: przemysł oraz usługi. Badania obejmują analizę następujących zmiennych ilustrujących GOW na szczeblu regionalnym w ujęciu dynamicznym, tj. w latach 2000 - 2005:

$X_1$  – udział pracujących z wyższym wykształceniem w ogólnej liczbie ludności w wieku 25-64 lat,

$X_2$  – udział ludności uczestniczącej w ustawicznym kształceniu w ogólnej liczbie ludności w wieku 25- 64 lat,

$X_3$  – liczba patentów przypadających na 1 mln siły roboczej,

$X_4$  – udział pracujących w usługach „opartych na wiedzy” (*knowledge-intensive services*),

$X_5$  – udział pracujących w przemyśle wysoko i średnio zaawansowanym technologicznie.

Do oceny stopnia rozwoju GOW na szczeblu regionalnym zostaną wykorzystane metody klasyfikacji obiektów hierarchicznych z uwzględnieniem dwóch aspektów zróżnicowania wewnętrznego:

- przestrzennego, z uwagi na obiekty niższego rzędu oraz

- dynamiki zmian w poziomie GOW.

Klasyfikacja dynamiczno-przestrzenna regionów państw UE szczebla NUTS 2 pozwoli na identyfikację czterech klas regionów:

- pierwszej – obejmującej regiony, które znalazły się w grupie o wysokim poziomie GOW tak przemysłu jak i w usługach, w każdym z badanych momentów czasowych,

- klasy drugiej wskazującej regiony, które były kreatorami GOW w przemyśle w każdym z rozpatrywanych momentów czasowych,

- trzecia klasa gromadzi regiony, które posiadały wysokie charakterystyki GOW w usługach w zadanych momentach czasowych,

- ostatnia klasa obejmuje regiony, w których nie odnotowano znaczących poziomów charakterystyk GOW ani w przemyśle ani w usługach w zadanych latach.

Wyniki klasyfikacji posłużą do budowy uogólnionej miary GOW w ujęciu dynamiczno-przestrzennym. Klasyfikacja ta pozwoli na pomiar GOW skumulowanej w przemyśle oraz w usługach na szczeblu obiektu hierarchicznego czyli krajowym, z uwzględnieniem dynamiki zmian w zakresie przyjętych zmiennych ilustrujących GOW w regionach.

9:15 Oral

#### SymScal: Metoda skalowania wielowymiarowego obiektów symbolicznych

[Marcin Pelka](mailto:Marcin.Pelka@ae.jgora.pl), Andrzej Dudek

*Akademia Ekonomiczna (AEWROCLAW), Komandorska 118/120, Wrocław 53-345, Poland*

*e-mail: marcin.pelka@ae.jgora.pl*

Celem artykułu jest zaprezentowanie metod skalowania wielowymiarowego, które można zastosować dla obiektów symbolicznych. Pierwszą prezentowaną metodą skalowania wielowymiarowego dla obiektów symbolicznych jest SymScal, którą zaproponowali Groenen, Winsberg, Rodríguez i Diday [2005]. Drugą z metod jest INTERSCAL, którą zaproponowali Deneux i Masson [2000]. Trzecią z metod jest I-SCAL opracowany przez zespół pod kierunkiem Groenena [2006]. Metoda ta jest rozwinięciem metody SYMSCAL. Wszystkie te metody skalowania wielowymiarowego dla obiektów symbolicznych (*symbolic multidimensional scaling*) wymagają, aby danymi wejściowymi była macierz minimalnych i maksymalnych odległości pomiędzy obiektami symbolicznymi. Celem skalowania wielowymiarowego dla obiektów symbolicznych jest przedstawienie relacji pomiędzy obiektami traktowanymi jako hiperprostokądnymi w przestrzeni wielowymiarowej. Artykuł w części empirycznej prezentuje wyniki skalowania wielowymiarowego uzyskane przy zastosowaniu metody SYMSCAL na przykładowych danych symbo-

licznych. Do obliczeń wykorzystano autorski plik źródłowy realizujący procedurę skalowania zgodnie z metodą SYMSCAL opracowaną dla pakietu statystycznego R.

**Zakres tematyczny: metody analizy danych symbolicznych – teoria**

#### Bibliografia

1. Bock H.-H., Diday E. (red.) (2000), *Analysis of symbolic data. Explanatory methods for extracting statistical information from complex data*, Springer Verlag, Berlin-Heidelberg.
2. Denoeux T., Masson M. (2000), *Multidimensional scaling of interval-valued dissimilarity data*, "Pattern Recognition Letters", vol. 21, issue 1, 83-92.
3. Groenen P. J. F., Winsberg S., Rodríguez O., Diday E. (2005), *SymScal: Symbolic Multidimensional Scaling of Interval Dissimilarities*, Econometric Report EI 2005-15, Erasmus University, Rotterdam.
4. Groenen P. J. F., Winsberg S., Rodríguez O., Diday E. (2006), *I-Scal: Multidimensional scaling of interval dissimilarities*, „Computational Statistics & Data Analysis” vol. 51, issue 1, 360-378.

9:30

Oral

### Badanie efektywności zmodyfikowanych metod aglomeracyjnych na zbiorach danych ze świata realnego

Jerzy M. Korzeniewski

Uniwersytet Łódzki, Wydział Ekonomiczno-Socjologiczny, P.O.W. 3/5, Łódź 90-255, Poland

e-mail: jurkor@wp.pl

W referacie kontynuowana jest tematyka badania efektywności aglomeracyjnych metod klasyfikacyjnych ze szczególnym uwzględnieniem metod zmodyfikowanych przez Autora. Rozważane są następujące metody : metoda pojedynczego łączenia, metoda kompletnego łączenia, metoda średniej klasowej, metoda ważonej średniej klasowej, metoda średniej odległości, metoda wewnątrzklasowej sumy kwadratów odległości, metoda powiększonej sumy kwadratów odległości, metoda środka ciężkości oraz metoda medianowa. Modyfikacje tych metod oparte są na połączeniu idei aglomeracji opartej na macierzy odległości z wykorzystaniem lokalnej gęstości punktów zbioru danych. Analizowane zbiory danych pochodzą z różnych dziedzin życia . Cechą wspólną ocenianych jakości klasyfikacji jest zarówno odniesienie otrzymanych klasyfikacji zbioru do klasyfikacji istniejących jak i ocenianie jakości klasyfikacji przy pomocy indeksów sylwetkowych oraz indeksów RAND. Przeprowadzone badania rzucają nowe światło na ocenę modyfikacji Autora dotychczas ocenianych na sztucznie generowanych zbiorach danych z przestrzeni euklidesowych z wykorzystaniem indeksów sylwetkowych do oceny jakości klasyfikacji.

9:45

Oral

### Zastosowanie metody boosting w indukcji reguł

Mariusz Kubus

Politechnika Opolska (PO), Mikołajczyka 5, Opole 45-271, Poland

e-mail: marikubu@wp.pl

Ze względu na możliwość analizy danych niemetrycznych, dużym powodzeniem wśród metod dyskryminacji cieszą się algorytmy indukcji reguł. Klasyczny już schemat indukcji reguł separacji-zwycięzaj (*separate-and-conquer*) został zaproponowany po raz pierwszy przez Michalskiego [1969]. Za pomocą przeszukiwania heurystycznego przestrzeni opisów klas znajduje się reguła maksymalizująca funkcję kryterium. Opisane przez nią obiekty usuwane są ze zbioru uczącego, po czym generowana jest następna reguła. Cohen i Singer [1999], w algorytmie SLIPPER, zaproponowali nieco inne podejście. Do indukcji reguł zaadoptowali metodę *boosting* [Schapire i Singer 1998]. Zamiast usuwać ze zbioru uczącego opisane przez regułę obiekty, nadawane są im wagi, co wpływa na generowanie następnej reguły.

Celem referatu jest zbadanie algorytmu SLIPPER. Zbadane będą: wpływ liczby iteracji na dokładność klasyfikacji oraz różne sposoby rozwiązania problemu wielu klas. Uzyskane wyniki zostaną porównane z dobrze znanym algorytmem drzew klasyfikacyjnych CART.

#### Literatura

Cohen W.W.(1995), *Fast effective rule induction*. In Prieditis A., Russell S. (Eds.), Proceedings of the 12th International Conference on Machine Learning

Cohen W.W., Singer Y. (1999), *A Simple, Fast, and Effective Rule Learner*. In Proceedings of Annual Conference of American Association for Artificial Intelligence (pp.335-342).

Michalski R.S. (1969), *On the quasi-minimal solution of the covering problem*. In Proceedings of the 5<sup>th</sup> International Symposium on Information Processing (FCIP-69), Vol. A3 (Switching Circuits), p.125-128 Bled, Yugoslavia.

Schapire R.E., Singer Y. (1998), *Improved boosting algorithms using confidence-rated predictions*. In Proceedings of the Eleventh Annual Conference on Computational Learning Theory, 80-91.

### Zakończenie obrad w sesji A

Piątek rano, 21 Wrzesień, 10:00



# Sesja B

Zastosowania w finansach, zastosowania międzynarodowe i zastosowania w analizach przestrzennych

## Program

Środa, 19 Wrzesień

### Sesja B1

Zastosowanie w Finansach I

Środa popołudniu, 19 Wrzesień, 15:00

15:00

Oral

### Modelowanie rozkładu strat w pomiarze ryzyka operacyjnego

Katarzyna Kuziak

Akademia Ekonomiczna (AEWROCLAW), Komandorska 118/120, Wrocław 53-345, Poland

e-mail: [katarzyna.kuziak@ae.wroc.pl](mailto:katarzyna.kuziak@ae.wroc.pl)

W pomiarze ryzyka operacyjnego wykorzystywane jest m.in. podejście modelowania stochastycznego, w którym estymacja parametrów rozkładów częstości i rozkładów wielkości (skali zdarzenia) następuje za pomocą estymacji empirycznej (dopasowanie empiryczne) oraz estymacji parametrycznej (Loss Distribution Approach, LDA). W referacie przedstawione zostanie podejście LDA. W przypadku LDA zakładamy typ rozkładu dla pojedynczego zdarzenia. Dla rozkładu częstości może to być np. rozkład Poissona natomiast w przypadku skali zdarzenia np. rozkład normalny. Do estymacji parametrów rozkładu, wykorzystywane są głównie metoda momentów i funkcja największej wiarygodności. Następnie ma miejsce agregacja strat wynikających z pojedynczych zdarzeń, rozpatrywanych dotąd osobno. Agregacja strat i analizowanie rozkładu zagregowanego wykorzystuje funkcje połączeń (w przypadku ryzyka operacyjnego może być: t-student copula, Gumbel copula). Parametryzacja funkcji powiązań jest przeprowadzana na podstawie danych historycznych lub analizy scenariuszy za pomocą symulacji. Finalnie analizowany jest cały rozkład strat (tj. zagregowany, wszystkich zdarzeń). Opis tego podejścia zilustrowany zostanie przykładem.

Cruz M. (2002): *Modeling, Measuring and Hedging Operational Risk*, John Wiley & Sons, New York.

Ebnoeter S., Vanini P., McNeil A., Antolinez P.: *Operational risk: A practitioner's view*, artykuł dostępny na: [www.gloriamundi.org](http://www.gloriamundi.org), 2002.

Embrechts P., Lindskog F., McNeil A. (2001): *Modelling dependence with copulas and applications to risk management*, raport, ETHZ Zurich.

Frachot A., Georges P., Roncalli T.: *Loss Distribution Approach for operational risk*, Artykuł dostępny na [gro.creditlyonnais.fr](http://gro.creditlyonnais.fr), 2001.

Powojowski M.R., Reynolds D., Tuenter H.J.H: *Dependent events and operational risk*. Artykuł dostępny na: [www.gloriamundi.org](http://www.gloriamundi.org), 2002.

15:15

Oral

### Analiza wrażliwości rozwiązań wyznaczonych dla różnej struktury próby kredytobiorców: analiza dyskryminacyjna i drzewa klasyfikacyjne.

Mariola Chrzanowska<sup>1</sup>, Dorota Witkowska<sup>2</sup>

1. Wyższa Szkoła Ekonomii i Administracji, Jagiellońska 109A, Kielce 25-734, Poland 2. Szkoła Główna Gospodarstwa Wiejskiego (SGGW), Nowoursynowska 166, Warszawa 02-787, Poland

e-mail: [mariola.chrzanowska@wseia.edu.pl](mailto:mariola.chrzanowska@wseia.edu.pl)

Dynamiczny wzrost zapotrzebowania na usługi bankowe, wzrost obrotów oraz rosnące wymagania klientów, co do jakości usług bankowych, wymagają doskonalenia narzędzi pracy w rozumieniu automatyzacji procedur i zwiększania wiarygodności ocen potencjalnych kredytobiorców. Jednym ze sposobów usprawnienia pracy banku jest wykorzystanie statystycznych metod klasyfikacji. Jednakże wybór odpowiednich metod musi uwzględniać specyfikę danych opisujących klientów banku. Innymi słowy metody wspomagające decyzje inspektorów kredytowych powinny być odporne na powszechnie występujące cechy zbioru informacji, do których należą (\*):

1. jakościowy charakter danych dotyczących indywidualnych kredytobiorców;
2. występowanie obserwacji nietypowych (w sensie statystycznym);
3. brak symetryczności klas, np. liczba kredytobiorców spłacających terminowo jest inna (zazwyczaj znacznie większa) niż liczba kredytobiorców zalegających ze spłatą kredytu.

Celem pracy jest badanie wpływu struktury próby na jakość klasyfikacji indywidualnych kredytobiorców, uzyskanej za pomocą analizy dyskryminacyjnej oraz drzew klasyfikacyjnych. Metody te zostały wybrane do porównań, ponieważ pierwsza z nich została zastosowana w praktyce jako pierwsza (Altman, 1968), natomiast druga z metod klasyfikacji wydaje się szczególnie godna polecenia w przypadku opisu obiektów za pomocą cech jakościowych.

Badaniem objęto 2576 osób, które zaciągnęły pożyczkę hipoteczną w okresie od kwietnia 2004 do lipca 2005. Spośród nich 419 kredytobiorców nie spłaca kredytu (termin spłaty kredytu dla większości klientów jeszcze nie upłynął). Po usunięciu obserwacji nietypowych (\*\*\*) z bazy wylosowano trzy próby:

- próbę symetryczną (tj. połowa kredytów jest spłacana terminowo),
- próbę, w której 16% klientów nie spłaca kredytu w terminie, bowiem taki procent osób nie spłaciło kredytu w analizowanej bazie danych;
- próbę zawierającą 9,4% kredytobiorców, którzy nie spłacają kredytu w terminie. Tak skonstruowana próba odpowiada sytuacji zaobserwowanej w kraju w czerwcu 2006 (*Raport NBP o stabilności 2006 I półrocze*).

Wyniki przeprowadzonych eksperymentów pozwalają ocenić jakość klasyfikacji, uzyskanych za pomocą obu metod i dla różnej struktury prób. Na tej podstawie można zbadać czy struktura prób uczących

(estymacyjnych) ma wpływ na efektywność klasyfikacji ocenianej w próbie testowej (weryfikacyjnej).

(\*) Pominiemy fakt, że posiadane przez bank informacje mogą nie być w pełni wiarygodne.

(\*\*) Większość banków przyjmuje jako minimalną wartość oferowanej pożyczki hipotecznej kwotę 30 tys. zł., zatem przypadki o mniejszej wartości kredytów hipotecznych potraktowano jako obserwacje nietypowe.

---

15:30 Oral

---

### **Badanie efektywności ekonomicznej portfeli losowych generowanych dla kompozycji wybranych indeksów GPW w latach 2000-2006**

Krzysztof Kompa

*Szkola Główna Gospodarstwa Wiejskiego (SGGW), Nowoursynowska 166, Warszawa 02-787, Poland*

*e-mail: kkompa@mors.sggw.waw.pl*

Inwestowanie na rynku kapitałowym postrzegane jest jako skuteczna alternatywa dla klasycznych metod pomnażania kapitału. W konsekwencji teoria i praktyka inwestowania koncentrują się na poszukiwaniu coraz bardziej złożonych metod analizy i modelowania szeregów finansowych, wskaźników fundamentalnych i technicznych przydatnych w doborze instrumentów finansowych do portfeli inwestycyjnych. Nadrzędnym celem tych prac jest optymalizacja stopy zwrotu lub ryzyka inwestycyjnego z portfela.

Zagadnienie budowy portfeli inwestycyjnych o różnym horyzoncie czasowym nabiera szczególnego znaczenia w warunkach rynku informacyjnie efektywnego, choćby w formie słabej. Dotyczy to także polskiego rynku kapitałowego - badania potwierdzają hipotezę słabej efektywności informacyjnej w zakresie podstawowych indeksów GPW.

Przedmiotem badania jest analiza porównawcza stóp zwrotu i ryzyka dla krótko- i średnioterminowych portfeli, generowanych na bazie instrumentów z kompozycji informacyjnie słabych indeksów GPW. Do wyboru spółek wykorzystane zostaną generatory liczb pseudolosowych oraz wybrane metody taksonomiczne.

---

15:45 Oral

---

### **Zastosowanie wielowymiarowych modeli GARCH do szacowania współczynnika zabezpieczenia dla kontraktów futures na WIG20**

Krzysztof L. Piontek

*Akademia Ekonomiczna (AEWROCLAW), Komandorska 118/120, Wrocław 53-345, Poland*

*e-mail: krzysztof.piontek@ae.wroc.pl*

zakres tematyczny: zastosowania - analiza danych finansowych

Koncepcja zarządzania ryzykiem rynkowym za pomocą kontraktów *futures* pozostaje ciągle w obszarze zainteresowania zarówno praktyków, jak i teoretyków rynku finansowego. Nie bez znaczenia pozostaje niewątpliwie fakt popularyzacji wyrafinowanych metod matematycznych, jak i rozwoju narzędzi informatycznych. Kluczowy-

mi elementami procesu zabezpieczenia portfela pozostają: wybór odpowiedniej miary efektywności zabezpieczenia oraz estymacja optymalnego (dla tej miary) współczynnika zabezpieczenia. Technika szacowania optymalnego współczynnika wynika bezpośrednio z przyjętego modelu opisującego szeregi cen (stóp zwrotu) instrumentu bazowego oraz kontraktu *futures*. Wartość współczynnika zabezpieczenia determinuje liczbę kupowanych lub wystawianych kontraktów *futures*.

W części teoretycznej pracy zaprezentowane zostaną skrótowo możliwe kryteria oceny efektywności zabezpieczenia ze szczególnym uwzględnieniem najpopularniejszego podejścia polegającego na minimalizacji wariancji stóp zwrotu zabezpieczonego portfela. Odpowiednią miarę efektywności zabezpieczenia zaproponowali Bera, Garcia i Roh jako modyfikację miary Ederingtona. Polega ona na ocenie procentowej redukcji wariancji stóp zwrotu portfela zabezpieczonego względem wariancji portfela niezabezpieczonego. Miara ta wykorzystana zostanie w części empirycznej pracy.

Tradycyjnie stosowane metody estymacji optymalnego współczynnika zabezpieczenia zakładają stałość wektora warunkowych wartości oczekiwanych oraz warunkowej macierzy kowariancji dla szeregów stóp zwrotu z zabezpieczanego portfela i kontraktu *futures*, co nie znajduje potwierdzenia w badaniach empirycznych. Zaproponowano więc szereg modeli ze zmiennymi w czasie warunkowymi wartościami oczekiwanymi i warunkowymi wartościami elementów macierzy kowariancjami. Prowadzi to do tzw. warunkowych (zmiennych w czasie) wartości współczynnika zabezpieczenia. W rozważaniach szczegółowych, ze względu na ograniczoną objętość pracy pominięte zostaną jednak modele opisujące potencjalne efekty wektorowej autoregresji czy kointegracji pomiędzy szeregami stóp zwrotu lub cen. Zagadnienia te staną się obszarem dalszych badań autora.

Najpopularniejszym podejściem w zakresie opisu warunkowych wartości macierzy wariancji-kowariancji pozostaje wielowymiarowy model *GARCH*, którego ogólna postać określa się jako *VECH-GARCH*. Pomimo faktu, iż model ten jest naturalnym rozszerzeniem jednowymiarowego modelu *GARCH*, jest on mało popularny ze względu na szereg trudności związanych z jego praktycznym zastosowaniem, a mianowicie z: dużą liczbą parametrów, koniecznością zapewnienia dodatniej określoności macierzy, koniecznością zapewnienia skończoności wartości bezwarunkowej macierzy kowariancji. Celem pracy w warstwie teoretycznej jest więc przede wszystkim przegląd, porównanie właściwości oraz ocena trudności praktycznego wykorzystania zaproponowanych modeli zawierających się w ogólnym modelu *VECH*, które to modele mogą posłużyć wprost do szacowania współczynnika zabezpieczenia. Porównane zostaną: modele diagonalne *DVECH* (pełne i skalarne), modele klasy *BEKK* (pełne, diagonalne i skalarne), wygładzania wykładniczego *EWMA* oraz najprostszy model ze stałymi wartościami elementów macierzy kowariancji. Wszystkie te modele zawierają się w ogólnej postaci modelu *VECH-GARCH*. Dodatkowo przedstawiony zostanie również popularny model stałej macierzy korelacji.

Najpopularniejszym instrumentem pochodnym na rynku polskim pozostaje kontrakt *futures* na *WIG20*. W badaniach empirycznych wykorzystany zostanie więc ten instrument, a portfelem zabezpieczanym będzie portfel odwzorowujący wprost indeks *WIG20* oraz wybrane portfele odbiegające składem od indeksu *WIG20*.

Wszystkie omówione modele klasy *GARCH* posłużą do wyznaczenia warunkowych współczynników zabezpieczenia, a następnie badania efektywności strategii. Podejście takie jest w ogólności strate-

gią dynamiczną, w której wymagana jest przebudowa portfela wraz ze zmianami współczynnika zabezpieczenia. Rozpatrzone zostaną więc różne modele szeregów stóp zwrotu kontraktów *futures* i instrumentów bazowych, różne okresy przebudowy portfeli oraz długości szeregów służących do estymacji parametrów modeli.

Celem części empirycznej jest odpowiedź na pytanie, czy istnieje jeden model, który dla rynku polskiego, dla różnych zabezpieczanych portfeli byłby (choćby w większości przypadków) rozwiązaniem optymalnym, stanowiącym konieczny zapewne kompromis pomiędzy ogólnością modelu a jakością dopasowania do danych, efektywnością zabezpieczenia oraz efektem ekonomicznym po uwzględnieniu kosztów transakcji. Uzyskanie twierdzącej odpowiedzi byłoby ważnym wnioskiem praktycznym.

Niniejsza praca jest kontynuacją badań autora zaprezentowanych w opracowaniach „Prognozowanie macierzy kowariancji i korelacji finansowych szeregów czasowych” oraz „Niektóre praktyczne wyzwania w modelowaniu wielowymiarowych procesów *GARCH*”.

### Przerwa

Środa popołudniu, 19 Wrzesień, 16:00

### Sesja B2

Zastosowania Międzynarodowe

Środa popołudniu, 19 Wrzesień, 16:15

16:15

Oral

### Wielowymiarowa analiza korespondencji jako narzędzie segmentacji międzynarodowej

Elżbieta Sobczak

*Akademia Ekonomiczna (AE), Nowowiejska3, Jelenia Góra 58-500, Poland*

*e-mail: sobczak@ae.jgora.pl*

Celem referatu jest propozycja zastosowania wielowymiarowej analizy korespondencji MCA (*multiple correspondence analysis*) do segmentacji rynku międzynarodowego na podstawie danych binarnych, identyfikujących konsumenckie łańcuchy celów i środków (*means-end chains*). MCA jest jedną z metod współwystępowania kategorii zmiennych mierzonych na skali nominalnej. Umożliwia graficzne przedstawienie poszczególnych kategorii zmiennych jako punktów na płaszczyźnie i określenie związków między nimi. Mapa punktów otrzymana z wykorzystaniem wielowymiarowej analizy korespondencji pozwala na identyfikację występujących skupień kategorii zmiennych, opisujących międzynarodowe segmenty konsumentów. Zastosowanie MCA pozwoli na określenie współwystępujących kryteriów segmentacji i zmiennych profilowych. Kryteria segmentacji w proponowanym podejściu stanowią elementy łańcuchów celów i środków (sposoby kojarzenia atrybutów produktu z korzyściami jakie może przynieść jego użytkowanie, oraz wartościami jakie niosą ze sobą te korzyści.). Wielowymiarowa analiza korespondencji umożliwi identyfikację łańcuchów celów i środków, charakterystycznych dla konsumentów o określonych cechach. Dostarczy zatem informacji o podziale rynku na segmenty i profilach konsumentów. Każdy z wyłonionych segmentów obejmować będzie konsumentów identyfikujących zbliżone struktury poznawcze pro-

duktu (łańcuchy celów i środków).

16:30

Oral

### Obszary wykorzystania miar syntetycznych i subsyntetycznych w kontekście ich własności formalnych

Tadeusz Borys, Małgorzata Markowska

*Akademia Ekonomiczna (AE), Nowowiejska3, Jelenia Góra 58-500, Poland*

*e-mail: mmarkowska@ae.jgora.pl*

Wskaźniki to podstawowe narzędzie monitoringu rozwoju, odsłaniające w sposób wymierny istotę różnych jego koncepcji czy paradygmatów. Poszukiwania syntetycznych miar oddających we właściwy sposób specyficzne cechy konkretnej koncepcji rozwoju jest nadal otwartym problemem badawczym.

Doświadczenia w tworzeniu i stosowaniu miar subsyntetycznych w ramach poszczególnych łańcuchów czy sfer są zachęcające. Dotyczy to ładu społecznego – wykorzystania miar HDI i pochodnych, środowiskowego - wykorzystania miary ESI oraz gospodarczego - wykorzystania miary ISEW. Największe problemy dotyczą pomiaru ładu przestrzennego zarówno w układzie analitycznym (zdezagregowanym), jak i subsyntetycznym.

Ważną własnością syntetycznych miar rozwoju powinien być wysoki stopień wrażliwości tych miar na wartości skrajne. Zbadanie tej własności ma istotne znaczenie w określaniu obszarów zastosowań miar syntetycznych o różnej konstrukcji formalnej. Dokonany w pracy przegląd tych funkcji umożliwi ocenę przydatności poszczególnych miar dla estymacji poziomu rozwoju w różnych obszarach ich implementacji.

16:45

Oral

### Propozycja metody klasyfikacji krajów z wykorzystaniem metod korekcji sezonowej

Dominik A. Rozkrut

*Uniwersytet Szczeciński, Jedności Narodowej 22, Szczecin 70-453, Poland*

*e-mail: dominic@univ.szczecin.pl*

W artykule przedstawiono propozycję wykorzystania metod korekcji sezonowych w klasyfikacji krajów, opartej o badanie cykliczności i sezonowości rozwoju gospodarczego tych krajów. Przedstawiona propozycja wpisuje się bezpośrednio w dynamiczny nurt badań związanych (w dużym uogólnieniu) z poszukiwaniem „wspólnego” cyklu koniunkturalnego dla obszaru Unii Europejskiej. Likwidacja barier gospodarczych wewnątrz Unii Europejskiej, zgodnie z przewidywaniami, może prowadzić do konwergencji ścieżek rozwoju gospodarczego, a co za tym idzie wyodrębnienia się wspólnego cyklu koniunkturalnego. W badaniach tej problematyki użyteczne okazują się metody statystycznej analizy wielowymiarowej, które, w odróżnieniu od istniejących badań, zaproponowano jednak wykorzystać łącznie z metodami korekcji sezonowej. Metody te wykorzystywane są do wyodrębniania składowych cyklicznych w szeregach czasowych agregatów ekonomicznych. Jednym z przykładów tej klasy metod, jest procedura korekcji sezonowej Census II w jej róż-

nych wersjach. Z zastosowaniem tej procedury wiąże się możliwość otrzymania wielu interesujących statystyk charakteryzujących cykliczność i sezonowość analizowanego szeregu. Statystyki te dotyczą np. relatywnej wagi poszczególnych składowych szeregu czasowego, i w związku z tym same w sobie mogą stanowić przedmiot odrębnych analiz. Proponowana procedura została zilustrowana na przykładzie badań dotyczących wybranych krajów OECD. Wyniki potwierdzają jej wysoką przydatność w kontekście poszukiwań obszarów charakteryzujących się podobieństwem przebiegu procesów ekonomicznych związanych z cyklem koniunkturalnym, zwłaszcza w porównaniu z podobnymi badaniami opartymi o inne procedury badawcze.

Zakres tematyczny: Analiza danych przestrzennych

### Zakończenie obrad sesji B2

Środa popołudniu, 19 Wrzesień, 17:15

## Czwartek, 20 Wrzesień

### Sesja B3

Zastosowania w Analizach Przestrzennych

Czwartek rano, 20 Wrzesień, 10:15

10:15

Oral

### Zróżnicowanie poziomu rozwoju gospodarczego województw w Polsce a wielkość migracji międzywojewódzkich. Zastosowanie wybranych procedur taksonomicznych

Barbara Dańska-Borsiak

Uniwersytet Łódzki, Wydział Ekonomiczno-Socjologiczny, P.O.W. 3/5, Łódź 90-255, Poland

e-mail: [danska@uni.lodz.pl](mailto:danska@uni.lodz.pl)

W poziomie rozwoju gospodarczego województw w Polsce zauważalne są duże różnice. Wydaje się, że zróżnicowanie to może być jednym z czynników, wpływających na wielkość migracji ludności między województwami.

Głównym celem referatu jest zbadanie, czy stopień rozwoju gospodarczego wpływa na liczbę osób migrujących między województwami, oraz określenie siły wpływu poszczególnych zmiennych, wyznaczających poziom rozwoju województwa na wielkość migracji.

W pierwszej części artykułu dokonano uporządkowania województw ze względu na zmienne, będące wyznacznikami stopnia rozwoju gospodarczego (np. PKB per capita, nakłady na BiR, stopa inwestycji, stopa bezrobocia, etc.). Posłużono się w tym celu miarą rozwoju. Następnie zbadano poziom korelacji miary rozwoju z wielkością migracji do i z danego województwa.

W dalszej części wyodrębniono grupy województw podobnych ze względu na poziom rozwoju gospodarczego. Następnie zbadano siłę wpływu poszczególnych determinant poziomu rozwoju gospodarczego na wielkość migracji międzywojewódzkich. Badania tego dokonano na podstawie danych ze wszystkich województw, oraz w obrębie każdej z wyodrębnionych grup.

10:30

Oral

### Typologia regionów Polski ze względu na dysproporcje rozwojowe

Monika Rozkrut

Uniwersytet Szczeciński, Jedności Narodowej 22, Szczecin 70-453, Poland

e-mail: [sobot@uoo.univ.szczecin.pl](mailto:sobot@uoo.univ.szczecin.pl)

Celem badania była delimitacja przestrzenna podregionów, umożliwiająca określenie typowych układów wartości wskaźników ekonomicznych – a co za tym idzie – próba określenia czynników na nie wpływających. Opracowana typologia podregionów Polski oparta została na wskaźnikach potencjału ekonomicznego i jego determinant. W szczególności więc, badania przeprowadzono w oparciu o analizę danych dotyczących następujących sfer: produkcji i dochodów, infrastruktury i inwestycji, rynku pracy i migracji, oraz kapitału ludzkiego, mierzonego aktywnością na polu badań i rozwoju oraz przedsiębiorczości. Zestaw wykorzystanych narzędzi obejmuje metody grupowania, skalowania wielowymiarowego i badania zależności. Analizy dotyczą okresu 2000-2005, umożliwiając dokonywanie porównań w czasie. Wyniki ujawniają znaczne dysproporcje pomiędzy podregionami dużych aglomeracji i podregionami leżącymi w ich pobliżu, a tymi od nich oddalonymi. Szczególne znaczenie ma znaczne zróżnicowanie dochodów, które zdaje się zależeć od tempa rozwoju gospodarczego. Rezultaty te wydają się potwierdzać wnioski z innych badań, sugerujących pogłębianie się dysproporcji pomiędzy najlepiej i najmniej zarabiającymi, właśnie w regionach najszybciej rozwijających się. Interesująca jest wyższa dynamika rozwoju regionów wschodnich w porównaniu z zachodnimi. Porównania w czasie ujawniają wpływ spadku bezrobocia na stopniową poprawę oceny sytuacji gospodarczej regionów. Niepokojący jest niewielki, w porównaniu z innymi krajami UE, poziom wynalazczości, nawet w regionach o wyższej intensywności inwestycji, co w przypadku wielu regionów może stanowić w przyszłości istotną barierę rozwoju. W artykule zawarto szczegółową charakterystykę otrzymanych grup, oraz zaprezentowano porównanie wybranych wartości z przeciętnymi dla całej Unii Europejskiej.

Zakres tematyczny: Analiza danych przestrzennych

10:45

Oral

### Taksonomiczna analiza porównawcza poziomu zrównoważonego rozwoju województw w Polsce

Elżbieta Wiszniewska

Uniwersytet Łódzki, Wydział Ekonomiczno-Socjologiczny, P.O.W. 3/5, Łódź 90-255, Poland

e-mail: [elka@o2.pl](mailto:elka@o2.pl)

Zrównoważony rozwój jest nadrzędną zasadą polityki państw – członków Unii Europejskiej, w tym Polityki Ekologicznej Polski. Zgodnie z zaleceniami Agendy 21 budowane są kompleksowe programy ekorozwoju na poziomie kontynentów, krajów, regionów, gmin oraz miejscowości. Ich celem jest stworzenie warunków dla wszechstronnego rozwoju, dla stworzenia harmonijnej koegzystencji człowieka i przyrody. Zwłaszcza w skali regionalnej w pełni manife-

stują się kompleksowe powiązania struktur przyrodniczych ze strukturami osadnictwa i gospodarki, stąd też opracowanie programu ekorozwoju jest szczególnie przydatne dla każdego regionu. W związku z wdrażaniem idei zrównoważonego rozwoju do regionalnych polityk społeczno – gospodarczych w Polsce nastąpiły dynamiczne zmiany w takich dziedzinach jak: ochrona środowiska naturalnego, ochrona zdrowia, wykluczenie społeczne i ubóstwo, zatrudnienie, wzorce produkcji i konsumpcji, transport, i w wielu innych.

Celem niniejszego artykułu jest ocena poziomu trwałego rozwoju województw Polski w oparciu o wybrane wskaźniki środowiskowe, zaczerpnięte ze Strategii Zrównoważonego Rozwoju UE oraz Polityk Ekologicznych Państwa. W celu uporządkowania województw ze względu na poziom ekorozwoju w latach 2002 – 2005 oraz na podstawie zaproponowanej listy cech diagnostycznych, stosując metodę unitaryzacji zerowej, zbudowano syntetyczne mierniki rozwoju zwane metacechami. Następnie uporządkowano województwa według ich „ekologicznej atrakcyjności”, wyznaczonej przez najwyższe wartości mierników syntetycznych. W celu zbadania podobieństwa województw, na podstawie metody wzorca rozwoju, zbudowano również miary rozwoju. Źródłem danych statystycznych był EUROSTAT oraz GUS. Zastosowanie metod taksonomicznej analizy porównawczej posłużyło uporządkowaniu a także utworzeniu grup województw o podobnych poziomie ekorozwoju w latach 2001 – 2005.

*słowa kluczowe:* ekorozwój, Agenda 21, metacechy, unitaryzacja zera, miary rozwoju, metody bezwzorcowe, metody wzorca rozwoju;

*Bibliografia:*

1. Jajuga K. (red.), M. Walesiak (red.), (2005), „*Taksonomia 12. Klasyfikacja i analiza danych – teoria i zastosowania*”, Prace Naukowe Akademii Ekonomicznej im. Oskara Langego, Wrocław;
2. Malina A., (2004), *Wielowymiarowa analiza przestrzennego zróżnicowania struktury gospodarki Polski według województw*, Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków;
3. Młodak A., (2006), *Analiza taksonomiczna w statystyce regionalnej*, Difin, Warszawa;
4. *Polityka Ekologiczna Państwa na lata 2003 – 2006 z uwzględnieniem perspektywy na lata 2007 – 2010*, Rada Ministrów, Warszawa, grudzień 2002;
5. Portal o Unii Europejskiej – [www.europa.eu/index\\_pl.htm](http://www.europa.eu/index_pl.htm);
6. [www.stat.gov.pl](http://www.stat.gov.pl) – strona Głównego Urzędu Statystycznego;
7. *Odnowiona strategia UE dotycząca zrównoważonego rozwoju*, Rada Unii Europejskiej, Bruksela, czerwiec 2006;

---

11:00

Oral

---

### **Wielowymiarowa analiza porównawcza poziomu rozwoju gmin woj. małopolskiego w latach 1996-2005**

Stanisław W. Matusik

*Akademia Wychowania Fizycznego, Zakład Statystyki i Informatyki, Instytut Turystyki i Rekreacji, al. Jana Pawła II 78, Kraków 31-571, Poland*

*e-mail: matusik@poczta.onet.pl*

Badania rozwoju jednostek administracyjnych różnych szczebli przeprowadzono analizując rozmaite jego aspekty. Stosunkowo mało prac jest poświęconych badaniu rozwoju gmin.

Rozwój lokalny bazuje najczęściej na rozwoju gospodarczym, choć nie jest z nim równoważny dzięki pozamaterialnym i społecznym korzyściom (w tym kulturowym, technologicznym, środowiskowym oraz wypływającym z podnoszenia kwalifikacji, poziomu wykształcenia i zdobywanej wiedzy). W niniejszej pracy pojęcie rozwoju regionalnego traktowane jest w sensie „trwałego rozwoju”, opartego na mierzalnych związkach ze wzrostem gospodarczym.

W analizach wzięto pod uwagę dane dla 181 gmin województwa małopolskiego, z wyłączeniem gminy Kraków, aby odmienny charakter stolicy województwa nie wpływał na statystyki gmin, w większości typu wiejskiego i miejsko-wiejskiego. Rozważono 12 zmienionych dobranych pod kątem ich znaczenia dla badań rozwoju społeczno-gospodarczego. Wskaźniki poziomu rozwoju społeczno-gospodarczego gmin oparto na zmiennych związanych z przedsiębiorczością, demografią, rolnictwem, turystyką, infrastrukturą społeczną i infrastrukturą techniczną. Aby uzyskać porównywalność w czasie oraz pomiędzy gminami, cechy wejściowe odniesiono do liczby ludności, grup wiekowych bądź powierzchni gminy lub posłużono się wartościami procentowymi.

Celem pracy było:

- 1) porównanie różnych wskaźników rozwoju gospodarczego dla gmin,
- 2) określenie na ich podstawie charakterystyki grup gmin o najwyższym i najniższym poziomie, uzyskanych metodami taksonometrii, drzew decyzyjnych i wielomianowej regresji logitowej (ang. multinomial logit regression),
- 3) analiza uzyskanych rezultatów dla roku 2005 w odniesieniu do roku 1996 oraz wskazanie metod dających porównywalne rezultaty.

Rozważane wskaźniki rozwoju mogą dostarczać informacji zarówno o aktualnym stanie, jak też o zmianach tego poziomu w poszczególnych gminach. Najwyższym poziomem rozwoju charakteryzowały się gminy miejskie i miejsko-wiejskie oraz nieliczne gminy wiejskie położone w bezpośrednim sąsiedztwie stolicy województwa. Najniższy poziom odnotowano w podgórskich gminach wiejskich.

Przestrzenna analiza badanego zjawiska wykazała znaczące, pozytywne oddziaływanie trzech największych miast województwa (a szczególnie Krakowa) na rozwój otaczających je gmin wiejskich. Wskazano na fakt, że blisko 2/3 gmin wymaga podniesienia poziomu rozwoju społeczno-gospodarczego, aby mogły one osiągnąć przeciętny poziom, charakterystyczny dla gmin w województwa małopolskiego w 2005 roku.

Uzyskane wyniki wskazują także na dywergencję rozwoju społeczno-gospodarczego gmin województwa małopolskiego w badanym okresie.

---

11:15

Oral

---

### **Wykorzystanie metod porządkowania liniowego do oceny polskiego rynku nieruchomości.**

Iwona A. Forys

*Uniwersytet Szczeciński, Jedności Narodowej 22, Szczecin 70-453, Poland*

*e-mail: iwonafor@poczta.onet.pl*

Rynek nieruchomości w Polsce jest rynkiem, który na przestrzeni

ostatnich lat rozwijał się z różną dynamiką, począwszy od hossy w latach 90-tych, po stagnację na przełomie wieków aż do ponownego rozkwitu w ostatnich dwóch latach. Jednak charakter zmian zachodzących na tym rynku wskazuje na duże zróżnicowanie przestrzenne. Są regiony, które silniej reagują na zmiany a także regiony, które dystansują się w rozwoju rynku nieruchomości od liderów.

Metody porządkowania liniowego mogą być pomocne w uszeregowaniu obiektów (w artykule województw) i ocenie poziomu rozwoju danego lokalnego rynku nieruchomości z uwagi na czynniki bezpośrednio wpływające na rynek nieruchomości takie jak zasób nieruchomości, ceny, liczba transakcji oraz czynniki zewnętrzne wynikające z otoczenia rynku nieruchomości: czynniki demograficzne, społeczne, ekonomiczne czy prawne.

W artykule podjęta zostanie próba porządkowania województw z uwagi na wspomniane grupy czynników wpływające na rozwój rynku nieruchomości, w latach 2003-2005. Ocena podobieństwa wyników porządkowania liniowego przeprowadzona zostanie w czasie, na podstawie wartości zmiennych syntetycznych (SMR).

Wydaje się, że proste metody porządkowania liniowego mogą mieć zastosowanie praktyczne, zwłaszcza w dobie poszukiwania nowych obszarów inwestowania na rynku nieruchomości i oceny ich potencjału przez inwestorów.

### Przerwa

Czwartek rano, 20 Wrzesień, 11:30

### Sesja B4

Zastosowania w Finansach II

Czwartek rano, 20 Wrzesień, 11:45

11:45

Oral

### Porównanie wyników segmentacji na różnych rynkach ubezpieczeń gospodarstw domowych

Katarzyna Wawrzyniak<sup>1</sup>, Barbara Batóg, Magdalena C. Mojsiewicz<sup>2</sup>

1. Akademia Rolnicza (AR), Janosika 8, Szczecin 71-424, Poland

2. Uniwersytet Szczeciński, Jedności Narodowej 22, Szczecin 70-453, Poland

e-mail: kwawrzyniak@e-ar.pl

Segmentację rynku ubezpieczeń na podstawie preferencji *klientów* i *nie-klientów* usług i produktów ubezpieczeniowych można scharakteryzować następująco: badanie dotyczy obiektów, jest częściej badaniem współwystępowania niż zależności, zmienne są zarówno niemetryczne, jak i metryczne, liczebność próby jest uzależniona od możliwości finansowych. W związku z powyższym, do segmentacji różnych rynków ubezpieczeń gospodarstw domowych na podstawie ich preferencji zostanie wykorzystane skalowanie wielowymiarowe.

Segmentacja będzie dotyczyła następujących rynków ubezpieczeń gospodarstw domowych: majątkowych z podziałem na ubezpieczenie mienia i samochodu, życiowych oraz zdrowotnych.

Celem artykułu jest weryfikacja tezy, że grupy klientów o podobnych zachowaniach wyodrębnione na różnych rynkach ubezpieczeń w

znacznym stopniu się pokrywają, czyli klient zachowujący się racjonalnie na rynku ubezpieczeń mienia, zachowuje się racjonalnie również na innych rynkach.

Dane statystyczne pochodzą z badania ankietowego nt.: *Sklonności do ubezpieczeń i preferencji klientów na rynku ubezpieczeń* przeprowadzonego w ramach grantu KBN<sup>1</sup> przez firmę TNS OBOP. Wywiady przeprowadzono z 500 respondentami – głowami gospodarstw domowych w lipcu/sierpniu 2006 roku.

Segmenty zostaną wyznaczone na bazie mapy preferencji gospodarstw domowych stworzonej dla zmiennych określających preferencje respondentów odnośnie poszczególnych rodzajów ubezpieczeń. Preferencje dotyczą m.in. samodzielności w podejmowaniu decyzji, motywu zakupu, kompletności usługi, ceny czy wydolności systemu opieki zdrowotnej. Po zidentyfikowaniu osi z wykorzystaniem metody korelacji wartości wymiaru z wartościami zmiennych wspomaganą wiedzą ekspercką z zakresu ubezpieczeń, zostaną określone dla każdego rynku segmenty zgodne z teorią Wagnera, czyli segmenty klientów racjonalnych oraz nieracjonalnych. Do pomiaru głębokości segmentów zostaną wykorzystane różne zmienne w zależności od rodzaju ubezpieczenia. Zmienne określające głębokość segmentów zostaną uwzględnione w sposób bezpośredni (trzeci wymiar jest tworzony na bazie rzeczywistych wartości przyjmowanych przez wybrane zmienne) oraz pośredni (trzeci wymiar jest generowany z wykorzystaniem skalowania wielowymiarowego).

<sup>1</sup>*Metody statystyczne w segmentacji rynku ubezpieczeń w Polsce*, nr 1H02B02827.

12:00

Oral

### Porównanie koncepcji zależności ekstremalnej i warunkowo zmiennej kowariancji w modelowaniu zależności między indeksami wybranych rynków akcji

Paweł Rokita

Akademia Ekonomiczna (AEWROCLAW), Komandorska 118/120, Wrocław 53-345, Poland

e-mail: pawel.rokita@ae.wroc.pl

Klasyczna metody dywersyfikacji portfela nie uwzględniają zależności między ekstremalnymi stopami zwrotu. Wyniki badań empirycznych publikowane w literaturze wielokrotnie wskazywały jednak na współwystępowanie nietypowych stóp zwrotu, a zwłaszcza ekstremalnych strat. Jeżeli takie zjawisko jest statystycznie istotne, może w znaczący sposób wpływać na ryzyko portfela. Dwa najczęściej wykorzystywane podejścia w analizie danych o takich własnościach to: (1) modelowanie zależności w ogonie wielowymiarowego rozkładu bezwarunkowego (podejście statyczne) oraz (2) wykorzystanie modeli wielowymiarowych procesów stochastycznych, w których rozkłady warunkowe stóp zwrotu mają co prawda cienkie ogony, ale występuje efekt grupowania kowariancji. W pierwszej części tego artykułu rozważane są podobieństwa między efektem zależności ekstremalnej w rozkładzie bezwarunkowym oraz grupowaniem klasycznej kowariancji. Na tym etapie wykorzystywane są dane generowane sztucznie. Następnie, w części empirycznej, oba podejścia są wykorzystane do modelowania stóp zwrotu z par portfeli replikujących indeksy wybranych rynków akcji. Celem tego badania jest stwierdzenie, czy efekty zależności ekstremalnej lub grupowania zależności liniowej występują oraz które z tych dwóch podejść lepiej

nadaje się do opisu stóp zwrotu ze zdywersyfikowanych międzynarodowo portfele akcji. Na potrzeby tego porównania wykorzystane zostały wyniki tzw. testów wstecznych (*back testing*) wartości zagrożonej (Value at Risk) obliczonej dla portfeli przy wykorzystaniu rozpatrywanych modeli. Wyniki przeprowadzonego porównania mogą mieć duże znaczenie dla konstrukcji modeli ryzyka dla inwestycji w akcje.

12:15 Oral

### Wykorzystanie wybranych metod analizy wielowymiarowej do klasyfikacji funduszy inwestycyjnych

Marcin Salamaga

*Cracow University of Economics (CUOE), Rakowicka 27, Kraków 31-510, Poland*

*e-mail: salamaga@ae.krakow.pl*

Klasyfikacje funduszy inwestycyjnych przeprowadza się dość powszechnie na podstawie rodzaju papierów wartościowych i innych praw majątkowych, będących przedmiotem lokat funduszy. Nie zawsze uwzględniają one wyniki pomiaru efektywności funduszy, strategię inwestycyjną czy wartość narażoną na ryzyko. W referacie zaproponowano klasyfikacje funduszy w grupy zawierające jednorodne, podobne do siebie obiekty wykorzystując metody skupień: metodę Warda oraz metodę *k*-średnich. Przedmiotem przeprowadzonej analizy jest zbiór dziewiętnastu funduszy inwestycyjnych, które zbadano pod względem 30 zmiennych opisujących funkcjonowanie funduszy inwestycyjnych w Polsce. Zmienne te opisują fundusze inwestycyjne pod względem doświadczenia w działalności na rynku finansowym, polityki wobec uczestników funduszy, wyników działalności finansowej, ryzyka portfela lokat funduszy, efektywności zarządzania portfelem lokat funduszu. Stopień zróżnicowania powstałych grup i „zdolność różnicowania” poszczególnych zmiennych oceniano za pomocą jednoczynnikowej analizy wariancji.

Otrzymane wyniki grupowania funduszy pozwoliły m.in. na wyodrębnienie funduszy „najlepszych” oraz „najsłabszych”.

W obliczeniach wykorzystano ogólnodostępne wskaźniki rynku kapitałowego oraz dane pochodzące z rocznych i półrocznych sprawozdań finansowych funduszy w latach 2000-2005. Wybór konkretnych obiektów do analizy był wynikiem kompromisu pomiędzy ich liczbą a długością szeregów czasowych cen jednostek uczestnictwa funduszy (niektóre fundusze działające od kilku miesięcy nie zostały uwzględnione w analizie z uwagi na krótkie szeregi czasowe wartości jednostek uczestnictwa).

### Zakończenie obrad w sesji B4

Czwartek popołudniu, 20 Wrzesień, 12:30

## Piątek, 21 Wrzesień

### Sesja B5

*Zastosowanie w Finansach III*

Piątek rano, 21 Wrzesień, 9:00

9:00 Oral

### Przykład zastosowania funkcji powiązań Farliego-Gumbela-Morgensterna w analizie związków zależności między wybranymi akcjami notowanymi na GPW w Warszawie

Daniel Papla

*Akademia Ekonomiczna (AEWROCLAW), Komandorska 118/120, Wrocław 53-345, Poland*

*e-mail: daniel.papla@ae.wroc.pl*

Celem artykułu jest zbadanie charakteru zależności występujących pomiędzy akcjami na polskiej giełdzie. Ze względu na czasochłonność i złożoność obliczeń ograniczono się jedynie do wybranych spółek z polskiej giełdy. Z tego samego powodu w obliczeniach wykorzystano jedynie portfele składające się z dwóch i trzech elementów. Dodatkowym ograniczeniem liczby składników portfela jest specyfika wykorzystanej przez autora funkcji powiązań.

Hipotezę badawczą jest następujący sposób: na polskiej giełdzie występują nie tylko zależności występujące między parami akcji (które można analizować za pomocą np. klasycznego współczynnika korelacji) ale występują również zależności między trójkami akcji, które można badać za pomocą wielowymiarowych (w tym przypadku trójwymiarowych) funkcji powiązań. Potwierdzenie tej hipotezy oznaczałoby, że zależność między akcjami polskich spółek jest nieco bardziej skomplikowana, niż to wynika z klasycznych modeli rynku kapitałowego i ma charakter nieliniowy.

W celu weryfikacji tej hipotezy autor wykorzystał funkcję powiązań Farliego-Gumbela-Morgensterna (FGM). Analiza składała się z dwóch etapów. W pierwszym estymowano współczynniki funkcji powiązań FGM dla par akcji, w drugim zaś dla trójek (czyli portfeli trójskładnikowych). Porównanie wyników tych estymacji stanowi podstawę weryfikacji hipotezy badawczej.

Praca składa się z trzech części. W pierwszej omówiono zastosowaną funkcję powiązań Farliego-Gumbela-Morgensterna. W drugiej przedstawiono wyniki badań. W ostatniej części przedstawiono ich interpretacje.

9:15 Oral

### Klasyfikacja regionów według stopnia specjalizacji i koncentracji sektorowej oraz zróżnicowania dochodowego

Jacek Batóg

*Uniwersytet Szczeciński, Jedności Narodowej 22, Szczecin 70-453, Poland*

*e-mail: batog@uoo.univ.szczecin.pl*

Podstawowym celem przeprowadzonego badania będzie wykorzystanie podejścia klasyfikacyjnego do zweryfikowania hipotezy o występowaniu istotnego związku między zmianami regionalnej struktury sektorowej Polski i zróżnicowaniem dochodów uzyskiwanych przez mieszkańców badanych regionów w latach 1995-2005. Tak sformułowana hipoteza jest odzwierciedleniem tezy o istniejących tendencjach migracji siły roboczej i kapitału z tzw. regionów peryfe-

ryjnych do regionów charakteryzujących się wyższym poziomem dochodów na 1 mieszkańca oraz wyższą wydajnością pracy. Wydaje się, że istnieje tu również zależność w postaci pewnego sprzężenia zwrotnego polegająca na dodatnim wpływie specjalizacji i koncentracji na wzrost wydajności pracy i wynagrodzeń poprzez wykorzystanie dodatnich efektów skali. Rezultatem powyższych procesów jest szybszy wzrost regionów uznanych za centralne, powiązany z wyższym tempem wzrostu zatrudnienia oraz większą konkurencyjnością tych regionów.

Obecnie znaczącą rolę w badaniach przestrzennych nierówności dochodowych odgrywają metody ekonometryczne ukierunkowane na weryfikację hipotezy o występowaniu zbieżności dochodowej (zob. na przykład prace takich autorów jak: Coulombe, Petrakos, Barro oraz Sala-i-Martin). Podejścia te wymagają jednak uzupełnienia szerszym zestawem miar, które opisać procesy przestrzenne w sposób bardziej kompleksowy.

Klasyfikacja regionów na podstawie analizy stanu i zmian regionalnej struktury sektorowej oraz ich powiązania ze zróżnicowaniem dochodowym regionów zostanie przeprowadzona poprzez zastosowanie szerokiego zestawu miar względnych oceniających stopień specjalizacji, koncentracji, konwergencji oraz trwałość dysproporcji dochodowych. Wykorzystane zostaną między innymi takie miary jak: indeks koncentracji, indeks dochodu, indeks położenia, indeks specjalizacji, miara grupowa, employment Location Quotient (Hoover-Balassa coefficient), employment Industry Quotient oraz miara entropii Theila, które zostaną uzupełnione o syntetyczne mierniki zmian struktur pozwalające dokonać nie tylko analizy ewolucji struktur, lecz również periodyzacji zachodzących procesów przestrzennych.

---

9:30 Oral

---

### Model analizy conjoint w ocenie oferty ubezpieczeń od ryzyka do kredytu hipotecznego

Paweł Żuraw, Piotr Tarka

*Akademia Ekonomiczna (AEWROCLAW), Komandorska 118/120, Wrocław 53-345, Poland*

*e-mail: piotr.tarka@op.pl*

Niniejszy artykuł powstał na podstawie wyników badań jakie przeprowadzono na kwotowo - losowej próbie klientów Banku Zachodniego WBK. Cel jaki przyświecał realizacji badań, wiązał się z określeniem stopnia zainteresowania (zarówno u potencjalnych jak i rzeczywistych klientów) w sferze proponowanych przez bank, poszczególnych pakietów ubezpieczeń od ryzyka do kredytu hipotecznego.

Przyjęty w badaniu proces doboru próby, zakładał w fazie pierwszej, oszacowanie procentowego rozkładu populacji w Polsce zgodnie z danymi dostarczonymi przez GUS według 16 województw. W fazie drugiej (na podstawie bazy danych klientów BZ WBK), dopasowano proporcjonalnie strukturę populacji w Polsce do struktury i wielkości wyznaczonej bezpośrednio w próbie. Klasyfikacja klientów banku w ramach konstrukcji próby, pozwoliła wyodrębnić w niej 3 niezależne względem siebie grupy. Do grupy pierwszej zaliczono klientów, posiadających od niedawna kredyt hipoteczny w BZ WBK. W grupie drugiej, klientów posiadających kredyt hipoteczny od ponad przeszło 5 lat oraz w grupie trzeciej, klientów rozważających doce-

lowo podjęcie kredytu hipotecznego w banku.

W badaniu wykorzystano wielowymiarowy model analizy conjoint, który pozwolił ustalić rzeczywistą i hipotetyczną użyteczność całkowitą profilu produktu bankowego (ocenianego przez respondenta) w zestawieniu z jego użytecznością cząstkową. Oceniane przez respondentów atrybuty w obrębie poszczególnych ofert, miały doprowadzić do podziału ich preferencji względem ofert i ich cech. W modelu tym wykorzystano: 1) metodę pełnych profilów, w zakresie której uwzględniono zbiór różnorodnych wariantów odpowiedzi stanowiących kombinację proponowanych atrybutów, 2) metodę porównywania parami oraz 3) metodę prezentacji dwóch atrybutów jednocześnie. Kierunkowy obszar badań i przeprowadzanych analiz sprowadzał się do: 1) określenia ważności czynników i stopnia atrakcyjności poszczególnych ryzyk w wyodrębnionych przez bank w pakietach ubezpieczeń, 2) ustalenia najbardziej preferowanych (przez klientów banku) wariantów ubezpieczeń, 3) określenia zdolności klientów do zapłaty za proponowany przez bank pakiet ubezpieczeń w zakresie kredytu hipotecznego, 4) identyfikacji możliwości doku-  
pienia przez klientów banku dodatkowych opcji do określonego pakietu przy założeniu wzrostu cen za dany pakiet oraz 4) przeprowadzenia symulacji w zakresie 20 najbardziej atrakcyjnych ofert ubezpieczeń spośród wszystkich możliwych ofert, jakie wzięto pod uwagę w badaniu.

---

9:45 Oral

---

### Zastosowanie sztucznych sieci neuronowych do oceny ryzyka kredytowego klienta w telefonii komórkowej

Tomasz Zabkowski

*Szkoła Główna Gospodarstwa Wiejskiego (SGGW), Nowoursynowska 166, Warszawa 02-787, Poland*

*e-mail: tzabkowski@mors.sggw.waw.pl*

Artykuł prezentuje przykład zastosowania sztucznych sieci neuronowych do klasyfikacji ryzyka kredytowego klienta telefonii komórkowej. Podstawą budowy modeli neuronowych był zbiór cech charakteryzujących indywidualnych użytkowników, natomiast klasyfikacja została oparta na informacji o rozwiązaniu umowy z klientem z powodu niewypłacalności. Ponadto, w badaniach został wyznaczony model logitowy w celu określenia punktu odniesienia dla trafności klasyfikacji modeli sieci neuronowych.

Dane wykorzystane w eksperymencie stanowiły losową próbkę 51733 klientów operatora telefonii komórkowej. Dla każdego klienta zmienna o nazwie BP jest wykorzystywana do stwierdzenia rozwiązania umowy z powodu braku płatności za dostarczone przez operatora usługi. Zmienna ta jest postaci zero-jedynkowej i przyjmuje wartość 1, jeśli nastąpił brak płatności, oraz 0, w przypadku przeciwnym. W badanej próbie odsetek klientów, u których stwierdzono brak płatności skutkujący rozwiązaniem umowy wynosił 9.33 %. Jako dane charakteryzujące klienta posłużyły zmienne transakcyjne (m.in. płatności, korzystanie z usług głosowych, tekstowych, ponaglenia płatności) oraz demograficzne (m.in. wiek, płeć, miejsce zamieszkania). Z powodu dużej liczby dostępnych zmiennych objaśniających (58), w eksperymencie zdecydowano się na zastosowanie procedury wstępnej eliminacji cech.

Przedstawiony problem oceny ryzyka kredytowego użytkownika telefonii komórkowej sprowadzał się do rozwiązania problemu klasy-



fikacyjnego, którego celem było przypisanie obiektom przynależności do danej klasy(grupy). W związku, z tym, że znane są wartości klas, z których pochodzą obiekty rozważany był przypadek tzw. uczenia z nauczycielem.

Eksperyment został zapoczątkowany przez zbudowanie modelu logitowego z wszystkimi dostępnymi zmiennymi. Model ten stanowił punkt odniesienia dla trafności klasyfikacji modeli sieci neuronowych. W przypadku sieci neuronowych badanie zostało przeprowadzone w ten sposób, że dla z każdej sieci typu perceptron wielowarstwowy, były tworzone układy z różną ilością neuronów w warstwie ukrytej. Badanie zostało rozpoczęte poprzez utworzenie sieci z 2 neuronami w warstwie ukrytej, a następnie w każdym kolejnym modelu zostały dodawane dodatkowe neurony. Proces został zakończony w momencie, gdy każde następne dodanie neuronów nie poprawiało wyników osiąganych przez ostatnią sieć, bądź też pojawiał się efekt przeuczenia, czyli nadmiernego dopasowania sieci do danych.

Uzyskane wyniki pozwoliły na sformułowanie następujących wniosków:

- 1) Sztuczne sieci neuronowe w zastosowaniu do klasyfikacji ryzyka kredytowego klienta telefonii komórkowej wydają się być skutecznym narzędziem, które w istotny sposób pozwalają na lepsze zabezpieczenie przychodów.
- 2) Porównując modele neuronowe z modelem logitowym, otrzymano zbliżony wynik klasyfikacji. Procent poprawnych klasyfikacji wynosił odpowiednio 91.2-91.4% dla sieci oraz 90.6% dla modelu logitowego.
- 3) Badając trafność modeli w poszczególnych percentylach, poszegowanych według malejącego prawdopodobieństwa przyjęcia wartości 1 przez zmienną oznaczającą brak płatności (BP), sieciowe modele klasyfikacyjne wykazują przewagę nad modelem logitowym. W tym wypadku różnica w zastosowaniu modelu logitowego wobec modelu neuronowego może sięgać nawet 11 pp. (odpowiednio, 38% wobec 49%), co w istotny sposób przekłada się na trafność klasyfikowania przypadków z dużym poziomem ryzyka, przy jednoczesnym ograniczeniu ponagrania użytkowników dobrych.

### **Zakończenie obrad w sesji B**

Piątek rano, 21 Wrzesień, 10:00



---

# Sesja C

Zastosowania w medycynie i pozostałe

## Program

Środa, 19 Wrzesień

### Sesja C1

*Rożne Zastosowania Metod Klasyfikacji i Analizy Danych*  
Środa popołudniu, 19 Wrzesień, 15:00

---

15:00 Oral

#### **Porównanie wyników klasyfikacji nauczycieli akademickich pod względem wpływu osobowości na popularność wykładów ze statystyki**

Małgorzata Machowska-Szewczyk

*Politechnika Szczecińska, Wydział Informatyki (PS), ul. Żołnierska 49, Szczecin 71-210, Poland*

*e-mail: mmachowska@wi.ps.pl*

Mozolne obliczenia, duża liczba pojęć, zasad, wzorów do zapamiętania – to skojarzenia studentów z przedmiotem statystyka, prowadzonym na wyższych uczelniach. Zazwyczaj osoby o wrodzonej zdolności do abstrakcyjnego myślenia, posługują się pojęciami statystycznymi bez większych problemów. Oprócz zdolności do abstrakcyjnego myślenia, niewątpliwie wpływ na „popularność” statystyki matematycznej wśród studentów mają również zdolności dydaktyczne osoby, prowadzącej przedmiot. Badanie popularności wykładów ze statystyki matematycznej przeprowadzone zostało na podstawie ankiet wypełnianych przez studentów szczecińskich uczelni. Ze względu na to, że pojedynczemu obiektowi przypisano może być więcej niż jeden wariant odpowiedzi na kolejne pytania zawarte w ankiecie, poszczególnym wariantom przyporządkowane zostały wagi, oznaczające procentowy udział osób wybierających dany wariant. W przestrzeni zmiennych, przyjętych jako czynniki wpływające na chęć uczestnictwa w wykładach, dokonano klasyfikacji obiektów symbolicznych. W poszczególnych klasach obiektów podobnych oceniono frekwencję na wykładach.

W artykule porównane zostaną wyniki klasyfikacji nauczycieli akademickich, otrzymane w wyniku wykorzystania zmiennych symbolicznych z wagami z rezultatami klasyfikacji, opartej na elementach zbiorów przybliżonych, zaprezentowanej na konferencji SKAD 2006 i opublikowanej w Taksonomii 14.

---

15:15

Oral

#### **Wykorzystanie analizy czynnikowej w segmentacji behawioralnej na przykładzie skali semiometrycznej**

Magdalena C. Mojsiewicz, Katarzyna Wawrzyniak, Barbara Batorów

*Uniwersytet Szczeciński, Jedności Narodowej 22, Szczecin 70-453, Poland*

*e-mail: mojsiv@uoo.univ.szczecin.pl*

Celem artykułu jest przedstawienie propozycji zastosowania metod wielowymiarowej analizy statystycznej w segmentacji behawioralnej rynku. W dotychczas stosowanych procedurach segmentacji rynku można wyodrębnić trzy etapy: segmentacji społeczno-demograficznej; segmentacji preferencji (w której analizuje się cechy psychograficzne, wzorce konsumpcji oraz preferencje w zakresie warunków zakupu); segmentacji behawioralnej (analizę skłonności pierwotnych). Zaprezentowana metoda wykorzystywana w badaniach TNS OBOP pozwala różnicować nabywców ze względu na trzy różne typy osobowości. Podejście behawioralne stosowane było w psychometrycznych badaniach rynku niemieckiego, obecnie badania w wielu krajach europejskich prowadzone są przy użyciu skali przedziałowej mierzącej odczucia respondentów budzące się odnośnie elementarnych pojęć. W Polsce badania te są wykorzystywane szczególnie w sondażach przedwyborczych, rzadziej w segmentacji rynku. Idea badań zaprojektowanych przy współudziale psychologów miała w pierwotnych założeniach wskazywać media i sposoby najskuteczniejszych kampanii reklamowych. W obecnych metodach rozpoznawania rynku i identyfikacji preferencji klientów coraz więcej zainteresowania budzi zaprezentowana metoda. Znaczenie jej wzrasta w dziedzinie tzw. *Knowledge Intensive Services*. W artykule metoda ta zostanie zilustrowana na przykładzie 500-elementowej próby gospodarstw domowych, na podstawie której badano przydatność metod wielowymiarowej analizy statystycznej w segmentacji rynku ubezpieczeń w Polsce.

---

15:30

Oral

#### **Eksploracyjna analiza polskojęzycznych ofert zatrudnienia specjalistów z zakresu informatyki**

Paweł Lula

*Akademia Ekonomiczna w Krakowie, Katedra Informatyki, Rakowicka 27, Kraków 31-510, Poland*

*e-mail: pawel.lula@post.pl*

Celem pracy jest prezentacja wyników badań dotyczących analizy ofert zatrudnienia specjalistów z zakresu informatyki. Badania obejmują oferty przygotowane w postaci tekstów polskojęzycznych. W trakcie badań wykorzystane zostały metody stosowane na gruncie eksploracyjnej analizy dokumentów tekstowych (*text mining*). Do najważniejszych zagadnień badawczych prezentowanych w pracy należy zaliczyć:

- opracowanie metody pozyskiwania w sposób automatyczny z polskojęzycznych ofert zatrudnienia najistotniejszych wymagań dotyczących poszukiwanych pracowników,
- klasyfikację ofert zatrudnienia specjalistów z zakresu informatyki

przygotowanych w języku polskim i dostępnych w wybranych portalach internetowych,

- przeprowadzoną za pomocą metod text miningowych analizę porównawczą ofert zatrudnienia ze standardami nauczania na studiach wyższych z zakresu informatyki (minima programowe).

Praca składa się z dwóch części - w pierwszej przedstawiono podstawy teoretyczne przeprowadzonych analiz, zaś w drugiej wyniki badań empirycznych przeprowadzonych przy wykorzystaniu dostępnych w Internecie ofert zatrudnienia.

Analiza przeprowadzona została przy użyciu programu *STATISTICA Text Miner* oraz języka *R*.

15:45 Oral

### Wykorzystanie metod eksploracyjnej analizy tekstu do identyfikacji kluczowych zagadnień zawartych w dużych zbiorach publikacji naukowych

Mariusz Grabowski

*Akademia Ekonomiczna w Krakowie, Katedra Informatyki, Rakowicka 27, Kraków 31-510, Poland*

*e-mail: Mariusz.Grabowski@ae.krakow.pl*

Obserwowany na przestrzeni ostatniej dekady rozwój metod i środków przechowywania i udostępniania danych zaowocował zwiększeniem dostępności zbiorów informacji, w szczególności dokumentów tekstowych. Prognozowany przez Ackoffa (1967) zalew informacji stał się faktem, a postulowana przez niego funkcjonalność systemów informacyjnych polegająca na zadaniach selekcji i kondensacji informacji, stała się koniecznością. Rozwój metod analizy danych, w szczególności narzędzi należących do ogólnej kategorii eksploracyjnej analizy tekstu (*Text Mining*), pozwala w wielu przypadkach na automatyczną realizację pierwszej z wyżej wymienionych funkcjonalności.

Lyytinen i King (2004) podjęli obszerną dyskusję na temat teoretycznych podstaw dziedziny systemów informacyjnych zarządzania (SIZ). Autor niniejszego artykułu podejmuje kontynuację tej dyskusji.

W artykule zostanie zaprezentowane wykorzystanie metod *Text Mining* do identyfikacji kluczowych obszarów rozważań teoretycznych dziedziny systemów informacyjnych zarządzania. Jako zbiór danych zostaną wykorzystane streszczenia oraz słowa kluczowe pochodzące z wszystkich (tj. prawie 700) artykułów opublikowanych na przestrzeni 30 lat (1977-2006) w renomowanym (Lista Filadelfijska) czasopiśmie naukowym dziedziny SIZ -- *MIS Quarterly*

#### Literatura:

Ackoff R. L., (1967). *Management Misinformation Systems*, "Management Science", Vol. 14, No. 4, December, 147-156.

Lyytinen K., King J. L., (2004). *Nothing At The Center?: Academic Legitimacy in the Information Systems Field*, "Journal of the Association for Information Systems", Vol. 5 No. 6, 220-246.

#### Przerwa

Środa popołudniu, 19 Wrzesień, 16:00

#### Sesja C2

*Zastosowania w Medycynie*

Środa popołudniu, 19 Wrzesień, 16:15

16:15

Oral

### Badanie efektywności służby zdrowia z wykorzystaniem metod wielowymiarowej analizy porównawczej

Iwona E. Laskowska, Karolina E. Lewandowska

*Uniwersytet Łódzki, Wydział Ekonomiczno-Socjologiczny, P.O.W. 3/5, Łódź 90-255, Poland*

*e-mail: ilaskow@uni.lodz.pl*

Stan zdrowia społeczeństw, jako istotny element jakości życia, w ostatnim czasie jest coraz wyżej ceniony. Jednocześnie starzenie się społeczeństw wielu krajów zmusza do zwiększania nakładów na ochronę zdrowia. Powstaje jednak pytanie, czy większe nakłady gwarantują wzrost efektywności służby zdrowia? Pojawienie się takich obszarów problemowych skłania do prowadzenia rozmaitych analiz umożliwiających podejmowanie jak najlepszych decyzji zmierzających do zagwarantowania społeczeństwu właściwej opieki medycznej.

Celem referatu jest analiza efektywności służby zdrowia w Polsce według województw w 2005 roku. Wskaźnik ten został zdefiniowany jako stosunek nakładów na służbę zdrowia (np. finansowanie, personel medyczny, infrastruktura medyczna) do efektów (np. wskaźniki epidemiologiczne, umieralność niemowląt). W przeprowadzonym badaniu wykorzystano metody wielowymiarowej analizy porównawczej do klasyfikacji województw pod względem efektywności opieki zdrowotnej. Metody te umożliwiły zaprezentowanie przestrzennego zróżnicowania badanego zjawiska w kraju. Wyróżniono grupy województw, w których efektywność opieki zdrowotnej jest najwyższa i te, w których jest stosunkowo niska, wskazując tym samym obszary problemowe. Wyniki analiz zaprezentowano w formie graficznej w postaci wykresów, map, twarzy Chernoffa.

Odpowiednie dane statystyczne zaczerpnięto z publikacji GUS.

#### Bibliografia

Jajuga K., (1993), *Statystyczna analiza wielowymiarowa*, Wydawnictwo Naukowe PWN, Warszawa.

Leowski J., (2004), *Polityka zdrowotna a zdrowie publiczne. Ochrona zdrowia Młoda gospodarce rynkowej*, CeDeWu, Warszawa

Młodak A., (2006), *Analiza taksonomiczna w statystyce regionalnej*, Difin, Warszawa.

Portal Internetowy Głównego Urzędu Statystycznego, [www.stat.gov.pl](http://www.stat.gov.pl).

Suchecka J., (1998), *Ekonometria ochrony zdrowia*, Absolwent, Łódź.

16:30

Oral

### Zagregowane modele dyskryminacyjne i regresyjne w prognozowaniu czasu pobytu na OIOM pacjentów z chorobą wieńcową

Małgorzata A. Misztal

Uniwersytet Łódzki, Katedra Metod Statystycznych (UE), Rewolucji 1905 r. nr 41, Łódź 90-214, Poland

e-mail: mmisztal@toya.net.pl

W pracy Misztal (2007) podjęto próbę wyodrębnienia przedoperacyjnych czynników ryzyka mających wpływ na przedłużenie czasu pobytu na OIOM pacjentów z chorobą wieńcową leczonych operacyjnie. Prognozowanie czasu pobytu pacjenta na OIOM nie należy do zadań łatwych, występują tu bowiem rozkłady o silnej asymetrii prawostronnej, z dużą liczbą wartości nietypowych.

W prowadzonych badaniach do prognozowania czasu pobytu na OIOM wykorzystano między innymi drzewa klasyfikacyjne i regresyjne, uzyskane wyniki nie były jednak całkowicie satysfakcjonujące. Wpływ na to mógł mieć między innymi brak stabilności utworzonych modeli. Obecnie coraz częściej dokonuje się agregacji indywidualnych modeli w jeden model zagregowany, dla którego używamy błąd predykcji mniejszy niż błąd każdego z modeli składowych.

W referacie dokonano próby zastosowania zagregowanych modeli drzew klasyfikacyjnych i regresyjnych (*Bagging* – Breiman 1996, *Boosting* - Freund i Shapire 1997, *Random forest* – Breiman 2001, *BART* - Chipman, George, McCulloch 2005) do poprawy dokładności predykcji czasu pobytu na OIOM pacjentów poddanych leczeniu operacyjnemu choroby wieńcowej.

#### Literatura

- Breiman L. (1996): Bagging predictors. *Machine Learning*, 24, 123-140.
- Breiman L. (2001): Random forests. *Machine Learning* 45, 5-32.
- Chipman H. A., George E. I., McCulloch R. E. (2006), *BART: Bayesian Additive Regression Trees*, Technical report, University of Chicago.
- Freund Y., Schapire R.E. (1997): A decision-theoretic generalization of on-line learning and an application to boosting, *Journal of Computer and System Sciences* 55, 119-139.
- Misztal M. (2007), *Wybrane metody analizy i prognozowania czasu pobytu na OIOM pacjentów z chorobą wieńcową*, (w:) **Taksonomia 14. Klasyfikacja i analiza danych – teoria i zastosowania**, pod redakcją K. Jajugi i M. Walesiaka, Prace Naukowe Akademii Ekonomicznej we Wrocławiu, Wydawnictwo AE we Wrocławiu, Wrocław (w druku).

16:45

Oral

### Badanie poziomu świadomości praw pacjenta w środowisku medycznym

Joanna Banaś, Małgorzata Machowska-Szewczyk

Politechnika Szczecińska, Wydział Informatyki (PS), ul. Żołnierska 49, Szczecin 71-210, Poland

e-mail: jbanas@wi.ps.pl

W artykule przedstawiono klasyfikację pracowników służby zdrowia w województwie zachodniopomorskim, wykorzystując metodę klasyfikacji opartą na uogólnionej mierze odległości GDM. Badanie respektowania praw dziecka przez pielęgniarki, uczestniczące w ankiecie zostało przeprowadzone w roku 2003. Do określenia globalnego poziomu świadomości w zakresie respektowania praw dziecka wykorzystano wyniki uzyskane z badania korelatów świadomości oraz metody przypadków. Wydzielone zostały następujące elementy świadomości: wiedza pielęgniarek o prawach dziecka, rozumienie i wartościowanie tych praw oraz działanie ukierunkowane na ich poszanowanie. Zastosowanie metody przypadków miało na celu zbadaanie respektowania praw dziecka przez pielęgniarki. W pracy zastosowano zaproponowaną przez Walesiaka metodę mierzenia odległości między obiektami, która może być zastosowana do zmiennych wyrażonych w skali ilorazowej, przedziałowej oraz porządkowej. Dla otrzymanej klasyfikacji przeprowadzono analizę zależności ze względu na cechy różnicujące (wiek, staż pracy, wykształcenie, miejsce pracy).

#### Zakończenie obrad w sesji C

Środa popołudniu, 19 Wrzesień, 17:15

## Czwartek, 20 Wrzesień

### Sesja C4

Analiza Rynku Pracy

Czwartek rano, 20 Wrzesień, 11:45

11:45

Oral

### Zmienna syntetyczna z medianą w analizie lokalnego rynku pracy

Aleksandra Witkowska, Marek Witkowski

e-mail: a.witkowska@ae.poznan.pl

Rynek pracy jest tą kategorią ekonomiczną, która znajduje się w centrum zainteresowania zarówno władz państwowych jak i samorządowych. Oceny sytuacji na tym rynku dokonuje się z reguły w oparciu o pojedyncze mierniki, z których podstawowymi są stopa bezrobocia, współczynnik aktywności zawodowej oraz wskaźnik zatrudnienia. Tymczasem rynek pracy jest kategorią, ze statystycznego punktu widzenia, bardzo złożoną i powinien być siłą rzeczy charakteryzowany przez wiele cech diagnostycznych. Dopiero wówczas można uzyskać kompleksowy jego obraz w przekroju różnych jednostek terytorialnych, w szczególności gmin i powiatów.

W pracy podejmujemy próbę zastosowania do oceny lokalnego rynku pracy zmiennej syntetycznej opartej na medianie. Sądzymy bowiem, biorąc pod uwagę, że dyferencjacja cech diagnostycznych opisujących ten rynek jest bardzo duża, że zmienna ta będzie do opisu takiego zjawiska użyteczna.

W związku z tym stawiamy w pracy następujące tezy:

1. Zmienna syntetyczna pozwala na dokonanie oceny stanu lokalnego rynku pracy na poziomie powiatów w jednoznaczny sposób.
2. Zmienna syntetyczna daje możliwość przeprowadzenia rankingu powiatów według stanu rynku pracy z uwzględnieniem zarówno jego strony popytowej jak i podażowej.
3. Na podstawie skonstruowanej zmiennej syntetycznej można dokonać klasyfikacji powiatów na grupy typologiczne według stanu badanego zjawiska złożonego.
4. Zmienna syntetyczna z medianą pozwala ocenić odległość danego powiatu od powiatów-sąsiadów.

Weryfikacji wymienionych tez dokonaliśmy przeprowadzając badania empiryczne, które objęło swoim zakresem powiaty województwa wielkopolskiego i dotyczyło lat 2000 - 2005. Źródłem danych do badania były publikacje Urzędu Statystycznego w Poznaniu: "Województwo wielkopolskie, podregiony, powiaty, gminy" oraz "Rynek pracy w województwie wielkopolskim".

---

12:00 Oral

---

### **Aktywność ekonomiczna ludności: klasyfikacja osób za pomocą wielomianowych modeli logitowych oraz jej związek z modelami hazardu dla czasów trwania.**

Joanna M. Landmesser

*Szkoła Główna Gospodarstwa Wiejskiego (SGGW), Nowoursynowska 166, Warszawa 02-787, Poland*

*e-mail: jmlg@poczta.onet.pl*

Za pomocą wielomianowego modelu logitowego dokonamy określenia przynależności osób do następujących klas: pracujący w pełnym wymiarze czasu, pracujący w niepełnym wymiarze czasu, bezrobotni, nieaktywni zawodowo (renciści, emeryci). Dobrane do modelu zmienne objaśniające będą dotyczyły takich charakterystyk osób, jak wiek, płeć, poziom wykształcenia, stan cywilny, wielkość miejscowości zamieszkania, typ wyuczonego zawodu, czy oczekiwane zarobki. Dodatkowo wprowadzimy zmienną oznaczającą aktualny czas pobytu w bieżącym stanie. Zweryfikujemy statystyczną istotność wpływu tych wielkości.

Wykażemy, iż uzyskane oceny parametrów w wielomianowym modelu logitowym są takie same, jak szacunki parametrów dla modelu hazardu z wieloma wyjściami, czyli tzw. modelu ryzyk konkurencyjnych. Zastosujemy modele hazardu o czasie dyskretnym – przynależne do metod analizy czasu przetrwania -, za pomocą których można szacować długość czasu trwania aktywności zawodowej, czy też długość czasu poszukiwania zatrudnienia na rynku pracy. Modele ryzyk konkurencyjnych pozwolą opisać warunkowe prawdopodobieństwa przejścia między poszczególnymi interesującymi nas stanami.

Estymacji modeli dokonamy na podstawie danych uzyskanych z bazy pewnego powiatowego urzędu pracy.

Literatura:

Allison P. (1982), Discrete time methods for the analysis of event histories, w: S. Leinhardt (red.) Sociological Methodology 1982, San Francisco, 61-98.

Cramer J.S. (2001), An introduction to the logit models for economists, Londyn.

Frątczak E., Gach-Ciepiela U., Babiker H. (2005), Analiza historii zdarzeń., Elementy teorii, wybrane przykłady zastosowań, Warszawa.

Yamaguchi K. (1991), Event History Analysis, Applied Social Research Methods Series, Vol.28, Sage Publ., London-New Delhi.

---

12:15 Oral

---

### **Analiza taksonomiczna ofert pracy pojawiających się w dodatku „Praca” w regionie Łódzkim w 2006 roku**

Iwona Staniec, Marek Martin

*Technical University of Łódź, Łódź, Poland*

*e-mail: istan@p.lodz.pl*

Rynek pracy charakteryzują dwa segmenty: podaż siły roboczej i popyt na pracę. Siła robocza określa liczbę ludności zainteresowaną i gotową wykonywać pracę. Natomiast popyt na pracę określa liczbę miejsc pracy, którą oferuje gospodarka w określonych warunkach społeczno-ekonomicznych. Na wzajemne relacje między podażą siły roboczej a popytem na pracę ogromny wpływ wywierają mechanizmy rynkowe. Wielkość i struktura popytu na pracę jest zależna od wielu czynników, m.in. od zapotrzebowania na produkty i usługi, kondycji ekonomicznej przedsiębiorstw, kosztów pracy.

Popyt na pracę określany jest poprzez:

- zagospodarowane miejsca pracy, określane aktualną liczbą pracujących,

- wolne miejsca pracy.

Celem analizy ofert pracy jest uzyskanie informacji na temat wolnych oraz nowo utworzonych miejsc pracy. Z badań GUS wynika, że w Polsce największą liczbą wolnych miejsc pracy dysponowały jednostki realizujące działalność w zakresie obsługi nieruchomości i firm, handlu i napraw – po 4,5 tys. wolnych miejsc pracy oraz budownictwa - 3,9 tys. wolnych miejsc pracy, najliczniejsza liczba miejsc pracy przeznaczona była dla pracowników z wykształceniem zasadniczym zawodowym (41,6%), a w drugiej kolejności dla pracowników z wykształceniem wyższym (24,0%).

Celem pracy jest przedstawienie przy użyciu metod taksonomicznych ofert pracy pojawiających się w dodatku „Praca” w regionie Łódzkim w 2006 roku, ukazanie ich zróżnicowania ze względu na czas oraz pojawiające się w okręgu łódzkim inwestycje. Jednym z dążeń w ramach prowadzonych badań jest wskazanie wymagań ogólnych i specjalizacyjnych stawianych przez pracodawców w poszczególnych grupach ofert oraz próba porównania uzyskanych wyników z danymi BAEL dla regionu łódzkiego.

Badania taksonomiczne zostaną wzbogacone o statystyczną analizę danych w układzie przestrzenno-czasowym oraz regresję logitową.

Praca naukowa finansowana ze środków na naukę w latach 2006-2008 jako projekt badawczy MNiSW  
N115 047 31/2071

## Piątek, 21 Wrzesień

## Sesja C5

*Analiza Gospodarstw Domowych*  
Piątek rano, 21 Wrzesień, 9:00

9:00 Oral

### Wielowymiarowe aspekty ubóstwa – statystyczna analiza zasobów dóbr trwałego użytku w gospodarstwach domowych rolników

Hanna Dudek

*Warsaw Agricultural University (SGGW), Nowoursynowska 166, Warszawa 02-787, Poland*

*e-mail: hdudek@mors.sggw.waw.pl*

W ujęciu klasycznym, identyfikacja sfery ubóstwa dokonywana jest wyłącznie ze względu na sytuację dochodową gospodarstw domowych. Taki sposób wyodrębniania rodzin ubogich jest pewnym uproszczeniem i może doprowadzić do pominięcia ważnych aspektów rzeczywistej sytuacji; stąd też w ostatnich latach coraz więcej zwolenników mają metody traktujące ubóstwo jako zjawisko wielowymiarowe. W podejściu wielowymiarowym rozważa się problematykę związaną z pozbawieniem możliwości zaspokojenia potrzeb gospodarstw domowych z powodu braku nagromadzenia odpowiednich zasobów materialnych (*deprivation*) oraz aspekty kulturowe i socjologiczne powodujące wykluczenie społeczne (*social exclusion*).

W pracy podjęto temat ubóstwa wielowymiarowego związanego z niewystarczającymi zasobami dóbr trwałego użytku. Analizę tego zjawiska przeprowadzono metodami, zaproponowanymi przez Deutscha i Silbera, wykorzystując wielomianowe modele logitowe. Ponadto podjęto próbę porównania wyników dotyczących identyfikacji gospodarstw domowych rolników narażonych na ubóstwo metodami klasycznymi oraz metodami przedstawionymi w pracy. Podstawą prezentowanych analiz są dane z badań budżetów gospodarstw domowych zrealizowanych przez GUS w 2004 roku.

#### Literatura:

1. Deutsch J., Silber J.: *The Order of Acquisition of Durable Goods and the Multidimensional Measurement of Poverty*. Conference Paper "The Many Dimension of Poverty", Brasilia, 29-31.08.2005.
2. Greene W.H.: *Econometric Analysis*. Prentice Hall Inc., Upper Saddle River, New Jersey, 2000.
3. Gruszczyński M.: *Modele i prognozy zmiennych jakościowych w finansach i bankowości*. Oficyna Wydawnicza SGH. Warszawa, 2002.
4. Panek T.: *Metody pomiaru ubóstwa*, Wiadomości Statystyczne, nr 7, str. 1-12, 2005.

9:15

Oral

### Klasyfikacja gospodarstw rolniczych w oparciu o dane w postaci szeregów czasowych

Joanna Kisielińska

*Szkoła Główna Gospodarstwa Wiejskiego (SGGW), Nowoursynowska 166, Warszawa 02-787, Poland*

*e-mail: jkisielinska@mors.sggw.waw.pl*

Celem badań, było zbudowanie funkcyjnych i sieciowych modeli dyskryminacyjnych umożliwiających prognozowanie sytuacji finansowej gospodarstw rolniczych. Podstawę klasyfikacji stanowił dochód rolniczy uzyskany w roku 2002. Jeśli dochód był mniejszy od mediany, gospodarstwo zaliczono do klasy I (gospodarstwa słabe), w przeciwnym wypadku do II (gospodarstwa dobre). Modele budowano w oparciu o szeroki zestaw wskaźników finansowych obejmujący wskaźniki płynności finansowej, rentowności, obrotowości, sprawności gospodarowania, wspomagania finansowego, oraz stanowiących charakterystykę majątku trwałego. Dodatkowo uwzględniono powierzchnię użytków rolnych (wyrażoną w hektarach przeliczeniowych), wiek osoby prowadzącej gospodarstwo, poziom intensywności produkcji, oraz poziom intensywności organizacji produkcji.

W literaturze znaleźć można różne metody uwzględniania w modelach klasyfikacyjnych danych w postaci szeregów czasowych. Jako przykłady można podać wielomianowe funkcje trendu, metodę polegającą na wprowadzaniu danych pochodzących nie z jednego, lecz kilku okresów (każda opóźniona zmienna traktowana jest wówczas jako pojedyncza zmienna niezależna), czy uwzględnienie w modelu zamiast danych z kilku okresów wcześniejszych obliczonej z nich średniej ruchomej.

W przeprowadzonych badaniach wykorzystano natomiast metodę ważonej średniej ruchomej, wychodząc z założenia, że w przypadku wskaźników finansowych wpływ ich na sytuację finansową maleje wraz ze wzrostem opóźnienia. Zastosowano dwie stałe wygładzania 2 i 3 lata, a dla każdej stałej wygładzania po trzy warianty wag. Zarówno funkcyjne jak i sieciowe modele klasyfikacyjne budowano w oparciu o dane pochodzące z jednego roku (poprzedzającego prognozę, czyli z roku 2001), oraz dane w postaci uśrednionego szeregu czasowego (z lat 1999-2001). W efekcie powstało 7 funkcyjnych i 7 sieciowych modeli klasyfikacyjnych, spośród których wybrano najlepsze. Modele te wykorzystano do klasyfikacji 704 gospodarstw rolniczych. Okazało się, że wszystkie modele, zarówno funkcyjne jak i sieciowe, zbudowane w oparciu o dane w postaci szeregów czasowych były lepsze od modeli opartych na danych jedynie z jednego roku – poprzedzającego prognozę. Stwierdzono ponadto, że wszystkie modele sieciowe były lepsze od wszystkich modeli funkcyjnych.

Najlepszy model funkcyjny prawidłowo zaklasyfikował 78,98% gospodarstw (79,83% należących do klasy I i 78,13% z klasy II). Natomiast najlepszy model sieciowy poprawnie rozpoznał 82,24% gospodarstw (80,97% należących do klasy I i 83,52% z klasy II). Wyniki te można uznać za zadowalające, choć prawdopodobnie rozszerzenie informacji o gospodarstwie mogłoby poprawić wyniki klasyfikacji.

9:30

Oral

---

### Wykorzystanie analizy koszykowej do identyfikacji zachowań konsumpcyjnych gospodarstw domowych w Polsce

Izabela Kurzawa, Feliks Wysocki

*Akademia Rolnicza (AR), ul. Wojska Polskiego 28, Poznań 60-637, Poland*

*e-mail: kurzawa@au.poznan.pl*

Eksploatacja danych (ang. *data mining*) jest jedną z najbardziej rozwijanych dziedzin naukowych, która integruje wiele dyscyplin takich jak: statystyka, ekonometria, systemy baz danych czy optymalizacja. Mając do dyspozycji bardzo wiele danych powstaje pytanie w jaki sposób efektywnie i racjonalnie wykorzystać te nagromadzone informacje. Jedną z metod eksploatacji danych jest klasa metod odkrywania asocjacji, które służą do badania różnego rodzaju zależności w bazie danych. Metody te obejmują głównie odkrywanie zależności pomiędzy obiektami.

Celem pracy jest próba wykorzystania metod asocjacji - analizy koszykowej do wykrycia zależności ukrytych w bazie danych zawierających informacje o konsumpcji w gospodarstwach domowych w Polsce i przedstawienie ich w postaci prostych reguł. Reguły te mogą dotyczyć na przykład zwyczajów zakupowych grup gospodarstw domowych lub prawidłowości w korzystaniu z usług wybranego typu. Otrzymane reguły noszą nazwę reguł asocjacyjnych.

Przeprowadzona analiza oprócz aspektu poznawczego będzie posiadać również wymiar praktyczny. Stanowić będzie cenne źródło informacji o zachowaniach konsumentów i może być wykorzystana do szybkiego budowania, aktualizowania i wdrażania skutecznych i trafnych strategii marketingowych w sektorze rolno-żywnościowym.

---

9:45

Oral

---

### Klasyfikacja rozkładów dochodów gospodarstw domowych

Piotr Łukasiewicz<sup>1</sup>, Grzegorz Koszela<sup>1</sup>, Arkadiusz J. Orłowski<sup>2,3</sup>

**1.** *Szkoła Główna Gospodarstwa Wiejskiego (SGGW), Nowoursynowska 166, Warszawa 02-787, Poland* **2.** *Warsaw Agricultural University, Department of Food Hygiene and Public Health (SGGW), Nowoursynowska 166, Warszawa 02-787, Poland* **3.** *Polish Academy of Sciences, Institute of Physics, al. Lotników 32/46, Warszawa 02-668, Poland*

*e-mail: plukasiewicz@mors.sggw.waw.pl*

W pracy przedstawiono klasyfikację rozkładów dochodów gospodarstw domowych z 16 województw Polski. Materiał badawczy stanowiły mikroobserwacje dotyczące dochodów gospodarstw domowych pochodzące z badań budżetów gospodarstw domowych prowadzonych przez GUS. Empiryczne rozkłady dochodów przybliżono trzyparametrycznym modelem Daguma. W ten sposób, każdy z 16 rozkładów dochodów utożsamiony został z pewnym punktem w trójwymiarowej przestrzeni. Otrzymane obiekty poddano klasyfikacji stosując wybrane aglomeracyjne metody taksonomiczne: metodę najdalszego sąsiedztwa oraz metodę Warda. Badając wewnątrzgru-

powe i międzygrupowe zróżnicowanie dokonano wyboru najlepszego podziału rozkładów dochodów.

### Zakończenie obrad w sesji C

Piątek rano, 21 Wrzesień, 10:00



# Sesja plakatowa

## Program

### Środa, 19 Wrzesień

#### Sesja Plakatowa

Środa popołudniu, 19 Wrzesień, 12:00

#### Zakończenie sesji plakatowej

Środa popołudniu, 19 Wrzesień, 13:00

## Plakaty

### Środa, 19 Wrzesień

#### Sesja Plakatowa

Środa popołudniu, 19 Wrzesień, 12:00

12:00	Poster	1
-------	--------	---

#### Estymacja typu Winsora w badaniach podmiotów gospodarczych

Grażyna Dehnel

*Akademia Ekonomiczna w Poznaniu, al. Niepodległości 10, Poznań 60-967, Poland*

*e-mail: g.dehnel@ae.poznan.pl*

Estymacja punktowa dla małych domen w statystyce gospodarczej, z uwagi na własności cech, nie należy do prostych. Jak wiadomo zmienne dotyczące podmiotów gospodarczych charakteryzują się silną prawostronną asymetrią, znacznym zróżnicowaniem i dużą koncentracją. Przyczyną niejednorodnych rozkładów podmiotów jest między innymi obecność obserwacji odstających i dość znaczny odsetek jednostek, dla których zmienne przyjmują wartości zerowe. Stąd też stosowanie klasycznych estymatorów jest niewłaściwe. Często bowiem nie zachowują swoich własności takich jak: nieobciążoność, czy duża efektywność.

Jedno z podejść do estymacji stosowanej w statystyce gospodarczej polega na dokonaniu pewnej modyfikacji, która ma na celu „uodpornienie” estymatora na duże reszty. Jednostki wylosowane do próby, u których wartości cechy wykraczają poza pewne ustalone wartości graniczne zostają zmienione. Ta zmiana może odbywać się poprzez zmodyfikowanie wartości wag związanych z obserwacjami odstającymi lub poprzez zmodyfikowanie wartości zmiennych badanych tych jednostek. Przykładem tego drugiego podejścia jest estymator Winsora. Proponuje dostosowanie jednostek znajdujących się poza wartościami granicznymi do jednostek bliskich wartościom granicznym. Dalej szacunek przebiega zgodnie z estymacją typu GREG, czy HT. Największą trudność w estymacji Windsora stwarza dobranie odpowiednich wartości granicznych pozwalających nad

dokonywanie podziału danych na dwie grupy. Jedną z nich stanowią dane wykorzystane do budowy modelu, drugą - obserwacje odstające.

W referacie przedstawione są rezultaty badania symulacyjnego, w którym ocenie poddano różne techniki stosowane do wyznaczania wartości granicznych. W obliczeniach wykorzystano wyniki badania reprezentacyjnego SP3 dotyczącego mikroprzedsiębiorstw wzbogacone o informacje pochodzące z rejestrów podatkowych Ministerstwa Finansów.

Należy podkreślić, że zastosowanie rejestrów administracyjnych, w tym także podatkowych w badaniach z zakresu statystyki gospodarczej jest w krajach Unii Europejskiej bardzo powszechne. Polska znajduje się na początkowym etapie przemian dotyczących wykorzystywania przez statystykę publiczną informacji ze źródeł administracyjnych. Warto zatem bliżej przyjrzeć się nowym technikom estymacji pośredniej, które korzystając z dodatkowych źródeł informacji dostarczają szacunków bardziej precyzyjnych, są bardziej odporne na wartości odstające.

12:00	Poster	2
-------	--------	---

#### Klasyfikacja danych rolniczych w oparciu o metody analizy skupień stosowane w genetyce

Stanisław K. Jabłonowski

*Szkoła Główna Gospodarstwa Wiejskiego (SGGW), Nowoursynowska 166, Warszawa 02-787, Poland*

*e-mail: sj@mors.sggw.waw.pl*

Metody klasyfikacji danych, taksonometria znajdują się w polu stałego zainteresowania badaczy. Proponuje się nowe podejścia, nowe algorytmy różnych typów klasyfikacji. Dotyczy to dużych zbiorów danych opisujących obiekty o dużej liczbie cech.

Takie dane występują np. w genetyce. Badania tych danych mogą mieć na celu przeprowadzenie analiz podobieństwa genetycznego różnych organizmów żywych, ich populacji itp. Metody stosowane w genetyce, mimo że specyficznie ukierunkowane są chyba jednak na tyle ogólne, że mogą być przeniesione też do innych dziedzin i do rozwiązywania innych problemów. Obiekty wieloatrybutowe związane z rolnictwem, takie jak np. gospodarstwa rolne, czy gminy wiejskie są klasyfikowane różnymi sposobami. Celem niniejszego opracowania jest sprawdzenie, czy pewnych metod taksonomicznych stosowanych w genetyce nie można przenieść do tak odległych dla genetyki dziedzin jak badania ekonomiczno-rolnicze. Danymi źródłowymi do stosowania tych metod, służących w efekcie do oceny podobieństw genetycznych organizmów żywych i populacji tych organizmów, mogą być wyniki reakcji łańcuchowej polimerazy, PCR (Polymerase Chain Reaction). Technika PCR umożliwia namnażanie fragmentów DNA. Elektroforegramy, które obrazują wyniki doświadczeń zawierają ścieżki, złożone z prążków sygnalizujących istnienie, często nie znanego, ale określonego fragmentu łańcucha DNA. Każda taka ścieżka odpowiada konkretnym próbkom materiału genetycznego, pochodzącym z konkretnego organizmu żywego. Prążki na tych ścieżkach można opisać liczbami określającymi ich położenie w ramach ścieżki. Posiadając takie zestawy liczb można poddać je czysto statystycznej obróbce, mającej na celu obliczanie odległości między ścieżkami, a zatem i badanymi organizmami żywymi. Otrzymane wyniki świadczą o mniejszym lub większym

zróznicowaniu genetycznym populacji, można wyróżniać skupienia. Występuje tu jednak szereg zagadnień logicznych, jak np. jak ustalić relację identyczności między prążkami z różnych ścieżek, gdyż często ten sam fragment DNA jest reprezentowany przez różne, choć bliskie sobie liczby. Zanim zatem będzie się obliczać odległości genetyczne między ścieżkami, czyli organizmami żywymi, trzeba prążki podzielić na grupy o dużym wewnętrznym podobieństwie. Zagadnienia te można różnie rozwiązywać, ale ostateczne decyzje są tylko w przybliżeniu ścisłe. Po podjęciu takich decyzji można stosować dalszą obróbkę statystyczną.

Wydaje się, że dane o pewnych obiektach związanych z rolnictwem, obiektach wielocechowych, można potraktować podobnie jak dane pochodzące z analizy PCR i podobnie je przetwarzać jak tamte. Każde np. badane gospodarstwo z pewnej grupy można przedstawić jako taką ścieżkę, czyli „zestaw prążków”. Wydaje się, że ostateczne zestawy liczbowe można przetwarzać statystycznie podobnie jak dane genetyczne.

W niniejszej pracy jest opisana metoda RAPD, czyli metoda amplifikacji losowej, stanowiąca pewną odmianę reakcji PCR, sposób obróbki statystycznej danych otrzymanych w ten sposób, a następnie próba przeniesienia takich sposobów obliczeniowych do danych rolniczych. Dane te opisują obiekty takie jak gospodarstwa rolne. Atrybutami obiektów są pewne cechy ekonomiczne.

12:00	Poster	3
-------	--------	---

### **Analiza ryzyka szacowania rezerw IBNR w zakładzie ubezpieczeniowym – modele stochastyczne dla danych zagregowanych i indywidualnych.**

Marek Karwański, Wiesław Szczesny

*SAS Institute Sp. z o.o., Gdańska 27/31, Warszawa 01-633, Poland  
Szkoła Główna Gospodarstwa Wiejskiego (SGGW), Nowoursynowska 166, Warszawa 02-787, Poland*

*e-mail: anstat@jeet.pl*

Jedną z podstawowych funkcji aktuariatu, w zakładzie ubezpieczeniowym, jest szacowanie rezerw. Część z tych rezerw – rezerwy na szkody zaistniałe lecz nie zgłoszone do dnia, na który tworzona jest rezerwa – nosi nazwę IBNR. Ich szacowanie ma charakter prognozy przyszłych wypłat i jest obliczane przy pomocy modeli stochastycznych. W praktyce stosowane są metody analizy danych zagregowanych w tzw. trójkąty szkodowe.

Mimo stosowania zaawansowanych modeli, uzyskiwane wyniki obarczone są dużym błędem. Szczególnie duże różnice występują przy szacowaniu rezerw dla szkód ostatnio zaistniałych – a są one najbardziej „ciekawe” z punktu widzenia zakładu ubezpieczeniowego.

W literaturze pojawiają się próby wprowadzenia modeli dla danych indywidualnych. Są one oparte na analizach danych obciętych. Poważną przeszkodą w stosowaniu tych metod jest koszt pozyskania danych indywidualnych oraz konieczność stosowania bardzo wyrafinowanych metod eliminujących dane nietypowe oraz uzupełnienie braków danych.

Celem niniejszej pracy było porównanie wyników szacowania rezerw metodami tradycyjnymi: model Macka, modele GLM z szacowaniem zmienności metodą bootstrap, metody Bayesa MCMC z me-

todami analizy regresji dla danych obciętych na indywidualnych rejestrach.

Analizy przeprowadzone zostały dla trzech zestawów danych:

- danych jednej z dużych amerykańskich firm ubezpieczeniowych (z uwagi na poufność zostały one nieznacznie zaburzone)
- dwu zestawów danych wygenerowanych na podstawie pewnych modeli, które uwzględniają segmentację klientów w zakresie zróżnicowania pod względem wartości „potencjalnej szkodowości”

Takie porównanie, może być jednym z głosów w dyskusji, w kontekście trwającego aktualnie procesu tworzenia standardów rachunkowości ubezpieczeniowej oraz systemów monitorujących adekwatność rezerw.

12:00	Poster	4
-------	--------	---

### **Zastosowanie wybranych modeli tobitowych do opisu tygodniowej liczby godzin pracy**

Jadwiga H. Kostrzewska

*Akademia Ekonomiczna w Krakowie, Katedra Statystyki (AE), Rakowicka 27, Kraków 31-510, Poland*

*e-mail: Jadwiga.Kostrzewska@ae.krakow.pl*

Ekonometryczne modele zmiennych cenzurowanych wywodzą się od standardowego modelu tobitowego – tj. pierwszego modelu, w którym zmienną cenzurowaną rozważano w kontekście regresji. Model ten został wprowadzony w 1958 r. przez J. Tobina, który zauważył, że wnioskowanie na podstawie danych bez uwzględnienia cenzurowania lub ucięcia w przypadku, gdy ono występuje, zawsze jest obciążone pewnym błędem. W oparciu o ten model zostały skonstruowane kolejne, bardziej złożone modele, tworzące rodzinę modeli zmiennej zależnej ograniczonej (*limited dependent variable model*, LDV). W zależności od tego czy zmienna zależna podlega ucięciu, cenzurowaniu czy tzw. samoselekcji (*self-selection*, *sample selection*) można rozważać: modele regresji uciętej (*truncated regression model*) oraz różne cenzurowane modele regresji (*censored regression model*) czyli modele tobitowe (*tobit model*).

Warto nadmienić, iż w 2000 r. laureatem nagrody Nobla w dziedzinie ekonomii został J. Heckman. Jego główne osiągnięcia naukowe stały się podstawą szerokich badań empirycznych w różnych dziedzinach społeczno-ekonomicznych. Prace J. Heckmana dotyczyły m.in. zagadnień rynku pracy, w tym analizy siły roboczej, wysokości zarobków, okresu pozostawania bez pracy, oceny programów rynku pracy, dyskryminacji płacowej itp. Modele ekonometryczne rozważane przez J. Heckmana to modele wywodzące się od modeli zmiennych cenzurowanych.

Celem prowadzonych prac badawczych, będących częścią opracowywanej rozprawy doktorskiej, jest prezentacja ekonometrycznych modeli zmiennych cenzurowanych oraz ich zastosowanie w przypadku danych dotyczących rynku pracy w Polsce. W referacie przedstawione zostaną wybrane modele tobitowe. Następnie podjęta będzie próba opisu, za pomocą omawianych modeli, tygodniowej liczby godzin pracy z uwzględnieniem cech wpływających na fakt posiadania pracy lub pozostawania bezrobotnym. Dokonane zostanie porównanie otrzymanych wyników z rezultatami zastosowania mo-

delu uciętego i klasycznego modelu regresji wielorakiej. Podstawą prowadzonych rozważań będą dane pochodzące z kwartalnego badania aktywności ekonomicznej ludności (BAEL) prowadzonego przez GUS.

12:00 Poster 5

### Ekonometryczne modelowanie asymetrii reakcji klienta na zmiany cen

Tadeusz Kufel<sup>1</sup>, Marcin Błażejowski<sup>2</sup>

1. Uniwersytet Mikołaja Kopernika w Toruniu, WNEiZ, Katedra Ekonometrii i Statystyki (UMK), Gagarina 13A, Toruń 87-100, Poland 2. Wyższa Szkoła Bankowa (WSB), Toruń 87-100, Poland

e-mail: tadeusz.kufel@umk.pl

Praca zawiera omówienie konstrukcji ekonometrycznych modeli popytu konsumpcyjnego budowanych na podstawie danych dziennych o transakcjach sprzedażowych w supermarkecie uwzględniających asymetrię reakcji klienta na zmiany cen. Wyznaczono współczynniki elastyczności cenowej uwzględniając asymetrię w reakcji na **wzrost ceny** i **spadek ceny** oraz zaproponowano kilka typów konstrukcji modeli do opisu tego typu zachowań konsumentów.

12:00 Poster 6

### Wykorzystanie fundamentalnych baz danych w analizie portfelowej

Małgorzata L. Łuniewska

Uniwersytet Szczeciński, Jedności Narodowej 22, Szczecin 70-453, Poland

e-mail: mmmlun@uoo.univ.szczecin.pl

Analiz portfelowa jest jedną z metod długoterminowego inwestowania na rynku kapitałowym. Budowa portfela papierów wartościowych nie jest prosta i wiąże się z wieloma czynnikami. W tej kwestii jednym z istotnych problemów jest dywersyfikacja ryzyka portfela w ujęciu poziomym i pionowym. Kluczowym staje się odpowiedni wybór walorów, na podstawie których portfel zostanie zbudowany. Ich jakość przesądza o wynikach portfela: ryzyku i stopie zwrotu. Z punktu widzenia długoterminowego inwestowania ważne jest uwzględnienie w portfelu akcji spółek o dobrych podstawach fundamentalnych. Pojawia się więc pytanie, czy fakt utrzymywania przez spółkę stabilności fundamentalnej w czasie ma wpływ na jakość inwestycji, a tym samym na poziom stopy zwrotu i ryzyko tak utworzonego portfela. W tym zakresie mogą być wykorzystane fundamentalne bazy danych, zbudowane przy wykorzystaniu metod wielowymiarowej analizy porównawczej. Zmienne wykorzystane w badaniu pochodzą z analizy fundamentalnej, a konkretnie z analizy wskaźnikowej.

Badaniem objęto spółki notowane na Gieldzie papierów Wartościowych w Warszawie. Okres analizy dotyczy lat 2002-2005.

12:00 Poster 7

### Analiza zapotrzebowania na kwalifikacje kadry inżynierskiej i menedżerskiej w regionie łódzkim na podstawie wyników badania ankietowego

Iwona Staniec, Anna Szmit, Jan Żółtowski

Technical University of Łódź, Łódź, Poland

e-mail: istan@p.lodz.pl

12:00 Poster 8

### Analiza czynnikowa z wykorzystaniem środowiska R

Mirosława Sztemberg-Lewandowska

Akademia Ekonomiczna (AE), Nowowiejska3, Jelenia Góra 58-500, Poland

e-mail: mirosława.sztemberg-lewandowska@ae.jgora.pl

Analiza czynnikowa służy do wykrywania optymalnej grupy czynników głównych, tłumaczących korelację między zmiennymi obserwowalnymi. Procedura analizy czynnikowej polega na dokonywaniu szeregu następujących po sobie operacji obliczeniowych w wyniku których otrzymuje się strukturę czynnikową badanych zmiennych. W trakcie wyodrębniania czynników badacz zmuszony jest do podejmowania szeregu arbitralnych decyzji, które w sposób oczywisty wpływają na końcowe rezultaty. Decyzje te dotyczą:

- selekcji zmiennych,
- metody wyodrębniania czynników wspólnych,
- liczby czynników wspólnych,
- rotacji czynników,
- identyfikacji i interpretacji czynników.

W artykule scharakteryzowane zostaną pakiety programu R, w których przeprowadza się kolejne etapy analizy czynnikowej.

12:00 Poster 9

### Determinanty głębokości ubóstwa gospodarstw domowych

Krzysztof Szwarz

Akademia Ekonomiczna w Poznaniu, al. Niepodległości 10, Poznań 60-967, Poland

e-mail: k.szwarz@ae.poznan.pl

Amartya Sen w książce „Rozwój i wolność” wlicza czynniki mogące wpływać na ubóstwo. Są to: wiek, płeć i rola społeczna, miejsce zamieszkania, zwiększona możliwość zachorowań, a także inne elementy, na które dana jednostka nie ma wpływu lub ma wpływ ograniczony (np. dostępność żywności). Niniejsza praca ma służyć wyjaśnieniu, które z czynników społecznych i demograficznych gospodarstw domowych decydują o rozmiarach głębokości ubóstwa.

Celem artykułu jest pomiar głębokości ubóstwa w zależności od następujących cech:

- płci i wieku głowy gospodarstwa domowego;

- stanu zdrowia oraz uzyskanego poziomu wykształcenia głowy gospodarstwa domowego;
- głównego źródła dochodu gospodarstwa domowego;
- liczby osób w gospodarstwie domowym;
- liczby osób zatrudnionych na podstawie umowy o pracę;
- stref miejsca zamieszkania, tzn. według tzw. terytorialnych jednostek pomocniczych (określonych rejonów badanych dzielnic);
- częstotliwości kontaktów z rodziną jako zmiennych warunkujących szeroko rozumianą pomoc świadczoną tym gospodarstwom domowym.

Wykładnikiem głębokości ubóstwa, a więc zmienną objaśnianą jest poziom dochodu, jakim dysponują poszczególne gospodarstwa domowe znajdujące się w strefie ubóstwa. Wskaźnik głębokości ubóstwa dla poszczególnych typów gospodarstw domowych szacuje się jako stosunek różnicy między poziomem dochodu wyznaczającego linię ubóstwa a posiadanym dochodem gospodarstwa domowego do poziomu dochodu wyznaczającego linię ubóstwa. Głębokość ubóstwa informuje zatem, jaką kwotę należałoby przekazać poszczególnym typom gospodarstw domowych, aby znalazły się one poza tą strefą.

Zebrano dane dotyczące 6336 gospodarstw domowych, które w latach 2001-2003 objęte były pomocą społeczną w Poznaniu w następujących filiach Miejskiego Ośrodka Pomocy Rodzinie: Nowe Miasto, Grunwald, Stare Miasto (wraz z filią na Piątkowie), Wilda, Jeżyce. Siłę wpływu wyróżnionych zmiennych na głębokość ubóstwa w gospodarstwach domowych ogółem oraz według cech demograficzno-społecznych tych gospodarstw określono za pomocą analizy ścieżkowej. Punktem wyjścia w tej metodzie jest określenie dróg, czyli ścieżek bezpośredniego i pośredniego oddziaływania poszczególnych czynników na głębokość ubóstwa. Relacje te ujmuje się w formie diagramu ścieżkowego. Związki pomiędzy poszczególnymi zmiennymi zostały skwantyfikowane w oszacowanych współczynnikach ścieżkowych i ujęte w modelach regresji. Współczynniki ścieżkowe określają oczekiwaną zmianę zmiennej zależnej (wyrażoną w jej odchyleniach standardowych), będącą efektem zmiany zmiennej objaśniającej o jedno odchylenie standardowe, przy założeniu stałości pozostałych zmiennych objaśniających. Ogólny wpływ poszczególnych zmiennych niezależnych na zmienność zmiennej zależnej, opisanej przez określone równanie regresji, jest sumą wpływów bezpośrednich oraz pośrednich dokonujących się przez inne zmienne uwzględnione w diagramie ścieżkowym.

12:00	Poster	10
-------	--------	----

### Awersja do nierówności ekonomicznych w społeczeństwie polskim

Paweł Ulman

Akademia Ekonomiczna w Krakowie, Katedra Statystyki (AE), Rakowicka 27, Kraków 31-510, Poland

e-mail: [ulmanp@ae.krakow.pl](mailto:ulmanp@ae.krakow.pl)

Nierówności w rozkładach dochodów (wydatków) są jednym z kluczowych zagadnień intrygujących badaczy co najmniej od końca XIX w. W tym czasie zaproponowano wiele miar odzwierciedlających lepiej lub gorzej zagadnienie nierówności ekonomicznych w

społeczeństwie. Okazało się także, że nierówności są kluczowym czynnikiem wpływającym na poziom dobrobytu społecznego. W uproszczony i najbardziej przejrzysty sposób można ten wpływ zauważyć przy rozpatrzeniu tzw. uproszczonych funkcji dobrobytu. Jednakże poziom nierówności nie mówi wszystkiego o nastawieniu społeczeństwa do nierówności dochodowych. Możemy się spodziewać, że niskim poziomem nierówności będą się charakteryzowały społeczeństwa (grupy społeczne) przejawiające wysoki zakres niechęci (awersji) do nierówności i przeciwnie wysoki poziom nierówności świadczyć może o niskiej awersji do nierówności.

W teorii dobrobytu do pomiaru awersji do nierówności wykorzystuje się funkcję zwaną relatywną awersją do ryzyka Arrowa-Pratta, która bazuje na określeniu postaci funkcji użyteczności. Bezpośrednia estymacja funkcji użyteczności, a więc i awersji do nierówności jest niemożliwa z uwagi na niemożność obserwacji użyteczności. W tej sytuacji Kot<sup>1</sup> zaproponował interesujący sposób rozwiązania tego problemu bazując na schemacie dedukcyjnym *modus tollens*, który w efekcie daje oszacowania nieobserwowalnej funkcji użyteczności. Celem proponowanego referatu jest analiza zróżnicowania poziomu awersji do nierówności w społeczeństwie polskim w różnych przekrojach ze względu na czynniki społeczno-ekonomiczne gospodarstw domowych. Estymacja rozkładów dochodów i awersji do nierówności – za pomocą wspomnianego schematu wnioskowania - została dokonana na podstawie danych z Budżetów Gospodarstw Domowych z roku 2005 przy wykorzystaniu metody największej wiarygodności.

<sup>1</sup>*Dobrobyt społeczny, nierówności i sprawiedliwość dystrybucyjna*, (red.) S.M. Kot, A. Malawski, A. Węgrzecki, Wydawnictwo AE w Krakowie, Kraków 2004, s. 194

12:00	Poster	11
-------	--------	----

### Zdolności dyskryminacyjne wskaźników finansowych w ocenie kondycji finansowej podmiotów gospodarczych

Mirosław Wójciak<sup>1</sup>, Aleksandra Wójcicka-Krenz<sup>2</sup>

1. Akademia Ekonomiczna im. K. Adamieckiego w Katowicach, Katowice 40-287, Poland 2. Akademia Ekonomiczna w Poznaniu, al. Niepodległości 10, Poznań 60-967, Poland

e-mail: [mwojciak@ae.katowice.pl](mailto:mwojciak@ae.katowice.pl)

W połowie lat 90-tych na zachodzie zaczęły powstawać nowe modele oceny ryzyka kredytowego jako odpowiedź na niewystarczającą dokładność modeli klasycznych. Jednakże zastosowanie nowych metod oceny ryzyka kredytowego na rynkach rozwijających się wiąże się z pewnymi ograniczeniami np. można mierzyć poziom ryzyka tylko spółek giełdowych (np. modele intensywności DS, JLT) lub podmiot musi posiadać rating (np. model struktury ratingowej ryzyka). W związku z tym w warunkach polskiej gospodarki do prognozowania bankructwa nadal mają zastosowanie klasyczne modele oceny ryzyka kredytowego, w których zmiennymi są najczęściej różne zestawy wskaźników finansowych. W literaturze przedmiotu często podaje się optymalne wartości wskaźników (np. dla wskaźnika bieżącej i szybkiej płynności czy rotacji należności). Jednak wartości te najczęściej ustalane są dla podmiotów gospodarczych działających na zagranicznych rynkach np. na rynku amerykańskim.

W związku z tym w artykule autorzy poszukują takich wskaźników

finansowych, które w najlepszy sposób dyskryminują podmioty gospodarcze ze względu na poziom ryzyka kredytowego. Dlatego autorzy przyjmują określone wartości w odniesieniu do poziomu prawdopodobieństwa niewypłacalności (PD) oszacowanego na podstawie modelu MKMV. W tym celu dla określonych poziomów PD zostaną wyznaczone kwantyle rozkładu. Przyjęto, że wskaźnik finansowy o charakterze stymulanty dobrze dyskryminuje przedsiębiorstwa w dobrej i złej kondycji ekonomiczno-finansowej, gdy wyznaczone wartości danego percentyla (np. P10) dla spółek w dobrej kondycji (górną granicą wskaźnika) były wyższe od wartości P90 dla spółek w trudnej sytuacji (dolną granicą wskaźnika).

Optymalne wartości wskaźników zostaną wyznaczone i porównane w różnych branżach, gdyż specyfika branży wpływa zasadniczo na preferowany poziom wskaźników finansowych. W badaniu uwzględniono spółki, które były notowane na Giełdzie Papierów Wartościowych w Warszawie w latach 1997-2006. Także fazy cyklu stopień niewypłacalności szczególnie w przypadku podmiotów o wysokim ryzyku kredytowym jest bardzo wrażliwy w odniesieniu do faz cyklu koniunkturalnego. Oznacza to, że poziom niewypłacalności będzie się zmieniał w zależności od cyklu koniunkturalnego, w okresie recesji będzie większy niż w okresie ekspansji. W związku z tym dla każdej branży optymalne wartości wskaźników zostaną wyznaczone dla tych dwóch stanów gospodarki.

Literatura:

- [1] Bień W. (1999): *Zarządzanie finansami przedsiębiorstwa*, Difin, Warszawa
- [2] Byström H. (2004): *A Flexible Way of Modelling Default Risk*, [http://www.business.uts.edu.au/qfrc/research/research\\_papers/rp112.pdf](http://www.business.uts.edu.au/qfrc/research/research_papers/rp112.pdf)
- [3] Friedlob G. T., Schleifer L.F. (2002): *Essentials of Financial Analysis*, John Wiley & Sons, New York
- [4] Sierpińska M., Jachna T (1997): *Ocena przedsiębiorstwa według standardów światowych*, PWN, Warszawa
- [5] Hull J., Nelken I., White A. (2004): *Merton's Model, Credit Risk and Volatility Skew*, working paper

12:00	Poster	12
-------	--------	----

### Wrażliwość modelu MKMV ze względu na metodę estymacji wzrostu wartości aktywów

Aleksandra Wójcicka-Krenz

Akademia Ekonomiczna w Poznaniu, al. Niepodległości 10, Poznań 60-967, Poland

e-mail: [aleksandra.wojcicka@ae.poznan.pl](mailto:aleksandra.wojcicka@ae.poznan.pl)

Tendencja wzrostu ryzyka kredytowego na rynkach finansowych, obserwowane w ostatnich latach sprawia, że coraz liczniejsze są próby tworzenia i ulepszania modelu oceny ryzyka kredytowego. Jednym z tzw. nowych modeli oceny ryzyka kredytowego jest model Moody's KMV (MKMV) oparty na modelu wyceny opcji. Model ten wykorzystuje zarówno dane księgowe jak i dane pochodzące z rynków finansowych. Przy szacowaniu prawdopodobieństw niewypłacalności (PD – *probability of default*) wykorzystuje zmienne bezpośrednio obserwowalne (np.  $E$  – *rynkową wartość kapitału własnego przedsiębiorstwa*,  $D$  – *nominalną wartość zadłużenia* itd.) oraz takie,

które nie są bezpośrednio obserwowalne ( $A$  – *wartość rynkowa aktywów*, - *zmiennosc wartości rynkowej aktywów*).

W modelu tym pojawia się jednak jeszcze inny bardzo ważny parametr, który w wielu modyfikacjach jest traktowany marginalnie lub nawet całkowicie pomijany. Parametrem tym jest parametr  $\mu$  (*średnia stopa zwrotu z aktywów przedsiębiorstwa*). Może on znacząco wpływać na poziom ryzyka kredytowego badanych spółek. Oszacowanie tego parametru może odbyć się w oparciu o jeden z najczęściej do tej pory stosowany model wyceny aktywów, którym jest model równowagi rynku kapitałowego CAPM (*Capital Assets Pricing Model*) oraz konkurencyjny model tzw. model teorii arbitrażu cenowego APT (*Arbitrage Pricing Theory*). Przy szacowaniu pojawia się jednak dodatkowy element niepewności.

Celem tego artykułu jest omówienie podejść stosowanych do szacowania parametru  $\mu$  w nowych modelach oceny ryzyka kredytowego, które są stosowane w literaturze i praktyce. Porównanie wartości PD oszacowanych na podstawie modelu MKMV zaprezentuje w jakim stopniu model ten jest wrażliwy na wybór metody estymacji parametru stopy zwrotu z aktywów.

Literatura:

1. Cochrane J. H. „Asset pricing”, Princeton, 2005
2. Deventer D., Imai K., Mesler M., *Advanced Financial Risk Management. Tools and Techniques for Integrated Credit Risk and Interest Rate Risk Management*, wyd. John Wiley & Sons (Asia) Pte Ltd., 2005
3. Gątarek D., Maksymiuk R., Krysiak M., Witkowski Ł., *Nowoczesne metody zarządzania ryzykiem finansowym*, WIG-Press, Warszawa 2001
4. Saunders A., *Metody pomiaru ryzyka kredytowego*, Oficyna Ekonomiczna, Kraków 2001



# Sesja warsztatowa

## Program

### Czwartek, 20 Wrzesień

#### Sesja Warsztatowa

Czwartek rano, 20 Wrzesień, 10:15

10:15

Oral

#### Modelowanie i prognozowanie kursu euro/dolar: modele autoregresyjne z rozkładami opóźnień i sztuczne sieci neuronowe

Aleksandra Matuszewska, Dorota Witkowska

*Szkoła Główna Gospodarstwa Wiejskiego (SGGW), Nowoursynowska 166, Warszawa 02-787, Poland*

*e-mail: amatuszewska@mors.sggw.waw.pl*

Zmiany, jakie zachodzą w funkcjonowaniu rynków finansowych powodują, że do analizy zachowań poszczególnych instrumentów wykorzystuje się nie tylko informacje dotyczące bezpośrednio danego waloru, ale też związane z jego otoczeniem. Do instrumentów, które w dużym stopniu reagują na zmiany otoczenia należą kursy walutowe. Na zmiany kursów walutowych wpływają różne czynniki zarówno ekonomiczne jak i pozaekonomiczne. Rozpoznanie tych czynników oraz analiza ich wpływu na kurs walutowy często pozwolą na wykorzystanie zdobytej wiedzy w procesie przewidywania zmian kursowych.

W prezentowanym badaniu analizie poddano kurs euro/dolar, którego wybór podyktowany jest tym, że obie waluty uczestniczą w największej liczbie transakcji zawieranych na rynku walutowym oraz coraz więcej gospodarek krajowych (regionalnych) w coraz większym stopniu staje się otwartych na międzynarodową wymianę handlową.

W pierwszym etapie badania ustalono, które zmienne z rynku finansowego i towarowego mają wpływ na zmiany kursu euro/dolar. W tym celu wykorzystano analizę przyczynowości Grangera. W drugim etapie zbudowano modele autoregresyjne z rozkładami opóźnień oraz skonstruowano modele sztucznych sieci neuronowych, które wykorzystano w procesie prognozowania analizowanego kursu.

10:25

Oral

#### Analiza porównawcza nowych metod oceny ryzyka kredytowego

Mirosław Wójciak<sup>1</sup>, Aleksandra Wójcicka-Krenz<sup>2</sup>

**1.** *Akademia Ekonomiczna im. K. Adamieckiego w Katowicach, Katowice 40-287, Poland* **2.** *Akademia Ekonomiczna w Poznaniu, al. Niepodległości 10, Poznań 60-967, Poland*

*e-mail: mwojciak@ae.katowice.pl*

Ze względu na wzrost liczby bankructw coraz większe znaczenie odgrywa poprawne oszacowanie ryzyka kredytowego. W literaturze istnieje wiele różnych modeli oceny ryzyka kredytowego m.in. modele strukturalne pierwszej i drugiej generacji, modele intensywności czy modele zredukowane. Jednakże wszystkie te modele wiążą się z pewnymi ograniczeniami i ich zastosowanie jest często ograniczone do pewnej grupy podmiotów.

W artykule podjęto próbę zaimplementowania w realiach polskiej gospodarki trzech nowych modeli oceny ryzyka kredytowego i porównania otrzymanych prawdopodobieństw niewypłacalności (*probability of default – PD*) wyników pomiędzy modelami. Pierwszy z nich to jeden z najszerszej stosowanych w praktyce modeli – model MKMV – który opiera się na oryginalnym modelu Mertona wyceny opcji. Drugi model to modyfikacja modelu MKMV zaproponowana przez Byström'a, która przy przyjęciu określonych założeń znacznie upraszcza sposób szacowania prawdopodobieństwa niewypłacalności stosując do jego szacowania tylko zmienne bezpośrednio obserwowalne. Jednym z elementów, który nie został w oryginalnym modelu Mertona dogłębnie zbadany to rola jaką model ten odgrywa w określaniu implikowanej zmienności kapitału własnego oraz jej skłonności. W ramach ideologii Mertona opcja wystawiona na wartość kapitału własnego firmy, która wygasa zanim dług osiągnie moment zapadalności jest opcją złożoną – opcją na Europejską opcję kupna (European call option). Aby oszacować wartość obu opcji możemy skorzystać z modelu zaproponowanego przez Geske, który jest trzecim badanym modelem.

W badaniu uwzględniono spółki z branży budowlanej, które były notowane na Giełdzie Papierów Wartościowych w Warszawie w latach 1997–2006. Wybór branży budowlanej jest podyktowany faktem, że w badanym okresie występowało wysokie zróżnicowanie kondycji ekonomiczno-finansowej spółek, wśród których jedenaście ogłosiło upadłość lub wszczęło postępowanie układowe z wierzycielami. Porównanie osiągniętych wyników pozwoli w znacznym stopniu określić przydatność zaprezentowanych modeli, które powstały w oparciu o w realiach polskiego rynku.

#### Literatura:

- Byström H. (2004): A Flexible Way of Modelling Default Risk, [http://www.business.uts.edu.au/qfrc/research/research\\_papers/rp112.pdf](http://www.business.uts.edu.au/qfrc/research/research_papers/rp112.pdf)
- Gątarek D., Maksymiuk R., Krysiak M., Witkowski Ł., Nowoczesne metody zarządzania ryzykiem finansowym, WIG-Press, Warszawa 2001
- Hull J., Nelken I., White A. (2004): Merton's Model, Credit Risk and Volatility Skew, working paper
- Saunders A., Metody pomiaru ryzyka kredytowego, Oficyna Ekonomiczna, Kraków 2001

10:40

Oral

### Badanie efektywności GPW na przykładzie wybranych indeksów: test autokorelacji

Dorota Żebrowska-Suchodolska<sup>1</sup>, Dorota Witkowska<sup>2</sup>

1. Wyższa Szkoła Finansów i Zarządzania w Białymstoku, Ciepła 40, Białystok 15-472, Poland 2. Szkoła Główna Gospodarstwa Wiejskiego (SGGW), Nowoursynowska 166, Warszawa 02-787, Poland

e-mail: [zdorota@o2.pl](mailto:zdorota@o2.pl)

Większość analiz finansowych opiera się na hipotezie rynku efektywnego (*Efficient Market Hypothesis, EMH*), zgodnie z którą rynki finansowe szybko i dokładnie odzwierciedlają dostępne publicznie informacje. Do oceny słabej efektywności informacyjnej rynku można użyć wielu narzędzi, do których należą m.in. testy na występowanie autokorelacji.

Celem pracy jest weryfikacja hipotezy o słabej formie efektywności informacyjnej polskiego rynku kapitałowego. Badania przeprowadzono dla indeksu WIG20 oraz wyznaczonego subindeksu zawierającego notowania spółek z sektora bankowego, wchodzących w skład indeksu WIG20. Wykorzystano w tym celu: test współczynnika autokorelacji *Qeunoille'a*, test łącznej autokorelacji ze statystyką *Boxa-Pierce'a* oraz *Ljunga-Boxa*.

Analizy przeprowadzono w oparciu o dzienne logarytmiczne stopy zwrotu wyznaczone dla notowań z okresu od 3 października 1994 roku do 29 grudnia 2006 roku. Rozważany okres podzielono na 19 podokresów, dla których datami podziału są sesje odnotowane w historii Giełdy Papierów Wartościowych w Warszawie jako „kamienie milowe”. Innymi słowy poszczególne podokresy zostały wyznaczone biorąc pod uwagę ważne wydarzenia dla polskiego rynku kapitałowego.

10:55

Oral

### Statystyczna analiza efektywności gospodarowania spółek notowanych na Giełdzie Papierów Wartościowych w Warszawie

Agnieszka Sompolska-Rzechuła, Beata Szczecińska

Akademia Rolnicza (AR), Janosika 8, Szczecin 71-424, Poland

e-mail: [arzechula@e-ar.pl](mailto:arzechula@e-ar.pl)

Ocena efektywności gospodarowania jest jednym z podstawowych źródeł informacji o przedsiębiorstwie dla kadry zarządzającej oraz obecnych i potencjalnych akcjonariuszy. Analiza finansowa dostarcza informacji o sytuacji finansowo-ekonomicznej przedsiębiorstw. Porównanie wyników analizy finansowej, w tym także wskaźnikowej, w stosunku do minionych okresów i z innymi obiektami daje pełniejszą wartość poznawczą otrzymanych wielkości. W opracowaniu poddano analizie spółki branży spożywczej notowane na GPW w Warszawie w latach 2002-2006. Analizowano następujące wskaźniki: wskaźnik ogólnego zadłużenia, wskaźnik obrotowości majątku, wskaźnik rentowności sprzedaży, wskaźnik rentowności majątku ROA, wskaźnik rentowności kapitału własnego ROE, zysk netto na akcję (w zł) EPS. Część tych wskaźników opiera się na wyniku fi-

ansowym. Wynik finansowy jest zmienną, która może przyjmować wartości ujemne (przejście od zysku do straty). W takim przypadku, gdy zmienna przyjmuje wartości ujemne, wartość indeksu także będzie ujemna. Należy wówczas zastosować uogólnienie interpretacji indeksów dla zmiennych przyjmujących dowolne wielkości, zaproponowane w [1]. W pracy badano także rozkłady zmiennych. Ich znajomość stanowi ważną informację przy podejmowaniu różnego rodzaju decyzji, np. dotyczących działalności przedsiębiorstwa. Ze względu na niewielką liczbę analizowanych spółek, co związane jest z dostępnością informacji statystycznych, zastosowano test normalności Shapiro-Wilka. Obliczenia zostały wykonane w programie Statistica.

1. Wawrzyniak K., Zwolankowska M., Dynamika zmiennych ekonomicznych, przyjmujących wartości ujemne, Wiadomości Statystyczne nr 4, Warszawa 1998.

11:10

Oral

### Selekcja i klasyfikacja spółek notowanych na Warszawskiej Giełdzie Papierów Wartościowych przy pomocy pakietu SAS

Urszula Grzybowska<sup>1</sup>, Arkadiusz J. Orłowski<sup>2,3</sup>, Piotr Jałowicki<sup>1</sup>

1. Szkoła Główna Gospodarstwa Wiejskiego (SGGW), Nowoursynowska 166, Warszawa 02-787, Poland 2. Warsaw Agricultural University, Department of Food Hygiene and Public Health (SGGW), Nowoursynowska 166, Warszawa 02-787, Poland 3. Polish Academy of Sciences, Institute of Physics, al. Lotników 32/46, Warszawa 02-668, Poland

e-mail: [orlow@ifpan.edu.pl](mailto:orlow@ifpan.edu.pl)

Analizą objęte zostaną spółki notowane na WGPW, charakteryzowane wskaźnikami ekonomiczno-finansowymi za rok 2006, publikowanymi przez Notoria Serwis. Przedstawione zostaną sposoby selekcji spółek do portfela inwestycyjnego, a także klasyfikacja rozważanego zbioru spółek w oparciu o analizę skupień i metodę składowych głównych. Obliczenia, w szczególności ustalenie optymalnej liczby skupień, wykonane zostaną w pakiecie SAS.

11:25

Oral

### Zastosowanie wybranych metod rangowania obiektów do oceny stopnia zróżnicowania starości demograficznej ludności w krajach UE i EFTA

Artur Mikulec

Uniwersytet Łódzki, Katedra Metod Statystycznych (UŁ), Rewolucji 1905 r. nr 41, Łódź 90-214, Poland

e-mail: [amikulec@poczta.onet.pl](mailto:amikulec@poczta.onet.pl)

W referacie omówione zostaną wybrane metody porządkowania obiektów wielocechowych: według cechy syntetycznej, macierzy odległości od wzorca i rzutowania ortogonalnego punktów na prostą.

Głównym celem podjętych badań jest zastosowanie przedstawionych algorytmów do oceny stopnia zróżnicowania starości i starzenia się ludności w krajach UE i EFTA – objętych wspólną otwartą metodą koordynacji systemów emerytalnych.

W dalszym kroku podjęta zostanie próba oceny jakości uzyskany-



ch wyników oraz przedstawione zostaną metody wyodrębniania skupień, przy ustalonym już wcześniej porządku liniowym (wyniku rangowania).

11:35 Oral

**Zastosowanie wybranych metod statystycznych i taksonomicznych do oceny sytuacji na rynku pracy ze szczególnym uwzględnieniem kształtowania się wynagrodzeń**

Danuta J. Rozpedowska-Matraszek, Barbara Dańska-Borsiak

Uniwersytet Łódzki, Wydział Ekonomiczno-Socjologiczny, P.O.W. 3/5, Łódź 90-255, Poland

e-mail: daroma@uni.lodz.pl

Celem referatu jest prezentacja wyników analizy rynku pracy w Polsce. W analizie rynku pracy kluczowe znaczenie ma kształtowanie się płac przeciętnych. Oddziałują one bowiem na zatrudnienie w różnych sektorach, oraz są ważnym składnikiem kosztów, a ich zmiany powodują presję do wzrostu cen. Wynagrodzenia determinują ponadto wielkość dochodów aktywnej zawodowo części społeczeństwa, a zatem ich poziom wpływa na nastroje społeczne.

W części pierwszej referatu poddano analizie zależności sezonowe na rynku pracy. Obliczono w tym celu kwartalne wskaźniki sezonowości wynagrodzeń ogółem i według sektorów. Badania empiryczne zostały przeprowadzone na podstawie informacji kwartalnych.

Kolejną częścią prezentacji jest przedstawienie analizy ogólnej sytuacji na rynku pracy w okresie próby, na podstawie miary rozwoju. Miara ta pozwala na syntetyczne opisanie tendencji na rynku pracy i wyodrębnienie okresów, w których sytuacja poprawiała się lub pogarszała.

W klasycznej analizie mechanizmów kształtowania płac zakłada się, że przedsiębiorstwa działają w warunkach konkurencji doskonałej, a ich głównym celem jest maksymalizacja zysku. Bada się także związki zachodzący pomiędzy kształtowaniem się płac przeciętnych a stopą bezrobocia, określającymi sytuację na rynku pracy. Mechanizmy wzajemnych powiązań między zmiennymi opisane zostały za pomocą modelu ekonometrycznego wynagrodzeń z wykorzystaniem krzywej Philipsa.

**Bibliografia:**

1. De Loo Ivo, Thomas Ziesemer, *Determinants Of Sectoral Average Wage Growth Rates in a Specific Factors Model with International Capital Movements*, Department of Economics and MERIT, University of Maastricht, April 1998.
2. Kabaj M., *Partycypacyjny system wynagrodzeń*, Wyd. IPiSS, Warszawa 2003.
3. Kryńska E., *Polski rynek pracy - niedopasowania strukturalne*, Wyd. IPiSS, Warszawa 2004.
4. Kryńska E., *Dylematy polskiego rynku pracy*, „Studia i Monografie”, Wyd. IPiSS, Warszawa 2001.
5. Ostasiewicz W. (redaktor.), *Statystyczne metody analizy danych*, Wydawnictwo Akademia Ekonomiczna im. Oskara Lanego, Wrocław 1999.
6. Pułaska-Turyńska B., *Statystyka dla ekonomistów*, Wyd. Difin, Warszawa 2005.

11:45

Oral

**Zastosowanie metod statystycznych i ekonometrycznych do analiz popytu i podaży pracy w sektorze ochrony zdrowia w Polsce**

Danuta J. Rozpedowska-Matraszek

Uniwersytet Łódzki, Wydział Ekonomiczno-Socjologiczny, P.O.W. 3/5, Łódź 90-255, Poland

e-mail: daroma@uni.lodz.pl

Celem referatu jest prezentacja wyników statystycznej i ekonometrycznej analizy popytu i podaży pracy w sektorze ochrony zdrowia w Polsce w ujęciu regionalnym.

Zdrowie staje się coraz ważniejszą wartością dla coraz szerszego kręgu Polaków. Standard opieki zdrowotnej jest uznawany jako użyteczny wskaźnik postępu społecznego, który zwłaszcza ostatnio nabrał szczególnego znaczenia. Na zdrowie publiczne można wpływać zarówno na poziomie globalnym, poziomie populacji, czy na poziomie społeczeństw, poprzez prowadzenie właściwej polityki zdrowotnej. W sektorze ochrony zdrowia w związku z rozwojem nowoczesnego systemu opieki zdrowotnej zachodzą zmiany na rynku pracy. W medycynie następuje zmiana na zapotrzebowanie wykwalifikowanej kadry pomocniczej i zmiana zapotrzebowania na lekarzy z zakresu coraz to innych specjalizacji. Poprawia się większość wskaźników zdrowotnych, między innymi zauważamy spadek liczby zgonów, nastąpiło odwrócenie trendu przedwczesnej umieralności dorosłych. Przeciętne trwanie życia wydłużyło się zarówno dla kobiet jak i dla mężczyzn i wydłuży nadal. To z kolei przekłada się na zwiększenie popytu w obsłudze tychże osób.

Wyniki analizy mogą być wykorzystane do poprawy poziomu efektywności systemu ochrony zdrowia. Prognozowanie w opiece zdrowotnej, projektowanie metod medycyny zapobiegawczej oraz systemów wczesnego ostrzegania jest najważniejszym celem systemu Ogólnopolskiego Systemu Ochrony Zdrowia (OSOZ).

**Bibliografia:**

1. Frąckiewicz-Wronka Aldona (red.), *Ochrona zdrowia w regionie*, Wyd. Akademii Ekonomicznej w Katowicach, Katowice 2005.
2. Getzen Thomas E., *Ekonomika zdrowia. Teoria i praktyka*, (Przekład) M. Jakubiak, T. Żukowski, Wyd. Naukowe PWN SA, Warszawa 2000.
3. Kryńska Elżbieta, Jadwiga Suchocka, Bogdan Suchecki, *Prognoza podaży i popytu na prace w Polsce do roku 2010*, Wyd. Instytut Pracy i Spraw Socjalnych, Warszawa 1998.
4. Newhouse J., *Health Economics and Econometrics*, *The American Economic Review*, Vol.77, No.2, Papers and Proceedings (May 1987) 269-274.
5. Suchecka Jadwiga, *Ekonomiczne aspekty ochrony zdrowia – wybrane problemy*, w: *Gospodarka w Praktyce i Teorii* Nr 2(19), Wyd. Instytut Ekonomii Uniwersytetu Łódzkiego, Łódź 2006.
6. Suchecki Bogdan, *Kompletne modele popytu*, Wyd. Polskie Wydawnictwo Ekonomiczne S.A., Warszawa 2006.

11:55

Oral

### Porównanie prognoz spożycia papierosów w Polsce w latach 2006-2010 na podstawie modeli wyrównywania wykładniczego oraz autoregresyjnych

Ewa Jałowiecka<sup>1</sup>, Piotr Jałowiecki<sup>1</sup>, Marek Karwański<sup>2</sup>

1. Szkoła Główna Gospodarstwa Wiejskiego (SGGW), Nowoursynowska 166, Warszawa 02-787, Poland 2. SAS Institute Sp. z o.o., Gdańska 27/31, Warszawa 01-633, Poland

e-mail: pjalowiecki@mors.sggw.waw.pl

Przemysł wyrobów tytoniowych jest jedną z najważniejszych gałęzi polskiej gospodarki, który generuje 7- 7,5% dochodów budżetu państwa. Jednocześnie jest to sektor, który od kilku lat podlega znaczącym przemianom związanym z dostosowywaniem polskich regulacji prawnych do prawodawstwa unijnego. W pracy podjęto próbę określenia kierunku zmian spożycia papierosów w Polsce do 2010 roku na podstawie szeregu czasowego obejmującego dane historyczne z lat 1970- 2005, w oparciu o wybrane modele wyrównywania wykładniczego oraz autoregresyjne. Identyfikację modeli przeprowadzono przy użyciu metod typu *corner*. Prognozy przygotowane na podstawie wybranych modeli obydwu kategorii były zbieżne oraz charakteryzowały się wyraźnie zaznaczonym trendem malejącym. Porównano prognozy przygotowane w oparciu o wybrane modele z uwzględnieniem wartości odstającej zidentyfikowanej dla 2000 roku, których wyniki również okazały się zgodne, wskazując na stabilizację spożycia papierosów w Polsce na poziomie 1950-1965 sztuk na osobę rocznie. Przygotowano ponadto prognozę z wykorzystaniem techniki *prewhitening* uwzględniającej dodatkowe zmienne.

12:10

Oral

### Modelowanie uzysku energii elektrycznej z modułów fotowoltaicznych na podstawie długoterminowych badań w warunkach naturalnych

Włodzimierz Kołodenny, Tadeusz Żdanowicz

Wrocław University of Technology, Wybrzeże Wyspiańskiego, Wrocław 50-370, Poland

e-mail: wlozdimierz.kolodenny@pwr.wroc.pl

Podczas szacowania kosztów inwestycji w system fotowoltaiczny potrzebne jest dopasowanie zakupionej mocy nominalnej systemu do warunków, w których będzie pracował. Produkowana energia zależy od warunków atmosferycznych wybranej lokalizacji i rodzaju modułów fotowoltaicznych. O ile pogoda jest specyficzna dla lokalizacji, to dla generatorów fotowoltaicznych można stworzyć zadowalające modele do oszacowania ilości produkowanej energii. Praca opisuje sposób zbierania danych, analizy i trzech różnych sposobów modelowania energii (prostej regresji liniowej, metody matrycowej i z użyciem sieci neuronowej). Modele są tworzone i weryfikowane przy użyciu oprogramowania analitycznego SAS Institute i danych mierzonych w SolarLab od stycznia 2003 do chwili obecnej, co 90 sekund dla parametrów elektrycznych i 60 sekund dla warunków meteorologicznych.

12:20

Oral

### Koncepcja systemu informatycznego do analizy danych regionalnych

Kesra Nermend

Uniwersytet Szczeciński, Szczecin 71-415, Poland

e-mail: kesra@szafir.univ.szczecin.pl

Artykuł przedstawia koncepcję systemu informatycznego przeznaczonego do analizy danych społeczno-gospodarczych ułatwiającego podejmowanie decyzji dotyczących rozwoju regionalnego. Dodatkowo zostanie dokonana analiza metod badania rozwoju regionalnego pod kątem ich możliwości manipulowania na dużych zbiorach danych w proponowanym systemie informatycznym.

### Zakończenie obrad w sesji warsztatowej

Czwartek popołudniu, 20 Wrzesień, 12:30

---

# Lista uczestników

## Andrzej Bąk

*andrzej.bak@ae.jgora.pl*

- Akademia Ekonomiczna (AE)  
*Nowowiejska3, Jelenia Góra 58-500, Poland*
- Akademia Ekonomiczna (AEWROCLAW)  
*Komandorska 118/120, Wrocław 53-345, Poland*

## Joanna Banaś

*jbanas@wi.ps.pl*

- Politechnika Szczecińska, Wydział Informatyki (PS)  
*ul. Żołnierska 49, Szczecin 71-210, Poland*

## Stanisława Bartosiewicz

*katarzyna.kuziak@ae.wroc.pl*

- Akademia Ekonomiczna (AEWROCLAW)  
*Komandorska 118/120, Wrocław 53-345, Poland*

## Jacek Batóg

*batog@uoo.univ.szczecin.pl*

- Uniwersytet Szczeciński  
*Jedności Narodowej 22, Szczecin 70-453, Poland*

## Zbigniew Binderman

*zbinderman@mors.sggw.waw.pl*

- Szkoła Główna Gospodarstwa Wiejskiego (SGGW)  
*Nowoursynowska 166, Warszawa 02-787, Poland*

## Marcin Błażejowski

*marcin@wrzosy.nsb.pl*

- Wyższa Szkoła Bankowa (WSB)  
*Toruń 87-100, Poland*

## Bolesław Borkowski

*bborkowski@mors.sggw.waw.pl*

- Szkoła Główna Gospodarstwa Wiejskiego (SGGW)  
*Nowoursynowska 166, Warszawa 02-787, Poland*

## Mariola Chrzanowska

*mariola.chrzanowska@wseia.edu.pl*

- Wyższa Szkoła Ekonomii i Administracji  
*Jagiellońska 109A, Kielce 25-734, Poland*

## Esteban A. Cortes

*IEsteban.Alfaro@uclm.es*

- Castilla-La Mancha University.  
*Plaza de la Universidad, 1, Albacete 02-071, Spain*

## Barbara Dańska-Borsiak

*danska@uni.lodz.pl*

- Uniwersytet Łódzki, Wydział Ekonomiczno-Socjologiczny  
*P.O.W. 3/5, Łódź 90-255, Poland*

## Reinhold Decker

*RDecker@wiwi.uni-bielefeld.de*

- Universitaet Bielefeld, Fakultaeet fuer Wirtschaftswissenschaften  
*Universitaetstr. 25, Bielefeld 33615, Germany*

## Grażyna Dehnel

*g.dehnel@ae.poznan.pl*

- Akademia Ekonomiczna w Poznaniu  
*al. Niepodległości 10, Poznań 60-967, Poland*

## Hanna Dudek

*hdudek@mors.sggw.waw.pl*

- Szkoła Główna Gospodarstwa Wiejskiego (SGGW)  
*Nowoursynowska 166, Warszawa 02-787, Poland*

## Iwona A. Foryś

*iwonafor@poczta.onet.pl*

- Uniwersytet Szczeciński  
*Jedności Narodowej 22, Szczecin 70-453, Poland*

## Eugeniusz Gatnar

*egatnar@ae.katowice.pl*

- Akademia Ekonomiczna im. K. Adameckiego w Katowicach  
*Katowice 40-287, Poland*

## Mirosława Gazińska

*mgazi@univ.szczecin.pl*

- Uniwersytet Szczeciński  
*Jedności Narodowej 22, Szczecin 70-453, Poland*

## Malgorzata Gliwa

*wojcikm8@wp.pl*

- Akademia Ekonomiczna im. K. Adameckiego w Katowicach  
*Katowice 40-287, Poland*

## Mariusz Grabowski

*Mariusz.Grabowski@ae.krakow.pl*

- Akademia Ekonomiczna w Krakowie, Katedra Informatyki  
*Rakowicka 27, Kraków 31-510, Poland*

## Stanisław K. Jabłonowski

*sj@mors.sggw.waw.pl*

- Szkoła Główna Gospodarstwa Wiejskiego (SGGW)  
*Nowoursynowska 166, Warszawa 02-787, Poland*

---

**Krzysztof Jajuga**

*krzysztof.jajuga@ae.wroc.pl*

- Akademia Ekonomiczna (AEWROCLAW)  
*Komandorska 118/120, Wrocław 53-345, Poland*

**Piotr Jałowiecki**

*pjalowiecki@mors.sggw.waw.pl*

- Szkoła Główna Gospodarstwa Wiejskiego (SGGW)  
*Nowoursynowska 166, Warszawa 02-787, Poland*

**Marek Karwański**

*anstat@jeet.pl*

- SAS Institute Sp. z o.o.  
*Gdańska27/31, Warszawa 01-633, Poland*
- Szkoła Główna Gospodarstwa Wiejskiego (SGGW)  
*Nowoursynowska 166, Warszawa 02-787, Poland*

**Iwona Kasprzyk**

*i\_kasprzyk@interia.pl*

- Akademia Ekonomiczna im. K. Adamieckiego w Katowicach  
*Katowice 40-287, Poland*

**Joanna Kisielńska**

*jkisielinska@mors.sggw.waw.pl*

- Szkoła Główna Gospodarstwa Wiejskiego (SGGW)  
*Nowoursynowska 166, Warszawa 02-787, Poland*

**Lukasz J. Kociuba**

*Lukasz.Kociuba@spl.sas.com*

- SAS Institute Sp. z o.o.  
*Gdańska27/31, Warszawa 01-633, Poland*

**Włodzimierz Kolodenny**

*wlodzimierz.kolodenny@pwr.wroc.pl*

- Politechnika Wrocławska (PWR)  
*Wybrzeże Wyspiańskiego, Wrocław 50-370, Poland*

**Krzysztof Kompa**

*kkompa@mors.sggw.waw.pl*

- Szkoła Główna Gospodarstwa Wiejskiego (SGGW)  
*Nowoursynowska 166, Warszawa 02-787, Poland*

**Jerzy M. Korzeniewski**

*jurkor@wp.pl*

- Uniwersytet Łódzki, Wydział Ekonomiczno-Socjologiczny  
*P.O.W. 3/5, Łódź 90-255, Poland*

**Jadwiga H. Kostrzevska**

*Jadwiga.Kostrzevska@ae.krakow.pl*

- Akademia Ekonomiczna w Krakowie, Katedra Statystyki (AE)  
*Rakowicka 27, Kraków 31-510, Poland*

**Grzegorz Koszela**

*gkoszela@mors.sggw.waw.pl*

- Szkoła Główna Gospodarstwa Wiejskiego (SGGW)  
*Nowoursynowska 166, Warszawa 02-787, Poland*

**Mariusz Kubus**

*marikubu@wp.pl*

- Politechnika Opolska (PO)  
*Mikolajczyka 5, Opole 45-271, Poland*

**Tadeusz Kufel**

*tadeusz.kufel@umk.pl*

- Uniwersytet Mikołaja Kopernika w Toruniu, WNEiZ, Katedra Ekonometrii i Statystyki (UMK)  
*Gagarina 13A, Toruń 87-100, Poland*

**Izabela Kurzawa**

*kurzawa@au.poznan.pl*

- Akademia Rolnicza (AR)  
*ul. Wojska Polskiego 28, Poznań 60-637, Poland*

**Katarzyna Kuziak**

*katarzyna.kuziak@ae.wroc.pl*

- Akademia Ekonomiczna (AEWROCLAW)  
*Komandorska 118/120, Wrocław 53-345, Poland*

**Joanna M. Landmesser**

*jmlg@poczta.onet.pl*

- Szkoła Główna Gospodarstwa Wiejskiego (SGGW)  
*Nowoursynowska 166, Warszawa 02-787, Poland*

**Beatriz Larraz Iribas**

*Beatriz.Larraz@uclm.es*

- University of Castilla-La Mancha, Faculty of Economics and Business Studies of Albacete (UCLM)  
*Plaza de la Universidad, Albacete 02071, Spain*

**Iwona E. Laskowska**

*ilaskow@uni.lodz.pl*

- Uniwersytet Łódzki, Wydział Ekonomiczno-Socjologiczny  
*P.O.W. 3/5, Łódź 90-255, Poland*

---

**Artur Lipieta**

*lipietaa@ae.krakow.pl*

- Akademia Ekonomiczna w Krakowie, Katedra Statystyki (AE)  
*Rakowicka 27, Kraków 31-510, Poland*

**Aleksandra Łuczak**

*luczak@au.poznan.pl*

- Akademia Rolnicza (AR)  
*ul. Wojska Polskiego 28, Poznań 60-637, Poland*

**Piotr Łukasiewicz**

*plukasiewicz@mors.sggw.waw.pl*

- Szkoła Główna Gospodarstwa Wiejskiego (SGGW)  
*Nowoursynowska 166, Warszawa 02-787, Poland*

**Paweł Lula**

*pawel.lula@post.pl*

- Akademia Ekonomiczna w Krakowie, Katedra Informatyki  
*Rakowicka 27, Kraków 31-510, Poland*

**Małgorzata L. Łuniewska**

*mmmlun@uoo.univ.szczecin.pl*

- Uniwersytet Szczeciński  
*Jedności Narodowej 22, Szczecin 70-453, Poland*

**Małgorzata Machowska-Szewczyk**

*mmachowska@wi.ps.pl*

- Politechnika Szczecińska, Wydział Informatyki (PS)  
*ul. Żołnierska 49, Szczecin 71-210, Poland*

**Anna Malina**

*malinaa@ae.krakow.pl*

- Akademia Ekonomiczna w Krakowie (AE)  
*Rakowicka 27, Kraków 31-510, Poland*

**Małgorzata Markowska**

*mmarkowska@ae.jgora.pl*

- Akademia Ekonomiczna (AE)  
*Nowowiejska3, Jelenia Góra 58-500, Poland*

**Stanisław W. Matusik**

*matusik@poczta.onet.pl*

- Akademia Wychowania Fizycznego, Zakład Statystyki i Informatyki, Instytut Turystyki i Rekreacji  
*al. Jana Pawła II 78, Kraków 31-571, Poland*

**Aleksandra Matuszewska**

*amatuszewska@mors.sggw.waw.pl*

- Szkoła Główna Gospodarstwa Wiejskiego (SGGW)  
*Nowoursynowska 166, Warszawa 02-787, Poland*

**Agnieszka Mazur**

*agnieszka@oizet.p.lodz.pl*

- Politechnika Łódzka, Katedra Zarządzania  
*Piotrkowska 266, Łódź 90-924, Poland*

**Artur Mikulec**

*amikulec@poczta.onet.pl*

- Uniwersytet Łódzki, Katedra Metod Statystycznych (UŁ)  
*Revolucji 1905 r. nr 41, Łódź 90-214, Poland*
- Urząd Statystyczny w Łodzi, Wydział Analiz (USWA)  
*Suwalska 29, Łódź 93-176, Poland*

**Małgorzata A. Misztal**

*mmisztal@toya.net.pl*

- Uniwersytet Łódzki, Katedra Metod Statystycznych (UŁ)  
*Revolucji 1905 r. nr 41, Łódź 90-214, Poland*

**Magdalena C. Mojsiewicz**

*mojsiv@uoo.univ.szczecin.pl*

- Uniwersytet Szczeciński  
*Jedności Narodowej 22, Szczecin 70-453, Poland*

**Jose Maria Montero Lorenzo**

*Jose.M.Lorenzo@uclm.es*

- University of Castilla-La Mancha, Faculty of Economics and Business Studies of Albacete (UCLM)  
*Plaza de la Universidad, Albacete 02071, Spain*

**Kamila Najman**

*kmn@panda.bg.univ.gda.pl*

- Uniwersytet Gdański Wydział Zarządzania (WZR)  
*Armii Krajowej 101, Sopot 81-824, Poland*

**Krzysztof Najman**

*knajman@wzr.pl*

- Uniwersytet Gdański Wydział Zarządzania (WZR)  
*Armii Krajowej 101, Sopot 81-824, Poland*

**Kesra Nermend**

*kesra@szafir.univ.szczecin.pl*

- Uniwersytet Szczeciński  
*Szczecin 71-415, Poland*

---

**Arkadiusz J. Orłowski**

*orlow@ifpan.edu.pl*

- Szkoła Główna Gospodarstwa Wiejskiego (SGGW)  
*Nowoursynowska 166, Warszawa 02-787, Poland*
- Polska Akademia Nauk (PAN)  
*Al. Lotników 32/46, Warszawa 02-668, Poland*

**Daniel Papla**

*daniel.papla@ae.wroc.pl*

- Akademia Ekonomiczna (AEWROCLAW)  
*Komandorska 118/120, Wrocław 53-345, Poland*

**Jan Paradysz**

*jan.paradysz@ae.poznan.pl*

- Akademia Ekonomiczna w Poznaniu  
*al. Niepodległości 10, Poznań 60-967, Poland*

**Barbara Pawelek**

*Barbara.Pawelek@ae.krakow.pl*

- Akademia Ekonomiczna w Krakowie, Katedra Statystyki (AE)  
*Rakowicka 27, Kraków 31-510, Poland*

**Roman Pawlukowicz**

*roman.pawlukowicz@ae.jgora.pl*

- Akademia Ekonomiczna (AE)  
*Nowowiejska3, Jelenia Góra 58-500, Poland*

**Marcin Pelka**

*marcin.pelka@ae.jgora.pl*

- Akademia Ekonomiczna (AEWROCLAW)  
*Komandorska 118/120, Wrocław 53-345, Poland*

**Roman Pielaszek**

*roman@pielaszek.net*

- Polish Academy of Sciences, Institute of High Pressure Physics (UNIPRESS)  
*Sokolowska 29/37, Warszawa 01-142, Poland*

**Krzysztof L. Piontek**

*krzysztof.piontek@ae.wroc.pl*

- Akademia Ekonomiczna (AEWROCLAW)  
*Komandorska 118/120, Wrocław 53-345, Poland*

**Iwona Roeske-Słomka**

*ksid@ae.poznan.pl*

- Akademia Ekonomiczna w Poznaniu  
*al. Niepodległości 10, Poznań 60-967, Poland*

**Paweł Rokita**

*pawel.rokita@ae.wroc.pl*

- Akademia Ekonomiczna (AEWROCLAW)  
*Komandorska 118/120, Wrocław 53-345, Poland*

**Dominik A. Rozkrut**

*dominic@univ.szczecin.pl*

- Uniwersytet Szczeciński  
*Jedności Narodowej 22, Szczecin 70-453, Poland*

**Monika Rozkrut**

*sobot@uoo.univ.szczecin.pl*

- Uniwersytet Szczeciński  
*Jedności Narodowej 22, Szczecin 70-453, Poland*

**Dorota E. Rozmus**

*drozmus@ae.katowice.pl*

- Akademia Ekonomiczna im. K. Adameckiego w Katowicach  
*Katowice 40-287, Poland*

**Danuta J. Rozpędowska-Matraszek**

*daroma@uni.lodz.pl*

- Uniwersytet Łódzki, Wydział Ekonomiczno-Socjologiczny  
*P.O.W. 3/5, Łódź 90-255, Poland*

**Marcin Salamaga**

*salamaga@ae.krakow.pl*

- Akademia Ekonomiczna w Krakowie (AE)  
*Rakowicka 27, Kraków 31-510, Poland*

**Elżbieta Sobczak**

*sobczak@ae.jgora.pl*

- Akademia Ekonomiczna (AE)  
*Nowowiejska3, Jelenia Góra 58-500, Poland*

**Agnieszka Sompolska-Rzechuła**

*arzechula@e-ar.pl*

- Akademia Rolnicza (AR)  
*Janosika 8, Szczecin 71-424, Poland*

**Iwona Staniec**

*istan@p.lodz.pl*

- Politechnika Łódzka, Katedra Zarządzania  
*Piotrkowska 266, Łódź 90-924, Poland*

**Beata Szczecińska**

*bszczecinska@e-ar.pl*

- Akademia Rolnicza (AR)  
*Janosika 8, Szczecin 71-424, Poland*

---

**Wiesław Szczesny**

*wszczesny@mors.sggw.waw.pl*

- Szkoła Główna Gospodarstwa Wiejskiego (SGGW)  
*Nowoursynowska 166, Warszawa 02-787, Poland*

**Anna Szmit**

*agorecka@p.lodz.pl*

- Politechnika Łódzka, Katedra Zarządzania  
*Piotrkowska 266, Łódź 90-924, Poland*

**Mirosław W. Szreder**

*mszreder@wzr.pl*

- Uniwersytet Gdański Wydział Zarządzania (WZR)  
*Armii Krajowej 101, Sopot 81-824, Poland*

**Mirosława Sztemberg-Lewandowska**

*mirosława.sztemberg-lewandowska@ae.jgora.pl*

- Akademia Ekonomiczna (AE)  
*Nowowiejska3, Jelenia Góra 58-500, Poland*

**Krzysztof Szware**

*k.szware@ae.poznan.pl*

- Akademia Ekonomiczna w Poznaniu  
*al. Niepodległości 10, Poznań 60-967, Poland*

**Marcin Szymkowiak**

*m.szymkowiak@ae.poznan.pl*

- Akademia Ekonomiczna w Poznaniu  
*al. Niepodległości 10, Poznań 60-967, Poland*

**Wldemar S. Tarczyński**

*wtarc@uoo.univ.szczecin.pl*

- Uniwersytet Szczeciński  
*Jedności Narodowej 22, Szczecin 70-453, Poland*

**Joanna Trzęsiok**

*trzesiok@ae.katowice.pl*

- Akademia Ekonomiczna im. K. Adamieckiego w Katowicach  
*Katowice 40-287, Poland*

**Michał Trzęsiok**

*trzesiok@ae.katowice.pl*

- Akademia Ekonomiczna im. K. Adamieckiego w Katowicach  
*Katowice 40-287, Poland*

**Paweł Ulman**

*ulmanp@ae.krakow.pl*

- Akademia Ekonomiczna w Krakowie, Katedra Statystyki (AE)  
*Rakowicka 27, Kraków 31-510, Poland*

**Marek Walesiak**

*marek.walesiak@ae.jgora.pl*

- Akademia Ekonomiczna (AEWROCLAW)  
*Komandorska 118/120, Wrocław 53-345, Poland*

**Katarzyna Wawrzyniak**

*kwawrzyniak@e-ar.pl*

- Akademia Rolnicza (AR)  
*Janosika 8, Szczecin 71-424, Poland*

**Elżbieta Wiszniewska**

*elka@o2.pl*

- Uniwersytet Łódzki, Wydział Ekonomiczno-Socjologiczny  
*P.O.W. 3/5, Łódź 90-255, Poland*

**Ewa M. Witek**

*eewwaaww@interia.pl*

- Akademia Ekonomiczna im. K. Adamieckiego w Katowicach  
*Katowice 40-287, Poland*

**Aleksandra Witkowska**

*a.witkowska@ae.poznan.pl*

- Akademia Ekonomiczna w Poznaniu  
*al. Niepodległości 10, Poznań 60-967, Poland*

**Dorota Witkowska**

*dwitkowska@mors.sggw.waw.pl*

- Szkoła Główna Gospodarstwa Wiejskiego (SGGW)  
*Nowoursynowska 166, Warszawa 02-787, Poland*

**Mirosław Wójciak**

*mwojciak@ae.katowice.pl*

- Akademia Ekonomiczna im. K. Adamieckiego w Katowicach  
*Katowice 40-287, Poland*

**Aleksandra Wójcicka-Krenz**

*aleksandra.wojcicka@ae.poznan.pl*

- Akademia Ekonomiczna w Poznaniu  
*al. Niepodległości 10, Poznań 60-967, Poland*

**Feliks Wysocki**

*wysocki@au.poznan.pl*

- Akademia Rolnicza (AR)  
*ul. Wojska Polskiego 28, Poznań 60-637, Poland*

**Tomasz Ząbkowski**

*tzabkowski@mors.sggw.waw.pl*

- Szkoła Główna Gospodarstwa Wiejskiego (SGGW)  
*Nowoursynowska 166, Warszawa 02-787, Poland*

---

**Artur P. Zaborski**

*artur.zaborski@ae.jgora.pl*

- Akademia Ekonomiczna (AE)  
*Nowowiejska3, Jelenia Góra 58-500, Poland*

**Kazimierz Zając**

*kazimierz.zajac@wszins.tychy.pl*

- Akademia Ekonomiczna w Krakowie, Katedra Statystyki (AE)  
*Rakowicka 27, Kraków 31-510, Poland*

**Dorota Żebrowska-Suchodolska**

*zdorota@o2.pl*

- Szkoła Główna Gospodarstwa Wiejskiego (SGGW)  
*Nowoursynowska 166, Warszawa 02-787, Poland*
- Wyższa Szkoła Finansów i Zarządzania w Białymstoku  
*Ciepła 40, Białystok 15-472, Poland*

**Paweł Żuraw**

*piotr.tarka@op.pl*

- Akademia Ekonomiczna (AEWROCLAW)  
*Komandorska 118/120, Wrocław 53-345, Poland*



---

# Indeks

## B

Banaś, Joanna, 33  
Batóg, Barbara, 26, 31  
Batóg, Jacek, 27  
Bąk, Andrzej, 14  
Błażejowski, Marcin, 39  
Borys, Tadeusz, 23

## C

Chrzanowska, Mariola, 21  
Cortes, Esteban A., 10

## D

Dańska-Borsiak, Barbara, 24, 45  
Decker, Reinhold, 9  
Dehnel, Grażyna, 37  
Dudek, Andrzej, 19  
Dudek, Hanna, 35

## F

Foryś, Iwona A., 25

## G

Gamez Martinez, Matias, 10  
Garcia Rubio, Noelia, 10  
Gatnar, Eugeniusz, 11  
Gazińska, Mirosława, 11  
Gaziński, Radosław, 11  
Gliwa, Małgorzata, 17  
Grabowski, Mariusz, 32  
Grzybowska, Urszula, 44

## J

Jabłonowski, Stanisław K., 37  
Jajuga, Krzysztof, 9  
Jałowiecka, Ewa, 46  
Jałowiecki, Piotr, 44, 46

## K

Karwański, Marek, 38, 46  
Kasprzyk, Iwona, 15  
Kisielińska, Joanna, 35  
Kociuba, Łukasz J., 9  
Kołodenny, Włodzimierz, 46  
Kompa, Krzysztof, 22  
Korzeniewski, Jerzy M., 20  
Kostrzewska, Jadwiga H., 38  
Koszela, Grzegorz, 36  
Kubus, Mariusz, 20  
Kufel, Tadeusz, 39  
Kurzawa, Izabela, 36

Kuziak, Katarzyna, 21

## L

Landmesser, Joanna M., 34  
Larraz Iribas, Beatriz, 10  
Laskowska, Iwona E., 32  
Lewandowska, Karolina E., 32  
Lula, Paweł, 31  
Łuczak, Aleksandra, 17  
Łukasiewicz, Piotr, 36  
Łuniewska, Małgorzata L., 10, 39

## M

Machowska-Szewczyk, Małgorzata, 31, 33  
Markowska, Małgorzata, 19, 23  
Martin, Marek, 34  
Matusik, Stanisław W., 25  
Matuszewska, Aleksandra, 43  
Mikulec, Artur, 44  
Miształ, Małgorzata A., 33  
Mojsiewicz, Magdalena C., 26, 31  
Montero Lorenzo, Jose Maria, 10

## N

Najman, Kamila, 15  
Najman, Krzysztof, 16  
Nermend, Kesra, 46

## O

Orłowski, Arkadiusz J., 36, 44

## P

Papla, Daniel, 27  
Paradysz, Jan, 14  
Pawełek, Barbara, 13  
Pełka, Marcin, 19  
Piontek, Krzysztof L., 22

## R

Rokita, Paweł, 26  
Rozkrut, Dominik A., 23  
Rozkrut, Monika, 24  
Rozmus, Dorota E., 17  
Rozpędowska-Matraszek, Danuta J., 45, 45

## S

Salamaga, Marcin, 27  
Sobczak, Elżbieta, 23  
Sompolska-Rzechuła, Agnieszka, 44  
Staniec, Iwona, 34, 39  
Strahl, Danuta, 19  
Szczecińska, Beata, 44  
Szczesny, Wiesław, 38  
Szmit, Anna, 39  
Sztemberg-Lewandowska, Mirosława, 39

---

Szwarc, Krzysztof, 39  
Szymkowiak, Marcin, 14

## **T**

Tarczyński, Wldemar S., 10  
Tarka, Piotr, 28  
Trzęsiok, Joanna, 18  
Trzęsiok, Michał, 18

## **U**

Ulman, Paweł, 40

## **W**

Walesiak, Marek, 9  
Wawrzyniak, Katarzyna, 26, 31  
Wisniewska, Elżbieta, 24  
Witek, Ewa M., 13  
Witkowska, Aleksandra, 33  
Witkowska, Dorota, 21, 43, 44  
Witkowski, Marek, 33  
Wójciak, Mirosław, 40, 43  
Wójcicka-Krenz, Aleksandra, 40, 41, 43  
Wysocki, Feliks, 17, 36

## **Z**

Zaborski, Artur P., 16  
Ząbkowski, Tomasz, 28  
Żdanowicz, Tadeusz, 46  
Żebrowska-Suchodolska, Dorota, 44  
Żółtowski, Jan, 39  
Żuraw, Paweł, 28