

SPRAWOZDANIA

ANNA WITASZCZYK, ALEKSANDRA KUPIS-FIJAŁKOWSKA

SPRAWOZDANIE Z XXX KONFERENCJI WIELOWYMIAROWA ANALIZA STATYSTYCZNA 2011

W dniach 7-9 listopada 2011 r. w Centrum Szkoleniowo-Konferencyjnym Uniwersytetu Łódzkiego odbyła się XXX Konferencja Wielowymiarowa Analiza Statystyczna (MSA 2011), organizowana przez Katedrę Metod Statystycznych Wydziału Ekonomiczno-Socjologicznego Uniwersytetu Łódzkiego. Współorganizatorami konferencji byli: Polskie Towarzystwo Statystyczne, Komitet Statystyki i Ekonometrii Polskiej Akademii Nauk. Przewodniczącym Komitetu Organizacyjnego był prof. dr hab. Czesław Domański, natomiast sekretarzami – mgr Aleksandra Kupis-Fijałkowska i dr Anna Witaszczyk. Tematyka konferencji obejmowała: rozkłady wielowymiarowe, wnioskowanie statystyczne, metody nieparametryczne, metody klasyfikacji i dyskryminacji danych, analizę głównych składowych, regresję, metody Monte Carlo, metody bayesowskie, metody bootstrapowe, planowanie eksperymentów, zastosowania statystyki w naukach przyrodniczych, ekonomiczno-społecznych, finansach i ubezpieczeniach, demografii, kontroli jakości, zarządzaniu ryzykiem.

W konferencji uczestniczyli przedstawiciele następujących uczelni: Szkoła Główna Handlowa w Warszawie, Szkoła Główna Gospodarstwa Wiejskiego, Uniwersytet im. Adama Mickiewicza w Poznaniu, Uniwersytet Ekonomiczny w Poznaniu, Uniwersytet Przyrodniczy w Poznaniu, Uniwersytet Ekonomiczny we Wrocławiu, Politechnika Opolska, Uniwersytet Ekonomiczny w Katowicach, Uniwersytet Szczeciński, Uniwersytet Łódzki. Uczestnikami konferencji były również osoby z następujących uczelni zagranicznych: Singapore Management University, Universite de Bourgogne, Mimar Sinan University of Fine Arts. W konferencji wzięło udział około 70 osób – pracowników i doktorantów uczelni. Ze względu na dydaktyczny charakter, słuchaczami byli również studenci kierunku studiów Informatyka i Ekonometria UŁ oraz doktoranci Wydziału Ekonomiczno-Socjologicznego.

Podczas sesji plenarnych, jak i równoległych, wygłoszono 46 referatów i dwa wykłady zaproszone.

Pierwsza z sesji plenarnych poświęcona została wybitnym polskim naukowcom związanym ze statystyką: Józefowi Kleczyńskiemu w 170 rocznicę urodzin (Czesław Domański, Uniwersytet Łódzki), Samuelowi Dicksteinowi w 160 rocznicę urodzin (Czesław Domański). Przedstawiono także sylwetki profesorów: Antoniego Łomnickiego (Miroslaw Krzyśko, Uniwersytet im. Adama Mickiewicza w Poznaniu), Zbi-

gniewa Michałkiewicza (Jerzy T. Kowaleski, Uniwersytet Łódzki), Wiesława Sadowskiego (Władysław Welfe, Uniwersytet Łódzki), Zofii Zarzyckiej (Wacława Starzyńska, Uniwersytet Łódzki), Piotra Chrzana (Janusz Wywiał, Uniwersytet Ekonomiczny w Katowicach), Zbigniewa Czerwińskiego (Grażyna Dehnel, Uniwersytet Ekonomiczny w Poznaniu), Wiesława Wagnera (Bronisław Ceranka, Uniwersytet Przyrodniczy w Poznaniu).

Wykłady zaproszone wygłosili: prof. dr hab. Stanisława Bartosiewicz (Wyższa Szkoła Bankowa we Wrocławiu) – *Kilka „nieuczestnych” myśli na temat traktatu o dynamice, stanie i perspektywach rozwoju polskiego rolnictwa*, prof. dr hab. Mirosław Krzyżko – *Wersja jądrowa funkcjonalnych składowych głównych*.

Zagraniczni uczestnicy konferencji wygłosili pięć referatów: Muhammad Aslam (Universite de Bourgogne, Dijon) – *Robust Inference for Linear Regression Model in the Presence of Heteroscedasticity and Multicollinearity*; Sajjad Haider Bhatti (Universite de Bourgogne, Dijon) – *A Panel Data Analysis of the Relationship between GDP per Capita and its Shares of Consumption, Government Expenditures and Investment*; Kian-Guan Lim (Singapore Management University) – *Statistical Tests of Conditional Shortfalls and Change-Point Estimation in Risk Management*; Samaradasa Weerahandi (Pfizer Inc., New York) – *Generalized Point Estimation with Application to Mixed Models*; Meral Yay (Mimar Sinan University of Fine Arts) – *A Content Analysis of the Representation of the European Uprising in Turkish News Media*.

Ponadto zaprezentowano następujące referaty (według nazwisk autorów w kolejności alfabetycznej:

Aleksandra Baszczyńska (Uniwersytet Łódzki), *Uwagi o jądrowym teście symetryczności* W pracy przedstawiono wybrane statystyczne testy wykorzystywane do weryfikacji hipotezy o symetryczności rozkładu zmiennej losowej. Szczegółowej analizie poddano test symetryczności oparty o metodę jądrową. Porównano własności zaprezentowanych testów symetryczności oraz zastosowano je go analizy rozkładu wskaźnika rozwoju społecznego (HDI).

Jacek Białek (Uniwersytet Łódzki), *Propozycja indeksu cen*. Przedstawiono i omówiono nową formułę indeksu cen. Pokazano, iż proponowana formuła spełnia podstawowe testy pochodzące z aksjomatycznej teorii indeksów. Na przykładzie porównano omawiany indeks z indeksami cen znanymi z literatury

Beata Bieszk-Stolorz, Iwona Markowicz (Uniwersytet Szczeciński), *Aktywność ekonomiczna kobiet i mężczyzn w Polsce. Interpretacja modelu logitowego z jedną i z wieloma zmiennymi objaśniającymi*. Celem pracy było zbadanie wpływu płci na szansę aktywności zawodowej oraz na szansę posiadania zatrudnienia. Badaniem objęto ludność Polski w wieku 15 lat i więcej w 2010 roku (31741 tys. osób). Punktem wyjścia analizy była budowa modeli uzależniających aktywność zawodową oraz posiadanie zatrudnienia jedynie od płci. Następnie rozważano modele, w których zmiennymi

objaśniającymi są następujące pary: płeć i miejsce zamieszkania, płeć i wykształcenie, płeć i wiek.

Justyna Brzezińska (Uniwersytet Ekonomiczny w Katowicach), *Kryteria wyboru modelu w analizie logarytmiczno-liniowej* Analiza logarytmiczno-liniowa jest metodą przeznaczoną do badania zależności pomiędzy zmiennymi niemetrycznymi w tablicy kontyngencji. Metoda ta pozwala na analizę dowolnej liczby zmiennych, a także na uwzględnienie interakcji zachodzących pomiędzy nimi. W analizie logarytmiczno-liniowej modelowane są liczebności w poszczególnych komórkach tablicy. Celem pracy była prezentacja i porównanie kryteriów wyboru modelu. Podstawowymi kryteriami wyboru modelu są statystyka chi-kwadrat, iloraz wiarygodności oraz kryteria informacyjne *AIC* i *BIC*. Zaprezentowana została także metoda Aitkina, która przeznaczona jest do porównywania jakości dopasowania modeli o dużej liczbie zmiennych.

Bronisław Ceranka, Małgorzata Graczyk (Uniwersytet Przyrodniczy w Poznaniu), *Konstrukcja E- optymalnego regularnego układu wagowego*. Rozważano model liniowy $y = Xw + e$, który opisuje w jaki sposób wyznaczyć nieznanne miary p obiektów w oparciu o n operacji pomiarowych w eksperymencie przeprowadzonym zgodnie z macierzą układu X o elementach $x_{ij} = 0, 1, i=1,2,\dots,n, j=1,2,\dots,p$, gdzie y jest losowym wektorem obserwowanych pomiarów, w jest wektorem nieznanymi miar obiektów. Zakłada się, że nie ma błędów systematycznych, że są nieskorelowane i mają różne wariancje, tzn. dla losowego wektora błędów e , $E(e) = \mathbf{0}_n$ i $E(ee') = \sigma^2 G$, gdzie G jest znaną, dodatnio określoną macierzą diagonalną. Regularny układ E- optymalny jest to układ, w którym maksymalna wartość własna macierzy kowariancji jest minimalna i osiąga dolne ograniczenie. W pracy przedstawiono nowe metody konstrukcji regularnego E- optymalnego sprężynowego układu wagowego w oparciu o macierze incydencji układów o grupach podzielnych o tym samym schemacie partnerstwa.

Chmielińska Magdalena (Uniwersytet Ekonomiczny w Katowicach), *Wpływ wyboru metody sprawdzania na całkowite koszty kontroli jakości*. Kontrola odbiorcza jest powszechnie stosowana w przedsiębiorstwach produkcyjnych, gdyż zmniejsza ona prawdopodobieństwo przekazania klientowi wadliwej partii produktów. Z wyróżnionych dwóch sposobów kontroli odbiorczej - według oceny alternatywnej i według oceny właściwości liczbowych- częściej wykorzystywana w praktyce gospodarczej jest kontrola według oceny alternatywnej. Celem pracy było podkreślenie wymiernych korzyści związanych z stosowaniem planów kontroli odbiorczej według liczbowej oceny właściwości, w szczególności w wymiarze ekonomicznym. Postawiona została hipoteza głosząca, iż plany odbiorcze oparte na ocenie właściwości liczbowych generują niższe ogólne koszty kontroli odbiorczej niż plany odbiorcze oparte na ocenie alternatywnej, przy tych samych wymaganiach producenta i odbiorcy. Rozważania zostały uzupełnione o symulacje komputerowe przeprowadzone w programie R.

Joanna Dębicka (Uniwersytet Ekonomiczny we Wrocławiu), *Macierzowe podejście do analizy portfela polis ubezpieczeń wielostanowych*. Celem pracy było przedstawienie ogólnego podejścia umożliwiającego wyznaczenie momentów zdyskontowanych przyszłych przepływów pieniężnych wynikających z zawarcia umowy ubezpieczenia wielostanowego ze stałą i stochastyczną stopą procentową (zarówno dla umowy indywidualnej jak i portfela polis). W celu uproszczenia zapisu i obliczeń numerycznych zaproponowana została notacja macierzowa. Wykorzystanie notacji macierzowej umożliwiło efektywny zapis w przypadku formułowania problemów aplikacyjnych. W przypadku ubezpieczeń grupowych, pokazano jak zapis macierzowy wykorzystuje się w analizie wpływu liczebności portfela polis na wysokość składki i narzutu bezpieczeństwa, zarówno dla grupy jednorodnej jak i niejednorodnej (w przypadku małego i dużego portfela polis). Teoria ilustrowana jest przykładami numerycznymi, w których struktura probabilistyczna modeli, związanych z analizowanymi ubezpieczeniami, wyznaczona została w oparciu o Tablice Trwania Życia, a proces stopy procentowej modelowany jest przez proces Wienera i proces Ornsteina-Uhlenbeck'a.

Andrzej Dudek (Uniwersytet Ekonomiczny we Wrocławiu), *Klasyfikacja dużych zbiorów danych. Problemy, metody, algorytmy*. Badacze analizujący przy pomocy metod analizy skupień duże (powyżej 10 000 obiektów) zbiory danych, stają często przed problemem złożoności obliczeniowej algorytmów, uniemożliwiającej niekiedy przeprowadzenie analizy w akceptowalnym czasie. Jednym z rozwiązań tego problemu jest stosowanie mniej złożonych obliczeniowo algorytmów (hierarchiczne aglomeracyjne, k-średnich), które z kolei mogą w wielu sytuacjach dawać zdecydowanie gorsze rezultaty niż np. algorytmy wykorzystujące dekompozycję względem wartości własnych. Rezultaty rzeczywistych analiz tego typu zbiorów są więc zazwyczaj kompromisem pomiędzy jakością a możliwościami obliczeniowymi komputerów. W pracy podjęto próbę przedstawienia aktualnego stanu wiedzy na temat klasyfikacji dużych zbiorów danych oraz wskazania dróg rozwoju i problemów otwartych.

Wojciech Gamrot (Uniwersytet Ekonomiczny w Katowicach), *O estymatorze Fattoriniego*. Empiryczny estymator Horwitza-Thompsona wartości globalnej cechy w populacji skończonej konstruowany jest poprzez zastąpienie nieznanymi prawdopodobieństw inkluzji rzędu pierwszego charakteryzujących schemat losowania próby oszacowaniami. Oszacowania te wyznaczone są w eksperymencie symulacyjnym polegającym na wielokrotnym losowaniu próby według tego samego schematu i zliczaniu wystąpień każdego z elementów w próbie. Zaproponowany przez Fattoriniego (2006) estymator prawdopodobieństwa inkluzji obliczany jest jako frakcja z próby zmodyfikowana poprzez dodanie do licznika i mianownika stałej równej jedności. W referacie rozważane były inne sposoby ustalania tej stałej i ich wpływ na własności oszacowań wartości globalnej.

Eugeniusz Gatnar (Uniwersytet Ekonomiczny w Katowicach), *Bank centralny w systemie statystyki publicznej*. W referacie została omówiona rola Narodowego Banku Polskiego, który jest bankiem centralnym Polski oraz członkiem Europejskiego Systemu Banków Centralnych (ESCB) w systemie statystyki publicznej w Polsce.

Elżbieta Gołata, Grażyna Dehnel (Uniwersytet Ekonomiczny w Poznaniu), *Społeczny wymiar procesów demograficznych*. Konsekwencje przemian demograficznych rozpatrywać można we wszystkich wymiarach, począwszy od najmniejszej jednostki społecznej jaką jest rodzina, gospodarstwo domowe, poprzez ujęcie lokalne, regionalne oraz w skali globalnej całej gospodarki łącznie z powiązaniem międzynarodowymi. Przedmiotem referatu była próba określenia reakcji samorządów na wyzwania wynikające ze zmieniającej się sytuacji demograficznej, w tym w szczególności na proces starzenia się społeczeństwa. Obok podstawowych metod analizy demograficznej, w badaniu wykorzystano metody wielowymiarowej analizy statystycznej, w tym porównawczej, w odniesieniu do danych różnych źródeł informacji, przede wszystkim ewidencji ruchu naturalnego ludności, Banku Danych Lokalnych GUS, ale także badań specjalnych, jak np. *Przemiany demograficzne – szczególne wyzwanie dla urzędów miast i gmin w Polsce*.

Małgorzata Graczyk (Uniwersytet Przyrodniczy w Poznaniu), *Regularny A- optymalny sprężynowy układ wagowy o skorelowanych błędach*. W pracy rozważana była tematyka estymacji nieznanymi miar obiektów w sprężynowym układzie wagowym. Założono, że błędy są jednakowo skorelowane. Podano warunki określające istnienie A- optymalnego regularnego sprężynowego układu wagowego oraz metodę konstrukcji. Do konstrukcji macierzy układu wykorzystano macierze incydencji układów zrównoważonych o blokach niekompletnych.

Wioletta Grzenda (Szkoła Główna Handlowa w Warszawie), *Znaczenie informacji a priori w bayesowskich parametrycznych modelach przeżycia*. W pracy przedstawiono parametryczne modele przeżycia w ujęciu bayesowskim. Podejście bayesowskie wymaga zadania rozkładów *a priori* dla szacowanych parametrów modelu. W celu pokazania znaczenia informacji *a priori* oraz jej wpływu na rozkład *a posteriori* oszacowano kilka parametrycznych modeli przeżycia przy różnych rozkładach *a priori*. Przedmiot badań stanowiły determinanty długości czasu pozostawania bez pracy. Estymację wszystkich modeli przeprowadzono w systemie SAS z wykorzystaniem metod Monte Carlo opartych na łańcuchach Markowa, w szczególności próbnika Gibbsa.

Joanna Kisielińska (Szkoła Główna Gospodarstwa Wiejskiego w Warszawie), *Doładna metoda bootstrapowa na przykładzie estymacji średniej i wariancji*. Metoda bootstrapowa polega na wtórnym próbkowaniu pierwotnej próby losowej pobranej z populacji o nieznanym rozkładzie. W referacie pokazano, że ze względu na postęp w technice komputerowej wtórne próbkowanie nie jest aktualnie konieczne, jeśli

rozmiar próby nie jest zbyt duży. Możliwe jest wówczas automatyczne wygenerowanie wszystkich możliwych prób wtórnych i obliczenie wszystkich realizacji wybranej statystyki. Uzyskany rozkład można wykorzystać do punktowej lub przedziałowej estymacji parametrów populacji, bądź testowania hipotez. Metodę zastosowano do szacowania średniej i wariancji. Porównanie uzyskanych rozkładów z rozkładami granicznymi potwierdziło poprawność dokładnej metody bootstrapowej.

Paweł Kobus (Szkoła Główna Gospodarstwa Wiejskiego w Warszawie), *Wykorzystanie funkcji kopula do modelowania łącznego rozkładu plonów roślin uprawnych*. Celem pracy było sprawdzenie możliwości wykorzystania funkcji kopula do modelowania zależności pomiędzy plonami i cenami roślin uprawnych, w kontekście ryzyka dochodowego w gospodarstwie rolnym. Brak precyzji w modelowaniu wspomnianych zależności może prowadzić do poważnych błędów w optymalizacji struktury gospodarstwa i wyższego ryzyka niż zakładano. W pracy porównano popularne, w modelach gospodarstw rolnych, podejście oparte na macierzy korelacji z podejściem wykorzystującym funkcje kopula.

Grzegorz Kończak (Uniwersytet Ekonomiczny w Katowicach), *O testowaniu kierunkowych hipotez w analizie jednorodności danych*. Test jednorodności chi-kwadrat jest wykorzystywany gdy ze względu na jedną zmienną jest analizowanych kilka prób. Jeżeli dane są jednorodne, to proporcje obserwacji dla j -tej kategorii będą jednakowe dla wszystkich wyróżnionych kategorii. W przypadku tablicy o wymiarach 2×2 możliwe jest rozważanie kierunkowej hipotezy alternatywnej. W przypadku tablic wielodzielczych o większych wymiarach możliwe jest rozważanie wielokrotnych alternatywnych hipotez kierunkowych. W pracy została przedstawiona propozycja testu pozwalającego na weryfikację takich hipotez. Analizę własności testu przeprowadzono z wykorzystaniem symulacji komputerowych.

Jerzy Korzeniewski (Uniwersytet Łódzki), *Ocena efektywności metody Talavery wybierania zmiennych w analizie skupień na empirycznych zbiorach danych*. Talavera zaproponował metodę wybierania zmiennych tworzących strukturę skupień w zbiorze danych dla zbiorów, w których występują tylko zmienne mierzone na skali nominalnej. W pracy zbadano tę metodę na kilku empirycznych zbiorach opierając ocenę na tym jak spisowała się metoda w połączeniu z ustalonym sposobem grupowania danych (algorytm COBWEB). W innych podejściach do tego samego zagadnienia autorzy starają się oprzeć wybór zmiennych na samym uporządkowaniu zbioru zmiennych bez odwoływania się do grupowania obserwacji. W referacie badano także efektywność metody, również w odniesieniu do empirycznych zbiorów danych, uzależnioną tylko od uporządkowania zmiennych, opartą na kryterium największego skoku. Rozważano ponadto zbiory z niektórymi zmiennymi mierzonymi na mocniejszych skalach po uprzedniej dyskretyzacji zmiennych.

Mariusz Kubus (Politechnika Opolska), *O wyborze postaci modelu w wybranych metodach regularyzowanej regresji liniowej*. W ostatnich latach można zaobserwować dynamiczny rozwój różnych postaci regularyzacji w modelach liniowych. Wprowadzenie kary za duże wartości współczynników skutkuje zmniejszeniem wariancji oraz eliminacją niektórych zmiennych. Dodatkową korzyścią tych metod jest możliwość zastosowania w szczególnym przypadku, gdy liczba zmiennych jest większa od liczby obserwacji. Selekcja zmiennych za pomocą regularyzowanych modeli liniowych jest w problemach wielowymiarowych preferowana wobec popularnego podejścia polegającego na przeszukiwaniu przestrzeni cech i ocenie podzbiorów zmiennych za pomocą kryterium jakości modelu (*wrappers*). Przyczyną są mniejsze koszty obliczeń i mniejsza podatność na nadmierne dopasowanie. Celem pracy była ocena empiryczna kryteriów wyboru postaci modelu.

Marta Małecka (Uniwersytet Łódzki), *GARCH process application in risk valuation for WIG20 index*. W pracy przedstawiono zastosowanie modelu GARCH w estymacji wartości narażonej na ryzyko dla indeksu WIG20. Metoda ta została porównana z pięcioma innymi wykorzystywanymi w tym zakresie: metodą wariancji-kowariancji, symulacją historyczną, RiskMetrics, symulacją Monte Carlo, metodą bootstrapową. Celem referatu było wykazanie wyższości modelu GARCH nad innymi metodami oceny ryzyka.

Magdalena Makuch (Uniwersytet Ekonomiczny w Katowicach), *O strategii losowania grupowego*. Schemat losowania grupowego jest jednym z ciekawszych schematów losowania. Zakłada się, że populacja jest podzielona na G rozłącznych grup, do próby losujemy grupy, a nie pojedyncze elementy. Wiadomo, że wariancja średniej z próby grupowej zależy od sposobu podziału populacji na grupy. Podziału dokonujemy na podstawie cechy dodatkowej, obserwowanej w całej populacji. Zakładamy, że cecha dodatkowa jest silnie skorelowana z cechą badaną, której średnią należy oszacować. Wybiera się ten podział, który prowadzi do minimalizacji wariancji średniej z próby grupowej. W pracy rozważono sześć sposobów doboru elementów do próby.

Małgorzata Misztal (Uniwersytet Łódzki), *Kilka uwag o imputacji danych z wykorzystaniem metody „missForest”*. W pracy Stekhovena i Bühlmana (2011) zaproponowano nową iteracyjną metodę imputacji (nazwaną „missForest”) opartą na metodzie *Random Forests* Breimana (2001). W pracy omówiono metodę „missForest” i porównano z nią kilka wybranych technik postępowania w sytuacji występowania braków danych. W tym celu wykorzystano podejście symulacyjne generując różne proporcje i mechanizmy powstawania braków danych w zbiorach danych pochodzących z repozytorium baz danych na Uniwersytecie Kalifornijskim w Irvine.

Dorota Pekasiewicz (Uniwersytet Łódzki), *Bayesowskie testy statystyczne dla wskaźnika struktury dla niezależnego i zależnego schematu losowania próby*. W wyniku

zastosowania bayesowskich testów statystycznych podejmuje się decyzję o akceptacji hipotezy, dla której ryzyko a posteriori jest mniejsze. Ryzyko a posteriori zależy od rozkładu a priori rozważanego parametru, funkcji straty i schematu losowania próby. W pracy przedstawiono bayesowskie testy statystyczne dla wskaźnika struktury, w przypadku różnych rozkładów a priori, przy niezależnym i zależnym schemacie losowania próby. Oprócz rozważań teoretycznych, zaprezentowano wyniki analiz symulacyjnych dotyczących własności tych testów.

Marcin Pełka (Uniwersytet Ekonomiczny we Wrocławiu), *Klasyfikacja danych symbolicznych z wykorzystaniem podejścia wielomodelowego*. Obiekty symboliczne, w odróżnieniu od obiektów w ujęciu klasycznym, mogą być opisywane przez wiele różnych typów zmiennych. Oprócz zmiennych w ujęciu klasycznym (metrycznych lub niemetrycznych) mogą być opisywane przez zmienne interwałowe, zmienne wielowariantowe i zmienne wielowariantowe z wagami, a także zmienne strukturalne. W zagadnieniu klasyfikacji danych symbolicznych może być z powodzeniem stosowane podejście wielomodelowe, tj. łączenie wyników wielu modeli w jeden model zagregowany. Celem pracy było przedstawienie wyników klasyfikacji danych symbolicznych z wykorzystaniem podejścia wielomodelowego na przykładzie różnych zbiorów danych symbolicznych. W części empirycznej przedstawiono wyniki klasyfikacji przy zastosowaniu różnorodnych modeli bazowych.

Robert Pietrzykowski (Szkoła Główna Gospodarstwa Wiejskiego), *Zastosowanie modyfikacji macierzy wag w badaniu zależności przestrzennych*. Podstawowym elementem analiz przestrzennych jest określenie interakcji pomiędzy badanymi jednostkami przestrzennymi: krajem, województwem, powiatem, gminą, miastem i itd. Zgodnie z prawem Toblera jednostki przestrzenne sąsiadujące ze sobą powinny bardziej oddziaływać na siebie niż te, które znajdują się dalej. Powiązania przestrzenne należy oprzeć o definicję sąsiedztwa, którą można określić przy pomocy tzw. macierzy wag. Określenie macierzy wag w analizach przestrzennych odgrywa kluczową rolę. Ponieważ macierz wag określa się a priori, a jakość dalszych analiz zależy od jej specyfikacji, dobór odpowiedniej macierzy wag stanowi poważny problem metodologiczny. W pracy zaproponowano modyfikację macierzy wag, która zawierałaby interakcje związane z położeniem badanej jednostki przestrzennej, ale również uwzględniałaby wagi związane z wybranymi wieloma cechami.

Krystyna Pruska (Uniwersytet Łódzki), *Modele probitowe i logitowe w procesie estymacji frakcji dla małego obszaru*. W pracy tej rozpatrywano estymację frakcji dla małego obszaru na podstawie modeli probitowych i logitowych. Przedstawiono wyniki eksperymentów symulacyjnych w przypadku różnych schematów losowania próby (losowanie niezależne i zależne). Próby losowano z całej populacji lub z podpopulacji. Porównywano błędy estymacji frakcji dla małego obszaru.

Dorota Raczkiewicz (Szkola Główna Handlowa w Warszawie), *Kariery: edukacyjna, zawodowa i rodzinna kobiet w Polsce*. Celem pracy była analiza statystyczna syntetycznych wskaźników na poziomie makro, opisujących kariery: edukacyjną, zawodową i rodzinną kobiet w Polsce, oraz analiza powiązań między tymi wskaźnikami karier na podstawie dostępnych danych statystycznych.

Dorota Rozmus (Uniwersytet Ekonomiczny w Katowicach), *Porównanie dokładności taksonomii spektralnej oraz zagregowanych algorytmów taksonomicznych opartych na macierzy współwystąpień*. Stosując metody taksonomiczne w jakimkolwiek zagadnieniu klasyfikacji ważną kwestią jest zapewnienie wysokiej poprawności wyników grupowania. Od niej bowiem zależy skuteczność wszelkich decyzji podjętych na ich podstawie. Stąd też w literaturze wciąż proponowane są nowe rozwiązania, które mają przynieść poprawę dokładności grupowania w stosunku do tradycyjnych metod. Przykładem mogą tu być metody polegające na zastosowaniu podejścia zagregowanego oraz algorytmy spektralne. Głównym celem tego referatu było porównanie dokładności zagregowanych i spektralnych algorytmów taksonomicznych. W pracy rozważano tylko specyficzną klasę metod agregacji, która oparta jest na macierzy współwystąpień.

Jacek Stelmach (Uniwersytet Ekonomiczny w Katowicach), *O testowaniu różnic pomiędzy populacjami za pomocą wektorów własnych*. Wykrywanie różnic pomiędzy kilkoma populacjami jest jednym z częściej rozważanych zagadnień w badaniach statystycznych. Najbardziej znaną i dobrze opisaną w literaturze jest wielowymiarowa analiza wariancji (MANOVA). Metoda ta, jako parametryczna wymaga spełnienia założeń o zgodności badanych populacji z rozkładem wielowymiarowym normalnym oraz homogeniczności macierzy wariancji/kowariancji wszystkich populacji. Oprócz trudności, jakie sprawia weryfikacja powyższych założeń, szczególnie dla prób o niewielkich licznosciach, bardzo często założenie o zgodności z rozkładem normalnym jest w rzeczywistości nierealne (na przykład w zjawiskach biologicznych lub fizycznych). W pracy zaprezentowano odmienne podejście, wykorzystujące testy permutacyjne – co zwalnia z weryfikacji wspomnianych założeń, gdyż możliwe jest stosowanie dowolnych, niestabilizowanych a dopasowanych do zjawiska statystyk testowych. W proponowanych statystykach wykorzystano fakt, że jeśli badane populacje nie różnią się od siebie, to także ich pierwsze wektory własne nie powinny być istotnie różne. Moc proponowanych testów porównano z mocą innych testów permutacyjnych wykorzystujących analizę odległości i analizę zmiennych oraz kilkoma najpopularniejszymi odmianami metody MANOVA. Do analiz wykorzystano symulację komputerową. Otrzymane wyniki zweryfikowano badając kilka przykładowych rzeczywistych zestawów prób wielowymiarowych.

Małgorzata Szerszunowicz (Uniwersytet Ekonomiczny w Katowicach), *O pewnej konstrukcji planów eksperymentów czynnikowych*. Metody planowania eksperymentów są istotnym narzędziem doskonalenia procesów produkcyjnych wykorzystującym meto-

dy probabilistyczne i statystyczne. Planowanie eksperymentów pozwala na odpowiednie ustawienie parametrów procesu produkcyjnego oraz określenie wpływu czynników na jego wyniki. Ponadto metody planowania eksperymentów umożliwiają poprawę ekonomicznych rezultatów badanego procesu. Przedmiotem pracy było zagadnienie wyboru optymalnego układu doświadczeń wówczas, gdy eksperymentator, ze względu na koszty bądź warunki, nie ma możliwości realizacji kompletnego planu doświadczeń. Zaproponowano metodę wyznaczania kolejnych punktów planu, w których należy przeprowadzić doświadczenie. Zastosowania prezentowanej metody przedstawione zostały dla wybranych planów eksperymentów czynnikowych. Własności metody zostały sprawdzone w analizach symulacyjnych.

Janusz Wywiół (Uniwersytet Ekonomiczny w Katowicach), *O estymacji średniej w domenie populacji*. W pracy średnią w domenie estymowano za pomocą statystyki Horvitz-Thompsona. Plan losowania próby był proporcjonalny do rozstępu z próby zmiennej dodatkowej. Badania symulacyjne wskazały, że rozważana strategia estymacji daje dokładniejsze oceny średniej w domenie niż zwykła średniej z próby prostej identyfikowanej w domenie.

Artur Zaborski, Marcin Pełka (Uniwersytet Ekonomiczny we Wrocławiu), *Geometryczna prezentacja preferencji z wykorzystaniem analizy PROFIT i programu R*. PROFIT jest przykładem „zewnętrznej” wektorowej metody map preferencji. Jest ona połączeniem skalowania wielowymiarowego i analizy regresji wielorakiej. Danymi wejściowymi w analizie PROFIT są zarówno współrzędne punktów reprezentujących obiekty na mapie percepcyjnej jak również oceny preferencji obiektów ze względu na wybrane zmienne. Dla konfiguracji punktów reprezentujących obiekty otrzymanej za pomocą skalowania wielowymiarowego przeprowadza się analizę regresji wielorakiej, w której zmiennymi objaśniającymi są współrzędne obiektów na mapie percepcyjnej, a zmiennymi zależnymi oceny marek ze względu na poszczególne cechy. Program dokonuje rozmieszczenia na mapie percepcyjnej zmiennych w postaci wektorów wskazujących kierunek maksymalnej preferencji ze względu na daną zmienną. Celem pracy była prezentacja analizy PROFIT oraz funkcji programu R, pozwalającej na jej realizację. Sposób użycia funkcji zilustrowano przykładem badania preferencji.

Wojciech Zieliński (Szkoła Główna Gospodarstwa Wiejskiego w Warszawie), *Porównanie przedziałów ufności dla wskaźnika struktury w skończonych populacjach*. W pracy porównano trzy przedziały ufności dla frakcji obiektów wyróżnionych w populacji skończonej: jeden wykorzystujący w swojej konstrukcji rozkład hipergeometryczny oraz dwa bazujące na przybliżeniach tego rozkładu – przybliżeniu rozkładem dwumianowym oraz przybliżeniu rozkładem normalnym.

Tomasz Żądło (Uniwersytet Ekonomiczny w Katowicach), *O estymacji parametrów pewnego modelu dla danych wielookresowych*. W pracy rozważano problem modelowania profili wielookresowych przy założeniu, że populacja i przynależność elementów

domen mogą zmieniać się w czasie. Proponowany model jest przypadkiem szczególnym ogólnego modelu liniowego i ogólnego mieszanego modelu liniowego. W modelu tym uwzględniono dwa wektory składników losowych spełniające odpowiednio założenia przestrzennego modelu autoregresyjnego i modelu autoregresyjnego rzędu pierwszego w czasie. W symulacji rozważano dokładność estymatorów parametrów modelu uzyskanych metodą największej wiarygodności z ograniczeniami.

Teksty referatów, po pozytywnej recenzji, zostaną opublikowane w *Acta Universitatis Lodzensis, Folia Oeconomica, „Multivariate Statistical Analysis”*, Wydawnictwo Uniwersytetu Łódzkiego.