

PRACE NAUKOWE

Uniwersytetu Ekonomicznego we Wrocławiu

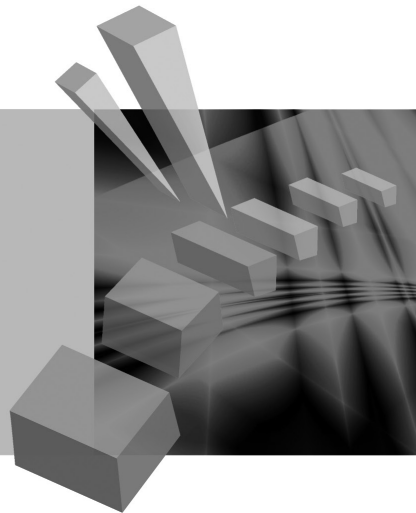
RESEARCH PAPERS

of Wrocław University of Economics

279

Taksonomia 21

Klasyfikacja i analiza danych – teoria i zastosowania



Redaktorzy naukowi

Krzysztof Jajuga

Marek Walesiak



Wydawnictwo Uniwersytetu Ekonomicznego we Wrocławiu
Wrocław 2013

Redaktor Wydawnictwa: Aleksandra Śliwka

Redaktor techniczny: Barbara Łopusiewicz

Korektor: Barbara Cibis

Łamanie: Małgorzata Czupryńska

Projekt okładki: Beata Dębska

Publikacja jest dostępna w Internecie na stronach:

www.ibuk.pl, www.ebscohost.com,

The Central and Eastern European Online Library www.ceeol.com,

a także w adnotowanej bibliografii zagadnień ekonomicznych BazEkon

http://kangur.uek.krakow.pl/bazy_ae/bazekon/nowy/index.php

Informacje o naborze artykułów i zasadach recenzowania znajdują się

na stronie internetowej Wydawnictwa

www.wydawnictwo.ue.wroc.pl

Tytuł dofinansowany ze środków Narodowego Banku Polskiego

oraz ze środków Sekcji Klasyfikacji i Analizy danych PTS

Kopiowanie i powielanie w jakiegokolwiek formie

wymaga pisemnej zgody Wydawcy

© Copyright by Uniwersytet Ekonomiczny we Wrocławiu

Wrocław 2013

ISSN 1899-3192 (Prace Naukowe Uniwersytetu Ekonomicznego we Wrocławiu)

ISSN 1505-9332 (Taksonomia)

Wersja pierwotna: publikacja drukowana

Druk: Drukarnia TOTEM

Spis treści

Wstęp	9
Sabina Denkowska, Kamil Fijorek, Marcin Salamaga, Andrzej Sokolowski: Sejm VI kadencji – maszynka do głosowania	11
Barbara Pawelek, Adam Sagan: Zmienne ukryte w modelach ekonomicznych – respecyfikacja modelu Kleina I	19
Jan Paradysz: Nowe możliwości badania koniunktury na rynku pracy	29
Krzysztof Najman: Samouczące się sieci GNG w grupowaniu dynamicznym zbiorów o wysokim wymiarze	41
Kamila Migdał-Najman: Zastosowanie jednowymiarowej sieci SOM do wyboru cech zmiennych w grupowaniu dynamicznym	48
Aleksandra Matuszewska-Janica, Dorota Witkowska: Zróżnicowanie płac ze względu na płeć: zastosowanie drzew klasyfikacyjnych	58
Iwona Foryś, Ewa Putek-Szeląg: Przestrzenna klasyfikacja gmin ze względu na sprzedaż użytków gruntowych zbywanych przez ANR w województwie zachodniopomorskim	67
Joanna Banaś, Małgorzata Machowska-Szewczyk: Klasyfikacja internetowych rachunków bankowych z uwzględnieniem zmiennych symbolicznych.....	77
Marta Jarocka: Wpływ metody doboru cech diagnostycznych na wynik porządkowania liniowego na przykładzie rankingu polskich uczelni	85
Anna Zamojska: Badanie zgodności rankingów wyznaczonych według różnych wskaźników efektywności zarządzania portfelem na przykładzie funduszy inwestycyjnych.....	95
Dorota Rozmus: Porównanie dokładności taksonomicznej metody propagacji podobieństwa oraz zagregowanych algorytmów taksonomicznych opartych na idei metody <i>bagging</i>	106
Ewa Wędrowska: Wrażliwość miar dywergencji jako mierników niepodobieństwa struktur.....	115
Katarzyna Wójcik, Janusz Tuchowski: Wpływ automatycznego tłumaczenia na wyniki automatycznej identyfikacji charakteru opinii konsumenckich ...	124
Małgorzata Misztal: Ocena wpływu wybranych metod imputacji na wyniki klasyfikacji obiektów w modelach drzew klasyfikacyjnych.....	135
Anna Czapkiewicz, Beata Basiura: Badanie wpływu wyboru współczynnika zależności na grupowanie szeregów czasowych	146
Tomasz Szubert: Czynniki różnicujące poziom zadowolenia z życia oraz wartości życiowe osób sprawnych i niepełnosprawnych w świetle badań „Diagnozy społecznej”	154

Marcin Szymkowiak: Konstrukcja estymatorów kalibracyjnych wartości globalnej dla różnych funkcji odległości	164
Wojciech Roszka: Szacowanie łącznych charakterystyk cech nieobserwowanych łącznie	174
Justyna Brzezińska: Metody wizualizacji danych jakościowych w programie R	182
Agata Sielska: Regionalne zróżnicowanie potencjału konkurencyjnego polskich gospodarstw rolnych w województwach po akcesji do Unii Europejskiej	191
Mariusz Kubus: Liniowy model prawdopodobieństwa z regularyzacją jako metoda doboru zmiennych	201
Beata Basiura: Metoda Warda w zastosowaniu klasyfikacji województw Polski z różnymi miarami odległości	209
Katarzyna Wardzińska: Wykorzystanie metody obwiedni danych w procesie klasyfikacji przedsiębiorstw	217
Katarzyna Dębowska: Modelowanie upadłości przedsiębiorstw oparte na próbach niezbilansowanych	226
Danuta Tarka: Wpływ metody doboru cech diagnostycznych na wyniki klasyfikacji obiektów na przykładzie danych dotyczących ochrony środowiska ..	235
Artur Czech: Zastosowanie wybranych metod doboru zmiennych diagnostycznych w badaniach konsumpcji w ujęciu pośrednim	246
Beata Bal-Domańska: Ocena relacji zachodzących między inteligentnym rozwojem a spójnością ekonomiczną w wymiarze regionalnym z wykorzystaniem modeli panelowych	255
Mariola Chrzanowska: <i>Ordinary kriging</i> i <i>inverse distance weighting</i> jako metody szacowania cen nieruchomości na przykładzie warszawskiego rynku	264
Adam Depta: Zastosowanie analizy wariancji w badaniu jakości życia na podstawie kwestionariusza SF-36v2	272
Maciej Beręsewicz, Tomasz Klimanek: Wykorzystanie estymacji pośredniej uwzględniającej korelację przestrzenną w badaniach cen mieszkań	281
Karolina Paradysz: Benchmarkowa analiza estymacji dla małych obszarów na lokalnych rynkach pracy	291
Anna Gryko-Nikitin: Dobór parametrów w równoległych algorytmach genetycznych dla problemu plecakowego	301
Tomasz Ząbkowski, Piotr Jałowiecki: Zastosowanie reguł asocjacyjnych do analizy danych ankietowych w wybranych obszarach logistyki przedsiębiorstw przetwórstwa rolno-spożywczego	311
Agnieszka Przedborska, Małgorzata Misztal: Zastosowanie metod statystyki wielowymiarowej do oceny wydolności stawów kolanowych u pacjentów z chorobą zwyrodnieniową leczonych operacyjnie	321
Dorota Perło: Rozwój zrównoważony w wymiarze gospodarczym, społecznym i środowiskowym – analiza przestrzenna	331

Ewa Putek-Szeląg, Urszula Gieraltowska, Analiza i diagnoza wielkości produkcji energii odnawialnej w Polsce na tle krajów Unii Europejskiej..	342
--	-----

Summaries

Sabina Denkowska, Kamil Fijorek, Marcin Salamaga, Andrzej Sokolowski: VIth-term Sejm – a voting machine	18
Barbara Pawelek, Adam Sagan: Latent variables in econometric models – respecification of Klein I model	28
Jan Paradysz: New possibilities for studying the situation on the labour market	40
Krzysztof Najman: Self-learning neural network of GNG type in the dynamic clustering of high-dimensional data.....	47
Kamila Migdał-Najman: Applying the one-dimensional SOM network to select variables in dynamic clustering	57
Aleksandra Matuszewska-Janica, Dorota Witkowska: Gender wage gap: application of classification trees.....	66
Iwona Foryś, Ewa Putek-Szeląg: Spatial classification of communes by usable land traded by the APA in the Zachodniopomorskie voivodeship...	76
Joanna Banaś, Małgorzata Machowska-Szewczyk: Classification of Internet banking accounts including symbolic variables	84
Marta Jarocka: The impact of the method of the selection of diagnostic variables on the result of linear ordering on the example of ranking of universities in Poland.....	94
Anna Zamojska: Empirical analysis of the consistency of mutual fund ranking for different portfolio performance measures.....	105
Dorota Rozmus: Comparison of accuracy of affinity propagation clustering and cluster ensembles based on bagging idea.....	114
Ewa Wędrowska: Sensitivity of divergence measures as structure dissimilarity measurements	123
Katarzyna Wójcik, Janusz Tuchowski: Machine translation impact on the results of the sentiment analysis	134
Małgorzata Misztal: Assessment of the influence of selected imputation methods on the results of object classification using classification trees ...	145
Anna Czapkiewicz, Beata Basiura: Simulation study of the selection of coefficient depending on the clustering time series.....	153
Tomasz Szubert: Factors differentiating the level of satisfaction with life and the life's values of people with and without disabilities in the light of the "Social Diagnosis" survey	162
Marcin Szymkowiak: Construction of calibration estimators of totals for different distance measures	173

Wojciech Roszka: Joint characteristics' estimation of variables not jointly observed.....	181
Justyna Brzezińska: Visualizing categorical data in \mathbf{R}	190
Agata Sielska: Regional diversity of competitiveness potential of Polish farms after the accession to the European Union	200
Mariusz Kubus: Regularized linear probability model as a filter	208
Beata Basiura: The Ward method in the application for classification of Polish voivodeships with different distances.....	216
Katarzyna Wardzińska: Application of Data Envelopment Analysis in company classification process.....	225
Katarzyna Dębowska: Modeling corporate bankruptcy based on unbalanced samples	234
Danuta Tarka: Influence of the features selection method on the results of objects classification using environmental data.....	245
Artur Czech: Application of chosen methods for the selection of diagnostic variables in indirect consumption research.....	254
Beata Bal-Domańska: Assessment of relations occurring between smart growth and economic cohesion in regional dimension using panel models	263
Mariola Chrzanowska: Ordinary kriging and inverse distance weighting as methods of estimating prices based on Warsaw real estate market	271
Adam Depta: Application of analysis of variance in the study of the quality of life based on questionnaire SF-36v2	280
Maciej Beręsewicz, Tomasz Klimanek: Using indirect estimation with spatial autocorrelation in dwelling price surveys.....	290
Karolina Paradysz: Benchmark analysis of small area estimation on local labor markets	300
Anna Gryko-Nikitin: Selection of various parameters of parallel evolutionary algorithm for knapsack problems	310
Tomasz Ząbkowski, Piotr Jałowiecki: Application of association rules for the survey of data analysis in the selected areas of logistics in food processing companies	320
Agnieszka Przedborska, Małgorzata Misztal: Using multivariate statistical methods to assess the capacity of the knee joint among the patients treated surgically for osteoarthritis	330
Dorota Perło: Sustainable development in the economic, social and environmental dimensions – spatial analysis.....	341
Ewa Putek-Szeląg, Urszula Gieraltowska: Analysis and diagnosis of the volume of renewable energy production in Poland compared to EU countries	352

Adam Depta

Politechnika Łódzka

ZASTOSOWANIE ANALIZY WARIANCJI W BADANIU JAKOŚCI ŻYCIA NA PODSTAWIE KWESTIONARIUSZA SF-36V2

Streszczenie: W pracy podjęto próbę aplikacji analizy wariancji do oceny jakości życia osób poddanych badaniom za pomocą kwestionariusza jakości życia SF-36v2™ Health Survey (Quality Metric Incorporated) w 8 podskalach, takich jak: sprawność fizyczna, ograniczenia aktywności z powodu zdrowia fizycznego, dolegliwości bólowe, ogólna percepcja zdrowia, vitalność, funkcjonowanie społeczne, zdrowie psychiczne oraz ograniczenie aktywności z powodu problemów emocjonalnych. Dane swoim zasięgiem obejmują lata 2011 i 2012.

Słowa kluczowe: jakość życia, analiza wariancji, kwestionariusz jakości życia SF-36v2™.

1. Wstęp

Jakość życia jest to stopień satysfakcji człowieka (społeczeństwa) z całej swojej egzystencji, zatem jest sumą indywidualnego lub zbiorowego odczucia istniejących warunków oraz ich oceną i często rozpatrywana jest w sensie medycznym.

Jakość życia według WHOQOL Group (*World Health Organization Quality of Life Group*) jest rozumiana jako „sposób postrzegania przez jednostki swoich pozycji w życiu w kontekście kultury i systemu wartości, w których egzystują w powiązaniu z własnymi celami, oczekiwaniami, standardami i obawami; jest to szeroko pojęta koncepcja, na którą w sposób kompleksowy wpływają zdrowie fizyczne jednostki, jej stan psychiczny, relacje społeczne, stopień niezależności oraz jej stosunek do znaczących cech otaczającego środowiska”¹.

W pracy podjęto próbę aplikacji analizy wariancji do oceny jakości życia osób poddanych badaniom za pomocą kwestionariusza jakości życia SF-36v2™ Health Survey², w 8 podskalach: sprawność fizyczna (*physical functioning* – F), ograniczenia aktywności z powodu zdrowia fizycznego (*role limitations due to physical problems* – R), dolegliwości bólowe (*bodily pain* – P), ogólna percepcja zdrowia

¹ WHOQOL Group. Study protocol for the World Health Organization project to develop a quality of life assessment instrument (WHOQOL). *Quality Life Research* 1993;2:153-159.

² Health Assessment Lab, Medical Outcomes Trust and Quality Metric Incorporated.

(*general health perception* – H), witalność (*vitality* – V), funkcjonowanie społeczne (*social functioning* – S), zdrowie psychiczne (*mental health* – W), oraz ograniczenie aktywności z powodu problemów emocjonalnych (*role limitation due to emotional problems* – E). Ponadto celem pracy jest również poszukiwanie metod statystycznych umożliwiających opracowanie wyników badań uzyskanych na podstawie wymienionego kwestionariusza.

W przeprowadzonych badaniach wykorzystano metody wnioskowania statystycznego, takie jak: test normalności Shapiro-Wilka, testy jednorodności wariancji – Hartleya, C Cochrańa oraz Bartletta oraz porównań wielokrotnych – test Newmana-Keulsa i test rozszerzonej istotnej różnicy Tukeya.

2. Materiał i wyniki badań

Badanie jakości życia przeprowadzono wśród pacjentów jękających się w Centrum Terapii Jękania (CTJ)³ w latach 2011-2012. Narzędzie badawcze, jakie zastosowano, to kwestionariusz jakości życia SF-36v2™ Health Survey, licencja Nr QM009973/2011. Kwestionariusz ten posiada właściwości psychometryczne i jest narzędziem wystandaryzowanym.

Do badania wylosowano 15 osób spośród 59 pacjentów centrum. W doborze do próby zastosowano metodę bez zwracania⁴. Do sprawdzenia założenia dotyczącego losowego charakteru próby posłużono się testem serii losowości próby, w którym sprawdzano hipotezę zerową, że próba ma charakter losowy, wobec hipotezy alternatywnej, która mówi, iż próba nie ma charakteru losowego. W analizie sprawdzono kolejno: losowość (test serii), normalność (test Shapiro-Wilka), jednorodność wariancji (testy: Hartleya, C Cochrańa oraz Bartletta).

Dla wylosowanej próby na poziomie istotności $\alpha = 0,05$ nie było podstaw do odrzucenia hipotezy zerowej, że próba ma charakter losowy ($p > 0,05$), zatem można przyjąć założenie o losowości próby.

W kolejnym etapie podjęto próbę ustalenia normalności rozkładu wyników dla analizowanych podskal (wyniki na skali 0-100), co umożliwi następnie stosowanie poszczególnych testów jednorodności wariancji.

W celu zweryfikowania hipotezy zerowej o normalności rozkładu wyników analizowanych podskal zastosowano test Shapiro-Wilka. W teście tym sprawdzono hipotezę zerową, że badana cecha populacji ma rozkład normalny wobec hipotezy alternatywnej, że cecha populacji nie ma rozkładu normalnego.

Z badania wyłączono podskale: sprawność fizyczna (F), ograniczenia aktywności z powodu zdrowia fizycznego (R) oraz ograniczenie aktywności z powodu pro-

³ Centrum Terapii Jękania jest prowadzone w Szamocinie przy ul. Parkowej 3 przez logopedę dypl. Bernadetę Dziekan-Standowicz – uczennicę prof. L. Arutiunian Z. (Лилия Зиновьевна Арутюнян) z Moskwy (twórczyni metody „trwałej normalizacji mowy jękających się”). Badanie zostało przeprowadzone za zgodą logopedy dypl. Bernadety Dziekan-Standowicz.

⁴ Wykorzystano generator liczb pseudolosowych w programie Excel.

blemów emocjonalnych (E), dla wyników których rozkład nie był normalny i z tego powodu nie mogły spełniać wymaganego założenia normalności rozkładu w analizie wariancji.

Wyniki podskal spełniających założenie ww. testu zestawiono w tab.1.

Tabela 1. Wartości statystyk empirycznych testu Shapiro-Wilka oraz poziomy p

	Podskala P	Podskala H	Podskala V	Podskala S	Podskala W
Statystyka Shapiro-Wilka	0,915	0,961	0,943	0,894	0,897
Poziom p	0,163	0,715	0,425	0,078	0,085

Źródło: opracowanie własne.

Na poziomie istotności $\alpha = 0,05$ brak jest podstaw do odrzucenia weryfikowanej hipotezy zerowej, można więc uznać, że rozkład wyników uzyskanych przez pacjentów w badanych podskalach jest normalny. Zatem godząc się na popełnienie błędów 5 razy na 100 przypadków, można powiedzieć, że wymagane założenie normalności w analizie wariancji jest spełnione.

Do weryfikacji jednorodności wariancji zastosowano testy: Hartleya, C Cochra-na oraz Bartletta. Wyniki zestawiono w tab. 2.

Tabela 2. Wyniki statystyk testowych jednorodności wariancji

	Hartleya F-maks	Cochrana C	Bartlett Chi-kw.	Liczba stopni swobody	Poziom p
Wynik	3,553724	0,330308	5,923432	4	0,204940

Źródło: opracowanie własne.

W tabeli 2 podano wartości poszczególnych testów oraz minimalny poziom prawdopodobieństwa p (dla ww. testów). Na poziomie istotności $\alpha = 0,05$ nie ma podstaw do odrzucenia sprawdzanej hipotezy zerowej o równości wariancji w podklasach ($p = 0,205$), zatem można powiedzieć, że wariancje są równe, czyli wymagane założenie jednorodności wariancji jest spełnione.

Po sprawdzeniu wymaganych założeń przystąpiono do analizy wariancji.

Wyniki analizy wariancji (tab. 3) potwierdzają przewidywania dotyczące różnicy w wartościach średnich wyników uzyskanych przez badanych – na poziomie istotności $\alpha = 0,05$ można odrzucić hipotezę zerową o równości średnich.

Wybrane statystyki opisowe dla każdej podskali, czyli średnie, odchylenia standardowe, błędy standardowe oraz 95-procentowe przedziały ufności, dla średniej liczby punktów każdej podskali są przedstawione w tab. 4. Wśród wszystkich podskal „wyróżnia” się podskala funkcjonowanie społeczne (S), która ma największą średnią. Najmniejsza zaś jest w podskali witalność (V).

Tabela 3. Wyniki testu analizy wariancji

Efekt	Stopnie swobody	Suma kwadratów	Średnie kwadraty	Statystyka empiryczna Fishera-Snedecora	Poziom <i>p</i>
Wyraz wolny	1	367 360,0	367 360,0	1614,529	0,000000
Łączny wynik w podskalach	4	3 189,9	797,5	3,505	0,011488
Błąd standardowy	70	15 927,4	227,5		
Ogółem	74	19 117,2			

Źródło: opracowanie własne.

Tabela 4. Wartości średnich dla poszczególnych podskal

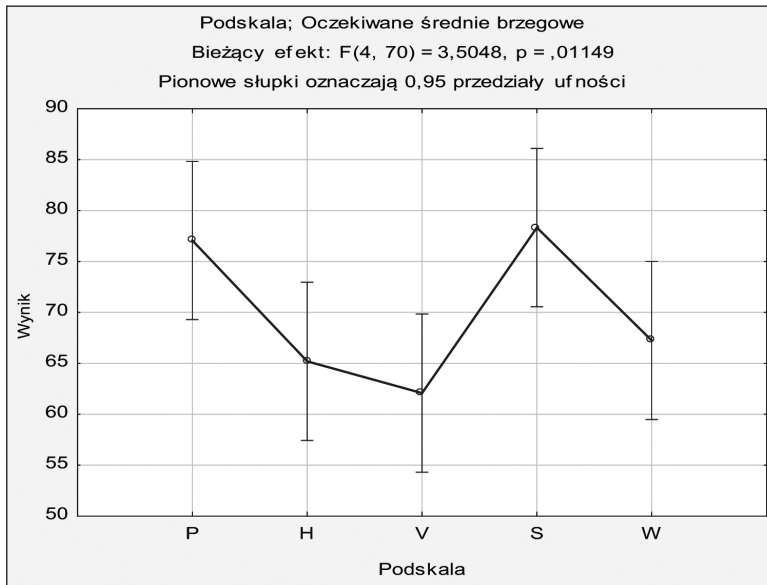
Efekt	Poziom czynnika	N	Wynik średnie	Wynik odch.st.	Wynik bł. std.	Wynik –95,00%	Wynik +95,00%
Ogół		75	69,98667	16,07299	1,855949	66,28861	73,68473
Podskala	P	15	77,06667	19,38507	5,005204	66,33157	87,80176
Podskala	H	15	65,20000	10,28313	2,655094	59,50539	70,89461
Podskala	V	15	62,08333	13,66587	3,528512	54,51543	69,65124
Podskala	S	15	78,33333	16,68154	4,307155	69,09540	87,57126
Podskala	W	15	67,25000	13,82447	3,569464	59,59426	74,90574

Źródło: opracowanie własne.

Powyżej prezentowane wyniki można przedstawić graficznie. Na rysunkach 1 i 2 na wykresie liniowym przedstawiono punkty, które odpowiadają wartościom średnich w poszczególnych podskalach, a słupki błędów obrazują granice przedziałów ufności dla wartości średnich. Opierając się na tab. 4, można zaobserwować, że badane osoby oceniają wysoko analizowane podskale. Pacjenci najwyższej oceniają podskalę funkcjonowanie społeczne, najniżej zaś witalność – najwyższej położony jest punkt odpowiadający podskali funkcjonowanie społeczne (S), a najniżej witalności (V). Oznacza to, że zdrowie fizyczne lub problemy emocjonalne nie mają dużego wpływu na aktywność towarzyską w kręgu m.in. rodziny, przyjaciół i sąsiadów.

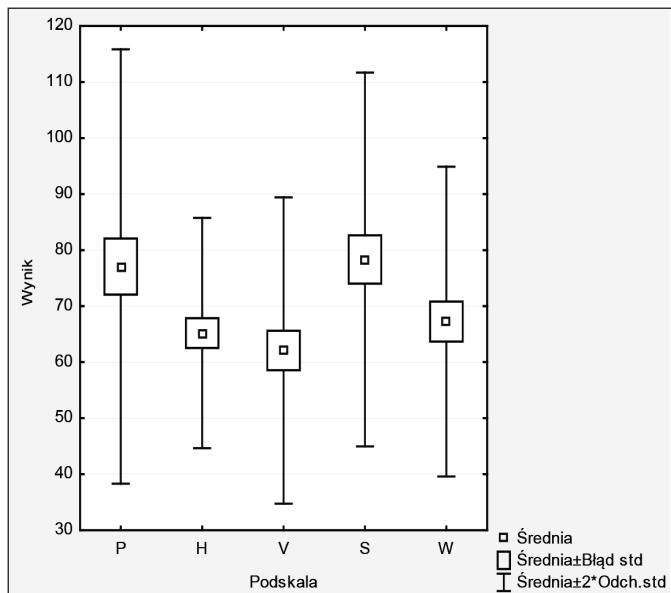
Ze względu na to, że analiza wariancji wykazała istotne różnice pomiędzy wynikami rozpatrywanych podskal, można przeprowadzić testy *post-hoc*: test NIR (najmniejszych istotnych różnic – *least significant differences* (LSD)), test Scheffego, test Newman-Keulsa, wielokrotny test rozstępu Duncana, test rozsądnej istotnej różnicy (RIR) Tukeya, zwane testami porównań wielokrotnych.

Ze względu na to, iż nie wszystkie testy porównań wielokrotnych dają gwarancję łącznego poziomu istotności dla wszystkich testowanych wyników podskal, w dalszym kroku wybrano tylko test Newman-Keulsa oraz test rozsądnej istotnej różnicy Tukeya.



Rys. 1. Interpretacja graficzna wyników analizy wariancji

Źródło: opracowanie własne.



Rys. 2. Skategoryzowany wykres ramkowy dla podskal

Źródło: opracowanie własne.

Wyniki tych testów pozwolą odpowiedzieć na pytanie, które z porównywanych podskal przyczyniły się do odrzucenia hipotezy zerowej w analizie wariancji. Zatem dadzą odpowiedź, które ze średnich różnią się między sobą, a które są równe. Ponadto umożliwią również tworzenie jednorodnych grup ze względu na średnie. Te z nich, które należą do tej samej grupy nie różnią się istotnie między sobą, średnie zaś należące do różnych grup istotnie się różnią.

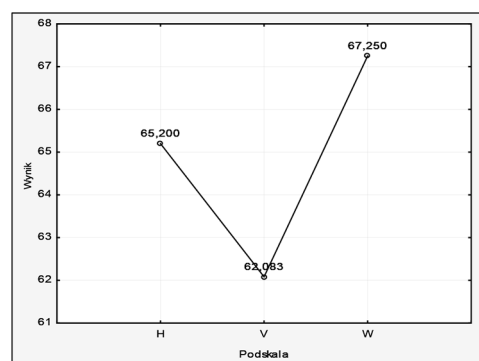
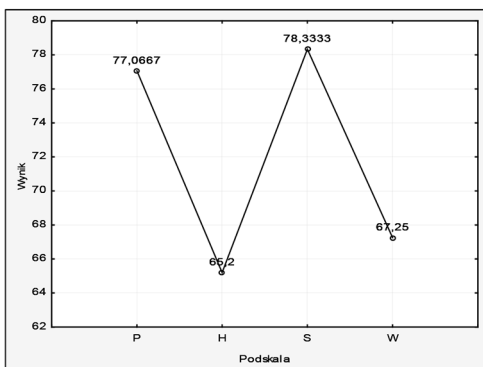
Ze względu na to, iż nie wszystkie testy porównań wielokrotnych dają gwarancję łącznego poziomu istotności dla wszystkich testowanych, w dalszym kroku posłużono się testem Newmana-Keulsa oraz testem rozsądnej istotnej różnicy Tukeya. Wymienione testy pozwalają również na tworzenie jednorodnych grup ze względu na średnie. Wyniki testów zestawiono w tab. 5 i 6.

Tabela 5. Wyniki testu *post hoc* Newmana-Keulsa

	Podskala	(1) 77,067	(2) 65,200	(3) 62,083	(4) 78,333	(5) 67,250
Podskala	P		0,086497	0,040188	0,818907	0,079144
Podskala	H	0,086497		0,573422	0,089596	0,711002
Podskala	V	0,040188	0,573422		0,034010	0,618305
Podskala	S	0,818907	0,089596	0,034010		0,116870
Podskala	W	0,079144	0,711002	0,618305	0,116870	

Źródło: opracowanie własne.

Wyniki testu Newmana-Keulsa pozwalają twierdzić, że odrzucenie hipotezy zerowej o równości średnich zostało spowodowane znaczną różnicą pomiędzy średnimi podskal P i V, czyli dolegliwości bólowych i zdrowia psychicznego, oraz V i S, czyli witalności i funkcjonowania społecznego.



Rys. 3. Średnie w grupie pierwszej i drugiej

Źródło: opracowanie własne.

W wyniku powyższej procedury utworzone zostały dwie grupy jednorodne ze względu na wartości średnich wyników.

W pierwszej grupie znalazły się następujące podskale: dolegliwości bólowe (P), ogólna percepcja zdrowia (H), funkcjonowanie społeczne (S) oraz zdrowie psychiczne (W), a w drugiej: HVW ogólna percepcja zdrowia (H), witalność (V) oraz zdrowie psychiczne (W).

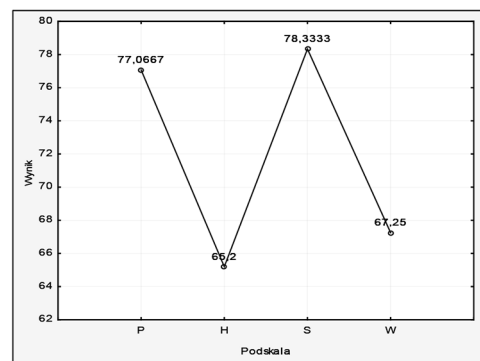
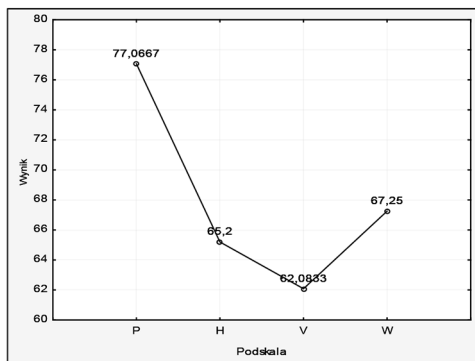
Na podstawie testu Tukeya można stwierdzić, że o odrzuceniu hipotezy zerowej o równości średnich zadecydowała znaczna różnica pomiędzy średnimi wynikami podskal S i V, czyli witalności oraz funkcjonowania społecznego.

Tabela 6. Wyniki testu *post hoc* Tukeya

	Podskala	(1) 77,067	(2) 65,200	(3) 62,083	(4) 78,333	(5) 67,250
Podskala	P		0,209421	0,061072	0,999421	0,392068
Podskala	H	0,209421		0,979609	0,131812	0,995892
Podskala	V	0,061072	0,979609		0,034010	0,881122
Podskala	S	0,999421	0,131812	0,034010		0,271077
Podskala	W	0,392068	0,995892	0,881122	0,271077	

Źródło: opracowanie własne.

Ww. procedura Tukeya utworzyła dwie grupy jednorodne ze względu na wartości średnich: w pierwszej grupie znalazły się następujące podskale: dolegliwości bólowe (P), ogólna percepcja zdrowia (H), witalność (V) oraz zdrowie psychiczne (W), a w drugiej: dolegliwości bólowe (P), ogólna percepcja zdrowia (H), funkcjonowanie społeczne (S) oraz zdrowie psychiczne (W).



Rys. 4. Średnie w grupie pierwszej i drugiej

Źródło: opracowanie własne.

3. Podsumowanie

Wyniki przeprowadzonej analizy wskazują na możliwość zastosowania analizy wariancji w ocenie jakości życia wśród jakążących się pacjentów badanych za pomocą kwestionariusza SF36, w takich podskalach jak sprawność fizyczna, ograniczenia aktywności z powodu zdrowia fizycznego, dolegliwości bólowe, ogólna percepcja zdrowia, witalność, funkcjonowanie społeczne, zdrowie psychiczne oraz ograniczenie aktywności z powodu problemów emocjonalnych. Na podstawie przeprowadzonego badania okazało się, że takie wyniki badanych podskal, jak: sprawność fizyczna, ograniczenia aktywności z powodu zdrowia fizycznego oraz ograniczenie aktywności z powodu problemów emocjonalnych, nie spełniały wymaganych założeń analizy wariancji, co oznacza nie można było zastosować równości wyników podskal dla średnich. Wyniki testu analizy wariancji potwierdzają istotne statystycznie różnice średnich w pozostałych analizowanych podskalach. Natomiast przeprowadzone testy porównań wielokrotnych pozwoliły wyodrębnić jednorodne grupy ze względu na wartości średnich w analizowanych podskalach.

Literatura

- Domański Cz. (1990), *Testy statystyczne*, PWE, Warszawa.
- Domański Cz., Parys D. (2007), *Statystyczne metody wnioskowania wielokrotnego*, Wydawnictwo Uniwersytetu Łódzkiego, Łódź.
- Greń J. (1987), *Statystyka matematyczna: podręcznik programowany*, PWN, Warszawa.
- Karris S.T. (2007), *Mathematics for Business, Science, and Technology*, Third Edition, Orchard Publications.
- Larose D.T. (2008), *Metody i modele eksploracji danych*, PWN, Warszawa.
- Morrison D.F. (1990), *Wielowymiarowa analiza statystyczna*, PWN, Warszawa.
- Ross S.M. (2004), *Introduction to Probability and Statistics for Engineers and Scientists*, Third Edition, Elsevier Academic Press.
- Walesiak M., Gatnar E. (red.) (2009), *Statystyczna analiza danych z wykorzystaniem programu R*, PWN, Warszawa, s. 104-107.
- WHOQOL Group, *Study protocol for the World Health Organization project to develop a quality of life assessment instrument (WHOQOL)*, Quality Life Research 1993; 2:153-159.

APPLICATION OF ANALYSIS OF VARIANCE IN THE STUDY OF THE QUALITY OF LIFE BASED ON QUESTIONNAIRE SF-36V2

Summary: Quality of life is a sum of individual or collective feeling of the existing conditions and their assessment. This paper attempts to apply a variance analysis to assess the quality of life of people interviewed with the quality of life questionnaire SF-36v2™ Health Survey, in 8 subscales: physical fitness, activity limitations due to physical health, pain, general health perceptions, vitality, social functioning, mental health, and activity limitation due to emotional problems. The data cover the years 2011 and 2012.

Keywords: quality of life, analysis of variance, the quality of life questionnaire SF-36v2™.