


Analizy IBE/01/2013

Ewa Matczak
Waldemar Kozłowski

Zastosowanie metody drzew klasyfikacyjnych w analizie aspiracji edukacyjnych rodziców



Kontakt do autorów:

Waldemar Kozłowski: w.kozlowski@ibe.edu.pl

Ewa Matczak: e.matczak@ibe.edu.pl

Wydawca:

Instytut Badań Edukacyjnych

ul. Górczewska 8

01-180 Warszawa

tel. (22) 241 71 00; www.ibe.edu.pl

© Copyright by: *Instytut Badań Edukacyjnych, Warszawa 2011*

Publikacja współfinansowana przez Unię Europejską ze środków Europejskiego Funduszu Społecznego w ramach projektu: Badanie jakości i efektywności edukacji oraz instytucjonalizacja zaplecza badawczego

Egzemplarz bezpłatny

Streszczenie

Drzewa klasyfikacyjne są metodą podziału populacji na klasy obserwacji o podobnych właściwościach. Tekst przedstawia próbę zastosowania tej metody w badaniach dotyczących aspiracji edukacyjnych rodziców uczniów klas IV szkoły podstawowej. Procedura oparta została na danych zebranych od 4931 rodziców z terenu całej Polski. Analiza miała dwa etapy: w pierwszym uwzględniono wybrane zmienne związane ze społeczną pozycją rodziców a w drugim dołączono do nich dwie inne zmienne. Wykorzystując pakiet „party” środowiska R otrzymano diagramy zwane drzewami klasyfikacji obserwacji statystycznych. Niezależnie od uwzględnionej liczby zmiennych wykształcenie najsilniej różnicuje badaną populację. W zależności od branych pod uwagę zmiennych oraz odmian aspiracji (maksymalnych, realnych, akceptowalnych) otrzymano różne konfiguracje czynników warunkujących aspiracje.

Streszczenie angielskie

ABSTRACT

Classification trees are a method of division of the population into observation classes with similar characteristics. The text presents an attempt to apply this method in studies of educational aspirations of parents of fourth grade pupils of primary schools. The procedure is based on data collected from 4931 parents from across the Poland. The analysis has two stages: the first includes selected variables related to the social position of the parents and the second is accompanied by two other variables. Using the package “party” in the R environment obtained diagrams called classification trees of statistical observation. The study population is most diverse by parents' education regardless of the number of variables being considered. Depending on the variables under consideration and aspirations dimension (maximum, realistic, acceptable) yielded different configurations of the determinants of aspirations.

Spis Treści

Streszczenie	3
Streszczenie angielskie	4
ABSTRACT	4
Spis Treści	5
1. Wprowadzenie	6
2. Procedura analizy danych	7
3. Charakterystyka osób badanych	8
4. Wyniki	9
4.1. Ranking ważności predyktorów	10
5. Wyniki analizy	13
5.1. Wyniki pierwszego etapu	13
5.2. Wyniki drugiego etapu	18
6. Podsumowanie	24
7. Literatura cytowana	25

1. Wprowadzenie

Drzewa klasyfikacyjne należą do metod statystycznych, w których dokonuje się podziału próby statystycznej na klasy obserwacji o podobnych właściwościach i stanowią odmianę hierarchicznej analizy skupień. W efekcie otrzymuje się diagramy zwane drzewami klasyfikacji obserwacji statystycznych. Metoda drzew klasyfikacyjnych znalazła zastosowanie m.in.: w badaniach marketingowych związanych z pozyskaniem nowych nabywców, w badaniu przyczyn nielojalności konsumentów (churn analysis), w wykrywaniu oszustw (fraud detection) oraz w analizie danych ankietowych.

Drzewa klasyfikacyjne ze względu na przejrzystość graficzną analizy upraszczają prezentację otrzymanych wyników i są łatwiejsze w interpretacji niż czysto liczbowe wyniki. W związku z podziałem badanej populacji na jednorodne klasy można sformułować reguły przynależności obiektów do danej klasy. Nie są jednak metodą, która zastępuje tradycyjny sposób analizy danych.

W naszym opracowaniu metodę tę zastosowano do analizy danych ankietowych dotyczących aspiracji edukacyjnych rodziców uczniów klas IV szkoły podstawowej wykorzystując pakiet „party” środowiska R. Do konstrukcji drzew użyto funkcji `cree`, pakiet „party” środowiska R (Hothorn, Hornik, and Zeileis, 2006). Wszystkie zmienne są traktowane jako jakościowe, przy czym wykształcenie rodzica/opiekuna wypełniającego ankietę pogrupowano w 3 kategorie: najniższe podstawowe (grupujące wykształcenie niepełne podstawowe, podstawowe oraz zasadnicze zawodowe); średnie grupujące średnie szkoły ogólnokształcące, zawodowe, policealne i pomaturalne oraz wyższe grupujące wyższe licencjackie, inżynierskie, magisterskie i wyższe ze stopniem co najmniej doktora. Dla czytelniejszej ilustracji wyników dane dotyczące aspiracji traktowane są jako skala ilościowa.

Drzewo klasyfikacyjne jest graficznym modelem powstałym w wyniku podziału zbioru obserwacji A na n rozłącznych podzbiorów $A_1, A_2, A_3, \dots, A_n$. Drzewo składa się z korzenia oraz gałęzi prowadzących z korzenia do kolejnych węzłów. Drzewo „rośnie” od góry do dołu. (od korzenia do liści). W korzeniu skupiona jest cała wybrana przez nas próba. Elementy tej próby są przesuwane w dół drzewa poprzez gałęzie do węzłów. W każdym węźle podejmowana jest decyzja o wyborze gałęzi (kierunek w którym przesuną się elementy). W ten sposób próba zostaje podzielona na podgrupy. Pod każdym węzłem wyszczególnione jest kryterium podziału dokonywanego w danym węźle, które jest jednakowe dla wszystkich elementów próby. Ostatecznym etapem budowy są liście drzewa, którym przypisane zostają etykiety (określona zostaje przynależność klasowa). Dla każdego liścia istnieje dokładnie jedna droga łącząca go z korzeniem (Koronacki, Ćwik 2005).

W praktyce najczęściej stosuje się drzewa binarne. W drzewie takim każdy węzeł ma po dwóch potomków tzn. podpróba docierająca do węzła dzielona jest na dwie części. W każdym węźle stawiana hipoteza H_0 , że nie ma zależności zmiennej zależnej od żadnej niezależnej. Jeśli H_0 nie może być odrzucona ze względu na $p > 0,05$ to jest to węzeł końcowy. Jeśli H_0 jest odrzucona to wybierana jest zmienna z „najsilniejszym” powiązaniem ze zmienną zależną – i według tej zmiennej

dokonywany jest kolejny podział grupy w węźle na dwie podgrupy. Celem budowy modelu jest uzyskanie podzbiorów maksymalnie jednorodnych z punktu widzenia wartości zmiennej zależnej. Proces ten jest wieloetapowy -w każdym kolejnym kroku może wykorzystywać inną zmienną niezależną. Na każdym etapie analizuje się bowiem wszystkie predyktory czyli zmienne pozwalające na przewidywanie późniejszych wyników (wartości zmiennej zależnej) i wybiera ten, który zapewnia najlepszy podział węzła.

Drzewa klasyfikacyjne wykorzystuje się do wyznaczania przynależności przypadków do klasy jakościowej zmiennej zależnej na podstawie pomiarów jednej lub więcej zmiennych objaśniających (predyktorów).

2. Procedura analizy danych

Przyjęto, że analiza danych za pomocą metody drzew klasyfikacyjnych przeprowadzona będzie dwukrotnie. Zabieg ten ma na celu nadanie analizie większej przejrzystości. Na pierwszym poziomie analizy zmienne niezależne ograniczymy do kilku wskaźników SES. W naszych badaniach będą to: płeć i wykształcenie osoby wypełniającej ankietę oraz miejsce zamieszkania i płeć dziecka; zmienną zależną – aspiracje edukacyjne rodziców. SES bywa mierzony na podstawie jednego wskaźnika (np. zawodu lub wykształcenia), bądź wskaźników złożonych, (np. zawodu i wykształcenia) ale ten ostatni nie daje jednoznacznej odpowiedzi na pytanie, od którego wskaźnika zależy zmienna zależna. Dlatego niektórzy autorzy, jak np. Entwisle i Astone (1994) postulują, aby oddzielenie analizować dochody rodziny, wykształcenie rodziców oraz strukturę rodziny. Są to, ich zdaniem, trzy odmiany kapitału mające dla dzieci istotne znaczenie. W przypadku aspiracji największe znaczenie ma wykształcenie rodzica, w związku z czym ten wskaźnik SES-u przyjęto jako główny. Badania Z.Sawińskiego (2008) pokazują, że wysoki poziom wykształcenia matek sprzyja osiągnięciu wysokiego poziomu wykształcenia przez dzieci.

Poziom aspiracji rodziców w stosunku do swoich dzieci kontrolowano za pomocą skali, na której zamieszczono 9 poziomów wykształcenia od podstawowego do wyższego ze stopniem co najmniej doktora.

W naszych badaniach aspiracje były mierzone w trzech aspektach. Badanych pytano o:

- a. najwyższy poziom wykształcenia, jakiego pragnie rodzic, który został określony jako „wymarzony”;
- b. poziom najniższy, ale jeszcze możliwy do zaakceptowania w sytuacji, gdy warunki kształcenia będą niesprzyjające („akceptowany”);
- c. realny poziom wykształcenia jaki dziecko przypuszczalnie osiągnie („realny”)

Uwzględnienie trzech aspektów aspiracji pozwoliło na pełniejszą charakterystykę oczekiwań rodziców dotyczących własnych dzieci

W analizie uwzględniliśmy tylko ankiety wypełniane przez matki i ankiety wypełniane przez ojców gdyż pozostałe osoby (opiekunowie prawni, babcie i dziadkowie oraz rodzeństwo) stanowią ok.2% osób badanych a ich udział, niewiele zmieniający wyniki, znacznie zakłócałby tworzone grafy.

Drugi etap to włączenie do analizy dwóch zmiennych, których znaczenie w kontekście aspiracji postanowiono sprawdzić: preferencji dotyczących ścieżki edukacyjnej dziecka oraz deklaracji rodziców na temat działań sprzyjających zdobywaniu wykształcenia odpowiadającego aspiracjom.

Jeden z dylematów rodziców dotyczy tego, czy dziecko powinno dążyć do zdobycia jak najwyższego wykształcenia a potem podjąć pracę zawodową czy najpierw zdobyć w miarę popłatny zawód a następnie podnosić lub uzupełniać swoje wykształcenie. W związku z tym rodzice proszeni byli o dokonanie wyboru spośród następujących możliwości:

- zdobycie przez dziecko jak najwyższego wykształcenia a potem pojęcie pracy zawodowej
- zdobycie zawodu a następnie uzupełnienie wykształcenia
- równoczesna praca i kształcenie się.

Jak wynika z danych pilotażowych preferencje dotyczące ścieżki edukacyjnej dziecka są ściśle związane z aspiracjami edukacyjnymi rodziców.(Kozłowski, Matczak 2012) Poziom wykształcenia, zwłaszcza wykształcenie matek, jest mocno związany z myśleniem o dziecku i działaniami podejmowanymi na jego rzecz (np. Richman, Miller i LeVine; 1992).

Myślenie o przyszłości dziecka związane jest z wczesnym podejmowaniem działań związanych z inwestowaniem w jego kształcenie (pomoc rodziców w nauce, zajęcia dodatkowe itp.). Przy interpretacji wyników bierzemy pod uwagę tylko deklaratywne stwierdzenie czy rodzina takie działania podejmuje czy nie. Nie zajmujemy się szczegółową kwestią jakiego typu są to działania choć pytanie takie zadawane było w ankiecie. W przypadku analizy przy pomocy drzew klasyfikacyjnych jest ono zbyt szczegółowe.

3. Charakterystyka osób badanych

Badanie zostało zrealizowane na reprezentatywnej próbie uczniów szkół podstawowych (ok. 5 tys.) w ramach badania podłużnego przeprowadzonego w Zespole Szkolnych Uwarunkowań Efektywności Kształcenia (SUEK) Instytutu Badań Edukacyjnych w projekcie „Entuzjaści Edukacji..Schemat doboru uczniów miał charakter doboru warstwowego z zastosowaniem losowania systematycznego.

W naszym badaniu uczestniczyło 4931 rodziców (wylosowanych uczniów). Ankiety wypełniały w większości matki (4274) oraz ojcowie (545) pozostałe osoby (ok.2%) to babcie, dziadkowie, rodzeństwo, które nie były brane pod uwagę w dalszych analizach. Badanie przeprowadzono na przełomie 2011 i 2012 r.

Na podstawie analizy zgodności parametrów osób wypełniających ankietę z danymi wszystkich rodziców w próbie można stwierdzić, że parametry, na które wpływ ma wykształcenie mogą być zawyżone w grupie ojców. W grupie ojców wypełniających ankietę wystąpiła niewielka nadreprezentacja mieszkańców miast i znaczna osób z wyższym wykształceniem. Grupa matek

wypełniających ankietę jest reprezentatywna dla całej populacji matek uczniów klas IV szkół podstawowych. Badanie grupy ojców może być ciekawe poznawczo mimo braku reprezentatywności tej grupy.

Wypełniający ankietę ojcowie to ojcowie synów a nie córek mieszkający w mieście i mający wyższy poziom wykształcenia niż matki. Jeśli wypełniającymi ankietę były matki to w 95% ich poziom wykształcenia był wyższy niż ojców.

4. Wyniki

Wstępnym wymogiem stosowania metody jest określenie zbioru zmiennych oraz kolejności, w jakiej będą one w niej uwzględniane. Decyzje w tych kwestiach podjęto na podstawie analizy statystycznej zebranych danych. Przewodujemy tylko główne jej efekty (szczegółowe wyniki przedstawione zostaną w osobnym opracowaniu).

- Zdecydowanie najmocniej spośród wszystkich zmiennych uwzględnionych w analizie aspiracje różnicuje wykształcenie rodzica wypełniającego ankietę. Wszystkie testy (analizy wariancji, regresji liniowej, testy χ^2) wskazują na dominujący wpływ tej zmiennej.
- Poziom wykształcenia rodzica jest istotnie związany z deklaracją rodziców na temat tego czy podejmują działania sprzyjające osiągnięciu upragnionego wykształcenia, preferencjami dotyczącymi ścieżki edukacyjnej dziecka oraz miejscem zamieszkania.
- Dane pokazują, że poziom aspiracji nie zależy od tego czy wypełniającym ankietę jest matka czy ojciec dziecka;decyduje przede wszystkim poziom wykształcenia respondenta.
- Analizowane zmienne najsilniej różnicują aspiracje realne. Model regresji liniowej wyjaśnia 30% zmienności aspiracji realnych. Najslabiej wyjaśniane zróżnicowane są aspiracje akceptowalne. Model regresji liniowej wyjaśnia 12% zmienności.
- Kolejną zmienną różnicującą aspiracje jest zmienna związana zpreferencjami dotyczącymi ścieżki edukacyjnej dziecka. Najwyższe poziom aspiracji występuje w grupie osób wybierających wariant zdobycia przez dziecko wykształcenia a następnie podjęcie pracy a najniższy w grupie rodziców wybierających wariant najpierw zdobycie zawodu a następnie wykształcenia.
- Test χ^2 wykazuje związek pomiędzy tym pytaniem a wykształceniem rodzica . Test Cochran–Mantel–Haenszel (CMH) potwierdził istnienie związku pomiędzy aspiracjami a pytaniem o preferencje niezależnego od wykształcenia rodzica wypełniającego ankietę,
- Zmienną wpływającą na aspiracje są deklaracje dotyczące działań. Osoby deklarujące podejmowanie działań mają wyższe aspiracje. Najmocniej wpływ ten zauważalny jest na poziomie aspiracji realnych. W przypadku aspiracji akceptowalnych deklaracje mają mniejsze niż płeć dziecka.

- Aspiracje rodziców różnicuje płeć dziecka. Są one wyższe w stosunku do dziewcząt niż chłopców szczególnie na poziomie aspiracji wymarzonych. Płeć dziecka stosunkowo silnie wpływa na aspiracje akceptowalne, w mniejszym stopniu na wymarzone.
- Miejsce zamieszkania oraz osoba wypełniająca ankietę w istotny, ale niezbyt silny, sposób różnicują aspiracje. Miejsce zamieszkania jest mocno związane z wykształceniem wypełniającego ankietę. Analiza wariancji wykazuje istotny wpływ miejscowości zamieszkania na aspiracje tylko wśród respondentów z wykształceniem średnim (na aspiracje wymarzone i realne) oraz wśród respondentów z wykształceniem wyższym na aspiracje wymarzone i akceptowalne.
- Analiza wariancji wykazuje istotny wpływ osoby wypełniającej ankietę: w grupie respondentów z wykształceniem podstawowym na aspiracje akceptowalne i realne, w grupie z wykształceniem średnim na wszystkie aspiracje i brak wpływu w grupie z wykształceniem wyższym. Ojcowie i matki z wyższym wykształceniem mają zbliżone aspiracje, podobne zjawisko występuje w przypadku aspiracji wymarzonych ojców i matek z wykształceniem podstawowym.

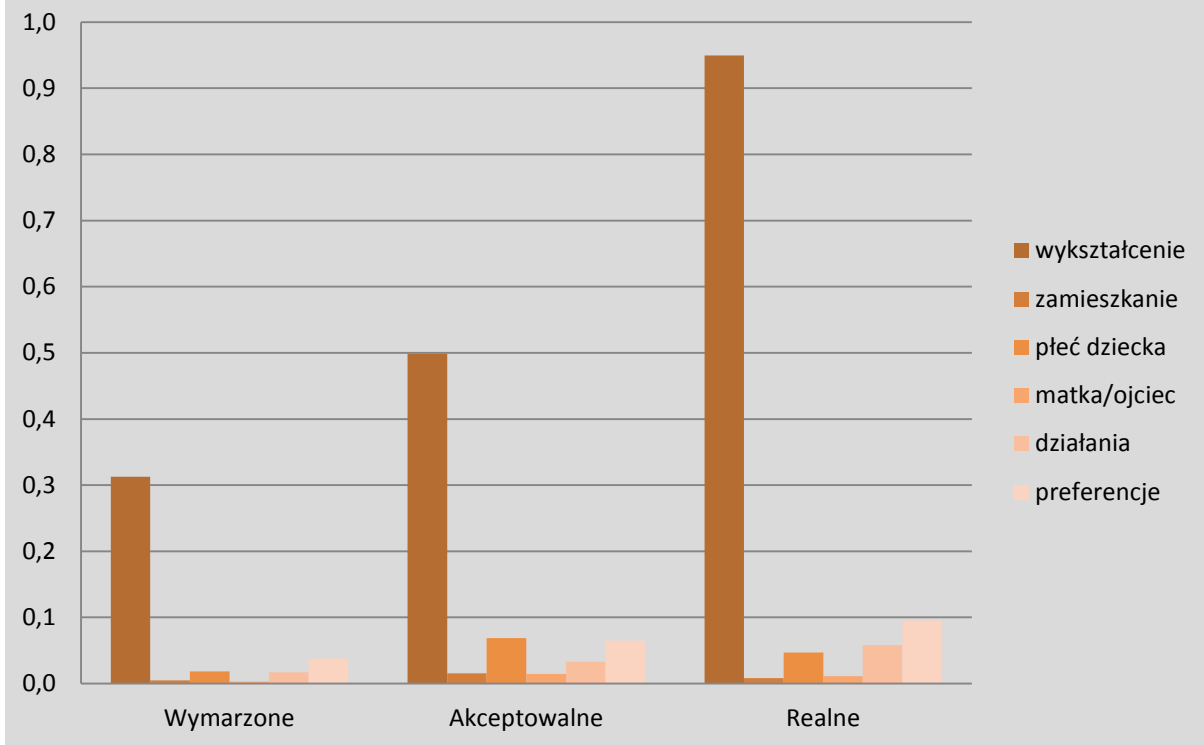
4.1. Ranking ważności predyktorów

Zgodnie z otrzymanymi wynikami przyjęliśmy, że zmiennymi niezależnymi (predyktorami) uwzględnionymi w analizie aspiracji będą: wykształcenie i miejsce zamieszkania, płeć dziecka, osoba wypełniająca ankietę (matka/ojciec), deklaracje rodziców na temat podejmowania działań sprzyjających zdobywaniu wykształcenia odpowiadającego aspiracjom oraz preferencje dotyczące wyboru ścieżki edukacyjnej dziecka.

Do oceny zdolności dyskryminacyjnych poziomu aspiracji dla poszczególnych zmiennych użyto funkcji cforest (tworzenie lasów losowych) oraz varimp pakietu „party” (Hothorn i inni 2006; Strobl i inni 2007; Strobl i inni 2008)

Ranking ważności predyktorów przedstawia uwzględnione zmienne niezależne - wartości najwyższe oznaczają największy wpływ danej zmiennej na zmienną zależną (aspiracje) (wykres 1).

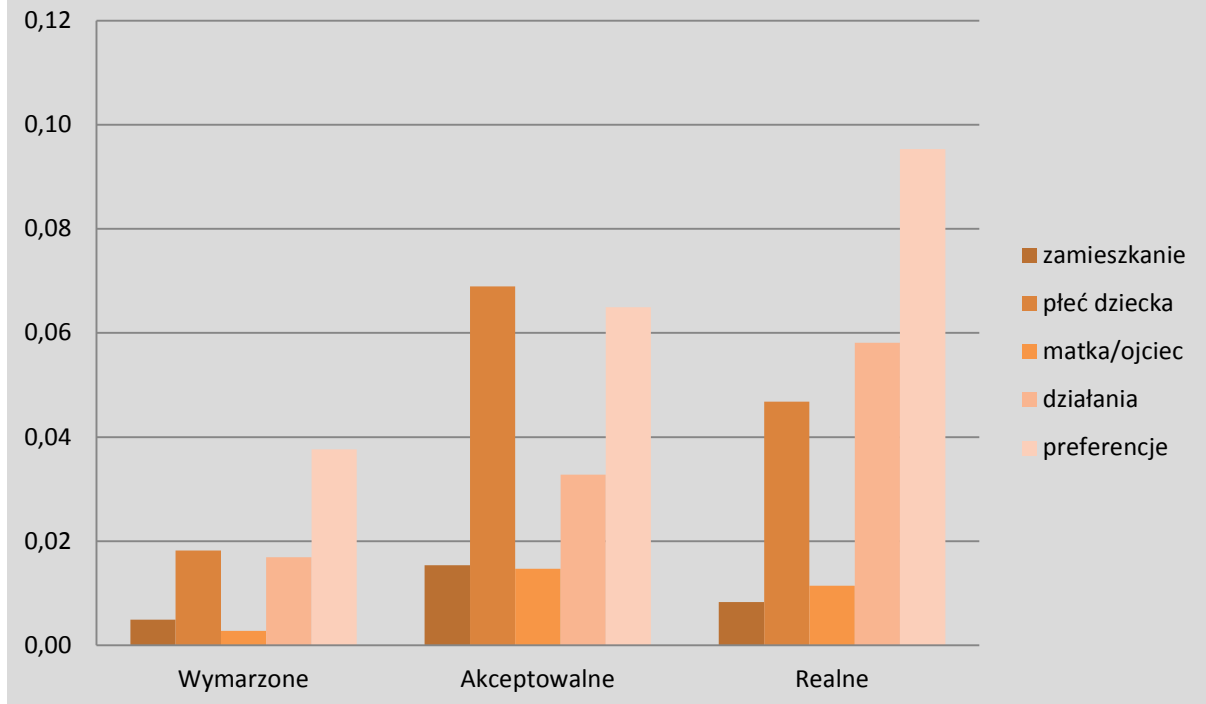
Wykres 1. Ranking ważności predyktorów



Jak pokazują wyniki zdecydowanie najważniejszą zmienną mającą wpływ na aspiracje (na wszystkich poziomach) jest wykształcenie. Zmienna ta najsilniej wpływa na aspiracje realne a najslabiej na wymarzone.

Na wykresie 2 przedstawiono te same zmienne z pominięciem wykształcenia aby pokazać słabo widoczny na wykresie 1 (w związku z dominacją wykształcenia) ranking ważności kolejnych zmiennych.

Wykres 2. Ranking ważności predyktorów z pominięciem wykształcenia



Następną w kolejności zmienną mającą wpływ na aspiracje jest ścieżka edukacyjna dziecka oraz deklaracja dotycząca podejmowanych przez rodziców działań wspomagających. W przypadku aspiracji akceptowalnych najważniejszą zmienną jest płeć dziecka. Małe znaczenie ma miejsce zamieszkania oraz płeć osoby wypełniającej ankietę.

Wyniki dotychczasowych analiz (ANOVA, testy χ^2 , regresja liniowa) wskazują, że wymienione wyżej dwie zmienne są silniejszymi predyktorami niż wskaźniki SES poza wykształceniem rodziców, które zdecydowanie najmocniej różnicuje aspiracje rodziców.

Dzięki rysunkom drzew możemy ustalić jaką pozycję zajmują poszczególne zmienne w poszczególnych odmianach aspiracji. Wiadomo, że najsilniejszą zmienną różnicującą jest wykształcenie rodziców a dzięki zastosowaniu metody drzew klasyfikacyjnych można stwierdzić jak wygląda ranking wszystkich zmiennych.

5. Wyniki analizy

5.1. Wyniki pierwszego etapu

Na pierwszym poziomie analizy zmiennymi niezależnym będą: osoba wypełniająca ankietę (matka/ojciec), wykształcenie i miejsce zamieszkania osoby wypełniającej ankietę oraz płeć dziecka; zmienną zależną - aspiracje mierzone w trzech wymiarach. Kolejno przedstawione zostaną drzewa klasyfikacyjne obrazujące wpływ wytypowanych zmiennych na poziom aspiracji oraz średnie wartości aspiracji w węzłach końcowych (liściach) służące tylko do porównania poziomów aspiracji w zależności od poszczególnych zmiennych. Pod rysunkami drzew dołączono wykresy, które są przybliżoną ilustracją zróżnicowania poziomu aspiracji w zależności od badanych zmiennych (rysunki 1-3)

Rysunek 1 pokazuje, że aspiracje wymarzone są w badanej próbie bardzo wysokie a najważniejszą zmienną różnicującą badaną populację jest wykształcenie osoby wypełniającej ankietę niezależnie od tego czy jest to matka czy ojciec. Ponieważ są to drzewa binarne, powstają dwie podgrupy: pierwsza złożona z osób z wykształceniem podstawowym i zasadniczym zawodowym (na rysunkach oznaczona jako podstawowe), druga gromadząca osoby z wykształceniem średnim i wyższym. W przypadku respondentów z wykształceniem wyższym wykształcenie to jedyna zmienna różnicująca – wartość średnia aspiracji jest tu najwyższa w całej populacji (węzeł 3). Aspiracje respondentów z wykształceniem średnim i podstawowym różnicuje kolejna zmienna płeć dziecka. Płeć dziecka, jako kolejna zmienna różnicująca wskazuje, że w przypadku dziewczynek średnie aspiracje są wyższe niż w przypadku chłopców. Najniższe aspiracje mają rodzice z wykształceniem podstawowym w stosunku do synów (węzeł 9)

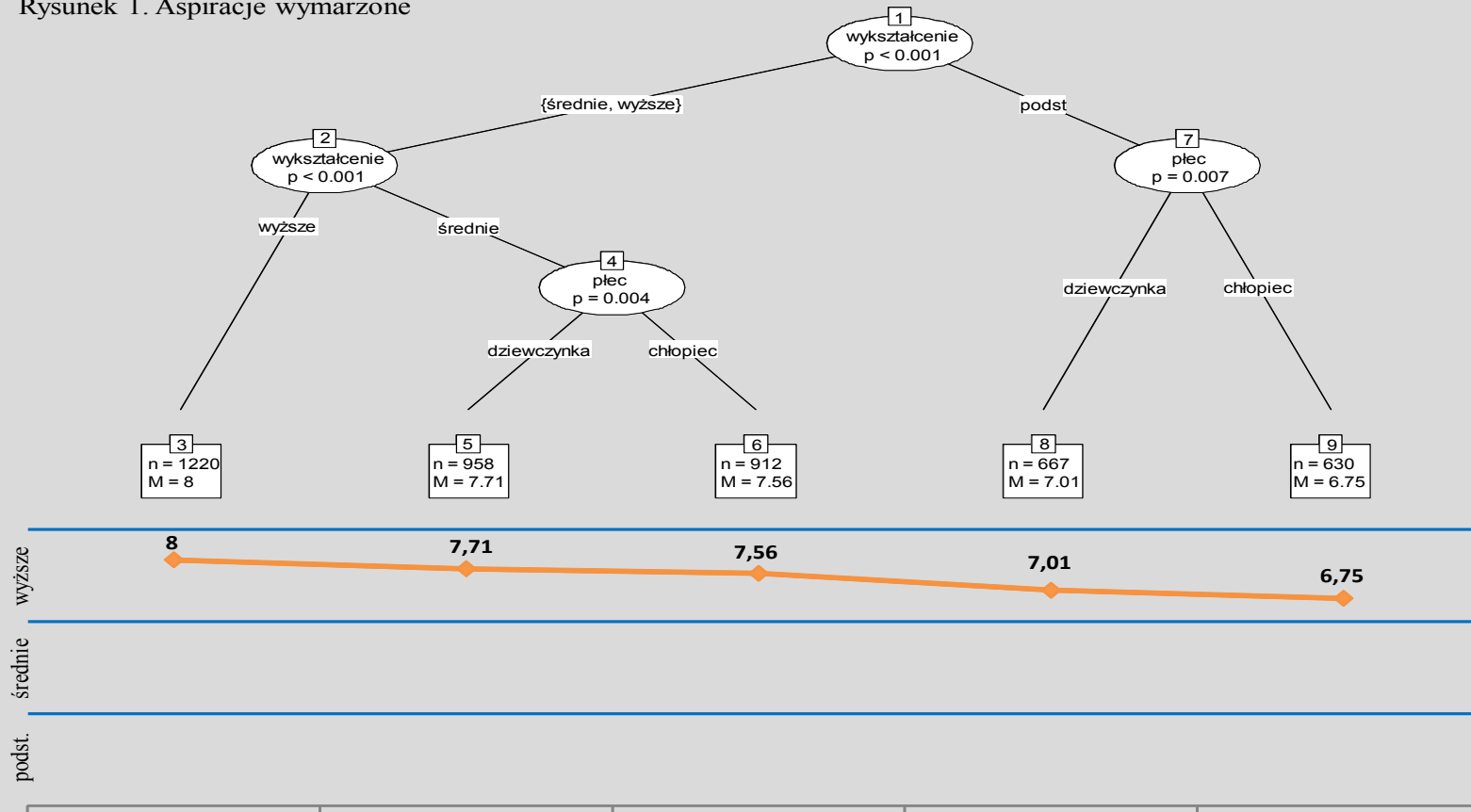
Wykształcenie osoby wypełniającej ankietę jest także najsilniejszym predyktorem w przypadku aspiracji realnych (rysunek 2). Na kolejnym poziomie zmienną różnicującą aspiracje, w grupach o różnym poziomie wykształcenia, jest płeć dziecka; są one wyższe w przypadku dziewczynek . Najwyższe aspiracje realne mają rodzice o najwyższym poziomie wykształcenia (węzeł 4 i 5) a najniższe rodzice z wykształceniem podstawowym (węzeł 12 i 13) .

Średni poziom aspiracji jest niższy niż w przypadku aspiracji wymarzonych ; tylko rodzice słabiej wykształceni sądzą, że dzieci realnie nie osiągną poziomu wykształcenia wyższego, jaki jest przez nich upragniony.

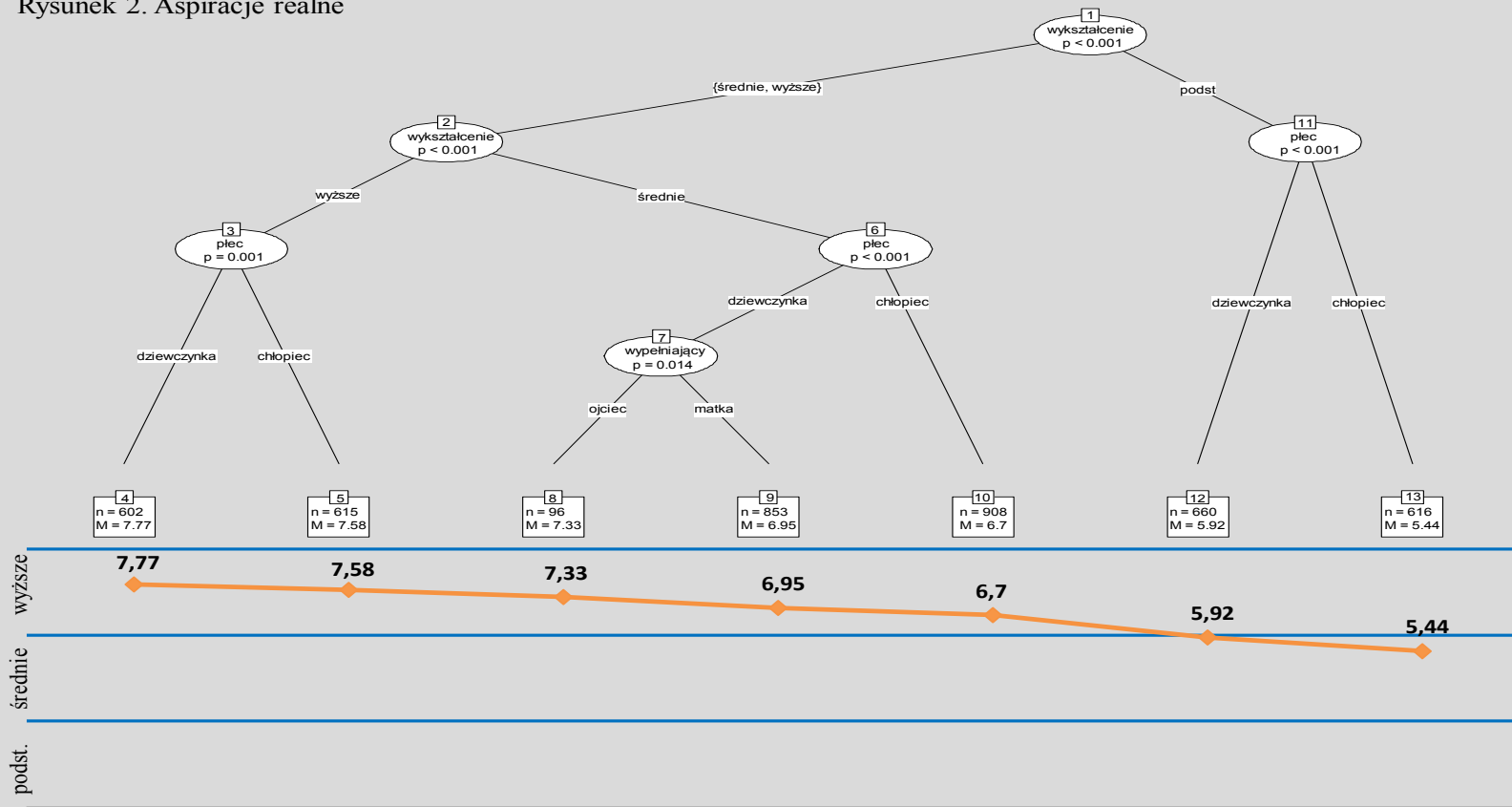
W odniesieniu do aspiracji akceptowalnych (rysunek 3) najsilniejszą zmienną jest również wykształcenie osoby wypełniającej ankietę. Dla osób z wykształceniem wyższym i średnim zmienne różnicujące te grupy są analogiczne jak w przypadku aspiracji wymarzonych. Inny jest tylko średni poziom aspiracji, najwyższy w grupie najwyższej wykształconych (węzeł 3). Inne zmienne różnicują wybory grupy z wykształceniem podstawowym. W tej grupie inaczej spostrzega się aspiracje w stosunku do dziewcząt niż do chłopców. W przypadku dziewcząt są one wyższe (węzeł 8) a dla

chłopców różnice wynikają z tego czy mieszkają w miastach i tu aspiracje są wyższe (węzeł 12) niż na wsi (węzeł 13).

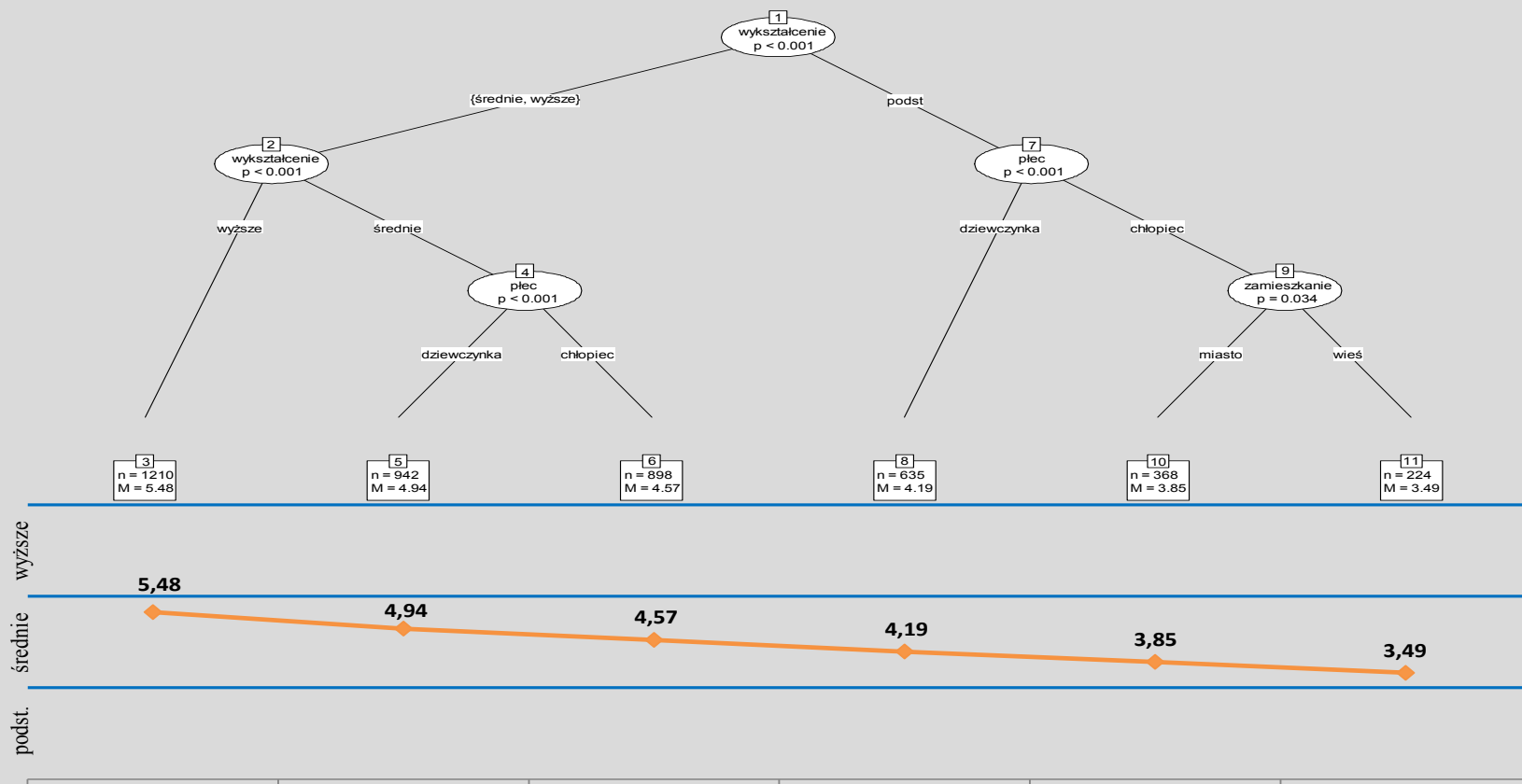
Rysunek 1. Aspiracje wymarzone



Rysunek 2. Aspiracje realne



Rysunek 3. Aspiracje akceptowalne



Podsumowanie pierwszego etapu analizy :

- Wykształcenie jest najważniejszym predykatorem.
- Najsilniejszy wpływ wykształcenia zaobserwować można w przypadku aspiracji realnych.
- W przypadku aspiracji akceptowalnych i wymarzonych te same zmienne różnicują grupy osób z wykształceniem wyższym i średnim.
- Aspiracje wymarzone i akceptowalne rodziców z wykształceniem wyższym nie są istotnie zróżnicowane przez żadną ze zmiennych, tylko aspiracje realne w tej grupie różnicuje płeć dziecka.
- Dla rodziców z najniższym poziomem wykształcenia najważniejszym predykatorem jest płeć dziecka (aspiracje wymarzone i realne) a w przypadku aspiracji akceptowalnych dodatkowe znaczenie ma miejsce zamieszkania

5.2. Wyniki drugiego etapu

W drugim etapie do analizy włączono dodatkowo dwie zmienne: poglądy na temat ścieżki edukacyjnej oraz deklaracje dotyczące działań związanych z aspiracjami. Wprowadzenie ich zmieniło konfigurację zmiennych.

Drugi etap analizy pokazał, że wykształcenie nadal jest zmienna najsilniej różnicującą aspiracje . Jednak dodanie dwóch zmiennych spowodowało, że zmieniła się hierarchia ważności predyktorów. Płeć dziecka, która różnicowała aspiracje rodziców ze średnim, zawodowym i podstawowym wykształceniem, stała się mniej ważna niż poglądy na temat przyszłości edukacyjnej dziecka.

Kolejno omówione zostaną rysunki przedstawiające zróżnicowania poziomów aspiracji ze względu na badane zmienne (rysunki 4-6).

Najwyższy poziom aspiracji wymarzonych (rysunek 4) mają rodzice z wyższym wykształceniem , którzy opowiadają się za wariantem zdobycia przez dziecko przede wszystkim wysokiego poziomu wykształcenia (węzeł 4) oraz ci, którzy preferują pozostałe warianty (węzeł 5). Fakt podejmowania działań różnicuje grupę rodziców z wykształceniem średnim, ale tylko tych, którzy preferują zdobywanie zawodu lub równoczesną pracę i kształcenie się . Ich poziom aspiracji jest jednak niższy (węzeł 9 i 10). Najniższy poziom aspiracji mają rodzice z najniższym wykształceniem, choć tu również różnicującą zmienną są preferencje związane ze ścieżką edukacyjną dziecka. Tu nieco wyższe aspiracje mają ci rodzice , którzy preferują wariant długiego kształcenia (węzeł 12) niż opowiadający się za pozostałymi wariantami (węzeł 13).

Wykształcenie najsilniej różnicuje populację w przypadku aspiracji realnych (rysunek 5). Poziom aspiracji rośnie wraz z poziomem wykształcenia rodziców choć inne zmienne je warunkuje Grupę rodziców wysoko wykształconych różnicują przede wszystkim poglądy na przyszłość edukacyjną dzieci. Wybierający wariant związany z długim cyklem kształcenia liczą się z koniecznością podejmowania działań wspierających edukację. Jeśli je podejmują, to ich poziom aspiracji różnicuje płeć dziecka; najwyższe aspiracje mają w stosunku do dziewcząt (węzeł 6) i nieco

niższe w stosunku do chłopców(węzeł 7). Ta grupa rodziców charakteryzuje się tym, iż nawet gdy nie deklaruje podejmowania działań to ma aspiracje wyższe niż [pozostałe grupy (węzeł 8)

Inna konfiguracja zmiennych występuje w grupie rodziców słabo wykształconych (podstawowe i zasadnicze zawodowe). Czynnikiem różnicującym jest przede wszystkim deklarowanie działań wspomagających edukację. Osoby nie podejmujące działań różnicują poglądy na przyszłość edukacyjną dzieci. Najniższy poziom aspiracji w całej populacji mają ci, którzy opowiadają się za zdobywaniem zawodu lub łączeniem pracy z kształceniem ; średnie aspiracji wskazują, że oceniają oni realne szanse edukacyjne dzieci na poziomie szkoły średniej zawodowej (węzeł 29) . Jeśli deklarują podejmowanie działań, to na ich aspiracje wpływa płeć dziecka, a w dalszej kolejności poglądy na temat przyszłości edukacyjnej. Niższe aspiracje mają rodzice chłopców, którzy wolą, aby synowie najpierw zdobywali zawód lub równocześnie uczyli się i pracowali (węzeł 26) niż ci , którzy opowiadają się za długim cyklu kształcenia (węzeł 25). Poziom aspiracji w stosunku do dziewcząt jest wyższy niż dla chłopców, mimo iż uwarunkowany jest tymi samymi zmiennymi. Oczywiście, jeśli rodzice dziewczynek optują za długim cyklem kształcenia lub łączeniem nauki i pracy to mają wyższe aspiracje (węzeł 22) niż wtedy, gdy preferują w pierwszej kolejności zdobycie zawodu i podjęcie pracy (węzeł 23)

Podobnie jak w pozostałych rodzajach aspiracji, w przypadku aspiracji akceptowalnych (najniższych możliwych do przyjęcia) głównym czynnikiem różnicującym populację jest wykształcenie rodziców (rysunek 6). Poziom aspiracji akceptowalnych jest jednak znacznie niższy niż w przypadku pragnień czy realnej oceny edukacyjnej przyszłości dziecka. Średnie wartości pokazują , że rodzice są w stanie zaakceptować wykształcenie na poziomie średnim, choć jest to znacznie zróżnicowane w zależności od tego, jaki poziom wykształcenia sami posiadają. Najwyższe aspiracje mają rodzice z wyższym wykształceniem preferujący zdobywanie najwyższego wykształcenia i podejmujący działania w tym kierunku (węzeł 5) . Rodzice najslabiej wykształceni mają najniższe aspiracje. a czynnikiem różnicującym jest płeć dziecka. W przypadku dziewczynek aspiracje są wyższe, jeśli rodzice podejmują działania wspierające edukację (węzeł 19) i niższe, jeśli takich działań nie deklarują (węzeł 20). Aspiracje w stosunku do dziewcząt są w tej grupie rodziców wyższe niż w stosunku do chłopców. W przypadku chłopców osiągają najniższy poziom i oscylują wokół wykształcenia zawodowego.

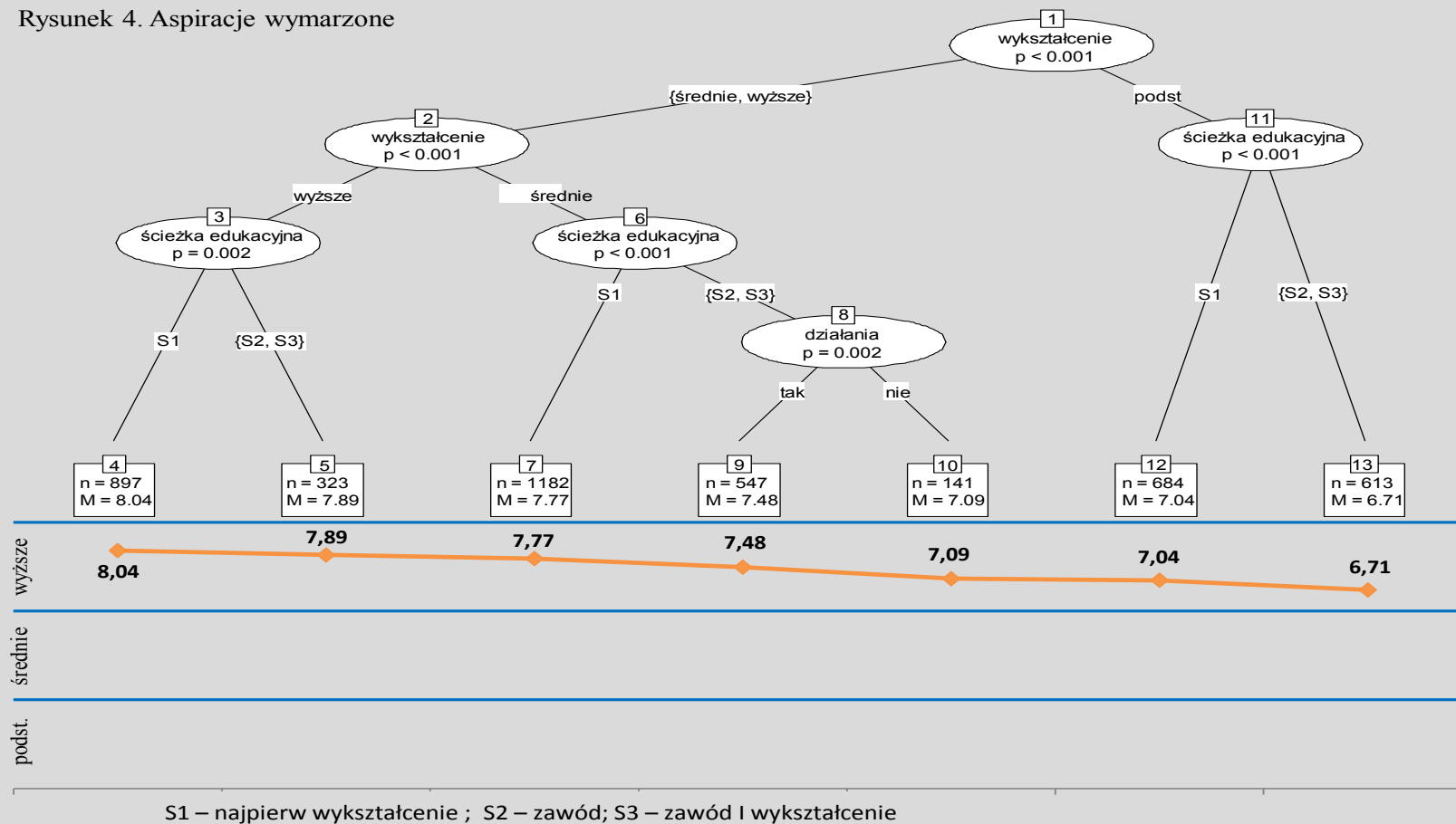
Przedstawione na wykresach pod rysunkami średnie wartości aspiracji służą tylko do porównań między grupami. Pokazują one, że na poziomie aspiracji upragnionych są bardzo wysokie niezależnie od poziomu wykształcenia rodziców oscylują w granicach wykształcenia wyższego.

Podobny trend występuje w przypadku aspiracji realnych, ale tylko w przypadku rodziców z wykształceniem wyższym. Im niższym poziom wykształcenia rodziców, tym niżej oceniają realne szanse zdobycia przez dziecko takiego wykształcenia, o jakim marzą. Rodzice najslabiej wykształceni sądzą, że będzie to poziom szkoły średniej lub co najwyżej szkoła pomaturalna.

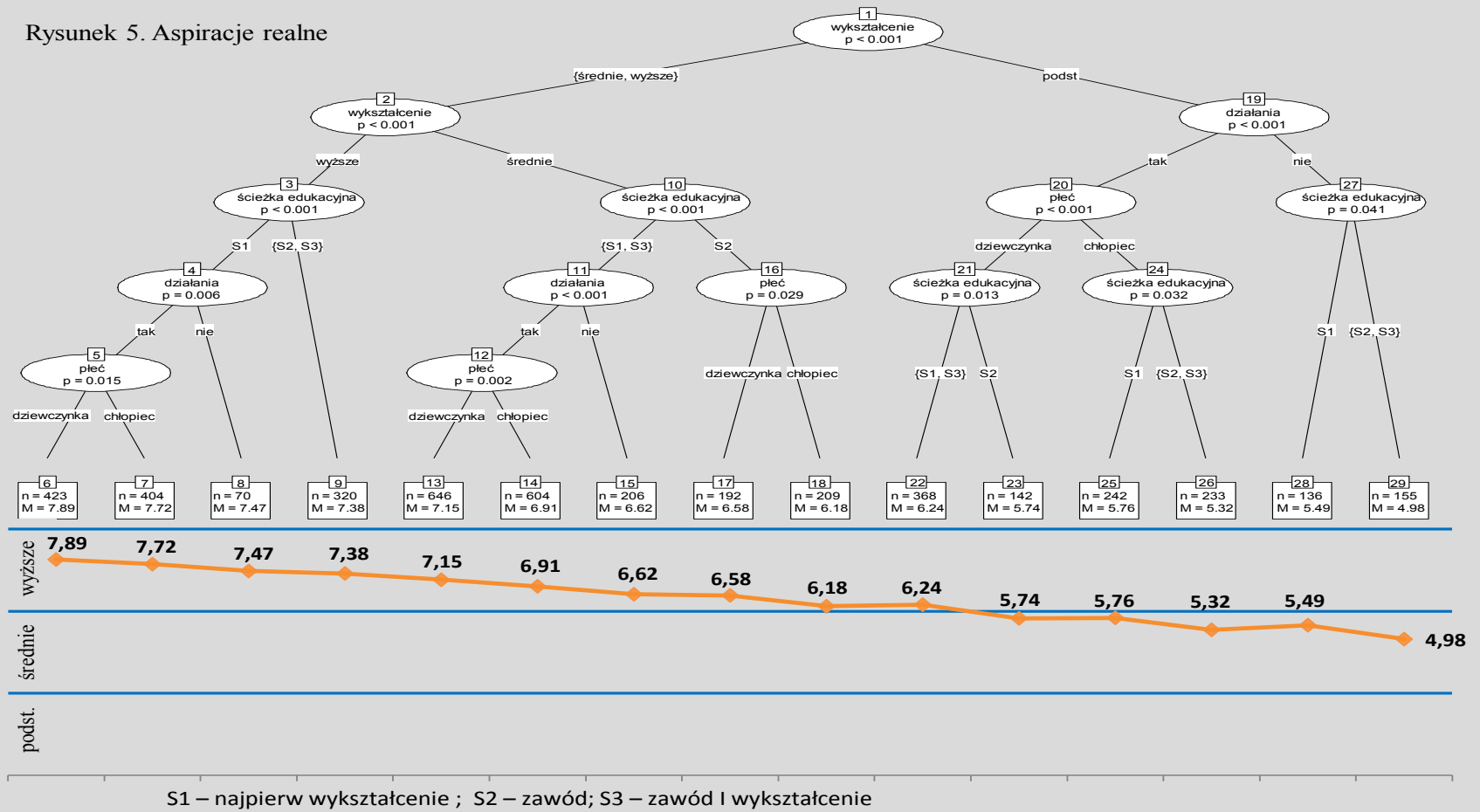
W przypadku aspiracji akceptowalnych wszyscy rodzice są skłonni obniżyć swoje aspiracje; najmniej rodzice z wyższym wykształceniem (o jeden poziom). Do rezygnacji z pragnień najbardziej gotowi są rodzice słabo wykształceni, którzy skłonni są zaakceptować wykształcenie dziecka na

poziomie zasadniczej szkoły zawodowej lub liceum zawodowego, tzn. akceptują takie wykształcenie, które daje dziecku konkretny zawód i umożliwia podjęcie pracy.

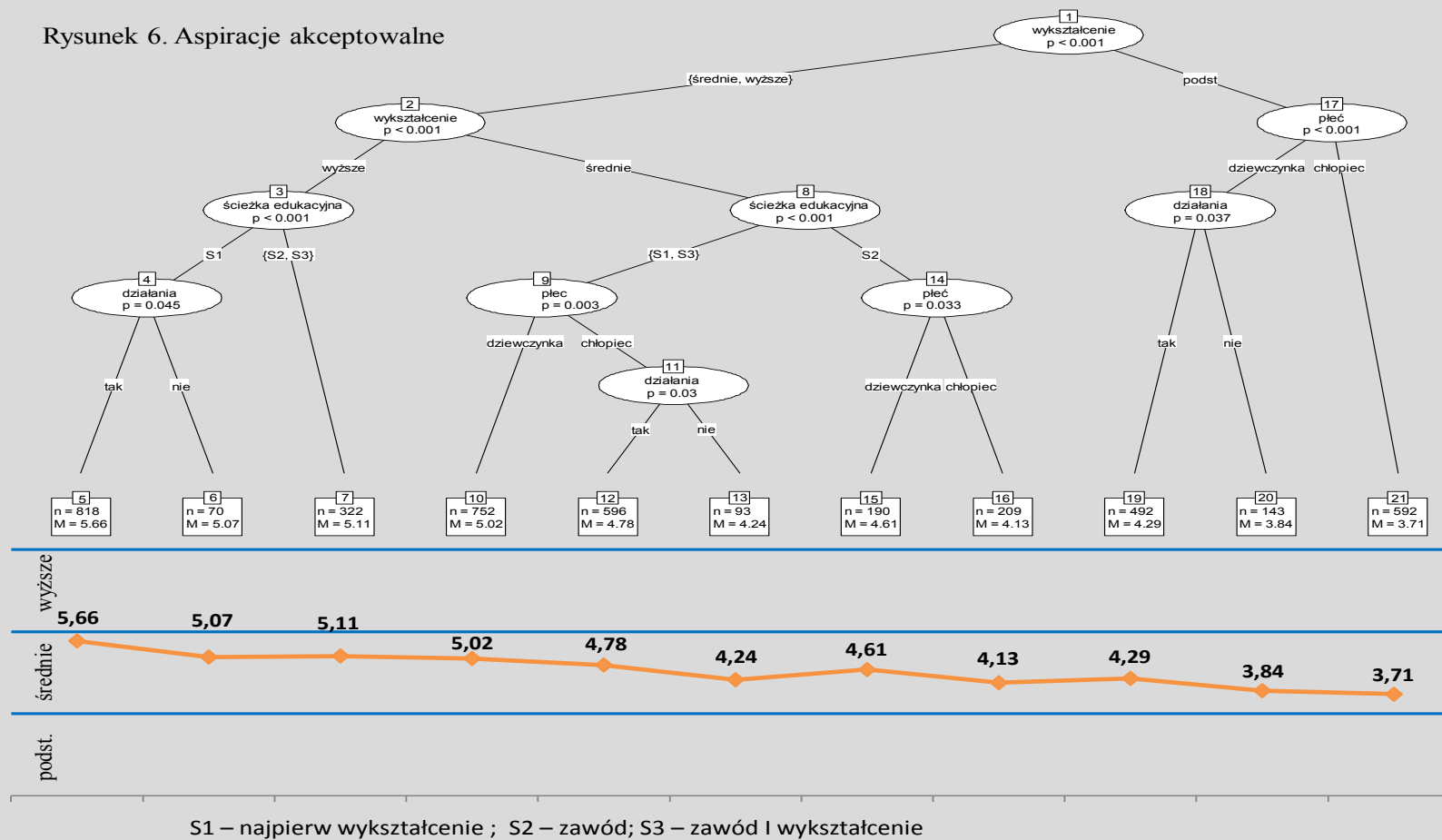
Rysunek 4. Aspiracje wymarzone



Rysunek 5. Aspiracje realne



Rysunek 6. Aspiracje akceptowalne



Podsumowanie drugiego etapu analizy :

- Podobnie, jak w pierwszym etapie analizy wykształcenie najsilniej różnicuje badaną populację.
- W przypadku aspiracji realnych rodzice z wyższym poziomem wykształcenia, którzy preferują długi cykl kształcenia dziecka, są gotowi podjąć działania wspomagające jego edukację. Ich aspiracje są najwyższe w populacji a różni je jedynie płeć dziecka; najwyższe aspiracje mają w stosunku do córek.
- Aspiracje realne słabiej wykształconych rodziców zależą przede wszystkim od tego, czy podejmują działania wspierające edukację dziecka. Najniższe, na tle innych grup, aspiracje realne mają w stosunku do synów, dla których wybierają wcześniejsze zdobycie zawodu i podjęcie pracy a potem ewentualne łączenie jej z dalszą nauką.
- Aspiracje akceptowalne różnicują dwie zmienne: płeć i poglądy na wybór ścieżki edukacyjnej. Dla wyżej wykształconych rodziców ważniejszą zmienną są preferencje dotyczące ścieżki edukacyjnej, słabiej wykształceni biorą pod uwagę przede wszystkim płeć dziecka.
- Aspiracje wymarzone różnicują jedynie poglądy dotyczące ścieżki edukacyjnej a ich poziom rośnie wraz ze wzrostem wykształcenia rodziców i jest najwyższy w przypadku osób z wykształceniem wyższym popierających długi cykl kształcenia dziecka. Najniższy poziom aspiracji wymarzonych występuje u osób z najniższym poziomem wykształcenia, dla których ważniejsze jest zdobycie przez dziecko zawodu, akceptują także wariant łączenie kształcenie z pracą zawodową.

6. Podsumowanie

Wykres drzewa przedstawia wszystkie informacje w prosty i łatwy do interpretacji sposób. Metoda ta nie zastępuje jednak podstawowej statystycznej analizy danych, a jedynie ją uzupełnia.

Na podstawie uzyskanych wykresów można sformułować następujące wnioski ogólne:

- Wykształcenie, niezależnie od liczby zmiennych branych pod uwagę, jest zmienną zdecydowanie najsilniej różnicującą populację i to we wszystkich odmianach aspiracji.
- W drugim etapie analizy, kiedy wzięto pod uwagę większą liczbę zmiennych, tracą na znaczeniu takie zmienne jak płeć czy miejsca zamieszkania a dominujące, po wykształceniu, stają się preferencje dotyczące ścieżki edukacyjnej dziecka oraz podejmowanie działań związanych z inwestowaniem w kształcenie dziecka.
- Drugą, po wykształceniu, zmienną najsilniej różnicującą aspiracje jest są preferencje dotyczące ścieżki edukacyjnej dziecka.

- Trzecią w kolejności zmienną wpływającą na aspiracje są deklaracje związane z podejmowaniem działań sprzyjających kształceniu dziecka. Wyjątkiem są aspiracje realne, gdzie trzecią w kolejności zmienną różnicującą jest płeć dziecka.
- Miejsce zamieszkania ma znaczenie tylko w przypadku aspiracji akceptowalnych.

7. Literatura cytowana

1. Entwisle D. R, Astone N. M. (1994) - *Some practical guidelines for measuring youth's race/ethnicity and socioeconomic status*. "Child Development" 65, 1521-1540
2. Hothorn T., Hornik K. and Zeileis A. (2006). *Unbiased Recursive Partitioning: A Conditional Inference Framework*. Journal of Computational and Graphical Statistics, 15(3), 651--674
3. Hothorn T., Peter P., Dudoit S., Molinaro A. and Van Der Laan M. (2006) - *Survival Ensembles*. Biostatistics, 7(3), 355--373
4. Koronacki J., Ćwik J. 2005. - *Statystyczne systemy uczące się*. Warszawa, Wyd. Naukowo-Techniczne.
5. Kozłowski W., Matczak E. *Aspiracje edukacyjne rodziców dzieci z klas III szkoły podstawowej (raport z badań pilotażowych)* link <http://eduentuzjasci.pl/pl/analizy.html> nr 04/2012
6. Matczak E., Kozłowski W.- *Aspiracje edukacyjne rodziców*, Ruch Pedagogiczny 1 2012, s.41-54
7. Richman A.M, Miller P.M, LeVine R.A (1992) - *Cultural and educational variations in maternal responsiveness*. "Child Development", 1992, 28, 614-621
8. Sawiński Z. (2008) – *Zmiany systemowe a nierówności w dostępie do wykształcenia*. W: *Zmiany stratyfikacji społecznej w Polsce*. Pod. red. H. Domańskiego. Warszawa; Wydawnictwo IFiS PAN.
9. Strobl C., Boulesteix A.L., Zeileis A. and Hothorn T. (2007). - *Bias in Random Forest Variable Importance Measures: Illustrations, Sources and a Solution*. BMC Bioinformatics, 8(25). URL <http://www.biomedcentral.com/1471-2105/8/25>.
10. Strobl C., Boulesteix A.L., Kneib T., Augustin T. and Zeileis A. (2008).- *Conditional Variable Importance for Random Forests*. BMC Bioinformatics, 9(307). URL <http://www.biomedcentral.com/1471-2105/9/307>.