

Uniwersytet Ekonomiczny w Krakowie
Wydział Zarządzania
Kierunek: Informatyka i Ekonometria
Specjalność: Modelowanie i prognozowanie procesów
gospodarczych

**Liniowa analiza dyskryminacyjna dotycząca posiadania konta
osobistego na podstawie badań ankietowych dla Polski z 2011
roku.**

Jakub Winter
Nr albumu: 161402

|

Wstęp	3
Rozdział I: Instytucje finansowe.....	5
1.1 Bank.....	5
a) Klasyfikacja banków ze względu na prawo.	6
b) Klasyfikacja banków ze względu na formę działalności.	7
1.2 Pozostałe instytucje finansowe.	8
1.3 Usługi bankowe.	10
Rozdział II: Zachowania gospodarstw domowych	11
2.1 Gospodarstwo domowe.	11
2.2 Psychologia finansowa.	12
2.2.1 Racjonalność podejmowania decyzji.	12
2.2.2 Zachowania oraz postawy finansowe.	13
2.2.3 Nastawienie do życia a zarządzanie finansami.	15
2.3 Rachunek bankowy.....	15
2.3.1 Konto osobiste w banku komercyjnym.	16
2.3.2 Rodzaje rachunków bankowych	17
2.3.3 Wybór konta przez konsumentów.	18
Rozdział III: Metodologia związana z analizą dyskryminacyjną.....	20
3.1 Przedstawienie zagadnienia analizy dyskryminacyjnej.	20
3.2 Funkcja dyskryminacyjna.	20
3.2.1 Zdolność dyskryminacyjna.....	22
3.3 Funkcja klasyfikacyjna.	22
3.4 Założenia analizy dyskryminacyjnej.....	27
Rozdział IV: Analiza dyskryminacyjna posiadania konta w banku w Polsce.....	28
4.1 Przedstawienie zmiennych.....	29
4.1.1 Tabele liczości.	29
4.1.2 Testy Chi-kwadrat.	37
4.2 Założenia.	Error! Bookmark not defined.
4.3 Analiza dyskryminacyjna.	Error! Bookmark not defined.
4.3.1 Dobór zmiennych.	Error! Bookmark not defined.
4.3.2 Funkcje dyskryminacyjne.....	45
4.3.3 Funkcje klasyfikacyjne.....	48
4.3.4 Prognoza.....	50
a) Pierwsza prognoza.	51
b) Druga prognoza.....	52
c) Trzecia prognoza.....	53
Zakończenie.....	54
Bibliografia.....	56
Aneksy.....	Error! Bookmark not defined.
Spis tabel i rysunków.....	57

Comment [A1]: Zmienić czcionkę

Wstęp.

Analizy dyskryminacyjne są często wykorzystywane w badaniu lojalności klientów, przyczyn decydujących o ich zadowoleniu, monitorowania zmian stopnia satysfakcji oraz segmentacji w celu uzyskania jednorodnych grup nabywczych. Jednakże warto również wiedzieć jaka osoba ze względu na posiadane cechy nie jest naszym klientem.

Jednym z najpopularniejszych segmentów, w których używana jest dyskryminacja to segment bankowy. Dlatego też w pracy zostanie poruszony temat dyskryminacji osób ze względu na posiadanie konta osobistego w banku. Jest to jeden z ciekawszych segmentów ze względu na typowe zachowania człowieka w stosunku do pieniądza. Ponadto, posiadanie konta w banku coraz częściej postrzegane jest jako kryterium wiarygodności, a jego brak uważany jako zacofanie społeczne. Ten problem jest szczególnie dla Polaków ponieważ poziom ubankowienia w Polsce jest jednym z najniższych w Europie.

Analiza zostanie przeprowadzona na danych ankietowych z 2011 roku zebranych przez Radę Monitoringu Społecznego, w ramach ogólnopolskiego projektu badawczego „Diagnoza społeczna”, który obejmuje najważniejsze aspekty życia gospodarstw domowych oraz ich członków. Ponieważ są one odpowiednio kompletne oraz reprezentatywne, możliwe jest skupienie się na wybranych przez mnie ważnych czynnikach. Diagnoza społeczna zakładała przebadanie wszystkich ważnych aspektów życia gospodarstwa domowego oraz jego członków, dzięki czemu dają one pełny obraz o potencjalnym kliencie usług bankowych.

Szczególnie przydatną literaturą podczas wyjaśniania zachowań konsumenckich wobec pieniądza okazały się „Psychologia finansowa” G. Wąsowicz-Kiryló oraz „Finanse osobiste” E. Bogackiej-Kisiel opisujące wiele ciekawych aspektów z psychologii finansowej, natomiast w przypadku analizy dyskryminacyjnej pomocna okazała się książka autorstwa A. Stanisza, „Przystępny kurs statystyki z zastosowaniem STATISTICA”, opisująca ciekawe metody dyskryminacji na przykładach w programie „Statistica”.

W mojej pracy zostanie pokazane, iż istnieją negatywne czynniki wpływające na to, że ludzie nie decydują się na prowadzenie konta, bądź sprawiają, że ich decyzje są nieracjonalne. Jednocześnie zakładam, że dyskryminacja posiadania konta w banku jest

trudna ze względu na niejednorodność posiadanych cech, co oznacza, że wyodrębnienie odpowiednich grup jest **trudne**.

Praca została ograniczona do liniowej analizy dyskryminacyjnej, która w prosty sposób rozdziela obiekty na grupy oraz jest znana z dogodnej interpretacji ekonomicznej. Praca składa się z 4 rozdziałów, zakończenia, bibliografii, aneksów oraz spisu tabel i rysunków. Pierwszy rozdział przedstawia teorię wprowadzającą czytelnika w podstawowe pojęcia dotyczące instytucji finansowych oraz ich usług, a w przypadku drugiego rozdziału, teoria dotycząca psychologii finansowej powiązanej z rachunkiem bankowym. Następne dwa rozdziały to metodologia przeprowadzonej analizy oraz jej właściwa część.

Rozdział I

Instytucje finansowe

Poniższy rozdział ma na celu przedstawienie czym jest instytucja bankowa, podział tych instytucji oraz odróżnienie instytucji bankowych od innych instytucji finansowych o podobnym charakterze działania, ale nieposiadających praw bankowych. Ważne jest **zrozumienie** sposobu działania banku oraz rozróżnienie oferowanych usług, aby lepiej **zrozumieć** pojęcie konta bankowego, którego posiadanie jest analizowane w dalszej części pracy.

1.1 Bank¹².

Początkowo bank był określany jako instytucja finansowa, która zajmuje się zarządzaniem pieniędzem albo inną formą wartości. Nazwa bank, wywodzi się z języka włoskiego ze słowa „banca”³, które to odnosiło się do wymiany pieniędzy na ławkach, najczęściej przed świątyniami. Należy jednak dodać, że bank „jest samodzielną jednostką organizacyjną”⁴, której celem jest powiększanie swojego kapitału oferując usługi finansowe przy użyciu charakterystycznych dla swojej działalności metod.

Ze względu na postęp technologiczny oraz coraz to częstszą dematerializację pieniądza, świadczenie usług finansowych można utożsamiać z zarządzaniem informacjami.⁵

¹ R. Patterson, *Compendium of Banking in Polish & English*, Wydawnictwo Zielona Sowa, Warszawa 2008, s. 54

² W. L. Jaworski, *Bankowość Podstawowe założenia*, Poltext, Warszawa 1998, s. 13.

³ Z języka włoskiego „banca” znaczy „ławka”.

⁴ S. Wojciechowska-Filipek, *Technologia informacyjna w usługach bankowości informacyjnej*, Difin, Warszawa 2010, s11.

⁵ W. L. Jaworski, *Bankowość Podstawowe założenia*, Poltext, Warszawa 1998, s. 13.

a) Klasyfikacja banków ze względu na prawo (postać/formę prawną?).⁶

Według dokumentów Unii Europejskiej, aby dana instytucja była bankiem musi świadczyć określone usługi. Występuje tutaj ścisły podział banków na banki depozytowo-kredytowe, banki inwestycyjne oraz banki uniwersalne.

Pierwszy **ana się tu przyczepi**, czyli bank depozytowo-kredytowy⁷ jest bankiem obsługującym klientów, których interesuje posiadanie konta osobistego oraz krótko i średnioterminowe pożyczki. Początkowo, w tym rodzaju banków, depozytorami byli głównie klienci indywidualni, natomiast pożyczkobiorcami klienci w rozwijających się sektorach, którzy borykali się z problemami utrzymania płynności finansowej. Dzisiaj już mało który bank można uznać za depozytowo-kredytowy. Stają się one bankami uniwersalnymi, poszerzając paletę swoich usług. Ze względu na to, że najważniejszą usługą banku depozytowo-kredytowego jest udzielanie pożyczek, jego głównym narzędziem jest zarządzanie ryzykiem kredytowym.

Z kolei bank inwestycyjny⁸ kieruje swoje usługi głównie do klientów biznesowych, którymi są przeważnie średnie oraz duże firmy. Są to głównie usługi doradcze, takie jak prywatyzacja, przejęcia czy fuzje, usługi związane z zarządzaniem aktywami oraz obrotem papierów wartościowych. Jedną z podstawowych cech tego rodzaju banku jest to, że nie przechowuje wkładów depozytowych swoich klientów.

Ostatnim rodzajem banku określanym przez prawo UE jest bank uniwersalny⁹, który oferuje usługi zarówno banku depozytowo-kredytowego, jak również banku inwestycyjnego. Taki model pozwala bankowi na lepsze relacje z klientem dzięki łączeniu usług finansowych oraz ich dywersyfikację, czyli rozkładanie możliwego ryzyka w taki sposób, aby ograniczać straty, a jest to możliwe dzięki usunięciu pośrednictwa innych banków.

⁶ W. L. Jaworski, *Bankowość Podstawowe założenia*, Poltext, Warszawa 1998, s. 13.

⁷ R. Patterson, *Compendium of Banking...*, prac. cyt., s. 139.

⁸ R. Patterson, *Compendium of Banking in Polish & English*, Wydawnictwo Zielona Sowa, Warszawa 2008, Investment Bank.

⁹ R. Patterson, *Compendium of Banking...*, prac. cyt., s. 131.

b) Klasyfikacja banków ze względu na formę działalności.¹⁰

Kolejny rodzaj klasyfikacji to podział banków ze względu na formę działalności. Omówione zostanie pojęcie banku centralnego, banków komercyjnych, banków specjalnych oraz internetowych. Otóż, różnią się one zarówno funkcjami jak i typami oferowanych usług. W zależności od tego co lub kogo osoba reprezentuje oraz jakie są jej potrzeby, wybierze dany rodzaj banku, bądź nie będzie mogła skorzystać z usług banków należących do pewnej grupy.

Pierwszym z omawianych banków jest bank centralny¹¹, który najczęściej jest najważniejszą instytucją finansową, stojącą na czele systemu bankowego w danym kraju. Decyzje banku centralnego w większości państw są niezależne od wpływów politycznych. Ma to na celu zwiększenie zaufania publicznego oraz ochronę wartości waluty przed nieodpowiednią polityką rządu. W zależności od kraju, bank centralny, może dzielić kilka bądź wszystkie funkcje z innymi instytucjami finansowymi. Jego głównym zadaniem jest ochrona bazy monetarnej¹² oraz stóp procentowych¹³. Jest również pożyczkodawcą ostatniej instancji, co oznacza, że w przypadku bankructwa innych banków zapewnia ochronę ich klientom. Co więcej, bank centralny bierze udział w nadzorze bankowym.

Następną oraz najliczniejszą grupę tworzą banki komercyjne.¹²¹³ Są to banki świadczące usługi skierowane głównie do klientów indywidualnych oraz małych przedsiębiorstw. Zazwyczaj oferują wiele niskooprocentowanych usług finansowych. Oferowane usługi są przeważnie łatwo dostępne i nieskomplikowane. Bank komercyjny czerpie zyski głównie dzięki efektowi skali¹⁴ oraz małym kosztom operacyjnym. Głównymi ograniczeniami dla takich banków są wymogi zachowania płynności finansowej oraz ograniczenia wynikające z przepisów rozliczeniowych i finansowych w danym kraju (również te nakładane przez bank centralny).

Banki specjalne tworzą ciekawą grupę banków ze względu na to, że ich działalność jest wyspecjalizowana w porównaniu do banków komercyjnych ze względu na branżę, oferując szczególną ofertę albo terytorium działania (np. bank zagraniczny czy regionalny).

¹⁰ W. L. Jaworski, *Bankowość Podstawowe założenia*, Poltext, Warszawa 1998, s. 16.

¹¹ Podstawowe stopy procentowe to: referencyjna, lombardowa, depozytowa oraz redyskontowa.

¹² R. Patterson, *Compendium of Banking...*, prac. cyt., s. 33.

¹³ W. L. Jaworski, *Bankowość Podstawowe założenia*, Poltext, Warszawa 1998, s. 16.

Najliczniejszą podgrupę banków specjalnych ze względu na oferowane usługi tworzą banki inwestycyjne.

Najmłodszą grupę tworzą banki internetowe, które można utożsamić z systemami udostępniającymi klientom indywidualnym usługi finansowe za pomocą sieci Internet. Większość banków internetowych to również banki komercyjne, podczas gdy nieliczne pozostają jedynie bankami internetowymi. Usługi oferowane przez bank internetowy są porównywalne do tych oferowanych przez bank komercyjny. Dzięki zmniejszonym kosztom obsługi w przypadku banków internetowych, które nie posiadają oddziałów, banki te mogą pozwolić sobie na zaoferowanie klientom znacznie lepsze warunki oszczędzania oraz stóp procentowych.

1.2 Pozostałe instytucje finansowe.

Obok bankowych instytucji finansowych istnieją jeszcze niebankowe instytucje finansowe, których polityka oraz działania nie są spójne z prawem bankowym, które to według ustawy z dnia 29 sierpnia 1997r o „Prawie bankowym”: „ (...) określa zasady prowadzenia działalności bankowej, tworzenia i organizacji banków, oddziałów i przedstawicielstw banków zagranicznych, a także oddziałów instytucji kredytowych oraz zasady sprawowania nadzoru bankowego, postępowania naprawczego, likwidacji i upadłości banków.” Dlatego też nie są one objęte nadzorem bankowym sprawowanym przez Ministerstwo Finansów, Bank Centralny, bądź organizację niezależną.¹⁴ Nadzór bankowy sprowadza się do wskazania właściwego funkcjonowania oraz kontroli działania zgodnie z prawem bankowym.

Pozostałe instytucje finansowe, których nie obejmuje prawo bankowe, można podzielić ze względu na to, czy przyjmują one wkłady od klientów. Jest to bardzo intuicyjny podział z punktu widzenia konsumenta.

a) Instytucje przyjmujące depozyty.

Grupę instytucji przyjmujących depozyty tworzą przedsiębiorstwa, które oferują usługi pozwalające na zarządzanie wcześniej zdeponowanymi środkami pieniężnymi, bądź ich oszczędzaniu. Zostaną teraz przedstawione dwa typy takich instytucji kropka

¹⁴ W. L. Jaworski, *Bankowość Podstawowe założenia*, Poltext, Warszawa 1998, s. 34.

Pierwszym z nich są spółdzielnie kredytowe¹⁵, które to można zdefiniować jako instytucje finansowe, które są prowadzone w formie banku uniwersalnego. Początkowo usługi finansowe w tych instytucjach były dostępne tylko dla członków danych spółdzielni. Dzisiaj oferują one również usługi dla klientów zewnętrznych.

W większości przypadków, dla członków spółdzielni dostępne są korzystniejsze oprocentowania niż w przypadku innych banków.

Również kasy oszczędnościowe¹⁶, prowadzone są w formie banku uniwersalnego, jednakże ich środki pochodzą głównie z funduszy publicznych, małych firm oraz samorządów. Organami założycielskimi kas oszczędnościowych są samorządy lokalne. Te instytucje zazwyczaj wypłacają więcej w dywidendach niż inne instytucje finansowe.

b) Instytucje nieprzyjmujące depozytów.¹⁷

Charakterystyczną cechą instytucji nieprzyjmujących depozytów jest to, że środki powierzane takim instytucjom są wykorzystywane przez usługi oraz instrumenty finansowe¹⁸ w specyficznym celu np. inwestowania, albo wykonywania płatności.

Jedną z takich instytucji są towarzystwa ubezpieczeniowe, które uzyskują wkłady od klientów, aby następnie zainwestować je w fundusze oraz instrumenty finansowe. Dzięki temu w przypadku zdarzeń losowych na zasadach opisanych w umowie klientowi może zostać wypłacona odpowiednia suma pieniędzy.

Innym przykładem są działające na podobnych zasadach fundusze emerytalne, które zbierają składki od osób pracujących i ich pracodawców. Zebrane środki są inwestowane w instrumenty finansowe według ustalonych przepisów. Część odkładanych składek zostaje wypłacana na zasadach i po czasie ustalonym w umowie.

Fundusze inwestycyjne zbierają wkłady zarówno od osób indywidualnych jak i przedsiębiorstw. Różnią się tym od poprzednio wymienionych instytucji, że polegają na zbiorowym lokowaniu środków pieniężnych. Fundusze inwestycyjne dzielą się na fundusze otwarte z nieokreśloną liczbą jednostek uczestnictwa oraz zamknięte z określoną z góry ich liczbą.

¹⁵ W. L. Jaworski, *Bankowość Podstawowe założenia*, Poltext, Warszawa 1998, s. 30.

¹⁶ W. L. Jaworski, *Bankowość Podstawowe założenia*, Poltext, Warszawa 1998, s. 31.

¹⁷ M. Macierzyński, *Public relations w bankach wirtualnych*, CeDeWu, Warszawa 2008, s. 22.

¹⁸ Instrumenty finansowe są formą zobowiązań pieniężnych jednego podmiotu względem drugiego.

Istnieją również domy maklerskie i firmy powiernicze, które oferują doradztwo inwestycyjne dla oszczędzających oraz inwestorów. Ponadto takie instytucje mogą oferować możliwość prowadzenia konta, które umożliwia inwestowanie wpłaconych tam środków.

Szczególą formę pożyczki oferują firmy leasingowe, udostępniając możliwość korzystania z dobra w zamian za określone w umowie opłaty ratalne. W umowie może być określone, że na koniec okresu tytuł własności przedmiocie o którym mowa w umowie zostanie przeniesiony. Leasing możemy podzielić na bezpośredni, czyli taki w którym biorą udział wyłącznie dwie strony oraz pośredni, charakteryzujący się mniejszą ilością podmiotów biorących udział w umowie.

Sieci obsługi płatności są to instytucje oferujące proste usługi płatnicze tj. płatności elektroniczne z wcześniej wpłaconych środków, czy też wykonywanie rozliczeń podatkowych. Za oferowane usługi płatnicze często pobierana jest niewielka opłata.

Ostatnia grupa, którą przedstawię to wyspecjalizowane w danym rodzaju usług finansowych instytucje niebankowe. Przykładem ich działalności może być pośredniczenie w zakupach ratalnych pomiędzy nabywcą a sprzedawcą. W większości przypadkach ich usługi zbliżone są do tych oferowanych przez banki komercyjne.

1.3 Usługi bankowe.

Przed opisaniem usług bankowych kluczowe jest zrozumienie pojęcia usługi. Należy zauważyć, iż usługa jest niematerialną formą produktu. Natomiast produkt tworzony jest przez zespół atrybutów niematerialnych oraz materialnych umożliwiających zaspokajanie potrzeb klientów i przynoszenie im pewnych korzyści z jego użytkowania. Idąc tym tokiem rozumowania usługi finansowe będą to wszystkie te usługi które zaspokajają potrzeby konsumentów związane z zarządzaniem oraz obrotem pieniężnym. Należy nadmienić także, że usługi świadczone przez bank określane są mianem usług bankowych.¹⁹

Ze względu na treść, usługi bankowe możemy skategoryzować według czterech grup: usługi kredytowe, usługi depozytowe, operacje obrotu płatniczego oraz usługi dodatkowe.²⁰ Poniższa tabela 1, dzięki której łatwiej jest zrozumieć strukturę usług oferowanych przez banki, prezentuje wybrane usługi bankowe .

¹⁹ P. Stodulny, *Analiza satysfakcji i lojalności klientów bankowych*, CeDeWu, 2007, s. 92

²⁰ W. L. Jaworski, *Bankowość Podstawowe...*, prac. cyt., opracowanie własne na podstawie rozdziału 3.

Tabela 1. Wybrane usługi bankowe

Usługi kredytowe	Usługi depozytowe	Operacje obrotu płatniczego	Usługi dodatkowe
-udzielenie kredytów, krótko, średnio oraz długoterminowych	-prowadzenie rachunków oszczędnościowych	-wykonywanie polecenia przelewu krajowego oraz zagranicznego	-doradztwo finansowe
-emisja papierów wartościowych	-emisja papierów wartościowych	-rozliczenia finansowe	-pośrednictwo w obiegu papierów wartościowych
	-pośrednictwo banku		

Opracowanie własne na podstawie: S. Wojciechowska-Filipek, *Technologia...*, prac. cyt., s.17

Rozdział II

Zachowania gospodarstw domowych.

Rozdział drugi poświęcony został opisaniu wpływowi posiadania rachunku bankowego jednostki indywidualnej na gospodarstwo. Ponadto poruszony zostanie w nim temat kształtowania się zachowań finansowych głównie w kontekście prowadzenia wspomnianego rachunku. Wszystko to zamyka **rozdział** dotyczący najpopularniejszego z dostępnych rachunków: rachunku oszczędnościowo-rozliczeniowego.

2.1 Gospodarstwo domowe.

Decyzje związane z zarządzaniem pieniędzem przez jednostki indywidualne często dotyczą całego gospodarstwa domowego. Stanowią one funkcję przedmiotową gospodarstwa, są to nie tylko decyzje, ale również środki finansowe umożliwiające zaspokajanie potrzeb. Funkcje gospodarcze gospodarstwa domowego często pokrywają się z funkcjami rodziny, ponieważ stanowi ona pewną grupę społeczną i spełnia również funkcje wychowawcze.²¹ Gospodarstwo domowe tworzą osoby, które wspólnymi siłami oraz dochodami zaspokajają indywidualne potrzeby. Jednakże należy **zauważyć, iż** gospodarstwo domowe może tworzyć jedna osoba, która jest nabywcą usług finansowych.²² W takim przypadku skupia się ona na zaspokajaniu potrzeb całego gospodarstwa domowego jak i zabezpieczaniu jego przyszłości, poprzez inwestowanie oraz zbieranie oszczędności. Najczęściej cele oszczędzania dotyczą wszystkich uczestników

²¹ *Finanse osobiste, zachowania-produkty-strategie*, prac. zbior. Pod red. E. Bogacka-Kisiel, Wydawnictwo Naukowe PWN SA, Warszawa 2012. s. 15.

²² D. Fatuła, *Zachowania polskich gospodarstw domowych na rynku finansowym*, Wydawnictwo Uniwersytetu ekonomicznego w Krakowie, Kraków 2010, s. 26.

gospodarstwa domowego np. zakup mieszkania. Jednakże, nawet w przypadku kiedy finansowana zostaje potrzeba tylko jednej lub wybranych osób z gospodarstwa domowego, to bierze udział w tym całe gospodarstwo domowe, ponieważ uczestniczą oni np. w ograniczonej konsumpcji na rzecz finansowania danej jednostki.

2.2 Psychologia finansowa.

Na decyzje podejmowane przez człowieka w ekonomii wpływa wiele czynników. Należy pamiętać o tym, że różnimy się od siebie, wobec czego zdarzają się jednostki podejmujące decyzje nieracjonalne.

Badaniem decyzji ekonomicznych podejmowanych przez ludzi zajmuje się ekonomia behawioralna.²³ Nauka ta, na podstawie bodźców społeczno-psychologicznych stara się wytłumaczyć zachowania, które z ekonomicznego punktu widzenia nie byłyby racjonalne. Dzięki temu w rzeczywistości jesteśmy w stanie urzeczywistniać sztywne modele ekonomiczne. Jednakże, należy zwrócić uwagę, iż jest to bardzo ogólna dziedzina i nie daje nam odpowiedzi na zachowania człowieka wobec pieniądza. Dlatego coraz częściej można natknąć się na definicję psychologii finansowej²⁴, która skupia się na zachowaniu człowieka wokół decyzji związanych z finansami.

2.2.1 Racjonalność podejmowania decyzji.

Założenie o racjonalności zachowań ludzkich jest podstawowym założeniem w ekonomii. Zakładało się niegdyś, iż „założenie racjonalności jest definicyjną cechą ekonomii”²⁵ oraz, że racjonalność podejmowania decyzji wynika z „maksymalizowania szans w celu osiągnięcia celu”²⁵, co oznacza, że racjonalny wybór zależy od wartości jakie niesie dla nas dane zdarzenie i prawdopodobieństwo jego wystąpienia. Jednakże, zauważono, że racjonalność może być postrzegana subiektywnie i w rezultacie interpretowana jako maksymalizowanie oczekiwanej satysfakcji.

Należy również zwrócić uwagę, że racjonalność w podejmowaniu decyzji przez ludzi może być ograniczona. Wynika to z tego, że ludzie nie są doskonali i w rezultacie ich możliwości przetwarzania informacji są niedoskonałe, a ponadto często realizują wiele celów jednocześnie. Implikacją tego jest to, że popełniają błędy oraz nie są w stanie

²³ G. Wąsowicz-Kiryło, *Psychologia finansowa*, Diffin, Warszawa 2008, s. 12.

²⁴ G. Wąsowicz-Kiryło, *Psychologia finansowa*, Diffin, Warszawa 2008, s. 12.

²⁵ G. Wąsowicz-Kiryło, *Psychologia finansowa*, Diffin, Warszawa 2008, s. 83.

dokonać właściwego wyboru poprzez kalkulację opcji. W rezultacie wiele podejmowanych decyzji nie jest racjonalnych z ekonomicznego punktu widzenia.

2.2.2 Zachowania oraz postawy finansowe.

Zachowania związane z rozporządzaniem pieniędzmi są swoistą konsekwencją zaplanowanych decyzji. Ze względu na racjonalność procesu decyzyjnego, często podatne są na błędy oraz obarczone pewnym stopniem ryzyka. Każde zachowanie dotyczy czynności fizycznych bądź psychicznych, związanych z wyborami dotyczącymi gospodarstwa domowego, jego użytkowaniem i utrzymaniem.²⁶

Szczególny rodzaj zachowań konsumenckich gospodarstw domowych stanowią zachowania finansowe. Zachowania finansowe można utożsamić ze wszystkimi działaniami podejmowanymi w celu zaspokojenia potrzeb finansowych gospodarstwa domowego. Dotyczą one każdego gospodarstwa domowego.²⁷

Zachowania finansowe można przedstawić za pomocą hierarchii co przedstawia rysunek 1. **Te**, które dotyczą szerszej grupy osób oraz mają bardziej podstawowy charakter, znajdują się na samym dole piramidy. Natomiast **te** bardziej celowe oraz dotyczące węższej grupy, na szczycie. Należy zauważyć, że hierarchie mogą się różnić między sobą w różnych gospodarstwach domowych kolejnością, a nawet występować na tym samym poziomie.

²⁶ *Finanse osobiste...*, prac. cyt., prac. zbior. Pod red. E. Bogacka-Kisiel, s. 38.

²⁷ D. Fatuła, *Zachowania polskich...*, prac. cyt., s. 34.

Rysunek nr.1 *Hierarchia zachowań finansowych*



Opracowanie własne na podstawie: *Finanse osob.*, prac. cyt., prac. zbior. Pod red. E. Bogacka-Kisiel, s. 39.

W zarządzaniu budżetem i płynnością najczęściej pomagają nam usługi bankowe. Najpopularniejszą w Polsce oraz jedną z najbardziej podstawowych usług jest ROR²⁸. Pozwolę sobie zwrócić uwagę na fakt, że posiadanie pieniędzy na rachunku bankowym nie jest jednoznaczne z oszczędzaniem²⁹. Oszczędzanie pojawia się tam, gdzie środki przechowywane na rachunku nie są wykorzystywane przez dłuższy okres czasu. W dzisiejszych czasach ROR, można śmiało stwierdzić, że zaspokaja wszystkie podstawowe potrzeby związane z zarządzaniem płynnością w gospodarstwie domowym. Przez płynność finansową należy rozumieć zdolność do wywiązywania się z krótkoterminowych zobowiązań finansowych³⁰. Utrzymanie płynności jest niezbędne do prawidłowego funkcjonowania gospodarstwa domowego, dlatego też stanowi trzon piramidy hierarchii.

Nadwyżka dochodu gospodarstwa domowego może zostać wykorzystana w celach oszczędnościowych niezwiązanych z konkretnym celem. Najczęściej korzysta się z nieskomplikowanych instrumentów finansowych, takich jak ROR, czy prosta lokata bankowa. Głównie chodzi tutaj o nieobowiązkowe ubezpieczenia oraz odkładanie pieniędzy tak, aby móc je wykorzystać podczas nieoczekiwanego zdarzenia³¹. Ten poziom

²⁸ ROR-rachunek oszczędnościowo-rozliczeniowy.

²⁹ D. Fatuła, *Zachowania polskich...*, prac. cyt., s. 32.

³⁰ *Finanse osobiste...*, prac. cyt., prac. zbior. Pod red. E. Bogacka-Kisiel, s. 83.

³¹ D. Fatuła, *Zachowania polskich...*, prac. cyt., s. 32 i 35.

w hierarchii jest bardziej skomplikowany, ponieważ wymaga utrzymywania wspomnianej już wyżej płynności oraz wyrzeczeń ze strony członków gospodarstwa domowego.

Możemy wyróżnić również finansowanie celów, które wiąże się z oszczędzaniem bądź różnymi formami kredytowania, ponieważ w celu osiągnięcia celu wymagane jest zgromadzenie odpowiedniej ilości środków pieniężnych. Cele finansowe wynikają z zachowań człowieka uwarunkowanych jego potrzebami.³² Oprócz samokontroli, wymagają one także ustalenia priorytetów w wykonywaniu zadań przez nabywcę. Większość dzisiejszych ROR-ów oferuje dodatkowe usługi przy okazji jego prowadzenia, dzięki czemu, możliwe jest zarówno oszczędzanie jak i pozyskiwanie pieniędzy za pomocą innych usług.

Ostatni i najbardziej wymagający poziom w hierarchii zajmuje inwestowanie. Wiąże się ono z maksymalizacją oczekiwanej użyteczności, czyli powiększaniu wartości swoich zasobów finansowych lub zapobieganiu ich zmniejszeniu w przypadku inflacji³³ oraz dotyczy tych, którym udało się już wcześniej zgromadzić oszczędności. Możliwości inwestycyjne na koncie osobistym są zazwyczaj bardzo ograniczone. Jednakże, tak jak w przypadku oszczędzania, wiele banków, oferuje dodatkowe usługi inwestycyjne dla posiadaczy rachunku.

2.2.3 Nastawienie do życia a zarządzanie finansami.

Postrzeganie własnej sytuacji materialnej jest subiektywnym pojęciem od którego zależy to jak ktoś patrzy na życie i jakie ma cele. Odczuwalna sytuacja finansowa osoby która jest nastawiona do życia pozytywnie będzie zdecydowanie lepsza niż w przypadku osoby o negatywnym nastawieniu.³³ Świadczy tutaj pozytywna korelacja pomiędzy dochodem a poczuciem szczęścia. Należy jednak pamiętać, że nie w każdym przypadku można zaobserwować taką zależność. Zawsze kiedy wydawanie pieniędzy wiąże się z doświadczeniem negatywnych emocji ta zależność będzie odwrotna. Należy również wspomnieć o tym, że relacja pomiędzy dobrostanem a dochodem jest zdecydowanie silniejsza w przypadku osób mieszkających w państwach słabiej rozwiniętych, gdzie

³² D. Fatuła, *Zachowania polskich...*, prac. cyt., s. 37.

³³ D. Maison, *Polak w świecie finansów*, Wydawnictwo naukowe PWN, Warszawa 2013, s.25

pieniądze są przeznaczane na rzeczy bardziej istotne ze względu na przetrwanie niż w państwach wysoko rozwiniętych.³⁴

Z pozoru może się zdawać, że odpowiednie zachowania finansowe, takie jak sukcesywne zarządzanie, oszczędzanie i inwestowanie pieniędzy prowadzi do wzrostu materialnego dobrostanu, który wpływa na nasze postrzeganie życia, a tym samym na nasze samopoczucie psychiczne.³⁵ Jednakże w coraz to większej ilości prac zauważa się odwrotną korelację. Domniema się, iż to od naszego dobrostanu psychicznego zależą nasze cele finansowe oraz umiejętność zarządzania budżetem. U osób szczęśliwych oraz bardziej optymistycznych obserwuje się lepszą sytuację materialną uzyskaną dzięki sukcesom finansowym. W szczególności mówi się tutaj o nastawieniu do życia oraz stabilności emocjonalnej.³⁶ Warto również zauważyć, że osoby usatysfakcjonowane swoim stanem materialnym cechują się bardziej realnymi potrzebami. Warto zatem zastanowić się jak poziom zadowolenia z życia może wpływać na umiejętność zarządzania budżetem, czego odzwierciedleniem może być posiadanie konta w banku.

2.3 Rachunek bankowy.

W przypadku kiedy gospodarstwo domowe korzysta z konta w banku, które pełni rolę konta głównego, bądź gospodarstwo domowe ma formę jednoosobową, konto może być utożsamiane z budżetem domowym. Ze względu na to jak ważną rolę odgrywa konto bankowe w gospodarstwie domowym w poniższym podrozdziale zostanie przedstawiona charakterystyka rachunku bankowego oraz co warto wziąć pod uwagę przy jego wyborze.

2.3.1 Konto osobiste w banku komercyjnym.

Najpopularniejszą formą zarządzania finansami w Polsce jest posiadanie rachunku oszczędnościowo-rozliczeniowego, oferowanego przez banki komercyjne. Najczęściej występuje on pod nazwą „konto osobiste”, podczas gdy wszystkie inne warianty tej nazwy są natury typowo marketingowej. Podczas, gdy każdy może, ale nie musi otwierać rachunku bankowego w dowolnym banku, to podmioty gospodarcze mają obowiązek jego posiadania.³⁷

³⁴ *Finanse osobiste...*, prac. cyt., prac. zbior. Pod red. E. Bogacka-Kisiel, s. 156 oraz D. Maison, *Polak w świecie finansów*, Wydawnictwo naukowe PWN, Warszawa 2013, s.26

³⁵ *Finanse osobiste...*, prac. cyt., prac. zbior. Pod red. E. Bogacka-Kisiel, s. 158

³⁶ D. Maison, *Polak w świecie finansów*, Wydawnictwo naukowe PWN, Warszawa 2013, s.33

³⁷ Z. Krzyżkiewicz *Operacje Bankowe*, Poltex, Warszawa 2001, s.39.

Otwarcie rachunku jest możliwe na podstawie umowy rachunku bankowego zawieranej pomiędzy bankiem a osobą prywatną bądź przedsiębiorstwem. Na podstawie umowy bank jest zobowiązany do przechowywania środków pieniężnych deponenta oraz świadczenia usług określonych w umowie, natomiast osoba bądź podmiot gospodarczy do przestrzegania warunków określonych w umowie oraz przeprowadzania rozliczeń.³⁸ Umowę dotyczącą rachunku bankowego regulują przepisy prawa bankowego oraz kodeksu cywilnego, jednakże ze względu na charakter umowy niezależnej, umowy poszczególnych banków mogą różnić się między sobą.

2.3.2 Rodzaje rachunków bankowych .³⁹

Aby zrozumieć ROR (nie znalazłam rozszyfrowania tego skrótu) warto przytoczyć rodzaje rachunków oferowanych przez banki komercyjne służące do przechowywania środków pieniężnych oraz sposobów dokonywania rozliczeń. Według polskiego prawa bankowego, możemy wyróżnić cztery rodzaje rachunków bankowych.

Pierwszym z nich są rachunki bieżące, które służą głównie do rozliczeń krajowych np. wypłat, płatności podatków, czy wykonywania przelewów pieniężnych oraz gromadzenia środków pieniężnych. Należy zauważyć, że rachunek ten może zostać przekształcony w „rachunek wyodrębnionych środków” jako rachunek przeznaczony na określone cele np. fundusz socjalny.

Kolejnym rachunkiem jest rachunek pomocniczy, którego umowa szczegółowo określa źródło z którego pochodzą wpływy na ten rachunek. Rachunek pomocniczy prowadzony jest w celu wykonywania tylko określonych rozliczeń. Ten rodzaj rachunku może być prowadzony w innym banku niż rachunek bieżący, ale wymaga on od właściciela posiadania takiego rachunku.

Trzecim z kolei jest rachunek lokat terminowych, który jak nazwa sugeruje służy do przechowywania środków pieniężnych na określony w umowie czas. Ze względu na to, że środki zazwyczaj nie mogą być podjęte, aż do określonego czasu, cechuje go wysokie oprocentowanie.

³⁸ *Bankowa obsługa firmy*, prac. zbior. pod red. R. Szewczyk, Oficyna ekonomiczna, Kraków 2002, s. 12.

³⁹ *Bankowa obsługa firmy*, prac. zbior. pod red. R. Szewczyk, Oficyna ekonomiczna, Kraków 2002, s. 13.

Ostatnim, ale nie najmniej ważnym rachunkiem jest rachunek oszczędnościowy. Przeważnie ten typ rachunku skierowany jest do osób indywidualnych, a więc dla gospodarstw domowych zdecydowanych na oszczędzanie.

W ramach krótkiego podsumowania, należy zauważyć że rachunek oszczędnościowo-rozliczeniowy, który jest połączeniem rachunku bieżącego oraz rachunku oszczędnościowy, który oferuje możliwości każdego z nich.

2.3.3 Wybór konta przez konsumentów.⁴⁰

Przed rozpoczęciem podejmowania decyzji finansowych związanych z prowadzeniem konta osobistego, należy podjąć decyzję dotyczącą wyboru odpowiedniego konta oraz banku. Banki różnicują swoją ofertę proponując odmienne oprocentowania i opłaty oraz oferują dodatkowe usługi w ramach prowadzenia konta osobistego. Warto zauważyć, że podstawowym czynnikiem przy podejmowaniu tej decyzji jest to, dla kogo to konto ma być przeznaczone. Otóż, kontem może zarządzać więcej niż jedna osoba, co jest regulowane przy podpisywaniu umowy. Prowadzenie konta wspólnie niesie za sobą pewne zalety takie jak większe obroty na koncie, czego wynikiem mogą być większe limity kredytowania. Należy też zwrócić uwagę na sposób dostępu do naszego konta osobistego. W dzisiejszych czasach bankowość internetowa jest już elementem niemal każdego banku, jednakże nie we wszystkich mamy możliwość wykonania tych samych operacji przy użyciu sieci. Pisząc o kanale dostępu, ważnym jest aby zwrócić uwagę na dostępność oferowanych usług i sposób w jaki mogą być zarządzane. Banki oferujące „umowy ramowe” pozwalają na aktywację dotykowych usług przez klienta bez potrzeby składania dodatkowego podpisu. Następnym interesującym oraz ważnym czynnikiem jest dostęp do bankomatów oraz bezgotówkowy obrót pieniężny. W zależności od banku, różni się liczba bezprowizyjnych bankomatów, natomiast opłaty w tych z prowizjami nie są jednakowe. Ponadto można się spotkać z korzystnymi usługami wynikającymi z obsługi bezgotówkowej tj., moneyback, polegającej na zwrocie części wydatków przy płatnościach kartą.

Ogólnie rzecz biorąc przy wyborze rachunku bankowego konsument musi brać pod uwagę zarówno wszystkie opłaty całościowo, różnorodność oferowanych usług powiązanych z kontem, jak również czynniki „miękkie”, takie jak sprawność personelu czy

⁴⁰. *Finanse osobiste*..., prac. cyt., prac. zbior. Pod red. E. Bogacka-Kisiel, cały rozdział 3.3.4. s. 93.

kanal dostępu do usług. Bez posiadania kompletnej wiedzy na temat większości z wymienionych czynników w rozważanych bankach ciężko jest podjąć racjonalną decyzję związaną z prowadzeniem konta osobistego. W dzisiejszych czasach banki oferują wiele skomplikowanych z punktu widzenia laika pakietów oraz ukrytych kosztów, które mogą być problematyczne.

Rozdział III

Metodologia związana z analizą dyskryminacyjną.

W rozdziale tym zostaną przedstawione zagadnienia związane z analizą dyskryminacyjną, jak również sposób jej przeprowadzenia. Ponadto zostaną omówione założenia o których należy pamiętać podczas wykonywania analizy.

3.1 Przedstawienie zagadnienia analizy dyskryminacyjnej.

Analiza dyskryminacyjna pozwala na skonstruowanie funkcji, które opisują przynależność obiektów badania do poszczególnych grup. Grupy te są określane przez zbiór wartości zmiennej objaśniającej biorącej udział w funkcjach dyskryminacyjnych⁴¹. Kwintesencją analizy jest badanie różnic między grupami oraz zrozumienie wpływu zmiennych objaśniających na te różnice.⁴²

Analizę dyskryminacyjną można podzielić na dwa etapy. Pierwszym z nich jest dyskryminacja, czyli tworzenie grup na podstawie badania wpływu zmiennych objaśniających (nazywanych również zmiennymi dyskryminacyjnymi) na objaśnianą. Drugi etap stanowi klasyfikacja czyli przydzielanie nowych obiektów do wcześniej ustalonych grup.⁴³

3.2 Funkcja dyskryminacyjna.

W pracy tej wykorzystana zostanie metoda liniowej analizy dyskryminacyjnej zaproponowana w 1936r. przez genetyka i statystyka, Ronalda Fishera.⁴⁴ Zakłada ona tworzenie liniowych funkcji dyskryminacyjnych, których liczba nie jest większa od liczby wariantów zmiennej objaśniającej minus jeden. Warto wspomnieć, że zaletą liniowej funkcji jest łatwość w jej interpretacji. Liniowa funkcja dyskryminacyjna jest dana wzorem:

⁴¹ M. Rószkiewicz, *Metody ilościowe w badaniach marketingowych*, Wydawnictwo naukowe PWN, Warszawa 2002, s. 170

⁴² G. A. Churchill, *Badania marketingowe, podstawy metodologiczne*, Wydawnictwo naukowe PWN, Warszawa 2002, s. 787

⁴³ E. Gatnar, *Symboliczne metody klasyfikacji danych*, Wydawnictwo naukowe PWN, Warszawa 1998, s. 44

⁴⁴ E. Gatnar, *Symboliczne metody klasyfikacji danych*, Wydawnictwo naukowe PWN, Warszawa 1998, s. 45

$$Y_{kj} = \beta_0 + \beta_1 x_{1kj} + \beta_2 x_{2kj} + \dots + \beta_p x_{pkj},^{45} \quad (3.1)$$

gdzie: $k = 1, \dots, n$; $j = 1, \dots, g$; $i = 1, \dots, p$ oraz

p –liczba zmiennych objaśniających,

n –liczba przypadków,

g –liczba grup,

Y_{kj} –wartość funkcji dyskryminacyjnej w j -tej grupie i k -tym przypadku,

β_i –parametr funkcji dyskryminacyjnej.

A co to x

Liniowe funkcje dyskryminacyjne mają za zadanie rozdzielać grupy obiektów w taki sposób, aby jak najlepiej je separować, a to oznacza konstruowanie funkcji w taki sposób, aby średnie grup były jak najbardziej odmienne od siebie przy zachowaniu braku korelacji między funkcjami.⁴⁶ Niestandaryzowane parametry mówią nam jak silnie dana zmienna objaśniająca oddziałuje na zmienną objaśnianą oraz jaki jest to kierunek oddziaływania. Jeżeli jest to wartość dodatnia, dana zmienna jest stymulantą, czyli przemawia za przynależnością zmiennej do wydzielonej grupy. W przeciwnym wypadku, jeżeli jest to wartość ujemna, dana zmienna jest destymulantą, czyli przemawia za nie przynależnością zmiennej do wydzielonej grupy.⁴⁷

Należy zauważyć, że obliczone parametry funkcji dyskryminacyjnej nie nadają się do interpretacji ze względu na różne średnie, a także na różne zakresy wartości zmiennych objaśniających. Dlatego też stosuje się standaryzację parametrów podaną wzorem (2):

$$\hat{\beta}_i = \beta_i \sqrt{n - g}, \quad \hat{\beta}_0 = -\sum_{i=1}^p \hat{\beta}_i \bar{x}_i \quad (3.2)$$

p –liczba zmiennych objaśniających,

n –liczba przypadków,

g – liczba grup,

\bar{x}_i – średnia i -tej zmiennej objaśniającej,

⁴⁵ A. Stanisławski, *Przystępny kurs statystyki z zastosowaniem STATISTICA PL na przykładach z medycyny, Tom 3. Analizy wielowymiarowe*, Statsoft Polska, Kraków 2007, s. 53

⁴⁶ A. Stanisławski, *Przystępny kurs ..., prac. cyt., s. 54*

⁴⁷ http://www.nbp.pl/publikacje/materialy_i_studia/ms286.pdf s15

β_i – parametr funkcji dyskryminacyjnej,

β_i – parametr funkcji dyskryminacyjnej po standaryzacji,

Dzięki standaryzacji, wyniki możemy interpretować jako odchylenie standardowe od średniej oraz łatwo porównywać je do innych. Natomiast procent wariancji międzygrupowej przedstawia udział danej funkcji dyskryminacyjnej w wyjaśnionej wariancji. Procent wariancji obliczamy ze wzoru:⁴⁸

$$V = \frac{\lambda_i}{\sum_{i=1}^n \lambda_i} * 100\%, \quad (3.3)$$

gdzie:

λ_i – wartość własna funkcji.⁴⁹ **A co to?**

3.2.1 Zdolność dyskryminacyjna

Gdy znamy już procentowy udział funkcji w wariancji, warto również zbadać zależność pomiędzy zmienną objaśnianą a poszczególnymi funkcjami dyskryminacyjnymi. Służy do tego współczynnik korelacji, który przyjmuje wartości od 0 do 1, gdzie 0 oznacza brak powiązania, a 1 bardzo dobre powiązania. W praktyce, oznacza on jak dobrze grupy są dzielone przez odpowiednią funkcję. Ta miara jest dana wzorem:⁵⁰

$$R_i = \sqrt{\frac{\lambda_i}{1+\lambda_i}}, \quad (4)$$

gdzie:

λ_i – wartość własna funkcji.

⁴⁸ A. Stanisławski, *Przystępny kurs ...*, prac. cyt., s 61

⁴⁹ Czyli pierwiastek równania charakterystycznego

⁵⁰ A. Stanisławski, *Przystępny kurs ...*, prac. cyt., s 62

Do badania istotności funkcji dyskryminacyjnej używana jest statystyka Lambda Wilksa.⁵¹

$$\lambda = \frac{\det(\sum_{ij} \frac{(y_{ij} - \bar{y})^2}{n-k})}{\det(\sum_{i=1}^k \frac{(\bar{y}_i - \bar{y})^2}{k-1} + \sum_{ij} \frac{(y_{ij} - \bar{y})^2}{n-k})}, \quad (5)$$

$H_0: \lambda = 0$ (a słownie co znaczy ta hipoteza?)

$H_1: \lambda < 1$

gdzie: $i=1, \dots, k$ oraz co to k ?

n -ilość przypadków

y_{ij} -ta obserwacja w i -tej grupie

\bar{y}_i -wartość średnia dla i -tej grupy

y z kreska

Wartość współczynnika λ zawiera się w zakresie od 1 do 0. Wartość niższa świadczy o lepszej mocy dyskryminacyjnej funkcji. Jeżeli jej wartość jest bliska jedności to możemy zakładać, że nie mamy podstaw do klasyfikowania obiektów według wcześniej wyprowadzonych funkcji dyskryminacyjnych. Aproksymację statystyki Lambdy Wilksa stanowi statystyki Chi-kwadrat o liczbie stopni swobody równej ilości zmiennych objaśniających.

Lambda częściowa jest to wartość wynikająca ze stosunku wartości współczynnika lambda po dodaniu odpowiedniej zmiennej do wartości lambda przed dodaniem odpowiedniej zmiennej.⁵²

$$partial \lambda = \frac{\lambda_{przed\ dodaniem\ zmiennej}}{\lambda_{po\ dodaniu\ zmiennej}}, \quad (6)$$

⁵¹ R. Jagiełło, *Analiza dyskryminacyjna i regresja logistyczna w procesie oceny zdolności kredytowej przedsiębiorstw*, nbp.pl, Warszawa 2013, http://www.nbp.pl/publikacje/materialy_i_studia/ms286.pdf, s22

⁵² Statsoft, *Discriminant Function Analysis Notes and Technical Information – Formulas*, documentation 2012, <http://documentation.statsoft.com/STATISTICAHelp.aspx?path=Discriminant/DiscriminantAnalysis/DiscriminantFunctionAnalysisNotesandTechnicalInformationFormulas>

W przypadku gdyby zmienność międzygrupowa dla grup zmiennej objaśniającej była bliska zmienności wewnątrzgrupowej odseparowanie grup od siebie mogłoby być niemożliwe. Do badania tej zależności służy test F Fishera:⁵³

$$F = \frac{\text{zmienność międzygrupowa}}{\text{zmienność wewnątrzgrupowa}} = \frac{\sum_i^k \frac{(y_i - \bar{y})^2}{k-1}}{\sum_{ij}^k \frac{(y_{ij} - \bar{y})^2}{n-k}} \quad (7)$$

H₀: (symbolami jeśli się da) zmienność międzygrupowa nie różni się istotnie statystycznie od zmienności wewnątrzgrupowej

H₁: zmienność międzygrupowa różni się istotnie statystycznie od zmienności wewnątrzgrupowej.

gdzie: i=1,...,k oraz

n-ilość przypadków

y_{ij}-j-ta obserwacja w i-tej grupie

\bar{y}_i -wartość średnia dla i-tej grupy (chyba kolejność odwrotna)

Tak jak test w przypadku statystyki Lambda Wilksa, dla każdej próby możliwe jest skonstruowanie statystyki o rozkładzie chi-kwadrat. Natomiast przypadku sprawdzania istotności pojedynczych zmiennych stosujemy wzór.⁵⁴

$$F = \frac{n-g-p}{g-1} * \frac{1-\text{partial}\lambda}{\text{partial}\lambda}, \quad (8)$$

gdzie:

n-liczba przypadków

g-liczba grup

⁵³ R. Jagiełło, *Analiza dyskryminacyjna i regresja logistyczna w procesie oceny zdolności kredytowej przedsiębiorstw*, nbp.pl, Warszawa 2013, http://www.nbp.pl/publikacje/materialy_i_studia/ms286.pdf, s19

⁵⁴ Statsoft, *Discriminant Function Analysis Notes and Technical Information – Formulas*, documentation 2012, <http://documentation.statsoft.com/STATISTICAHelp.aspx?path=Discriminant/DiscriminantAnalysis/DiscriminantFunctionAnalysisNotesandTechnicalInformationFormulas>

p-liczba zmiennych objaśniających

3.3 Funkcja klasyfikacyjna.

Istnieje kilka metod klasyfikacji obiektów do wcześniej utworzonych grup. Jednym z najpopularniejszych podejść jest to przedstawione przez Fishera w którym zaproponował, że klasyfikacja powinna bazować na liniowej kombinacji zmiennych objaśniających. Zgodnie z tym klasyfikacja do każdej z grup powinna być wyliczana ze wzoru:

$$K_i = c_{i0} + c_{i1}x_1 + c_{i2}x_2 + \dots + c_{ip}x_p,^{55} \quad (5)$$

$j = 0, \dots, p; i = 0, \dots, n$

gdzie:

n – liczba funkcji klasyfikacyjnych,

c_{ij} – specjalne współczynniki zmiennych objaśniających,

K_i – wartość i-tej funkcji klasyfikacyjnej.

Aby prawidłowo sklasyfikować obiekt, należy obliczyć wartość wszystkich funkcji. Funkcja o najwyższej wartości będzie wyznaczała grupę, do której z największym prawdopodobieństwem należy dany obiekt.

Trafność klasyfikacji możemy badać za pomocą tablica trafności przedstawia stosunki wartości teoretycznych do wartości empirycznych funkcji dyskryminacyjnej. Pozwala ona na zbadanie ilości poprawnie zklasyfikowanych przypadków dla wszystkich wyróżnionych grup, a także:

- a) R zliczeniowe, przedstawiające liczbę poprawnie zklasyfikowanych przypadków w stosunku do liczby przypadków ogółem, liczone wzorem:

$$\text{Zliczeniowe } R^2 = \frac{\text{liczba poprawnie zklasyfikowanych przypadków}}{\text{liczba przypadków ogółem}}$$

⁵⁵ A. Stanisław, *Przystępny kurs ... , prac. cyt., s 64*

- b) iloraz szans, przedstawiający stosunek poprawnie zklasyfikowanych przypadków do liczby błędnie zklasyfikowanych przypadków. W przypadku kiedy wartość tej miary przekroczy 1, umiejętność klasyfikacyjna modelu jest lepsza od losowej. Dany jest on wzorem:

$$IS = \frac{\text{liczba poprawnie zklasyfikowanych przypadków}}{\text{liczba błędnie zklasyfikowanych przypadków}}$$

W przypadku próby niebilansowej, dokładność obliczonego prawdopodobieństwa błędnej klasyfikacji można skorygować wzorem **podany** niżej, uwzględniając prawdopodobieństwo a priori.⁵⁶

$$\theta = \sum_{i=1}^g p_i \sum_{k=1, i \neq k}^g P(k|i). \quad (6)$$

Które wykorzystuje Twierdzenie Bayesa dane wzorem:

$$P(G_k|X) = \frac{p_k * P(X|G_k)}{\sum_{i=1}^g p_i P(k|i)}, \quad (7)$$

gdzie:

g – liczba grup,

p_i – prawdopodobieństwo a priori,

$P(G_k|X)$ – prawdopodobieństwo, że dany obiekt należy do k -tej grupy pod warunkiem, że znamy zmienne objaśniające,

$P(X|G_k)$ – prawdopodobieństwo, otrzymania danych zmiennych objaśniających pod warunkiem, że wiemy iż należy do k -tej grupy.

⁵⁶ A. Stanisławski, *Przystępny kurs ... , prac. cyt., s 67*

3.4 Założenia analizy dyskryminacyjnej.

Słowno ewstępu vo tu będzxie

Na samym początku należy wspomnieć o rozmiarze badanej próby. Otóż, w zależności od literatury mówi się tutaj o co najmniej około pięć razy większej liczbie przypadków niż liczba zmiennych dyskryminujących. Ponadto, najmniej liczna grupa powinna być nie mniejsza niż liczba zmiennych poddanych analizie (wyłączając zmienną objaśnianą) minus dwa.⁵⁷

Istnieje również założenie techniczne dotyczące odwracalności macierzy wariancji i kowariancji. Aby macierze te były odwracalne, zmienne objaśniające muszą być liniowo niezależne. Oznacza to, że żadna ze zmiennych po przemnożeniu nie może mieć takich samych wartości jak inna zmienna oraz nie może być liniową kombinacją innych zmiennych.⁵⁸

Analiza dyskryminacyjna zakłada wielowymiarową normalność zmiennych objaśniających.⁵⁹ Oznacza to, że wartości zmiennych opisujących zmienną objaśnianą powinny tworzyć wielowymiarowy rozkład normalny, czyli że prawdopodobieństwo występowania wartości zmiennych powinno tworzyć krzywą Gaussa. Koniecznym warunkiem do istnienia wielowymiarowej normalności jest podleganie rozkładowi normalnemu każdej zmiennej z osobna. Założenie to można w praktyce uchylić ze względu zarówno na dużą wielkość próby oraz pozostałe testy badające istotność funkcji dyskryminacyjnych.⁶⁰

Następnym założeniem, jest jednorodność macierzy wariancji i kowariancji dla poszczególnych klas. W przypadku prób o małej liczności jest to bardzo ważne **założenie i może** to rzutować na niewłaściwy podział grup przez funkcje dyskryminacyjne. Równość macierzy wariancji i kowariancji może być weryfikowana przez statystykę M-boxa, która ma postać ma postać:⁶¹

$$M = \sum_{i=1}^k n_i \ln|S| - \sum_{i=1}^k n_i \ln |S_i|, \quad (8)$$

⁵⁷ A. Stanisławski, *Przystępny kurs ...*, prac. cyt., s 109

⁵⁸ A. Stanisławski, *Przystępny kurs ...*, prac. cyt., s 109

⁵⁹ E. Gatnar, *Symboliczne metody klasyfikacji danych*, Wydawnictwo naukowe PWN, Warszawa 1998, s 54

⁶⁰ A. Stanisławski, *Przystępny kurs ...*, prac. cyt., s 108

⁶¹ E. Gatnar, *Symboliczne metody klasyfikacji danych*, Wydawnictwo naukowe PWN, Warszawa 1998, s 55

Dla estymatora macierzy wariancji i kowariancji liczonego jako: ⁶²

$$S = \frac{1}{\sum_{i=1}^k n_i} \sum_{i=1}^k n_i S_i, \quad (9)$$

$i = 0, \dots, k$

gdzie:

n_i – liczba przypadków dla i -tej zmiennej

S_i – estymator i -tej macierzy wariancji i kowariancji

Statystyka M-boxa jest bardzo wrażliwa na wartości odstające, ponieważ silnie oddziałują na średnią. Dlatego też dodatkowym założeniem jakie się podaje, często jest brak wartości odstających dla zmiennych dyskryminujących. ⁶³

Warto również zwrócić uwagę na korelację średnich oraz wariancji wewnątrzgrupowych. ⁶⁴ Wysoka korelacją tych dwóch miar może wystąpić w przypadku występowania wspomnianych już wartości odstających. Ze względu na to, że testy istotności biorą pod uwagę łączną wariancję, dlatego też badanie istotności przy wysokich średnich dla poszczególnych zmiennych i niskich wariancji mogą być mylące.

Rozdział IV

Analiza dyskryminacyjna posiadania konta w banku w Polsce

W tym rozdziale zostanie przeprowadzona analiza dyskryminacyjna posiadania konta w banku dla danych ankietowych z 2011r z „Diagnozy Społecznej” (czsem jest DS. czasem Ds.) przeprowadzonej przez Radę Monitoringu Społecznego. Na samym początku zostanie przeprowadzona estymacja funkcji dyskryminacyjnych, na podstawie, których w dalszym

⁶² A. Stanisławski, *Przystępny kurs ... , prac. cyt.*, s. 109

⁶³ A. Stanisławski, *Przystępny kurs ... , prac. cyt.*, s. 111

⁶⁴ E. Gatnar, *Symboliczne metody klasyfikacji danych*, Wydawnictwo naukowe PWN, Warszawa 1998, s. 54

etapie zostaną skonstruowane funkcje klasyfikacyjne. Na końcu zostanie przeprowadzona prognoza w trzech wariantach.

4.1 Przedstawienie zmiennych.

Celem analizy jest wytworzenie grup o charakterystycznych cechach dla osób posiadających oraz nieposiadających konta w banku w celu późniejszej klasyfikacji obiektów do utworzonych grup. Próba badawcza przed usunięciem braków danych przypadkami wynosiła 65373. Po usunięciu braków danych przypadkami oraz nieistotnego pod względem liczności wariantu zmiennej objaśnianej, próba została zredukowana do 24453 przypadków. Zmienną objaśniającą jest zmienna „Konto w banku” natomiast zmiennymi objaśniającymi są: „wiek”, „płeć”, „wielkość miejsca zamieszkania”, „edukacja”, „bezrobotny”, „ocena dotychczasowego życia”, „zadowolenie z perspektyw na przyszłość”, „oszczędności gospodarstwa domowego”, „sytuacja finansowa”, „zaufanie do banków komercyjnych”, „brak zaufania do banków komercyjnych”, „zaufanie do rządu” oraz „brak zaufania do rządu”. W pierwszej kolejności zostaną przedstawione rozkłady zmiennych wziętych do badania. Postać tabelaryczna pozwala na łatwe przeanalizowanie zmienności

w cechach oraz deklaracjach ankietowanych.

4.1.1 Tabele licznosci.

W poniższym rozdziale znajdują się tabele licznosci dla wszystkich zmiennych. Tabele licznosci pozwalają na dogodną analizę rozkładu wartości zmiennych biorących udział w analizie dyskryminacyjnej.

a) Posiadanie konta w banku.

Zmienna objaśniana „Konto w banku” przedstawia odpowiedzi ankietowanych na pytanie dotyczące posiadania konta w banku. Możliwymi odpowiedziami na to pytanie było: „Tak”, „Nie” lub „Nie wiem”, jednakże przypadki zawierające wartość „Nie wiem” dla tej zmiennej zostały usunięte z analizy ze względu na małą licznosc wynoszącą 35 przypadków. Rozkład zmiennej objaśnianej prezentuje tabela 2.

Tabela 2. Tabela licznosci: Konto w banku.

	Tabela licznosci: "Konto w banku"
--	-----------------------------------

Kategoria	Liczebność	Liczebność skumulowana	Procent
Tak	17496	17496	71,55%
Nie	6957	24453	28,45%

Opracowanie własne przy pomocy programu „Statistica 10 EN”.

Na podstawie tabeli można stwierdzić, że większość osób poddanych badaniu, tj. 71,55%, posiadało konto w banku. Należy jednak zauważyć, że osoby nieposiadające konta bankowego stanowią niemałą grupę liczącą 6957 osób spośród 24453. Ciekawym jest również fakt, że 35 osób z badanych nie wie, czy posiada konto bankowe. Ze względu na znikomą licznosc odpowiedzi „Nie wiem” przez ankietowanych, zmienna ta została usunięta z dalszej analizy. Przez to, próba badawcza zmniejszyła się z **xxxx do xxxx**.

Zostaną teraz omówione zmienne opisujące posiadanie konta w banku.

b) Wiek.

Pierwszą zmienną objaśniającą jest „wiek” osoby badanej. W celu lepszego zobrazowania zostały wydzielone widoczne w tabeli 3 przedziały grupowe.

Tabela 3. *Tabela licznosci: Wiek.*

Tabela licznosci: "Wiek"				
Kategoria	Liczebność	Liczebność skumulowana	Procent	Procent skumulowany
18-24 lat	2992	2992	12,2%	12,2%
25-34 lat	3714	6706	15,2%	27,4%
35-44 lat	3689	10395	15,1%	42,5%
45-55 lat	4480	14875	18,3%	60,8%
55-64 lat	4861	19736	19,9%	80,7%
64 lat i więcej	4717	24453	19,3%	100,0%

Opracowanie własne przy pomocy programu „Statistica 10 EN”.

Można zauważyć, że wyznaczone przedziały grupowe są bardzo podobne licznosciowo. Największy z nich stanowią osoby w wieku 55 do 64 lat, należy zatem zauważyć, że dyskryminacja w największym stopniu będzie się opierała na opinii osób z tej grupy wiekowej. Natomiast najmniejszym jest przedział osób w wieku od 18 do 24 lat. Do analizy zostały wybrane jedynie osoby, które ukończyły 18 rok życia, ze względu na to, że osoby młodsze nie są odbiorcami usług bankowych.

c) Płeć.

Następną zmienną objaśniającą jest „płeć” respondenta, której rozkład przedstawia poniższa tabela.

Tabela 4. *Tabela licznosci: Płeć.*

Tabela licznosci: "Płeć"				
Kategoria	Liczebność	Liczebność skumulowana	Procent	Procent skumulowany
Kobieta	13597	13597	55,60%	55,60%
Mężczyzna	10856	24453	44,40%	100,00%

Opracowanie własne przy pomocy programu „Statistica 10 EN”.

Na podstawie tej tabeli widać różnice w liczbie mężczyzn i kobiet biorących udział w badaniu. Otóż mężczyzn, którzy odpowiedzieli na wszystkie interesujące ze względu tej analizy pytania, było o 25,25% mniej niż kobiet. Płeć” kobieta” kodowana jest jako 0, natomiast „mężczyzna” jako 1.

d) Wielkość miejsca zamieszkania.

Zmienna dotycząca wielkości miejsca zamieszkania grupuje wielkość miejscowości zamieszkałej przez ankietowanego według liczby mieszkańców. Została ona podzielona na wsie, miasta poniżej 20 tys., miasta od 20 do 100 tys., miasta od 100 do 200 tys., miasta od 200 do 500 tys., oraz miasta powyżej 500 tys. mieszkańców. Rozkład tej zmiennej prezentuje poniższa tabela.

Tabela 5. *Tabela licznosci: Wielkość miejsca zamieszkania..*

Tabela licznosci: "Wielkość miejsca zamieszkania"				
Kategoria	Liczebność	Liczebność skumulowana	Procent	Procent skumulowany
Wieś	11283	11283	46,14%	46,14%
Miasta < 20 tys.	3172	14455	12,97%	59,11%
Miasta 20-100 tys.	4471	18926	18,28%	77,40%
Miasta 100-200 tys.	1541	20467	6,30%	83,70%
Miasta 200-500 tys.	2114	22581	8,65%	92,34%
Miasta > 500 tys.	1872	24453	7,66%	100,00%

Opracowanie własne przy pomocy programu „Statistica 10 EN”.

53,86% ankietowanych pochodziło z miasta, a pozostałe 46,14% ze wsi. Biorąc pod uwagę osoby z miasta, najwięcej z nich należało do grupy osób zamieszkujących miasta o wielkości 20-100 tysięcy mieszkańców (tj. 33,94%), natomiast najmniej do grupy osób zamieszkujących miasta o wielkości od 100 do 200 tysięcy mieszkańców (tj. 11,70%). Zmienna przyjmuje wartości od 0 do 5, gdzie 0 to „wieś”, a 5 „miasta powyżej 500 tys.”.

e) Edukacja.

Zmienna „edukacja” dotyczy wykształcenia badanej osoby w dniu badania. Możliwymi odpowiedziami na to pytanie były: „podstawowe i niższe”, „zasadnicze zawodowe lub gimnazjalne”, „średnie” oraz „wyższe i policealne” Tabela 6. przedstawia poziom edukacji ankietowanych.

Tabela 6. *Tabela licznosci: Edukacja.*

Kategoria	Tabela licznosci: "Edukacja"			
	Liczebność	Liczebność skumulowana	Procent	Procent skumulowany
Podstawowe i niższe	4516	4516	18,47%	18,47%
Zasadnicze zawodowe lub gimnazjum	7802	12318	31,91%	50,37%
Średnie	7488	19806	30,62%	81,00%
Wyższe i policealne	4647	24453	19,00%	100,00%

Opracowanie własne przy pomocy programu „Statistica 10 EN”.

Najwięcej osób stanowiły osoby z wykształceniem „zasadniczym zawodowym lub gimnazjalnym”, jednakże była to grupa o podobnej liczności co osoby o wykształceniu „średnim” (co łącznie dawało około 62,53%). Również podobnie liczne były dwie pozostałe grupy stanowią osoby o wykształceniu „podstawowym i niższym” oraz „wyższym i policealnym”. Zmienne były kodowane liczbami całkowitymi od 0, rosnąco, zaczynając od wykształcenia „podstawowego i niższego”.

f) Status zawodowy.

Status zawodowy dotyczy statusu na rynku pracy, wyróżniając osoby posiadające status osoby bezrobotnej oraz osoby, które nie są zaliczane do tej kategorii. Kolejna tabela to tabela przedstawia liczbę ankietowanych w porównaniu do osób niebędących osobami bezrobotnymi.

Tabela 7. *Tabela liczności: Status zawodowy.*

Tabela liczności: "Bezrobotny"				
Kategoria	Liczebność	Liczebność skumulowana	Procent	Procent skumulowany
Nie	22778	22778	93,15%	93,15%
Tak	1675	24453	6,85%	100,00%

Opracowanie własne przy pomocy programu „Statistica 10 EN”.

Na podstawie powyższych danych, zdecydowana większość osób to te , które nie były zaliczane do grupy osób bezrobotnych, stanowiących ponad 93,15%, co całkowicie dominowało 6,85% osób bezrobotnych. Zmienna „bezrobotny” przyjmuje wartości 0-1.

g) Ocena dotychczasowego życia.

Kolejną wprowadzoną do analizy zmienną jest „ocena dotychczasowego życia”. Przedstawia ona subiektywną ocenę dotychczasowego życia badanej osoby. Możliwymi odpowiedziami były: „nieszczęśliwe”, „okropne”, „niezbyt udane”, „ani dobre, ani złe”, „dość dobre”, „udane”, „wspaniałe”. Wymienione warianty zostały posegregowane od najbardziej negatywnej do najbardziej pozytywnej.

Tabela 8. *Tabela liczności: Ocena dotychczasowego życia.*

Tabela liczności: "Ocena dotychczasowego życia"				
Kategoria	Liczebność	Liczebność skumulowana	Procent	Procent skumulowany
Nieszczęśliwe	197	197	0,81%	0,81%
Okropne	76	273	0,31%	1,12%
Niezbyt udane	1204	1477	4,92%	6,04%
Ani dobre, ani złe	4054	5531	16,58%	22,62%
Dość dobre	8538	14069	34,92%	57,53%
Udane	9512	23581	38,90%	96,43%
Wspaniałe	872	24453	3,57%	100,00%

Opracowanie własne przy pomocy programu „Statistica 10 EN”.

Można stwierdzić, że większość osób pozytywnie oceniało swoje dotychczasowe życie, tj. około 77,39%. Najmniej odpowiedzi przypadło na stwierdzenie „okropne” (tj. 76), natomiast najwięcej „udane” (tj. 9512). Neutralnie odpowiadało 4054 ankietowanych co stanowi 16,58% z ogółu. Zmienna ta przyjmuje wartości od 0 (nieszczęśliwe) do 6 (wspaniałe).

h) Zadowolenie z perspektyw na przyszłość.

Następną zmienną opisującą posiadanie konta w banku jest „zadowolenie z perspektyw na przyszłość”. Zmienna ta, nawiązuje do wpływu optymistycznego nastawienia do życia na decyzje finansowe opisanego w rozdziale 2.3.3. Posiada ona rozkład odpowiedzi, podobny do zmiennej opisującej ocenę dotychczasowego życia, co widać na poniższej tabeli, która posiada następujące kategorie: „bardzo niezadowolony”, „niezadowolony”, „dosyć niezadowolony”, „dosyć zadowolony”, „zadowolony”, „bardzo zadowolony”.

Tabela 9. Tabela licznosci: Zadowolenie z perspektyw na przyszłość.

Kategoria	Tabela licznosci: "Zadowolenie z perspektyw na przyszłość"			
	Liczebność	Liczebność skumulowana	Procent	Procent skumulowany
Bardzo niezadowolony	1915	1915	7,83%	7,83%
Niezadowolony	3950	5865	16,15%	23,98%
Dosyć niezadowolony	5147	11012	21,05%	45,03%
Dosyć zadowolony	7740	18752	31,65%	76,69%
Zadowolony	4994	23746	20,42%	97,11%
Bardzo zadowolony	707	24453	2,89%	100,00%

Opracowanie własne przy pomocy programu „Statistica 10 EN”.

Podobnie jak poprzednio, większość osób jest pozytywnie nastawiona w stosunku do swoich perspektyw. Najwięcej osób w liczbie 7740 opowiedziało się za odpowiedzią mówiącą, że są dość zadowoleni ze swoich perspektyw. Warto jednak zauważyć, że było prawie 3 razy więcej osób bardzo niezadowolonych w porównaniu do osób bardzo zadowolonych. Zmienna ta przyjmuje wartości od 0 (bardzo niezadowolony) do 5 (bardzo zadowolony).

i) Posiadanie oszczędności.

Kolejna zmienna „oszczędności gospodarstwa domowego” odpowiada na pytanie dotyczące spełniania przez gospodarstwo domowe drugiego stopnia hierarchii zachowań finansowych opisanych w rozdziale 2.2., czyli „Czy gospodarstwo domowe posiada jakieś oszczędności?”. Tym razem były możliwe dwie odpowiedzi: „Tak” oraz „Nie”.

Tabela 12. *Tabela liczności: oszczędności gospodarstwa domowego.*

Tabela liczności: "oszczędności gospodarstwa domowego"				
Kategoria	Liczebność	Liczebność skumulowana	Procent	Procent skumulowany
Nie	15506	15506	63,41%	63,41%
Tak	8947	24453	36,59%	100,00%

Opracowanie własne przy pomocy programu „Statistica 10 EN”.

Można zauważyć, że 15506, czyli większość ankietowanych (tj. 63,41%) żyje w gospodarstwach domowych, które nie posiadają oszczędności. Pozostałe 8947 z 24453 osób, mieszka w gospodarstwach domowych posiadających oszczędności. Jest to zmienna zero-jedynkowa.

j) Sytuacja finansowa.

Następna zmienna wprowadzana do modelu opisuje sytuację finansową zadeklarowaną przez respondenta w jego gospodarstwie domowym. Jej kategorie zostały stworzone ze względu na możliwości zaspokajania potrzeb oraz oszczędzania. Kategorie te zostały posortowane w tabeli 13, od najmniej wymagających do najbardziej wymagających.

Tabela 13. *Tabela liczności: Sytuacja finansowa.*

Tabela liczności: "Sytuacja finansowa"				
Kategoria	Liczebność	Liczebność skumulowana	Procent	Procent skumulowany
nie starcza nawet na najtańsze jedzenie	316	316	1,29%	1,29%
starcza na najtańsze, ale nie starcza na ubranie	925	1241	3,78%	5,08%
starcza na najtańsze, ale nie starcza na opłatę za mieszkanie	540	1781	2,21%	7,28%
starcza na najtańsze, ale nie starcza na spłatę kredytu	350	2131	1,43%	8,71%
starcza na najtańsze jedzenie, ubranie, opłaty i kredyt	3073	5204	12,57%	21,28%
żyjemy bardzo oszczędnie, aby odłożyć na poważniejsze zakupy	5752	10956	23,52%	44,80%
żyjemy oszczędnie i dzięki temu starcza na wszystko	9272	20228	37,92%	82,72%
starcza na wszystko, lecz nie oszczędzamy na przyszłość	1976	22204	8,08%	90,80%
starcza na wszystko i jeszcze oszczędzamy na przyszłość	2249	24453	9,20%	100,00%

Opracowanie własne przy pomocy programu „Statistica 10 EN”.

Najwięcej osób tj. 37,92% stanowią respondenci, których gospodarstwo domowe żyje oszczędnie, dzięki czemu starcza na wszystko. Ankietowanych, którzy żyją w gospodarstwach domowych którym nie starcza na jedzenie, ubranie albo opłatę za mieszkanie jest około 8,71%. Zmienna jakościowa została zakodowana od 0 („nie starcza nawet na najtańsze jedzenie”) do 8 („starcza na wszystko i jeszcze oszczędzamy na przyszłość”).

k) **Zaufanie do banków komercyjnych.**

Kolejna zmienna dotyczy zaufania ankietowany do banków komercyjnych. W analizie została ona podzielona na dwie zmienne: „zaufanie do banków komercyjnych” oraz „brak zaufania do banków komercyjnych” ze względu na występowania odpowiedzi „Nie wiem”. Dzięki poniższej tabeli x i y możemy przyjrzeć się rozkładowi odpowiedzi. Ankietowani mieli do wyboru trzy odpowiedzi: „Tak”, „Nie” oraz „Nie mam zdania”.

Tabela 10. Tabela liczości: Zaufanie do banków komercyjnych.

	Tabela liczości: "Zaufanie do banków komercyjnych"			
Kategoria	Liczebność	Liczebność skumulowana	Procent	Procent skumulowany
Tak	19806	19806	81,00%	81,00%
Nie	4647	24453	19,00%	100,00%

Opracowanie własne przy pomocy programu „Statistica 10 EN”.

Tabela

	Tabela liczości: "Brak zaufania do banków komercyjnych"			
Kategoria	Liczebność	Liczebność skumulowana	Procent	Procent skumulowany
Tak	17036	17036	69,67%	69,67%
Nie	7417	24453	30,33%	100,00%

Bardzo dużo osób, bo aż 12414 z 24453 zadeklarowało, iż nie ma zdania na ten temat. Nieco ponad 30 % z pozostałych osób zadeklarowało brak zaufania do banków komercyjnych, a jedynie 19,00%, że posiada zaufanie do wspomnianych już banków. Zmienna ta została zakodowana jako: 1 – „Tak, oraz 0– „Nie”.

l) **Zaufanie do rządu.**

Ostatnia zmienna odpowiada na pytanie czy ankietowany posiada zaufanie do rządu, a jej rozkład przedstawia następna tabela. Tak jak w przypadku poprzedniego pytania

możliwy był wybór jednego spośród trzech opcji. Zaufanie do rządu w przeciwieństwie do zaufania do banków komercyjnych wpływa na ogólne poczucie bezpieczeństwa finansowego, będącego wynikiem decyzji podejmowanych przez rząd.

Tabela 11. *Tabela licznosci: Zaufanie do rządu.*

Tabela licznosci: "Zaufanie do rządu"				
Kategoria	Liczebność	Liczebność skumulowana	Procent	Procent skumulowany
Nie	10862	10762	44,01%	44,01%
Tak	13691	24453	55,99%	100,00%

Opracowanie własne przy pomocy programu „Statistica 10 EN”.

Tabela licznosci: "Brak zaufania do rządu"				
Kategoria	Liczebność	Liczebność skumulowana	Procent	Procent skumulowany
Nie	20379	20379	83,34%	83,934%
Tak	4078	24453	16,65%	100,00%

W tym przypadku mniejszość, czyli 4078 co stanowi 16,65% ogółu oświadczyła, iż nie ma zaufania do rządu. Osób darzących rząd zaufaniem było ponad połowa, czyli 55,99%, a osób niemających zdania na ten temat 27,36%. Zmienna została zakodowana w taki sam sposób jak poprzednia.

4.1.2 Testy Chi-kwadrat.

Test zgodności Pearsona jest testem zgodności weryfikującym hipotezę mówiącej o istnieniu korelacji między zmiennymi.⁶⁵

Ze względu na założenie o istnieniu co najmniej 5 obserwacji w każdej rozłącznej klasie, zmienna „wiek” poddana została grupowaniu według schematu przedstawionego w rozdziale 4.4.1.

Statystyka Chi-kwadrat ma postać:

⁶⁵ M. Sobczyk, *Statystyka matematyczna*, C.H. Beck, Warszawa 2010, s. 139

$$\chi^2 = \sum_{i=1}^r \frac{(n_i - np_i)^2}{np_i}$$

n_i -empiryczne liczebności poszczególnych wariantów cechy

np_i -wyliczone liczebności teoretyczne

H_0 : nie ma zależności pomiędzy posiadaniem konta w banku, a daną zmienną objaśniającą.

H_1 : istnieje zależność pomiędzy posiadaniem konta w banku, a daną zmienną objaśniającą.

Tabela 14. Testy Chi-kwadrat dla poszczególnych zmiennych.

	"Konto w banku" * wszystkie zmienne		
Statystyka	Chi-kwadrat	stopnie swobody	p-value
Wiek	3245,239	10	0
Płeć	11,27663	2	0
Wielkość miejsca zamieszkania	868,1349	10	0
Edukacja	3722,083	6	0
Bezrobotny	225,3386	2	0
Ocena dotychczasowego życia	483,4049	12	0
Zadowolenie z perspektyw na przyszłość	277,5060	10	0
Zaufanie do banków komercyjnych	761,7095	4	0
Zaufanie do rządu	62,89333	4	0
Czy gospodarstwo domowe posiada jakieś oszczędności	759,5034	2	0
Sytuacja finansowa	1300,041	16	0

Opracowanie własne przy pomocy programu „Statistica 10 EN”.

Dla wszystkich zmiennych jesteśmy w stanie stwierdzić, że na podstawie p-value mniejszego niż 0,05 odrzucamy hipotezę zerową przemawiającą za brakiem zależności między posiadaniem konta w banku a daną zmienną objaśniającą, na rzecz hipotezy alternatywnej mówiącej, że dla każdej z wprowadzanych do modelu zmiennych istnieje zależność ze zmienną objaśniającą, determinującą posiadanie konta w banku.

4.2 Omówienie próby badawczej poddawanej analizie.

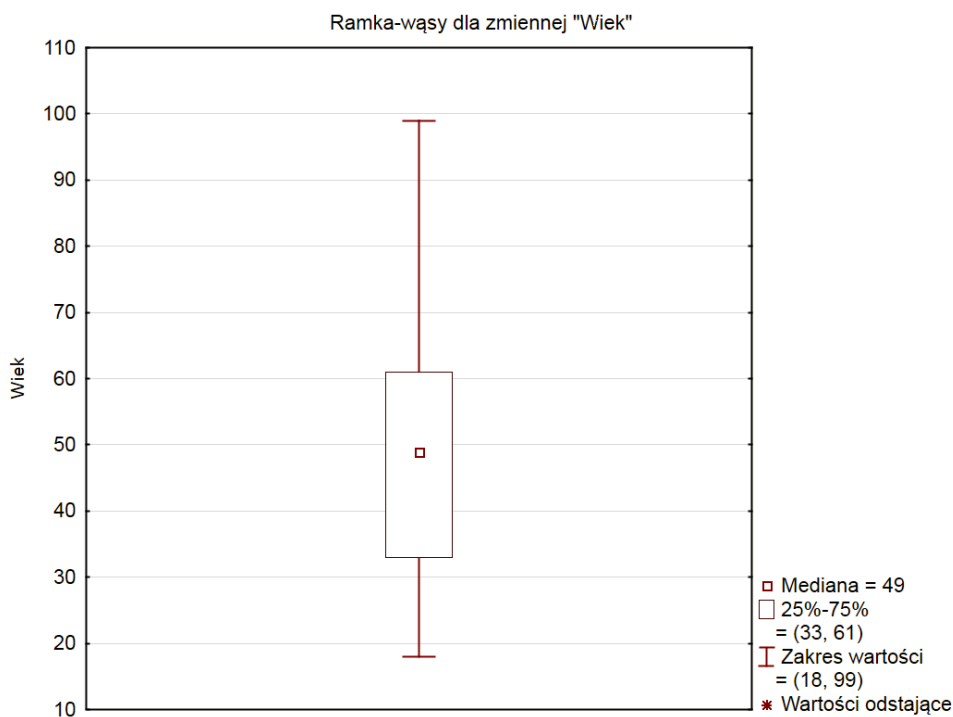
Przed wykonaniem analizy dyskryminacyjnej metodą liniową, należy omówić dobraną próbę badawczą oraz przyjrzeć się bliżej omówionym w rozdziale 3.4 założeniom, a w przypadku jej braku spełnienia któregoś z nich, podanie odpowiedniej przyczyny oraz

wyjaśnienia. W następnym etapie zostaną wydzielone grupy na podstawie wyprowadzonych funkcji dyskryminacyjnych, po których zostaną przedstawione funkcje klasyfikacyjne wraz z podsumowującą je prognozą.

a) Charakter badanej próby

Pozyskana z „Diagnozy Społecznej” próba została poddana obróbce. W przypadku występowania braków danych zostawały one usuwane przypadkami. W rezultacie analizowani byli respondenci, którzy odpowiedzieli na wszystkie istotne z punktu widzenia przeprowadzanej analizy pytania. W wyniku redukcji liczby przypadków w próbie, liczyła ona przed rozpoczęciem badania 24809 przypadków. Wynika z tego również spełnienie założenia o liczebności próby, która jest ponad 2000 większa od liczby zmiennych, których liczba wynosi 11 (później 10), co tym samym zapewnia spełnienie założenia o liczebności próby. Na podstawie zmiennych (z wyjątkiem zmiennej „wiek”) przedstawionych w rozdziale 4.1.1 możemy stwierdzić, iż nie posiadają one wartości odstających. Rozkład wartości zmiennych „wiek” prezentuje się następująco.

Wykres 1. *Ramka-wąsy dla zmiennej Wiek.*



Opracowanie własne przy pomocy programu „Statistica 10 EN”.

Z wykresu ramka-wąsy, możemy wywnioskować, że wszystkie możliwe wartości są reprezentowane przez co najmniej jedną osobę. Tym samym nie ma wartości odstających. Największa liczba osób spośród ankietowanych posiada 49 lat, a ich rozkład jest prawostronnie asymetryczny co oznacza, że większość osób to osoby o wieku niższym niż wynosiła średnia wieku respondentów.

Podsumowując, zmienne wzięte do badania dyskryminacji posiadania konta w banku, nie posiadają wartości odstających, dzięki czemu spełnione jest założenie o niewystępowaniu takich wartości.

Należało również przyjrzeć się warunkowi o podleganiu zmiennych wielowymiarowemu rozkładowi normalnemu. Statistica pozwala nam jedynie zbadać normalność rozkładu pojedynczych zmiennych, jednakże jak zostało to wcześniej wspomniane, warunkiem koniecznym do istnienia wielowymiarowego rozkładu normalnego jest istnienie jednowymiarowego rozkładu normalnego każdej zmiennej z

osobna. Większość zmiennych wziętych do analizy nie spełnia jednak tego założenia ze względu na wąskie zakresy wartości oraz zróżnicowane wartości średnich. W przypadku dysponowania próbą o liczbie 24453 przypadków można to założenie uchylić co nie będzie wpływało na jakość analizy.

Kolejnym założeniem wymagającym sprawdzenia było założenie o równości macierzy wariancji i kowariancji. Do sprawdzenia tego założenia został wykorzystany test M-boxa sprawdzający jednorodność macierzy wariancji i kowariancji dla wszystkich zmiennych zależnych. ODWOŁAĆ SIĘ DO METODOLOGII wzór 8 .

~~TAKE~~
 H_0 : Macierze wariancji i kowariancji dla poszczególnych klas są sobie równe.

H_1 : Co najmniej jedna macierz wariancji i kowariancji klasy jest inna niż pozostałe.

$$M = \sum_{i=1}^k n_i \ln|S| - \sum_{i=1}^k n_i \ln|S_i| = 5309,237$$

Tabela 15. Test M-Boxa.

	Test M-boxa dla zmiennej "Konto w banku"			
Statystyka	M-box	Chi-kwadrat	df	p-value
Test M-boxa	5309,237	5010,494	132	0,00

Opracowanie własne przy pomocy programu „Statistica 10 EN”.

Ze względu na to, że prawdopodobieństwo testowe jest mniejsze niż 0,05, odrzucona zostaje hipoteza zerowa na rzecz hipotezy alternatywnej mówiącej o tym, że co najmniej jedna macierz wariancji i kowariancji klasy jest inna niż pozostałe.

Założenie to nie zostało spełnione, ponieważ test ten jest bardzo wrażliwy na różnice w średnich dla poszczególnych zmiennych jak i na wartości odstające. Ze względu na duży rozmiar próby oraz wykorzystanie innych testów badających istotność w dalszej części pracy, założenie dotyczące powyższego testu może zostać uchylone.

4.3 Dyskryminacja osób ze względu na posiadanie konta w banku.

W tym rozdziale przedstawiona zostanie analiza dyskryminacyjna dotycząca wyróżnienie cech osób nieposiadających konta bankowego. Obejmuje ona estymację funkcji

dyskryminacyjnych, funkcji klasyfikacyjnych oraz przedstawienie przykładowej prognozy na podstawie wyznaczonych funkcji.

4.3.1 Sprawdzanie istotności zmiennych.

Pierwszym etapem analizy jest dobór zmiennych. Na samym początku analizie zostały poddane wszystkie zmienne. Dzięki temu możliwe było odczytanie łącznej siły dyskryminacyjnej oraz przejrzenie istotności poszczególnych zmiennych w celu ich ewentualnego usunięcia. Poniższa tabela przedstawia mierniki łącznej istotności dla wszystkich zmiennych.

Tabela 16. Mierniki łącznej istotności I.

	Istotność zmiennych w modelu (liczba zmiennych: 13)		
Statystyka	Lambda Wilksa	Test F	p-value
Lambda Wilksa	0,8157702	F (13,24439) = 424,5532	0,00

Opracowanie własne przy pomocy programu „Statistica 10 EN”.

Łącznie w modelu było 13 zmiennych objaśniających, dla których Lambda Wilksa wynosiła 0,8157702, która została obliczona zgodnie ze wzorem (5). Oznacza to, że wybrane zmienne bardzo słabo dyskryminują zmienną objaśnianą, jednocześnie statystyka F, obliczona ze wzoru (7) wskazywała na to, że dyskryminacja posiadania konta w banku jest istotna.

Statystyki dla poszczególnych zmiennych prezentowały się następująco:

Tabela 17. Analiza funkcji dyskryminacyjnych I.

	Podsumowanie analizy funkcji dyskryminacyjnych					
Liczba przypadków = 24453	Lambda Wilksa	Wartość własna	Wartość graniczna F	p-value	Współczynnik tolerancji	1-współczynnik tolerancji
„wiek”	0,815833	0,999923	1,882	0,170116	0,822262	0,177738
"płeć"	0,816524	0,999077	22,579	0,000002	0,974435	0,025565
"wielkość miejsca zamieszkania"	0,818398	0,996790	78,713	0,000000	0,907798	0,092202
"edukacja"	0,885152	0,921616	2078,540	0,000000	0,772173	0,227827
"bezrobotny"	0,820015	0,994824	127,165	0,000000	0,936751	0,063249
"ocena dotychczasowego życia"	0,816176	0,999503	12,149	0,000492	0,837855	0,162145
"zadowolenie z perspektyw na przyszłość"	0,815779	0,999989	0,271	0,602705	0,817797	0,182203
"oszczędności gospodarstwa domowego"	0,817241	0,998201	44,056	0,000000	0,836263	0,163737

"sytuacja finansowa"	0,820711	0,993980	148,019	0,000000	0,760324	0,239676
"zaufanie do banków komercyjnych"	0,824070	0,989929	248,640	0,000000	0,819543	0,180457
"brak zaufania do banków komercyjnych"	0,819354	0,995626	107,363	0,000000	0,823781	0,176219
"zaufanie do rządu"	0,816016	0,999699	7,353	0,006699	0,680237	0,319763
"brak zaufania do rządu"	0,815992	0,999728	6,657	0,009882	0,687783	0,312217

Opracowanie własne przy pomocy programu „Statistica 10 EN”.

Należy zauważyć, iż p-value dla statystyki F jest największe dla zmiennej „zadowolenie z perspektyw na przyszłość” oraz że wynosi ono aż 0,602705. Oznacza to, że zmienna ta nie jest istotna w tym modelu. Istotność nie została również spełniona w przypadku zmiennej „wiek”, dla którego wartość p-value wynosi 0,170116.

Na tym etapie, możemy również zauważyć, że żadna ze zmiennych nie jest współliniowa z inną. Świadczy o tym wysoki współczynnik tolerancji dla wszystkich zmiennych, który w przypadku „zadowolenia z perspektyw na przyszłość” wynosi 0,817797. W tym przypadku należało usunąć zmienną „zadowolenie z perspektyw na przyszłość” oraz na ponownie zbadać łączną istotność. Testy łącznej istotności dla modelu bez zmiennej „zadowolenie z perspektyw na przyszłość” prezentują się następująco:

Tabela 18. *Mierniki łącznej istotności II.*

Statystyka	Istotność zmiennych w modelu (liczba zmiennych: 12)		
	Lambda Wilksa	Test F	p-value
Lambda Wilksa	0,8157792	F (12,24440) = 459,9239	0,00

Opracowanie własne przy pomocy programu „Statistica 10 EN”.

Można zauważyć, że Lambda Wilksa pozostała na podobnym poziomie (tj. 0,8157792), co mówi nam o tym, że zmienne badane w modelu wciąż są słabo dyskryminujące. Podobnie jest w przypadku statystyki F, która przemawia za istotnością wszystkich zmiennych łącznie w modelu.

Poniższa tabela prezentuje statystyki po pierwszej redukcji.

Tabela 19. *Analiza funkcji dyskryminacyjnych II.*

	Podsumowanie analizy funkcji dyskryminacyjnych
--	--

Liczba przypadków = 24453	Lambda Wilksa	Wartość własna	Wartość graniczna F	p-value	Współczynnik tolerancji	1-współczynnik tolerancji
"wiek"	0,815837	0,999929	1,728	0,188718	0,837248	0,162752
"płeć"	0,816531	0,999079	22,521	0,000002	0,974593	0,025407
"wielkość miejsca zamieszkania"	0,818418	0,996776	79,040	0,000000	0,908515	0,091485
"edukacja"	0,885199	0,921577	2079,758	0,000000	0,772309	0,227691
"bezrobotny"	0,820016	0,994834	126,926	0,000000	0,940270	0,059730
"ocena dotychczasowego życia"	0,816180	0,999509	12,006	0,000531	0,892911	0,107089
"oszczędności gospodarstwa domowego"	0,817243	0,998210	43,839	0,000000	0,837838	0,162163
"sytuacja finansowa"	0,820778	0,993909	149,767	0,000000	0,779711	0,220289
"zaufanie do banków komercyjnych"	0,824089	0,989917	248,951	0,000000	0,824933	0,175067
"brak zaufania do banków komercyjnych"	0,819361	0,995628	107,318	0,000000	0,823801	0,176199
"zaufanie do rządu"	0,816034	0,999687	7,642	0,005707	0,684965	0,315035
"brak zaufania do rządu"	0,816006	0,999722	6,785	0,009196	0,689060	0,310940

Opracowanie własne przy pomocy programu „Statistica 10 EN”.

Na podstawie tabeli można stwierdzić, że po usunięciu zmiennej „zadowolenie z perspektyw na przyszłość”, p-value dla testu F pozostałych zmiennych uległo zmianie. W przypadku zmiennej „wiek” pogorszyło się o prawie 0,02, przez co, nie ma podstaw do wnioskowania o istotności tej zmiennej w rozważanym modelu, zatem również tą zmienną należy usunąć.

Następna tabela przedstawia skutek usunięcia zmiennej „wiek” z badanego modelu.

Tabela 18. Mierniki łącznej istotności II.

	Istotność zmiennych w modelu (liczba zmiennych: 11)		
Statystyka	Lambda Wilksa	Test F	p-value
Lambda Wilksa	0,8158368	F (11,2441) = 501,5633	0,00

Opracowanie własne przy pomocy programu „Statistica 10 EN”.

Po tej redukcji, zarówno Lambda Wilksa jak i wartość statystyki F dla modelu wciąż pozostaje na podobnym poziomie. Wartości te wynoszą odpowiedni: 0,8158368 w

przypadku Lambdy Wilksa oraz 501,5633 w przypadku testu F dla którego p-value jest mniejsze niż 0,00. Statystyki dla poszczególnych zmiennych przedstawia tabela x.

Nazwa tabeli

	Podsumowanie analizy funkcji dyskryminacyjnych					
	Lambda Wilksa	Wartość własna	Wartość graniczna F	p-value	Współczynnik tolerancji	1-współczynnik tolerancji
Liczba przypadków = 24453						
"płeć"	0,816626	0,999034	23,632	0,000001	0,980387	0,019613
"wielkość miejsca zamieszkania"	0,818419	0,996845	77,350	0,000000	0,923112	0,076889
"edukacja"	0,894573	0,911985	2358,779	0,000000	0,851094	0,148906
"bezrobotny"	0,820030	0,994887	125,617	0,000000	0,969463	0,030537
"ocena dotychczasowego życia"	0,816281	0,999455	13,318	0,000263	0,905968	0,094032
"oszczędności gospodarstwa domowego"	0,817266	0,998252	42,801	0,000000	0,842432	0,157568
"sytuacja finansowa"	0,820802	0,993951	148,745	0,000000	0,780912	0,219088
"zaufanie do banków komercyjnych"	0,824206	0,989846	250,721	0,000000	0,825901	0,174099
"brak zaufania do banków komercyjnych"	0,819387	0,995668	106,349	0,000000	0,825350	0,174651
"zaufanie do rządu"	0,816086	0,999695	7,453	0,006338	0,685462	0,314538
"brak zaufania do rządu"	0,816096	0,999682	7,773	0,005306	0,699666	0,300334

Opracowanie własne przy pomocy programu „Statistica 10 EN”.

4.3.2 Funkcje dyskryminacyjne.

Poniższy podrozdział poświęcony jest analizie funkcji dyskryminacyjnej. W przypadku badanego modelu została utworzona jedna funkcje dyskryminacyjne, rozdziałająca badane obiekty na dwie grupy. Testy istotności dla tej funkcji prezentują się następująco:

Tabela 20. Analiza istotności funkcji dyskryminacyjnych.

Ilość usuniętych pierwiastków	Chi-kwadrat z usuniętymi pierwiastkami					
	Wartość własna	R kanoniczne	Lambda Wilksa	Chi-kwadrat	stopnie swobody	p-value
0	0,225735	0,429142	0,815837	4975,658	11	0,00

Opracowanie własne przy pomocy programu „Statistica 10 EN”.

Dla badanej funkcji wartość p-value jest bliska 0,00 co oznacza, że jest ona statystycznie istotna. Współczynnik korelacji R wynoszący 0,429142 mówi o tym, że istnieje słabe powiązanie pomiędzy zmienną objaśnianą posiadanie konta w banku a funkcjami dyskryminacyjnymi oraz, że dyskryminacja na podstawie prawdopodobieństwa a-priori mogłaby być bardziej trafna.

Poniższa tabela **x** przedstawia zarówno surowe estymowane ze wzoru (2), standaryzowane według wzoru (3) współczynniki funkcji dyskryminacyjnych, wartości własne oraz wariancję wyjaśnioną przez daną funkcję.

Tabela 21. *Współczynniki funkcji dyskryminacyjnych..*

zmienna	Współczynniki funkcji dyskryminacyjnych	
	Surowe	Standaryzowane
"płeć"	-0,147247	-0,073146
"wielkość miejsca zamieszkania"	-0,083789	-0,136226
"edukacja"	-0,812218	-0,749356
"bezrobotny"	0,673042	0,169233
"ocena dotychczasowego życia"	-0,056848	-0,057132
"oszczędności gospodarstwa domowego"	-0,223820	-0,106148
"sytuacja finansowa"	-0,128199	-0,205088
"zaufanie do banków komercyjnych"	-0,666097	-0,258377
"brak zaufania do banków komercyjnych"	-0,367519	-0,168827
"zaufanie do rządu"	-0,099034	-0,049142
"brak zaufania do rządu"	0,133333	0,049674
Stała	2,636404	
Wartość własna	0,225735	0,225735
Wariancja skumulowana	1,000000	1,000000

Opracowanie własne przy pomocy programu „Statistica 10 EN”.

Na podstawie tabeli możliwe jest wyznaczenie funkcji dyskryminacyjnych zarówno dla surowych jak i dla standaryzowanych parametrów. Poniżej zostały przedstawione obie funkcje.

$$\begin{aligned}
 F(x) = & 2,636404 - 0,147247 * \text{płeć} - 0,083789 * \text{wielkość} \dots - 0,812218 * \text{edukacja} + 0,673042 \\
 & * \text{bezrobotny} \dots - 0,056848 * \text{ocena} \dots - 0,223820 * \text{oszczędności} \dots - 0,128199 \\
 & * \text{sytuacja} \dots - 0,666097 * \text{zaufanie do banków} \dots - 0,367519 \\
 & * \text{brak zaufania do banków} \dots - 0,099034 * \text{zaufanie do rządu} + 0,133333 \\
 & * \text{brak zaufania do rządu}
 \end{aligned}$$

$F_{2\text{standaryzowana}}(x)$

$$= 0,073146 * \text{płeć} - 0,136226 * \text{wielkość ...} - 0,749356 * \text{edukacja} + 0,169233 \\ * \text{bezrobotny ...} - 0,057132 * \text{ocena ...} - 0,106148 * \text{oszczędności ...} - 0,205088 \\ * \text{sytuacja ...} - 0,258377 * \text{zaufanie do banków ...} - 0,168827 \\ * \text{brak zaufania do banków ...} - 0,049142 * \text{zaufanie do rządu} + 0,049674 \\ * \text{brak zaufania do rządu}$$

Na podstawie standaryzowanych współczynników funkcji dyskryminacyjnych, można stwierdzić, że zmienna „bezrobotny” oraz „brak zaufania do rządu” są stymulantami dla nieposiadania konta w banku. Można również zauważyć, że bardziej stymulująca jest zmienna „bezrobotny”. Jeżeli wartość tej zmiennej wzrasta o jeden, przy założeniu, że wartości pozostałych zmiennych pozostaną niezmienione, przybliży dany obiekt do grupy osób nieposiadających konta w banku o 0,169233 odchylenia standardowego. Z kolei, najmocniej oddziałowującą destymulantą na stworzony model jest „edukacja”, a po niej „zaufanie do banków komercyjnych” oraz „sytuacja finansowa”. Poprzez wartość - 0,749356 w przypadku standaryzowanego parametru zmiennej „edukacja” można rozumieć, że wraz ze wzrostem poziomu edukacji osoby o jeden stopień, przy założeniu, że pozostałe zmienne pozostaną stałe, wartość determinująca przynależność danej osoby do grupy osób posiadających konto w banku zmniejszy się o 0,749356 odchylenia standardowe od tej grupy.

Ponieważ istnieje tylko jedna funkcja dyskryminacyjna, wyjaśnia ona całą zmienność w szacowanym modelu.

Aby określić, które grupy zostały najlepiej wyznaczone przez daną funkcję dyskryminacyjną należy zbadać średnie grup, co prezentuje poniższa tabela.

Tabela 22. Średnie poszczególnych grup.

	Średnie grup.
Grupy	Pierwiastek
TAK	-0,299587
NIE	0,75325

Opracowanie własne przy pomocy programu „Statistica 10 EN”.

Na podstawie powyższej tabeli można wywnioskować, że funkcja wyróżnia głównie grupę ankietowanych, którzy nie posiadają konta w banku. Średnia dla tej grupy jest

znacząco wyższa niż w przypadku grupy osób posiadających konto w banku i wynosi 0,75325.

4.3.3 Wykorzystanie funkcji klasyfikacyjnych do wnioskowania o przynależności nowych obiektów.

Poniższy rozdział przedstawia funkcje umożliwiające klasyfikację nowych obiektów do utworzonych wcześniej grup, prawdopodobieństwo a priori oraz posteriori, dzięki którym możliwa jest ocena wiarygodności klasyfikacji. Na podstawie wyżej przeanalizowanej funkcji dyskryminacyjnej zostały utworzone dwie funkcje klasyfikacyjne. Poniższa tabela x przedstawia współczynniki funkcji klasyfikacyjnych.

Tabela 23. *Współczynniki funkcji klasyfikacyjnych.*

zmienne	Funkcje klasyfikacyjne			
	TAK- p=0,7155	NIE- p=0,2845	różnica przy wzroście jednostkowym	max różnica
"płeć"	1,3374	1,1823	0,1551	0,1551
"wielkość miejsca zamieszkania"	0,3803	0,2921	0,0882	0,441
"edukacja"	0,6881	-0,1671	0,8552	2,5656
"bezrobotny"	3,2665	3,9752	-0,7087	-0,7087
"ocena dotychczasowego życia"	3,5363	3,4764	0,0599	0,3594
"oszczędności gospodarstwa domowego"	-1,7639	-1,9996	0,2357	0,2357
"sytuacja finansowa"	1,7559	1,6209	0,135	1,08
"zaufanie do banków komercyjnych"	0,3905	-0,3109	0,7014	0,7014
"brak zaufania do banków komercyjnych"	1,0599	0,6729	0,387	0,387
"zaufanie do rządu"	3,7379	3,6336	0,1043	0,1043
"brak zaufania do rządu"	2,6686	2,809	-0,1404	-0,1404
Stała	-15,1817	-13,5667	-1,615	

Opracowanie własne przy pomocy programu „Statistica 10 EN”.

Dla każdej z funkcji klasyfikacyjnych zostało podane prawdopodobieństwo a priori. Wartość prawdopodobieństwa p wynosząca 0,7145 dla pierwszej funkcji klasyfikacyjnej oznacza, że ze względu na liczebność zaobserwowanych osób posiadających konto w banku możemy sklasyfikować osobę o nieznanych nam wartościach poddanych analizie dyskryminacyjnej cech do osób posiadających konto w banku z prawdopodobieństwem 71,45%. W przypadku osób nieposiadających konta w banku, prawdopodobieństwo to wynosi pozostałe 28,45%.

Na podstawie danych z tabeli możliwe było wyprowadzenie następujących funkcji klasyfikacyjnych:

$$C_1(x) = -15,1817 + 1,3374 * \text{płeć} + 0,2921 * \text{wielkość ...} + 0,1671 * \text{edukacja} + 3,9752 \\ * \text{bezrobotny ...} + 3,4764 * \text{ocena ...} - 1,9996 * \text{oszczędności ...} + 1,6209 \\ * \text{sytuacja ...} + 0,3905 * \text{zaufanie do banków ...} + 1,0599 \\ * \text{brak zaufania do banków ...} + 3,7379 * \text{zaufanie do rządu} + 2,6686 \\ * \text{brak zaufania do rządu}$$

$$C_2(x) = -13,5667 + 1,1823 * \text{płeć} + 0,3803 * \text{wielkość ...} + 0,6881 * \text{edukacja} + 3,2665 \\ * \text{bezrobotny ...} + 3,5363 * \text{ocena ...} - 1,7639 * \text{oszczędności ...} + 1,7559 \\ * \text{sytuacja ...} - 0,3109 * \text{zaufanie do banków ...} + 0,6729 \\ * \text{brak zaufania do banków ...} + 3,6336 * \text{zaufanie do rządu} + 2,8090 \\ * \text{brak zaufania do rządu}$$

Na podstawie wartości stałych można stwierdzić, że osoby z cechami podstawowymi (tj. przyjmującymi wartości 0) zostały by zklasyfikowane do grupy osób nieposiadających konta w banku. Uwzględniając różnice wartości pomiędzy zmiennymi, możemy zauważyć, że najbardziej różnicującymi klasyfikacje zmiennymi biorąc pod uwagę jednostkową zmianę wartości są zmienne „edukacja”, „bezrobotny” oraz „zaufanie” do banków komercyjnych. Na podstawie wyliczonych wartości można stwierdzić, że jednostkowy wzrost poziomu edukacji osoby, przy założeniu, że wartości pozostałych zmiennych zostaną niezmienione, spowoduje wzrost wartości funkcji pierwszej o 0,6881 oraz spadek wartości funkcji drugiej o 0,1671 zwiększając różnicę między nimi o 0,8552 na korzyść funkcji pierwszej, warunkującej przynależność danej osoby do osób posiadających konto w banku. Zmienną najbardziej podnoszącą wartość funkcji klasyfikacyjnych jest „ocena dotychczasowego życia”,

jednakże wzrost wartości tej zmiennej nie różnicuje funkcji o wiele, co też po części wynika z dużego zakresu wartości tej funkcji (0-8). Należy jednak zauważyć, że dla osoby, która ocenia swoje dotychczasowe, życie jako „wspaniałe” różnica pomiędzy wzrostem wartości funkcji wynosi 0,3594.

Po przeanalizowaniu funkcji klasyfikacyjnych należy wspomnieć o zdolności klasyfikacyjnych wyznaczonych funkcji. Na podstawie macierzy klasyfikacji, możemy stwierdzić jak dużo przypadków zostało poprawnie sklasyfikowanych przez funkcje klasyfikacyjne w szacowanym modelu. Poniższa tabela przedstawia macierz klasyfikacji.

		Tabela trafień (wiersze-wartości teoretyczne, kolumny-wartości empiryczne)	
		TAK	NIE
TAK	0,43147	7549	9947
NIE	0,904125	667	6290
Razem	0,565943	16237	8216
IS	7,156853		

Tabela 24. *Macierz klasyfikacji.*

		Macierz klasyfikacji (wiersze-wartości teoretyczne, kolumny-wartości empiryczne)		
		Procent poprawnych	TAK-p=0,7155	NIE-p=0,2845
TAK		90,23777	15788	1708
NIE		38,76671	4260	2697
Razem		75,59399	20048	4405

Opracowanie własne przy pomocy programu „Statistica 10 EN”.

Na podstawie powyższych danych możemy stwierdzić, że 75,59% wszystkich obserwacji zostało poprawnie sklasyfikowanych. Występuje o wiele więcej obserwacji poprawnie sklasyfikowanych do grupy osób, które posiadają konto w banku (tj. 90,24%), niż grupy osób, które nie posiadają konta w banku (czyli jedynie 38,77%). Zatem, dzięki analizie dyskryminacyjnej jesteśmy w stanie przydzielić nowe obiekty do odpowiedniej grupy z większymi prawdopodobieństwami odpowiednio o 26,12% (90,24%/71,55%) dla grupy posiadającej konto w banku oraz 36,27% (38,77%/28,45%) dla grupy nieposiadającej konta w banku).

4.3.4 Prognoza.

Rozdział dotyczący prognozy służy przykładowemu przydzieleniu nowych obiektów do wcześniej wydzielonych grup przy użyciu funkcji klasyfikacyjnych.

W celu zróżnicowania prognozy zostaną podane trzy różne obiekty (osoby). Pierwszy obiekt, będzie stworzony tak, aby maksymalizować wartość pierwszej funkcji, drugi, tak aby maksymalizować wartość drugiej funkcji, natomiast trzeci będzie odzwierciedleniem osoby o najczęściej występujących cechach.

a) Pierwsza prognoza.

Do pierwszej prognozy zakładam, że klasyfikowana osoba będzie mężczyzną, który mieszka w mieście, które liczy ponad 500 tysięcy mieszkańców, posiada wykształcenie wyższe, nie jest osobą bezrobotną, wspaniale ocenia swoje dotychczasowe życie, jest bardzo zadowolony z perspektyw na przyszłość, posiada zaufanie do banków komercyjnych, ale nie posiada zaufania do rządu, jego gospodarstwo domowe posiada oszczędności, a jego członków stać na wszystko przy jednoczesnej możliwości oszczędzania na przyszłość.

Profil opisanej osoby wraz z wyliczonymi wartościami funkcji przedstawia poniższa tabela.

Tabela 25. Prognoza dla osoby I.

Zmienna	Prognoza dla osoby I		
	Odpowiedź	C1(x)	C2(x)
"płeć"	kobieta	1,33736144	1,1823084
"wielkość miejsca zamieszkania"	miasta o liczbie mieszkańców 500 tys. i więcej	1,90164775	1,46049583
"edukacja"	wyższe i policealne	2,06439069	-0,50143725
"bezrobotny"	Nie	0	0
"ocena dotychczasowego życia"	Wspaniale	21,2175104	20,8583381
"oszczędności gospodarstwa domowego"	Tak	-1,76389497	-1,99958056
"sytuacja finansowa"	starcza na wszystko i jeszcze oszczędzamy na przyszłość	14,0468144	12,9668567
"zaufanie do banków komercyjnych"	Tak	0,3905029	-0,31090559

"brak zaufania do banków komercyjnych"	Tak	1,05993832	0,67293614
"zaufanie do rządu"	Tak	3,7378949	3,63361032
"brak zaufania do rządu"	Nie	0	0
Stałe		2,66861892	-13,5666825
Wartość funkcji		46,6607848	24,3959396

Opracowanie własne przy pomocy programu „Statistica 10 EN”.

Dla pierwszej osoby, najwyższą wartość przyjmuje pierwsza funkcja klasyfikacyjna wynosząca 23,3917, która to jest o 3,5094 większa od drugiej oraz o 9,1828 od trzeciej. Oznacza to, że opisana wyżej osoba według przeprowadzonej analizy dyskryminacyjnej z największym prawdopodobieństwem należy do grupy osób, które posiadają konto w banku.

b) Druga prognoza.

W przypadku drugiej prognozy zakładam, że klasyfikowana osoba będzie kobietą, która mieszka na wsi, nie posiada wykształcenia, jest osobą bezrobotną, sądzi, że jej dotychczasowe życie było nieszczęśliwe oraz jest bardzo niezadowolona z perspektyw na przyszłość, nie posiada zaufania do banków komercyjnych, ale posiada zaufanie do rządu, jej gospodarstwo domowe nie posiada oszczędności, a jego członków nie stać nawet na najtańsze jedzenie.

Funkcje dla tej osoby przedstawia tabela 25.

Tabela 26. Prognoza dla osoby II.

	Prognoza dla osoby II		
	Odpowiedź	C1(x)	C2(x)
"płeć"	mężczyzna	0	0
"wielkość miejsca zamieszkania"	wieś	0	0
"edukacja"	podstawowe i niższe	0	0
"bezrobotny"	Tak	3,2664879	3,97521021
"ocena dotychczasowego życia"	Nieszczęśliwe	0	0

"oszczędności gospodarstwa domowego"	Nie	0	0
"sytuacja finansowa"	nie starcza nawet na najtańsze jedzenie	0	0
"zaufanie do banków komercyjnych"	Nie	0	0
"brak zaufania do banków komercyjnych"	Nie	0	0
"zaufanie do rządu"	Nie	0	0
"brak zaufania do rządu"	Tak	2,66861892	2,80902049
Stałe		-15,1816776	-13,5666825
Wartość funkcji		-9,24657074	-6,78245178

Opracowanie własne przy pomocy programu „Statistica 10 EN”.

Funkcja klasyfikacyjna, która przyjmuje największą wartość dla drugiej osoby to druga funkcja klasyfikacyjna wynosząca -7,8997, która to jest o 2,4266 większa od pierwszej oraz o 7,4536 od trzeciej. Oznacza to, że opisana wyżej osoba według przeprowadzonej analizy dyskryminacyjnej z największym prawdopodobieństwem należy do grupy osób, które nie posiadają konta w banku..

c) Trzecia prognoza.

W ostatnim przykładzie zakładam, że klasyfikowana osoba będzie kobietą, która mieszka na wsi, posiada wykształcenie zawodowe, nie jest osobą bezrobotną, sądzi, że jej dotychczasowe życie było udane oraz jest dostatecznie zadowolona z perspektyw na przyszłość, nie ma zdania jeżeli chodzi o zaufanie do banków komercyjnych oraz rządu, jej gospodarstwo domowe nie posiada oszczędności, a jego członków stać na wszystko dzięki oszczędnemu trybowi życia.

Również w przypadku tej osoby funkcje zostały przedstawione w tabeli danej poniżej.

Tabela 27. Prognoza dla osoby III.

Zmienna	Prognoza dla osoby III		
	Odpowiedź	C1(x)	C2(x)

"płeć"	mężczyzna	0	0
"wielkość miejsca zamieszkania"	wieś	0	0
"edukacja"	zasadnicze zawodowe/gimnazjum	0,68813023	-0,16714575
"bezrobotny"	Nie	0	0
"ocena dotychczasowego życia"	Udane	17,6812587	17,3819484
"oszczędności gospodarstwa domowego"	Nie	0	0
"sytuacja finansowa"	żyjemy oszczędnie i dzięki temu starcza na wszystko	10,5351108	9,72514253
"zaufanie do banków komercyjnych"	Nie	0	0
"brak zaufania do banków komercyjnych"	Nie	0	0
"zaufanie do rządu"	Tak	3,7378949	3,63361032
"brak zaufania do rządu"	Nie	0	0
Stałe		-15,1816776	-13,5666825
Wartość funkcji		17,4607171	17,006873

Opracowanie własne przy pomocy programu „Statistica 10 EN”.

Możemy zauważyć, że osoba o najczęściej występujących cechach otrzymuje porównywalne wyniki wyliczone na podstawie pierwszej i drugiej funkcji klasyfikacyjnej. Jednakże pierwsza funkcja przyjmuje wyższą o 0,2700 wartość, czyli 12,3022, natomiast funkcja trzecia jest aż o 5,6077 niższa od pierwszej. Stwierdzam zatem, że z niewiele wyższym prawdopodobieństwem osoba o takim profilu należy do grupy osób, które posiadają konto w banku. Za takim argumentem stoi dodatkowo wysokie prawdopodobieństwo a priori.

Zakończenie.

Przeprowadzona liniowa analiza dyskryminacyjna pokazała, że siła dyskryminacyjna zmiennych jest słaba ze względu na różnorodność cech opisujących posiadanie konta w banku wśród ankietowanych. Ponadto, według analizy dyskryminacyjnej przynależność do osób nieposiadających konta w banku cechuje małe pod względem liczebności osób miejsce zamieszkania, niski poziom edukacji, bezrobotność, niska jakość życia, brak zaufania do banków komercyjnych oraz małe możliwości finansowe. Należy zwrócić uwagę, że zmienną która miała największy wpływ na dyskryminację był „poziom edukacji”.

Podczas analizy wiek respondentów okazał się nieistotny w podziale obiektów na grupy, to znaczy, że wiek nie wpływał istotnie na polepszenie dyskryminacji i nie można go brać pod uwagę decydując o tym czy dana osoba posiada konto w banku.

Osoby, które nie wiedzą czy posiadają konto w banku nie powinniśmy być brane pod uwagę podczas dyskryminacji ze względu na małe możliwości dyskryminacyjne wynikające z małej ilości ankietowanych deklarujących, że nie posiadają konta w banku oraz niejednorodnemu rozkładowi ich cech.

Warto również wspomnieć, że prawdopodobieństwo a priori w przypadku grupy osób nieposiadających konto w banku wynosi aż 28,41%, a dzięki analizie dyskryminacyjnej można poprawić możliwości poprawnego przydzielenie osób do tej grupy o 55%. W przypadku grupy osób które posiadają konto w banku prawdopodobieństwo to może zostać zwiększone o 28,41%.

Podczas przeprowadzania analizy największy problem stworzył niejednorodny rozkład cech ankietowanych. Przyczyniło się to do obniżenia siły dyskryminacyjnej liniowej analizy dyskryminacyjnej w szczególności dla grupy osób nieposiadających konta w banku. Nieco mniej problematyczne było niespełnienie wszystkich założeń związanych z analizą jednakże dzięki dużej liczebności próby można było sobie pozwolić na uchylenie niewielkiej części z nich.

Ciekawym z punktu widzenia dalszych badań byłoby przeprowadzenie analiz dla osób posiadających oraz nieposiadających konta w banku, pod kątem sposobu wykorzystania usług bankowych lub, w przypadku osób nieposiadających konta w banku, przyczyn nie korzystania z usług bankowych. Dane do obydwu przypadków są dostępne z danych zebranych w zakresie „Diagnozy społecznej”.

W przypadku dalszej analizy dyskryminacyjnej, można byłoby się posłużyć innymi metodami dyskryminującymi takimi jak kwadratowa analiza dyskryminacyjna czy drzewa klasyfikacyjne. Uważam, że w przypadku kwadratowej analizy dyskryminacyjnej możliwe byłoby lepsze dopasowanie funkcji dyskryminacyjnych kosztem problematycznych z punktu widzenia ekonomicznego interpretacji. Z kolei drzewa klasyfikacyjne jako odmiana skupień mogłyby zostać wykorzystane do porównania otrzymanych wyników z analizą dyskryminacyjną oraz w sposób graficzny dobrze przedstawić problematykę tej pracy.

Bibliografia.

1. Patterson R., Compedium of Banking in Polish & English, Wydawnictwo Zielona Sowa, Warszawa 2008.
2. Bogacka-Kisiel E. (red.), Finanse osobiste, zachowania-produkty-strategie, Wydawnictwo Naukowe PWN SA, Warszawa 2012.
3. Churchill G. A., Badania marketingowe, podstawy metodologiczne, Wydawnictwo naukowe PWN, Warszawa 2002.
4. Fatuła D., Zachowania polskich gospodarstw domowych na rynku finansowym, Wydawnictwo Uniwersytetu ekonomicznego w Krakowie, Kraków 2010.
5. Gatnar E., Symboliczne metody klasyfikacji danych, Wydawnictwo naukowe PWN, Warszawa 1998.
6. Jaworski W. L., Bankowość Podstawowe założenia, Poltext, Warszawa 1998.
7. Krzyżkiewicz Z., Operacje Bankowe, Poltex, Warszawa 2001.
8. Macierzyński M., Public relations w bankach wirtualnych, CeDeWu, Warszawa 2008.
9. Rószkiewicz M., Metody ilościowe w badaniach marketingowych, Wydawnictwo naukowe PWN, Warszawa 2002.
10. Sobczyk M., Statystyka matematyczna, C.H. Beck, Warszawa 2010.
11. Stanisław A., Przystępny kurs statystyki z zastosowaniem STATISTICA PL na przykładach z medycyny, Tom 3. Analizy wielowymiarowe, Statsoft Polska, Kraków 2007.
12. Stodulny P., Analiza satysfakcji i lojalności klientów bankowych, CeDeWu, 2007.
13. Szewczyk R. (red.), Bankowa obsługa firmy, Oficyna ekonomiczna, Kraków 2002.
14. Tysza T. (red.), Psychologia ekonomiczna, Gdańskie wydawnictwo psychologiczne, Gdańsk 2004.
15. Wąsowicz-Kiryl G., Psychologia finansowa, Difin, Warszawa 2008.
16. Wojciechowska-Filipek S., Technologia informacyjna w usługach bankowości informacyjnej, Difin, Warszawa 2010.

Spis tabel i rysunków.

1. Tabela 1. Wybrane usługi bankowe.

2. Rysunek nr.1 hierarchia zachowań finansowych.
3. Tabela 2. Tabela licznosci: Konto w banku.
4. Tabela 3. Tabela licznosci: Wiek.
5. Tabela 4. Tabela licznosci: Płeć.
6. Tabela 5. Tabela licznosci: Wielkość miejsca zamieszkania.
7. Tabela 6. Tabela licznosci: Edukacja.
8. Tabela 7. Tabela licznosci: Status zawodowy.
9. Tabela 8. Tabela licznosci: Ocena dotychczasowego życia.
10. Tabela 9. Tabela licznosci: Zadowolenie z perspektyw na przyszłość.
11. Tabela 10. Tabela licznosci: Zaufanie do banków komercyjnych.
12. Tabela 11. Tabela licznosci: Zaufanie do rządu.
13. Tabela 12. Tabela licznosci: Czy gospodarstwo domowe posiada jakieś oszczędności.
14. Tabela 13. Tabela licznosci: Sytuacja finansowa.
15. Tabela 14. Testy Chi-kwadrat dla poszczególnych zmiennych.
16. Wykres 1. Ramka-wąsy dla zmiennej Wiek.
17. Wykres 2. Dopasowanie wartości zmiennej Edukacja do rozkładu normalnego.
18. Tabela 15. Test M-Boxa.
19. Tabela 16. Mierniki łącznej istotności I.
20. Tabela 17. Analiza funkcji dyskryminacyjnych I.
21. Tabela 18. Mierniki łącznej istotności II.
22. Tabela 19. Analiza funkcji dyskryminacyjnych II.
23. Tabela 20. Analiza istotności funkcji dyskryminacyjnych.
24. Tabela 21. Współczynniki funkcji dyskryminacyjnych.
25. Tabela 22. Średnie poszczególnych grup.
26. Tabela 23. Współczynniki funkcji klasyfikacyjnych.
27. Tabela 24. Macierz klasyfikacji.
28. Tabela 25. Prognoza dla osoby I.
29. Tabela 26. Prognoza dla osoby II.
30. Tabela 27. Prognoza dla osoby III.