

EKONOMICKÁ UNIVERZITA V BRATISLAVE

FAKULTA HOSPODÁRSKEJ INFORMATIKY

Evidenčné číslo: 103005/I/2018/36069194041218308

**PRÍJMOVÁ CHUDOBA, MATERIÁLNA
DEPRIVÁCIA A KONZISTENTNÁ CHUDOBA
SLOVENSKÝCH DOMÁCNOSTÍ**

Diplomová práca

2018

Veronika Straňáková

EKONOMICKÁ UNIVERZITA V BRATISLAVE
FAKULTA HOSPODÁRSKEJ INFORMATIKY

**PRÍJMOVÁ CHUDOBA, MATERIÁLNA
DEPRIVÁCIA A KONZISTENTNÁ CHUDOBA
SLOVENSKÝCH DOMÁCNOSTÍ**

Diplomová práca

Študijný program: Štatistické metódy v ekonómii

Študijný odbor: Kvantitatívne metódy v ekonómii

Školiace pracovisko: Katedra štatistiky

Vedúci záverečnej práce: doc. Mgr. Erik Šoltés, PhD.

Bratislava 2018

Veronika Straňáková

Čestné vyhlásenie

Čestne vyhlasujem, že záverečnú prácu som vypracovala samostatne a že som uviedla všetku použitú literatúru.

Dátum: 13.4.2018

.....

Pod'akovanie

Týmto by som sa chcela pod'akovať vedúcemu mojej záverečnej práce doc. Mgr. Erikovi Šoltésovi, PhD., za všetku pomoc pri písaní tejto práce a svojim priateľom a rodine za rady a podporu.

ABSTRAKT

STRAŇÁKOVÁ, Veronika: *Príjmová chudoba, materiálna deprivácia a konzistentná chudoba slovenských domácností*. – Ekonomická univerzita v Bratislave. Fakulta hospodárskej informatiky; Katedra štatistiky. – Vedúci záverečnej práce: doc. Mgr. Erik Šoltés, PhD. – Bratislava: FHI, 2018, 77 s.

Hlavným cieľom záverečnej práce je na základe poznatkov nadobudnutých počas štúdia rôznych relevantných zdrojov a vlastných štatistických analýz priblížiť situáciu v oblasti príjmovej chudoby, materiálnej deprivácie a konzistentnej chudoby na Slovensku. Práca je obsahovo rozdelená do štyroch kapitol a obsahuje 15 tabuliek, 15 obrázkov a 5 príloh. Prvá kapitola sa venuje súčasnému stavu problematiky na Slovensku a v Európskej únii. V druhej kapitole popisujeme ciele práce, kde dosiahnutie hlavného cieľa podkladáme aj čiastkovými cieľmi. V tretej kapitole sa venujeme metodike potrebnej na konštrukciu vybraných štatistických analýz. Napokon štvrtá kapitola obsahuje aplikáciu teoretických poznatkov do praxe pomocou vybraných štatistických metód spolu s interpretáciami a názornými výstupmi zo softvéru. Výsledkom riešenia danej problematiky je oboznámenie sa s problémami príjmovej chudoby, materiálnej deprivácie a konzistentnej chudoby a využitím metód logistickej regresie priblíženie súčasnej situácie v slovenských domácnostiach.

Kľúčové slová: Stratégia Európa 2020, EU-SILC, príjmová chudoba, materiálna deprivácia, konzistentná chudoba, chudoba a sociálne vylúčenie, logistická regresia

ABSTRACT

STRAŇÁKOVÁ, Veronika: *Income poverty, material deprivation and consistent poverty of Slovak households*. – University of Economics in Bratislava. Faculty of Economic Informatics; Department of Statistics. – Supervisor of the thesis: doc. Mgr. Erik Šoltés, PhD. – Bratislava: FHI, 2018, 77 p.

The main aim of the diploma thesis is based on knowledge gained during the study of relevant sources and own statistical analysis to enclose the situation in area of income poverty, material deprivation and consistent poverty in Slovakia. The thesis is divided into four chapters and contains 15 tables, 15 pictures and 5 appendices. First chapter is about the current problematic in the field in Slovakia and European Union. In the second chapter we describe the goals of the thesis, where we underline the achievement of the main goal by smaller partial goals. Chapter three is dedicated to methodology needed to construct chosen statistical analysis. At last, the fourth chapter contains application of theoretical knowledge into practice with help of chosen statistical analysis along with interpretations and schematic outputs from used software. The result of the thesis is familiarization with the subject of income poverty, material deprivation and consistent poverty and with usage of methods of logistic regression showing the current situation in Slovak households.

Key words: Strategy Europe 2020, EU-SILC, income poverty, material deprivation, consistent poverty, poverty and social exclusion, logistic regression

Obsah

Úvod	8
1 Súčasný stav problematiky doma a v zahraničí.....	9
1.1 Výberové zisťovanie EU-SILC	10
1.2 Stratégia Európa 2020	10
1.3 Príjmová chudoba, materiálna deprivácia a konzistentná chudoba	13
1.3.1 Príjmová chudoba na Slovensku a v krajinách Európskej Únie	15
1.3.2 Materiálna deprivácia na Slovensku a v krajinách Európskej Únie	17
1.3.3 Konzistentná chudoba na Slovensku a v krajinách Európskej Únie.....	18
2 Cieľ práce	20
3 Metodika a metódy skúmania záverečnej práce	21
3.1 Logistická regresia	21
3.2 Logistická regresia s binárnou vysvetľovanou premennou.....	22
3.2.1 Odhad parametrov pre model logistickej regresie.....	26
3.2.2 Testovanie štatistickej významnosti modelu ako celku.....	27
3.2.3 Testovanie štatistickej významnosti vplyvu vysvetľujúcich premenných	29
3.2.4 Intervalové odhady pre parametre modelu a pomer šancí.....	30
3.3 Logistická regresia s multinomickou vysvetľovanou premennou	30
3.3.1 Hodnotenie kvality modelu	32
3.4 Závislosť dvoch kategoriálnych znakov	33
3.4.2 Sila závislosti dvoch kategoriálnych znakov.....	34
4 Výsledky záverečnej práce a diskusia.....	36
4.1 Vplyv relevantných faktorov na konzistentnú chudobu.....	37
4.2 Analýza konzistentnej chudoby využitím logistickej regresie.....	38
4.2.1 Interpretácia parametrov logistickej regresie.....	44
4.3 Logistická regresia ako nástroj na výpočet pravdepodobnosti	51
4.3.1 Model s vysvetľujúcimi premennými EAS a Education	52
4.3.2 Model s vysvetľujúcimi premennými EAS a Health.....	53
4.3.3 Model s vysvetľujúcimi premennými EAS a Household Type.....	54
4.3.4 Model s vysvetľujúcimi premennými EAS a Marital Status.....	55

4.3.5 Model s vysvetľujúcimi premennými EAS a Region.....	56
4.3.6 Model s vysvetľujúcimi premennými EAS a Urbanisation.....	58
Záver	59
Použitá literatúra	62
Zoznam príloh	64

Úvod

Aj v 21. storočí je chudoba problémom na celom svete a stále dokonca aj vo vyspelých krajinách tejto planéty. Preto je neustále potrebné sa zameriavať nielen na jej monitorovanie, ale takisto aj riešenie a prijímanie opatrení v súvislosti s chudobou. Nejde totiž len o nepriaznivé príjmové podmienky, ale takisto vylúčenie zo spoločnosti a ťažkosti navrátiť sa do jej normálneho fungovania. Hlavným cieľom tejto práce je čitateľovi priblížiť problematiku chudoby a to hlavne príjmovej chudoby, materiálnej deprivácie a konzistentnej chudoby na Slovensku. Toto zahŕňa nielen objasnenie súčasného stavu a teoretických poznatkov o týchto javoch, ale takisto aj použitie praktických analýz na kvantifikáciu rozdielností v slovenských domácnostiach.

V prvej kapitole sa zaoberáme súčasným stavom problematiky na Slovensku a v Európskej únii, a to jednak teoretickými poznatkami o chudobe a neskôr samotnými riešeniami a nástrojmi, ktoré implementuje Európska únia pri riešení problému chudoby a sociálneho vylúčenia v rámci *Stratégie Európa 2020*. Napokon v predmetnej oblasti mapujeme situáciu na Slovensku.

V druhej kapitole uvádzame ciele práce – hlavný cieľ a čiastkové ciele, pomocou ktorých ho budeme postupne naplňať. V tretej kapitole opisujeme a definujeme teoretický aparát potrebný na zvládnutie problematiky logistickej regresie, ktorá je použitá ako štatistická metóda pre analýzu v poslednej časti práce.

V poslednej kapitole sa nachádzajú štatistické analýzy príjmovej chudoby a materiálnej deprivácie na Slovensku, teda analýzy 2 dimenzií *chudoby a sociálneho vylúčenia* v zmysle 3-dimenzionálneho konceptu, ktorý EÚ využíva v rámci stratégie Európa 2020. Špeciálnu pozornosť venujeme analýze konzistentnej chudoby, ktorá sa týka príjmovovo chudobných domácností, ktoré sú súčasne závažne materiálne deprivované. Pomocou údajov z databázy *EU-SILC 2016* a softvéru *SAS Enterprise Guide 5.1* využívame metódy logistickej regresie na analýzu týchto dimenzií chudoby a sociálneho vylúčenia a hlavne na kvantifikáciu vplyvu relevantných faktorov na konzistentnú chudobu. Model logistickej regresie využívame jednak na odhad pomerov šanci, ale takisto aj pravdepodobnosti ohrozenia sledovanými 2 dimenziami príjmovej chudoby a sociálneho vylúčenia. Vzhľadom na rozsiahlosť niektorých analýz sme zložitejšie výstupy zo softvéru uviedli v prílohách nachádzajúcich sa na konci práce.

1 Súčasný stav problematiky doma a v zahraničí

Aj keď krajiny Európskej Únie patria medzi najvyspelejšie na svete, problém chudoby sa týka značnej časti populácie EÚ. Pritom na definíciu chudoby je priam nespočítateľné množstvo názorov, z ktorých každý istou mierou zapadá do niektorého z oficiálnych konceptov.

Chudoba je totiž zložitý, interdisciplinárny jav, ktorý neskúmajú len ekonómovia, ale takisto aj sociológovia, psychológovia, štatistickí a ďalšie množstvo odborníkov z iných vedných disciplín. Na jednej strane teda neexistuje jediná definícia chudoby, ktorá by bola všeobecne akceptovaná, ale rovnako tak neexistuje jediná správna definícia – vždy vieme skúmať a merať len jeden z konceptov chudoby.

Za jednu z najvšeobecnejšie prijímaných myšlienok o chudobe sa považuje definícia z publikácie *Chudoba v spojenom kráľovstve* Roberta M. Townsenda z roku 1979, ktorá opisuje základy z konceptu relatívnej materiálnej deprivácie. Townsend hovorí, že jednotlivci sú chudobní vtedy, ak nemajú dosť zdrojov na zaobstaranie jednak základných potrieb (potravy a oblečenia) ale aj pracovných a vzdelávacích podmienok a takisto podmienok životného štýlu a účasti na aktivitách obvyklých pre spoločnosť, do ktorej patria. Jednoduchšie povedané, jednotlivci sú chudobní ak ich zdroje sú výrazne nižšie ako zdroje priemerných členov spoločnosti, do ktorej patria – to ich vylučuje zo vzorov správania typických pre túto spoločnosť.

Väčšina používaných definícií chudoby je založená na určení *indikátora blahobytu* (príjem, spotreba, výdavky a pod.) a určení *hranice chudoby*, ktorá rozdeľuje obyvateľstvo na časť chudobnú a nechudobnú. (Želinský, T. – 2014). Medzi takéto definície patria aj:

- **Absolútny a relatívny koncept** – pri absolútnom koncepte sa za hranicu chudoby považuje taký nedostatok zdrojov, kedy si jednotlivec nevie zabezpečiť základné potreby na jeho prežitie – ide o problém chudoby najmä v rozvojových spoločnostiach. Pri relatívnom koncepte už chápeme chudobu ako Townsend – ide aj o zabezpečenie účasti na aktivitách obvyklých pre spoločnosť, ide o porovnávanie k istej strednej hodnote. Takéto chápanie je najbližšie k chápaniu chudoby podľa Európskej Únie.

- **Priamy a nepriamy koncept** – nepriamy koncept stanovuje hranicu chudoby pomocou príjmu, pričom ale nezohľadňuje spotrebu. Naopak priamy koncept je založený na meraní príjmu a takisto jeho transformácie do spotreby – zohľadňuje aj tzv. sekundárnu chudobu, teda že chudobu nespôsobuje len nedostatok príjmov, ale aj zlé hospodárenie s nimi.
- **Objektívny a subjektívny koncept** – subjektívny koncept je založený na hodnotení chudoby podľa názoru jednotlivcov – či už jednoduchého posúdenia, či sa človek cíti chudobný, alebo aj zložitejších názorov. Objektívny koncept potom meria chudobu bez akéhokoľvek vplyvu zo strany jednotlivcov, ktorých sa výskum týka – využívajú sa štatistické metódy a expertné názory. (Michálek A., 2000)

1.1 Výberové zisťovanie EU-SILC

Vo všetkých ostatných krajinách Európskej Únie a tiež v Nórsku, Švajčiarsku, Islande a Srbsku sa dlhodobo a každoročne uskutočňuje Zisťovanie o príjmoch a životných podmienkach domácností EU-SILC (z angl. *European Union Statistics on Income and Living Conditions*) na základe nariadenia Európskeho parlamentu a Rady Európskej Únie. Na Slovensku sa toto harmonizované zisťovanie uskutočňuje už od roku 2005 a tvorí základ nielen pre monitorovanie príjmov, chudoby a sociálneho vylúčenia, ale takisto pre zlepšovanie životnej úrovne obyvateľstva a kvality života na národnej úrovni. (Vlačuha, R. – Kováčová. Y., 2017)

Referenčným obdobím pre zisťovanie je vždy predchádzajúci kalendárny rok. Výber štatistických jednotiek pre zisťovanie sa realizuje dvojstupňovým stratifikovaným výberom podľa dvoch kritérií – stratifikácia podľa krajov a podľa stupňa urbanizácie (počet obyvateľov v obci). Štatistickou jednotkou je hospodáriaca domácnosť s jej členmi, ktorí bývajú na území SR. Indikátory v zisťovaní EU-SILC zo spomínaných konceptov chudoby merajú koncept relatívny, nepriamy a objektívny. (Želinský, T. – 2014)

1.2 Stratégia Európa 2020

V období, kedy sa Európska Únia zameriavala na spoluprácu a spamätanie sa z hospodárskej krízy, bol stále 1 z 5 občanov EÚ ohrozený chudobou alebo sociálnym

vylúčením. Tieto neuspokojivé čísla viedli v roku 2010 k vytvoreniu plánu na riešenie chudoby na nasledujúce desaťročie – Stratégie Európa 2020. Tento plán bol a je veľmi ambiciózny, nakoľko jeho splnenie sa dostáva do rozporu jednak s ťažkou ekonomickou situáciou niektorých krajín, ale aj relatívnou protichodnosťou niektorých z jeho cieľov.

Hlavnými piliermi tejto stratégie je zabezpečiť inteligentný, udržateľný a inkluzívny rast:

- **Inteligentný rast** – rast hospodárskej úrovne založený na znalostiach a inováciach,
- **Udržateľný rast** – rast hospodárskej úrovne, nie však na úkor ekológie – vytvorenie hospodárstva, ktoré je konkurencieschopné a efektívne využíva zdroje,
- **Inkluzívny rast** – rast hospodárskej úrovne, nie však na úkor najslabších – podpora vysokej miery zamestnanosti, aby sa zabezpečila sociálna a územná súdržnosť.

V rámci týchto pilierov bolo stanovených základných 5 cieľov, ktoré sú rozdielne na celoeurópskej a národnej úrovni, kvôli rozdielnostiam jednotlivých krajín:

- 1.) **Zamestnanosť:** miera zamestnanosti obyvateľstva medzi vekom 20-64 rokov sa má zvýšiť v EÚ na 75%, na Slovensku na 72%,
- 2.) **Výskum a vývoj:** investície do oblasti výskumu a vývoja sa majú v EÚ zvýšiť na úroveň 3% z celkového HDP, na Slovensku 1% z HDP SR (neskôr zmena na 1,2%),
- 3.) **Klimatické zmeny a energia** - „Zelený cieľ“: zvýšiť energetickú efektívnosť minimálne o 20% (SR 13%), takisto znížiť emisie skleníkových plynov aspoň o 20% (SR 14%) a zvýšiť využívanie obnoviteľných zdrojov energie o 20% (SR 11%),
- 4.) **Vzdelanie:** jednak znížiť počet osôb, ktoré predčasne ukončia štúdium až na 10% (SR 6%) a zvyšovanie podielu obyvateľov, ktoré majú ukončenú vysokoškolské vzdelanie na minimálnu úroveň 40% (SR rovnako),
- 5.) **Chudoba a sociálna inklúzia:** zníženie počtu obyvateľov EÚ, ktorí sú ohrození chudobou a sociálnym vylúčením o minimálnu hodnotu 20 miliónov obyvateľov, čo predstavuje zníženie o 25% oproti stavu v roku 2008 (SR proporcionálne 170 tisíc obyvateľov). (Kormančíková, B., 2013)

Najhlavnejším a najodvážnejším cieľom je práve cieľ v oblasti boja proti chudobe a sociálnemu vylúčeniu, kde v súčasnosti nastáva výrazné zhoršenie, napriek všetkým zavedeným opatreniam – od roku 2010 po rok 2015 vzrástol počet obyvateľov ohrozených

chudobou a sociálnym vylúčením už o vyše 1 milión, v súčasnosti má však tento počet obyvateľov klesajúci trend.

Na monitorovanie stavu a naplnenia týchto cieľov v oblasti chudoby a sociálneho vylúčenia využíva Stratégia Európa 2020 svoj vlastný trojdimenzionálny koncept merania, ktorý zachytáva ukazovateľ **AROEPE – miera rizika chudoby alebo sociálneho vylúčenia** (z angl. *at risk of poverty or social exclusion*). Tento ukazovateľ je k dispozícii práve v databáze EU-SILC a skladá sa z troch čiastkových ukazovateľov príjmovej chudoby, materiálnej deprivácie a vylúčenia z pracovného trhu:

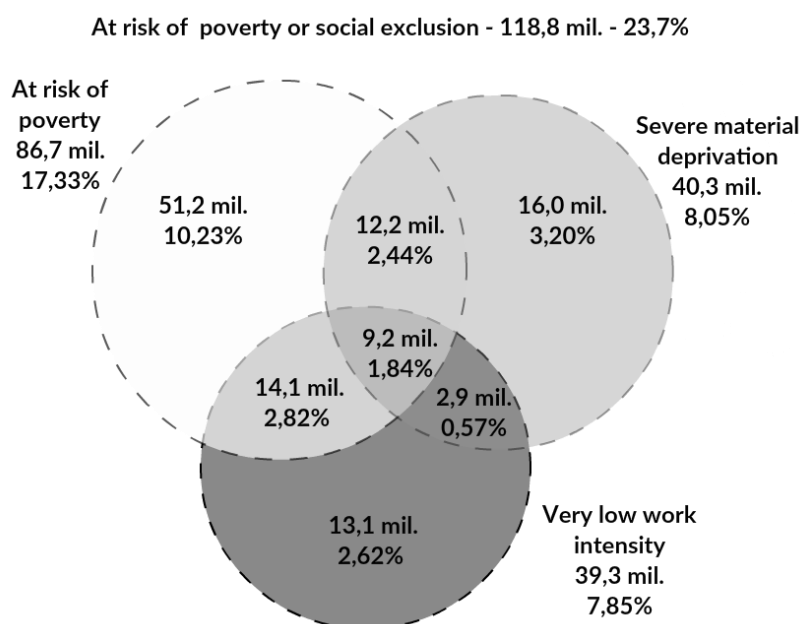
- **Miera rizika chudoby (AROP – at risk of poverty)**, ktorá predstavuje podiel obyvateľstva uvádzaný v percentách, ktorých ekvivalentný disponibilný príjem domácnosti je pod hranicou chudoby, ktorá predstavuje 60% ekvivalentného disponibilného príjmu danej krajiny. Hranica chudoby sa často udáva v parite kúpnej sily, aby bola medzinárodne porovnateľná. Ekvivalentný disponibilný príjem domácnosti predstavuje disponibilný príjem, ktorý je vydelený ekvivalentnou veľkosťou domácnosti. Na jej výpočet sa používa modifikovaná škála vytvorená OECD,
- **Miera závažnej materiálnej deprivácie (SMD – severe material deprivation)** vyjadruje podiel obyvateľstva uvádzaný takisto v percentách, ktoré majú vynútený nedostatok v najmenej štyroch z deviatich tzv. depriváčnych položiek, pretože si tieto položky nemôžu finančne dovoliť,
- **Miera veľmi nízkej pracovnej intenzity (VLWI – very low work intensity)**, ktorú definujeme ako podiel obyvateľstva, ktoré žije v domácnostiach s veľmi nízkou pracovnou intenzitou (menej než 20%) k počtu obyvateľstva vo veku od 0 do 59 rokov, pričom pracovnú intenzitu vyjadrujeme ako podiel medzi počtom mesiacov, ktoré členovia domácnosti vo veku od 18 do 59 rokov odpracovali k počtu mesiacov, ktoré sú celkovo schopní odpracovať.

Tieto dimenzie napomáhajú lepšiemu pochopeniu problému chudoby, nakoľko domácnosti sa často nachádzajú v ohrození chudoby alebo sociálneho vylúčenia vo viacerých alebo aj všetkých z týchto dimenzií. (Šoltés, E. – Šoltésová, T., 2016)

1.3 Príjmová chudoba, materiálna deprivácia a konzistentná chudoba

Za konzistentne chudobných považujeme také osoby, ktoré sú súčasne v riziku príjmovej chudoby aj závažnej materiálnej deprivácie. Tento koncept sa v Európskej únii začína sledovať čoraz viac ako pokus o zlepšenie a preklopenie doteraz dominantného vnímania chudoby len pomocou príjmovej chudoby. Materiálna deprivácia totiž v sebe zahŕňa nielen pohľad na to, koľko ľudí „zarábajú“ ale aj na to, ako sú schopní s príjmom pracovať a takisto aj isté subjektívne vnímanie chudoby.

Domácnosti, ktoré žijú v riziku konzistentnej chudoby, sú oproti tým, ktoré žijú len v riziku príjmovej chudoby alebo materiálnej deprivácie väčšinou z väčších rodín, majú nižšiu úroveň vzdelania a takisto veľmi slabú alebo žiadnu pracovnú intenzitu. Tiež subjektívne vnímajú svoje riziko chudoby horšie, ako iné domácnosti v tých istých podmienkach. Je teda vidieť, že je potrebné skúmať tento viacdimeznionálny koncept chudoby a zamerať sa aj na riešenie faktorov, ktoré zhoršujú nielen príjmovú chudobu, ale aj schopnosť domácností fungovať v normách spoločnosti. (Kis, A. – Gábos, A., 2016)



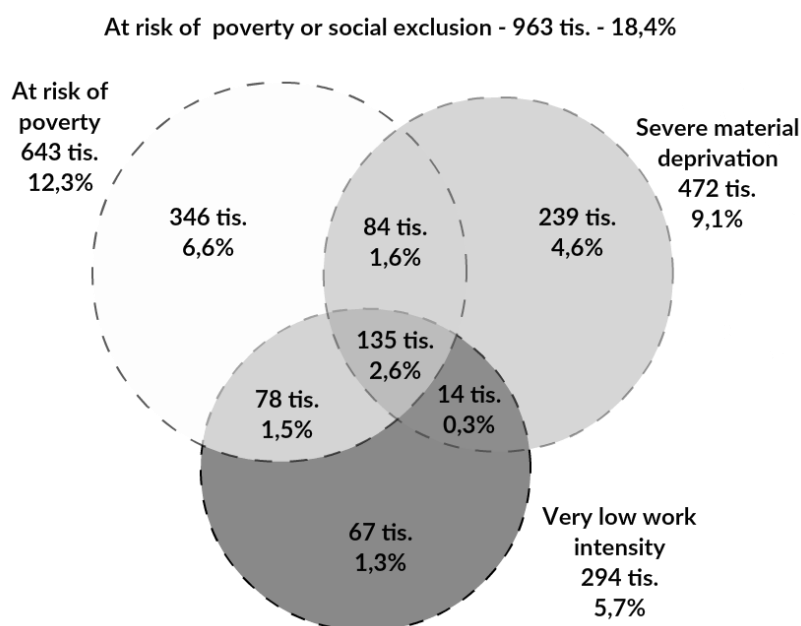
Obrázok 1.3.1 – Miera rizika chudoby a sociálneho vylúčenia v krajinách EÚ-28 (2015)

Zdroj: Eurostat, vlastné spracovanie (Creately)

Na obrázku 1.3.1 sú zobrazené jednotlivé dimenzie a ich prieniky súhrnne za 28 členských krajinách Európskej únie v roku 2015. Do Vennovho diagramu sú nanesené počty obyvateľov v miliónoch osôb, ktoré trpia jednotlivými dimenziami chudoby a sociálneho vylúčenia a takisto podiely obyvateľstva trpiace týmito dimenziami na

celkovom počte obyvateľov. V roku 2015 z celkového počtu 500,4 milióna obyvateľov bolo rizikom chudoby ohrozených 23,7% obyvateľstva, čo predstavuje 118,8 miliónov osôb. Viacerými dimenziami chudoby súčasne pritom bolo ohrozených až 38,4 milióna osôb, z toho všetkými tromi dimenziami súčasne 9,2 milióna osôb.

V riziku príjmovej chudoby sa nachádzalo až 86,7 miliónov obyvateľov (17,33% z celkového počtu obyvateľstva) a v riziku závažnej materiálnej deprivácie 40,3 miliónov obyvateľov (8,05%). Obidvoma týmito rizikami, teda konzistentnou chudobou, pritom bolo ohrozených 21,4 miliónov obyvateľov, čo tvorilo 4,28% populácie EÚ.



Obrázok 1.3.2 – Miera rizika a chudoby sociálneho vylúčenia na Slovensku (2015)

Zdroj: Eurostat, vlastné spracovanie (Creately)

Na obrázku 1.3.2 vidíme, že pomery na Slovensku sú iné – celkovo je v riziku chudoby alebo sociálneho vylúčenia menší pomer obyvateľstva, a to 18,4%, čo predstavuje 963 tisíc osôb. Takisto je iné rozloženie v rámci jednotlivých dimenzií. Menší podiel obyvateľstva je ohrozených príjmovou chudobou – 12,3% a naopak na Slovensku je v riziku závažnej materiálnej deprivácie väčší podiel osôb – 9,1%. V riziku konzistentnej chudoby je pritom 4,2% obyvateľstva, čo predstavuje 219 tisíc osôb a všetkými dimenziami chudoby a sociálneho vylúčenia je ohrozených 2,6% obyvateľstva, teda až 135 tisíc osôb.¹

¹ EUROSTAT. 2017. *Europe 2020 indicators - poverty and social exclusion*. [online]. Dostupné na internete: http://ec.europa.eu/eurostat/statistics-explained/index.php?title=Europe_2020_indicators_-_poverty_and_social_exclusion

1.3.1 Príjmová chudoba na Slovensku a v krajinách Európskej Únie

V tejto časti sa zameriame na dimenziu príjmovej chudoby. V nasledujúcej tabuľke 1.3.1.1 ilustrujeme hranicu chudoby jednotlivcov na Slovensku a v okolitých krajinách (okrem Ukrajiny, ktorá nie je súčasťou EÚ) a jej vývoj od roku 2010. Tieto hranice chudoby sú v Eurách a nemá veľký zmysel ich porovnávať v rôzne ekonomicky vyspelých krajinách vzhľadom na kúpnu silu. Môžeme vidieť, že najnižšiu hranicu chudoby má Maďarsko a naopak najvyššiu Rakúsko, čo je ale spôsobené omnoho väčšími nákladmi na bývanie, služby a pod.

Tabuľka 1.3.1.1 – Hranica chudoby vo vybraných krajinách EÚ

Country/Year	2010	2011	2012	2013	2014	2015
Hungary	2 544	2 696	2 818	2 670	2 707	2 734
Poland	2 643	3 015	3 036	3 098	3 202	3 333
Slovakia	3 670	3 784	4 156	4 042	4 086	4 158
Czech Republic	4 235	4 471	4 675	4 616	4 573	4 454
Austria	12 635	12 878	13 084	13 244	13 926	13 956

Zdroj: Eurostat, Vlastné spracovanie (MS Excel)

Oveľa väčšiu výpovednú hodnotu má miera rizika chudoby. V nasledujúcej tabuľke 1.3.1.2 vidíme mieru rizika chudoby po sociálnych transferoch na Slovensku a v okolitých krajinách od roku 2010 v percentách. Z týchto údajov už môžeme posúdiť stav ohrozenia príjmovou chudobou – zatiaľ čo Česká republika mala v roku 2015 najmenší podiel obyvateľov ohrozených touto dimenziou chudoby a sociálneho vylúčenia, najhoršie na tom bolo Poľsko, pričom ale obidve krajiny v posledných rokoch zaznamenávajú nárast miery rizika chudoby. Naopak na Slovensku a v Rakúsku nastáva dlhodobý pokles tejto miery.

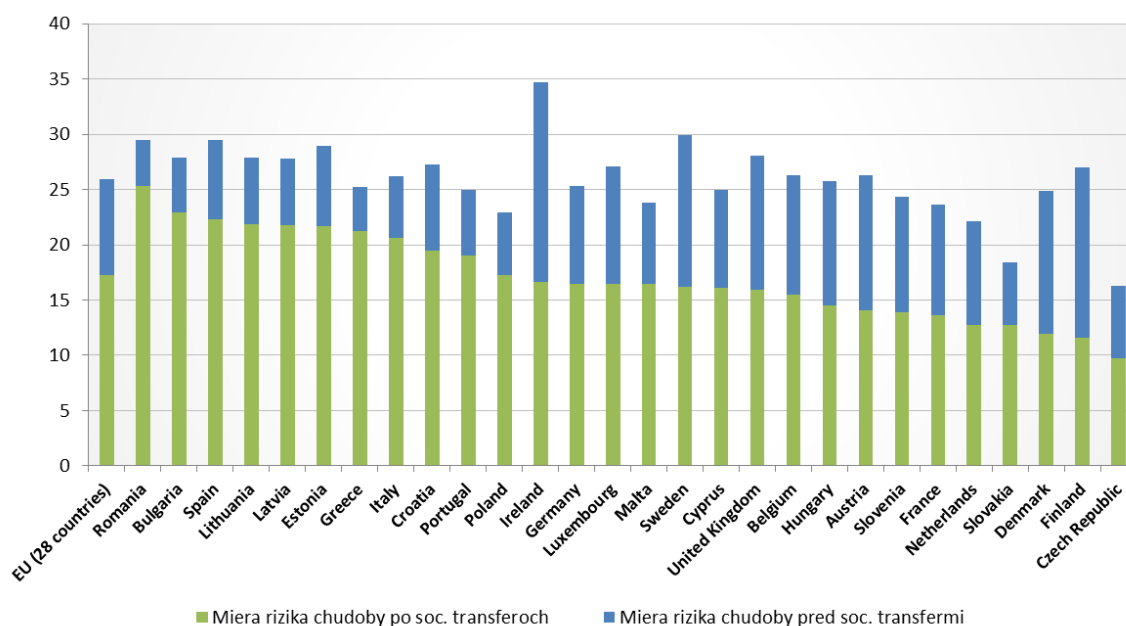
Tabuľka 1.3.1.2 – Miera rizika chudoby vo vybraných krajinách EÚ

Country/Year	2010	2011	2012	2013	2014	2015
Czech Republic	9	9,8	9,6	8,6	9,7	9,7
Slovakia	12	13	13,2	12,8	12,6	12,3
Austria	14,7	14,5	14,4	14,4	14,1	13,9
Hungary	12,3	14,1	14,3	15	15	14,9
Poland	17,6	17,7	17,1	17,3	17	17,6

Zdroj: Eurostat, Vlastné spracovanie (MS Excel)

Zatiaľ čo v predchádzajúcej tabuľke sú uvedené miery rizika chudoby po sociálnych transferoch, je takisto zaujímavé sledovať, aké boli miery rizika chudoby pred sociálnymi transfermi a ako opatrenia jednotlivých krajín v oblasti sociálnej politiky

zachytené v sociálnych transferoch pomohli týmto krajinám redukovať riziko príjmovej chudoby. Toto je ilustrované na nasledujúcom grafe. Vo všetkých krajinách došlo k výraznému zníženiu miery rizika chudoby, pričom najviac tieto opatrenia pomohli v Írsku – jedná sa najmä o dôchodky, dávky v nezamestnanosti, rodinné dávky, príspevky na bývanie či vzdelanie a inú sociálnu pomoc. Na Slovensku tieto opatrenia pomohli znížiť mieru chudoby až o 5,7 percentuálneho bodu – z 18,4% na 12,3%.

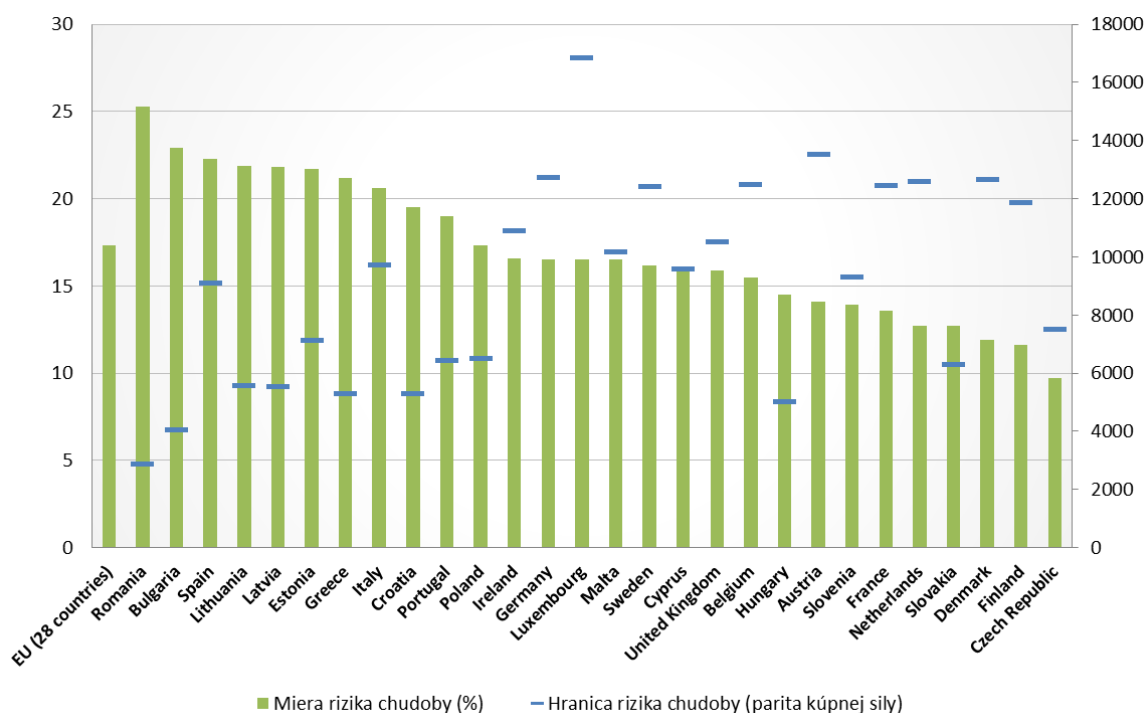


Obrázok 1.3.1.1 – Miera rizika chudoby po sociálnych transferoch v krajinách EÚ (2015)

Zdroj: Eurostat, vlastné spracovanie (MS Excel)

Napokon v obrázku 1.3.1.2 môžeme pozorovať miery rizika chudoby pre jednotlivé krajiny vzhľadom na národné hranice rizika chudoby, ktoré sú už vyjadrené v parite kúpnej sily. Takéto porovnanie je omnoho užitočnejšie z hľadiska národného porovnania. Pozorujeme, že Slovensko má vzhľadom na nižšiu hranicu rizika chudoby vysoký podiel obyvateľstva, ktoré je chudobou ohrozené. V najhoršej situácii sa nachádza Rumunsko – aj napriek nízkej hranici rizika chudoby má veľmi vysoký podiel obyvateľstva ohrozeného príjmovou chudobou. Naopak v Luxembursku je hranica rizika chudoby vysoká, napriek tomu podiel obyvateľov ohrozených chudobou je vzhľadom na túto hranicu nízky.²

² EUROSTAT. 2017. *Europe 2020 indicators - poverty and social exclusion*. [online]. Dostupné na internete: http://ec.europa.eu/eurostat/statistics-explained/index.php?title=Europe_2020_indicators_-_poverty_and_social_exclusion



Obrázok 1.3.1.2 – Miera a hranica chudoby v krajinách EÚ (2015)

Zdroj: Eurostat, vlastné spracovanie (MS Excel)

1.3.2 Materiálna deprivácia na Slovensku a v krajinách Európskej Únie

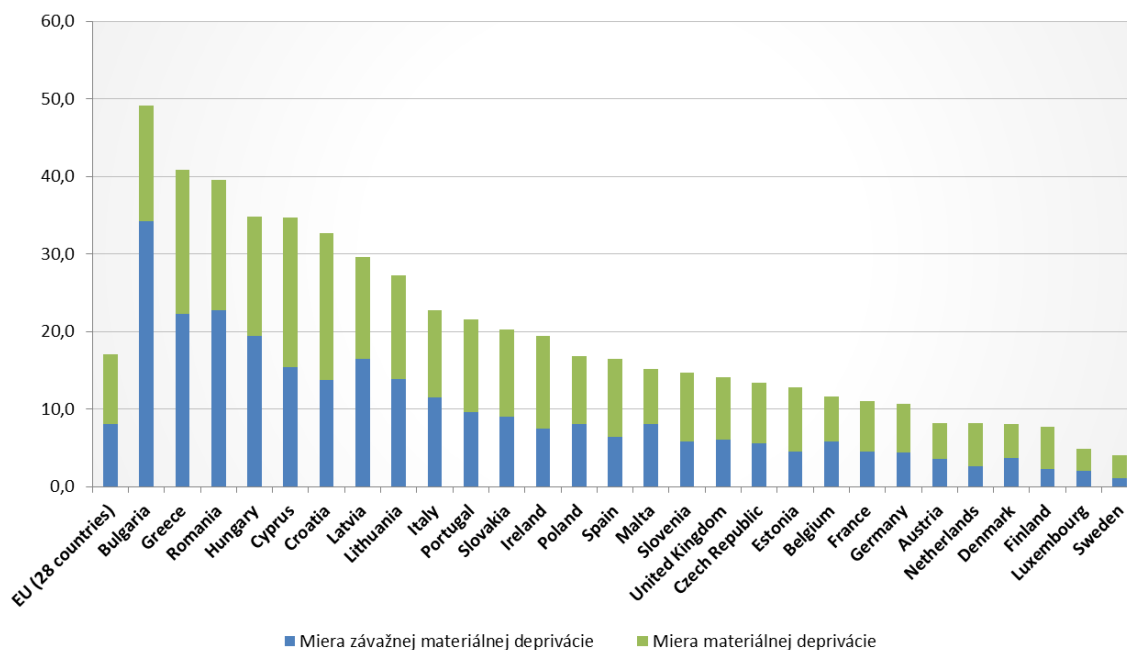
V tabuľke 1.3.2.1 vidíme mieru závažnej materiálnej deprivácie v percentách. Na základe hodnôt tejto miery a poradia krajín z hľadiska ohrozenia touto dimenziou môžeme vidieť, že sa jedná o odlišný problém ako je problém ohrozenia príjmovou chudobou. Zatiaľ čo Slovensko bolo v rámci ohrozenia príjmovou chudobou na 2. mieste z okolitých krajín, ohrozenie závažnou materiálnou depriváciou je tu podstatne väčšie. Najhoršie z vybraných krajín je na tom Maďarsko a to s veľkým odskokom – v roku 2015 až 19,4% obyvateľstva bolo závažne materiálne deprivované. Naopak veľmi dobre je na tom Rakúsko, kde je závažnou materiálnou depriváciou ohrozených len 3,6% obyvateľov.

Tabuľka 1.3.2.1 – Miera závažnej materiálnej deprivácie vo vybraných krajinách EÚ

Country/Year	2010	2011	2012	2013	2014	2015
Austria	4,3	4,0	4,0	4,2	4,0	3,6
Czech Republic	6,2	6,1	6,6	6,6	6,7	5,6
Poland	14,2	13,0	13,5	11,9	10,4	8,1
Slovakia	11,4	10,6	10,5	10,2	9,9	9,0
Hungary	21,6	23,4	26,3	27,8	24,0	19,4

Zdroj: Eurostat, vlastné spracovanie (MS Excel)

Miera závažnej materiálnej deprivácie je z krajín Európskej Únie najvyššia v Bulharsku a naopak najlepšie je na tom v tomto ohľade Švédsko. Na Slovensku je miera závažnej materiálnej deprivácie vyššia ako priemer EÚ, v posledných rokoch však má klesajúci trend a pomaly sa približuje k vyspelejším krajinám. Na nasledujúcom obrázku je ilustrovaná jednak miera závažnej materiálnej deprivácie, kedy si jednotlivec nemôže dovoliť viac ako 4 depriváčné položky a miera materiálnej deprivácie, kedy si jednotlivec nemôže dovoliť 3 depriváčné položky.³



Obrázok 1.3.2.1 – Miera materiálnej deprivácie v krajinách EÚ (2015)

Zdroj: Eurostat, vlastné spracovanie (MS Excel)

1.3.3 Konzistentná chudoba na Slovensku a v krajinách Európskej Únie

Slovensko je na tom z pohľadu ohrozenia konzistentnou chudobou spomedzi susedných krajín dobre – ohrozených je 4,2% obyvateľstva, rovnako ako v Poľsku. O 1,4 percentuálneho bodu je na tom lepšie Česká Republika a o 2,3 percentuálneho bodu Rakúsko. Najhoršie je na tom Maďarsko, kde je v riziku konzistentnej chudoby omnoho väčší podiel obyvateľov – až 8,1% osôb.⁴

³ EUROSTAT. 2017. *Material deprivation and low work intensity statistics*. [online]. Dostupné na internete: http://ec.europa.eu/eurostat/statistics-explained/index.php/Material_deprivation_and_low_work_intensity_statistics

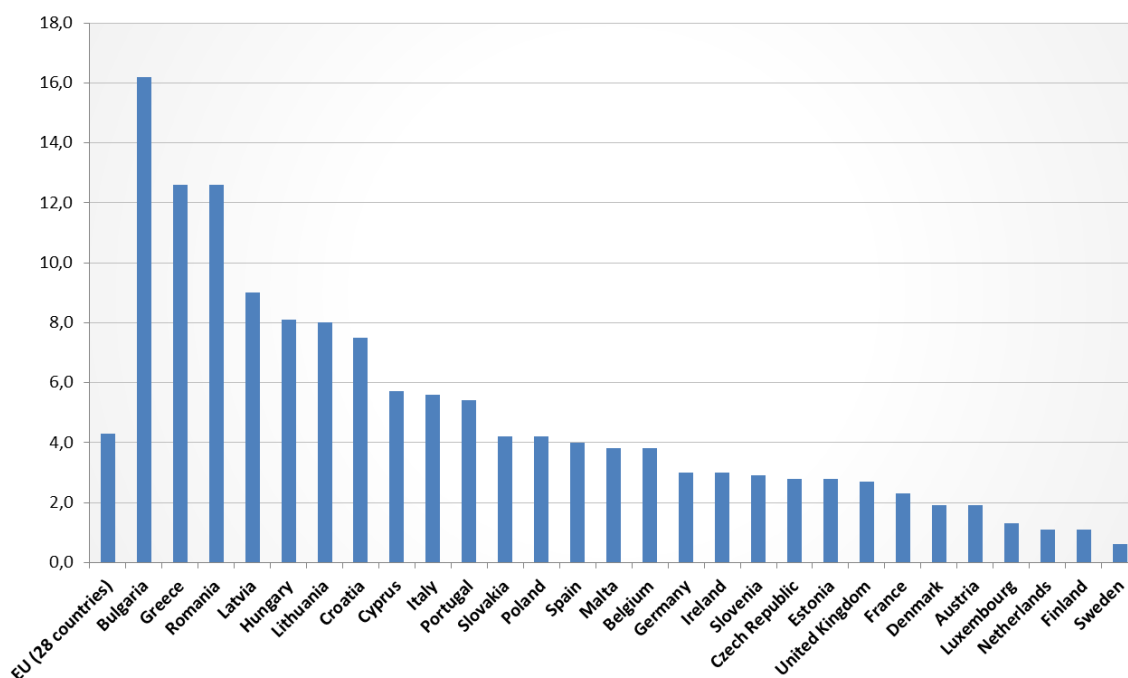
⁴ EUROSTAT. 2017. *Europe 2020 indicators - poverty and social exclusion*. [online]. Dostupné na internete: http://ec.europa.eu/eurostat/statistics-explained/index.php?title=Europe_2020_indicators_-_poverty_and_social_exclusion

Tabuľka 1.3.3.1 – Miera konzistentnej chudoby vo vybraných krajinách EÚ

Country/Year	2010	2011	2012	2013	2014	2015
Austria	2,3	2,2	2,2	2,3	1,8	1,9
Czech Republic	2,7	2,4	2,8	2,7	3,1	2,8
Poland	6,1	5,5	5,6	5,1	4,7	4,2
Slovakia	4,6	4,4	4,6	4,7	5,0	4,2
Hungary	6,7	8,7	9,6	10,3	9,4	8,1

Zdroj: Eurostat, vlastné spracovanie (MS Excel)

Aj v rámci ostatných krajín Európskej Únie je na tom Slovensko v rámci konzistentnej chudoby veľmi dobre, čím spolu s Českom tvorí výnimku v rámci nových členských krajín, ktoré tvoria skupinu krajín s najvyšším ohrozením touto kombináciou príjmovej chudoby a materiálnej deprivácie. Alarmujúco vysoké hodnoty tohto ukazovateľa môžeme vidieť v pobaltských krajinách, hlavne Bulharsku a Rumunsku a zvýšené hodnoty tiež môžeme pozorovať v ďalších južných štátoch – Cyprus, Taliansko a Grécko. (Kis, A. – Gábos, A., 2016)



Obrázok 1.3.3.1 - Miera konzistentnej chudoby v krajinách EÚ (2015)

Zdroj: Eurostat, vlastné spracovanie (MS Excel)

2 Cieľ práce

Hlavným cieľom práce je priblížiť problematiku príjmovej chudoby, materiálnej deprivácie a konzistentnej chudoby na Slovensku. Môžeme totiž vidieť, že na Slovensku je chudoba pomerne rozsiahlym a takisto diverzifikovaným problémom – výrazné sú hlavne rozdiely medzi jednotlivými krajmi. Takisto musíme abstrahovať od vnímania chudoby len ako problému nedostatočného príjmu, ale musíme sa zamerať aj na vnímanie ostatných konceptov ako je konzistentná chudoba, ktorá v sebe zahŕňa aj materiálnu depriváciu či istý subjektívny pohľad domácností na svoju chudobu.

Preto sme si na naplnenie hlavného cieľa zvolili nasledujúce čiastkové ciele:

1. **Opísať problém chudoby** a rôzne spôsoby jej vnímania a merania a takisto **priblížiť piliere a ciele**, ktoré si Európska Únia či Slovensko v súčasnosti kladie v oblasti riešenia problémov chudoby a sociálneho vylúčenia a **opísať súčasnú situáciu na Slovensku** a v krajinách Európskej Únie,
2. Pomocou aparátu logistickej regresie uskutočniť **štatistické analýzy** príjmovej chudoby, materiálnej deprivácie a konzistentnej chudoby slovenských domácností,
3. **Kvantifikovať vplyv relevantných faktorov na konzistentnú chudobu** slovenských domácností a ukázať tak rozdielnosti panujúce v chudobe a materiálnej deprivácii na našom území. Odhadnúť **pravdepodobnosť ohrozenia konzistentnou chudobou** a jej zložiek v rôznych skupinách slovenských domácností.

Dúfame, že naplnenie týchto cieľov pomôže čitateľovi pochopiť dôležitosť danej problematiky, či osvetlí situáciu v tejto oblasti na Slovensku alebo aj v Európskej únii. Výsledky diplomovej práce je možné využiť v sociálnej politike štátu, ktorá by sa mala sústrediť na rizikové skupiny domácností identifikované v diplomovej práci. Odhadnuté modely logistickej regresie budú vhodné na predikciu pravdepodobností hrozby predmetnými dimenziami chudoby a sociálneho vylúčenia, či konzistentnou chudobou v rôznych skupinách domácností. Výslednými modelmi bude možné porovnať rizikovosť najohrozenejších skupín domácností s ostatnými skupinami domácností.

3 Metodika a metódy skúmania záverečnej práce

Metódy regresnej a korelačnej analýzy sú používané už od 19. storočia, kedy Carl Friedrich Gauss publikoval metódu najmenších štvorcov a ku koncu storočia bol pojem regresia oficiálne zavedený Francisom Galtonom. V štatistike sa totiž ukázalo ako veľmi užitočné skúmať, ako je jedna premenná spojená s inými premennými. Práve regresná analýza skúma vzťah, podľa ktorého sa mení závislá (vysvetľovaná) premenná pri zmenách hodnôt nezávislých (vysvetľujúcich premenných) a snaží sa navrhnúť čo najlepší, najúspornejší a zároveň zmysluplný model tohto vzťahu. Kedysi trvalo dostať výsledky z jednej regresie až 24 hodín, ale dnes vďaka pokroku vo výpočtových technológiách patrí regresná a korelačná analýza aj napriek existencii pokročilejších metód k nenahraditeľným metódam v štatistike. (Terek, M. - Horníková, A. - Labudová, V., 2010)

Regresné metódy sa dajú využiť na analýzu problémov najrôznejšej povahy. Najčastejšou metódou, ktorá nám napadne pri regresii je lineárna regresia, ktorá skúma lineárny vzťah medzi premennými. V tejto práci sa budeme zaoberať logistickou regresiou, ktorá bola vyvinutá v roku 1958 anglickým štatistikom Davidom Coxom, aj keď už predtým sa touto tézou zaoberali aj iné veľké mysle ako Adolphe Quetelet alebo Ronald Fisher. Logistická regresia nám umožňuje skúmať prípad, kedy vysvetľovaná premenná je kategóriálna. Pričom kedysi logistická regresia začala ako metóda, ktorá sa využívala len v medicíne, dnes patrí medzi štandardné a často využívané metódy v mnohých oblastiach západoeurópskej a americkej vedy, ako je riadenie kreditného rizika, poisťovníctvo, marketing a výskumy trhu, ale aj spoločenské vedy. (Cramer, J. S., 2003)

3.1 Logistická regresia

Logistická regresia je metóda, pomocou ktorej môžeme skúmať jednostrannú závislosť medzi vysvetľujúcimi premennými a kategóriálnou vysvetľovanou premennou. Vysvetľujúce premenné môžu byť kategóriálneho aj spojitého typu. Odlišnosťou od lineárneho regresného modelu je práve charakter vysvetľovanej premennej (pri lineárnej regresii je vysvetľovaná premenná spojitá, teda zatiaľ čo pri lineárnej regresii vysvetľovaná premenná nadobúda neobmedzený počet hodnôt, pri logistickej regresii ide o obmedzený počet hodnôt). (Řeháková, B., 2000)

Práve z charakteru vysvetľovanej premennej vyplýva hlavný rozdiel lineárnej a logistickej regresie – model logistickej regresie odhaduje pravdepodobnosť toho, či sa nejaká udalosť stane, alebo nie. To platí pre všetky typy logistickej regresie, ktoré rozlišujeme podľa typu závislej kategoriálnej premennej:

- a) **Binárna logistická regresia** – ide o typ logistickej regresie, kedy kategoriálna vysvetľovaná premenná nadobúda len dve obmeny – absencia alebo prítomnosť javu (0/1) – splatenie/nesplatenie úveru, zákazník kúpi/nekúpi produkt a pod.
- b) **Multinomická logistická regresia** – pri tomto type kategoriálna vysvetľovaná premenná nadobúda viac ako dve obmeny, medzi ktorými existuje odlišnosť – splatenie/oneskorené splatenie/nesplatenie úveru a pod.
- c) **Ordinárna logistická regresia** – v tomto prípade kategoriálna vysvetľovaná premenná nadobúda viaceré obmeny, medzi ktorými nielen existuje odlišnosť, ale takisto má zmysel zoradenie týchto obmien, napr. spokojnosť zákazníka – nespokojný, skôr nespokojný, skôr spokojný, spokojný.

V ďalších častiach tejto práce najprv vysvetlíme základné princípy logistickej regresie, ktoré je užitočné ilustrovať pomocou binárneho modelu a napokon načrtneme aj prechod na multinomický model, ktorý je prirodzeným zovšeobecnením binárneho. (Pecáková, I., 2007)

3.2 Logistická regresia s binárnou vysvetľovanou premennou

V prípade lineárnej regresie modelujeme strednú hodnotu spojitej náhodnej premennej Y , ktorá je vyjadrená lineárnou funkciou vysvetľujúcich premenných X_j s parametrami β_j pre $j = 0, 1, 2, \dots, k$ (k – počet regresných koeficientov modelu):

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} = \beta_0 + \sum_{j=1}^k \beta_j x_{ij}, i = 1, 2, \dots, n$$

Výraz $\beta_0 + \sum_{j=1}^k \beta_j x_{ij}$ nadobúda hodnoty z intervalu $(-\infty; \infty)$, čo je pre strednú hodnotu spojitej premennej v poriadku.

Pri logistickej regresii so závislou premennou, ktorá je binárneho typu, však uvažujeme s premennou, ktorá nadobúda len dve hodnoty. Takáto náhodná premenná má Alternatívne rozdelenie pravdepodobnosti $A(\pi)$. Stav, kedy náhodný jav A nastane, vyjadrujeme hodnotou náhodnej premennej rovnajúcu sa 1. Naopak, ak náhodný jav A nenastane, náhodná premenná nadobúda hodnotu 0. Ďalej pravdepodobnosť, že jav A nastane, označujeme parametrom π . Pre náhodnú premennú Y potom platí, že $P(1) = P(y_i = 1) = \pi_i$ a analogicky $P(0) = P(y_i = 0) = 1 - \pi_i$. (Terek, M. - Horníková, A. - Labudová, V., 2010)

Pre strednú hodnotu náhodnej premennej s alternatívnym rozdelením platí vzťah $E(y_i) = \pi_i \cdot 1 + (1 - \pi_i) \cdot 0 = \pi_i$. Tu môžeme vidieť, že pre regresný model binárnej závislej premennej je vhodné uvažovať s pravdepodobnosťou π_i . Tu však ale narážame na problém, pre ktorý nie je možné pre takúto analýzu použiť lineárnu regresiu. Keďže hodnota pravdepodobnosti leží v intervale $\langle 0; 1 \rangle$, tento interval nemôže pokryť všetky hodnoty, ktoré môže dosahovať výraz $\beta_0 + \sum_{j=1}^k \beta_j x_{ij}$. (Allison, P. D., 2012)

Tento problém rieši väzbová funkcia $\text{logit}(\pi_i)$, ktorá transformuje hodnoty pravdepodobnosti tak, aby neboli z obmedzeného intervalu, preto sa model logistickej regresie zvykne nazývať aj logitovým modelom. Prvým krokom je odstránenie horného obmedzenia pomocou transformácie pravdepodobnosti na *šancu* (anglicky *odds*), ktorú definujeme ako podiel pravdepodobnosti, že sledovaný jav nastane a pravdepodobnosti, že sledovaný jav nenastane:

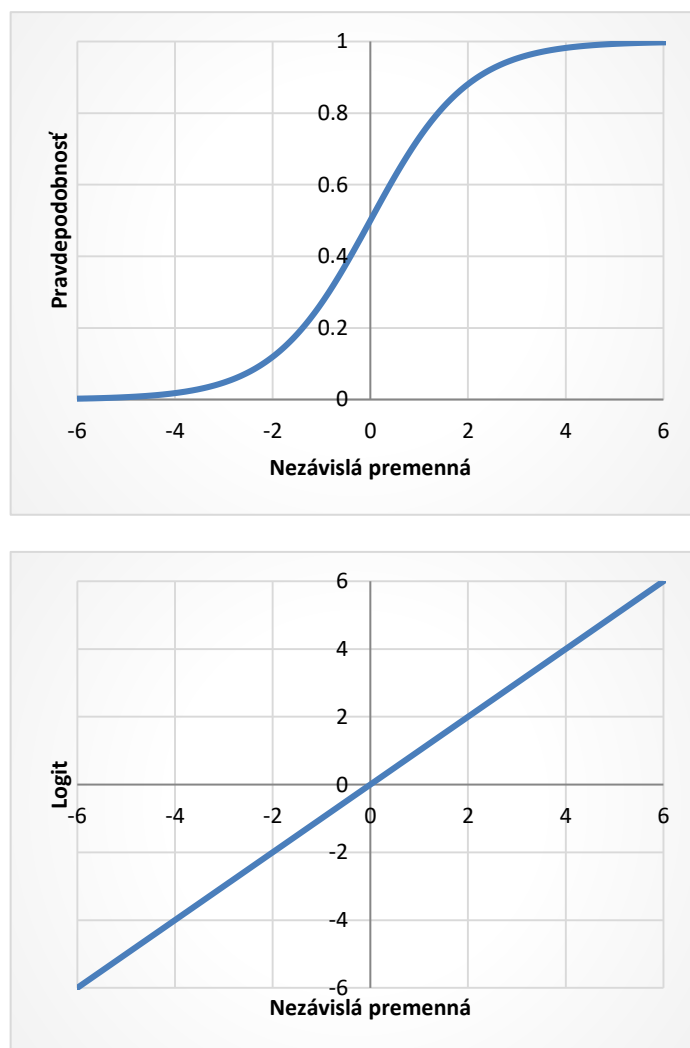
$$\text{šanca} = \frac{\pi_i}{1 - \pi_i}$$

Následne uvažujeme s prirodzeným logaritmom šance, čo nám umožňuje odstrániť aj dolné obmedzenie a výraz $\text{logit}(\pi_i) = \ln \frac{\pi_i}{1 - \pi_i}$ už nadobúda hodnoty z intervalu $\langle -\infty; \infty \rangle$. V praxi sa používajú aj iné väzbové funkcie (probit, complementary log-log), ďalej však budeme pracovať práve s funkciou logit, pre ktorú *model logistickej regresie* môžeme vyjadriť nasledovne:

$$\ln \frac{\pi_i}{1 - \pi_i} = \beta_0 + \sum_{j=1}^k \beta_j x_{ij}$$

Logitová funkcia takisto rieši problém lineárneho vzťahu ľavej a pravej časti vzťahu pre model regresie. Transformuje nelineárny vzťah pravdepodobnosti a hodnôt vysvetľujúcich

premenných (v prípade jednej vysvetľujúcej premennej je vyjadrením tohto vzťahu symetrická s-krivka) na lineárny vzťah logaritmu šance s hodnotami vysvetľujúcich premenných. (Rublíková, E. – Labudová, V. – Sandtnerová, S., 2009)



Obr. 3.2.1 – Logitová transformácia

Zdroj: Vlastné spracovanie (MS Excel)

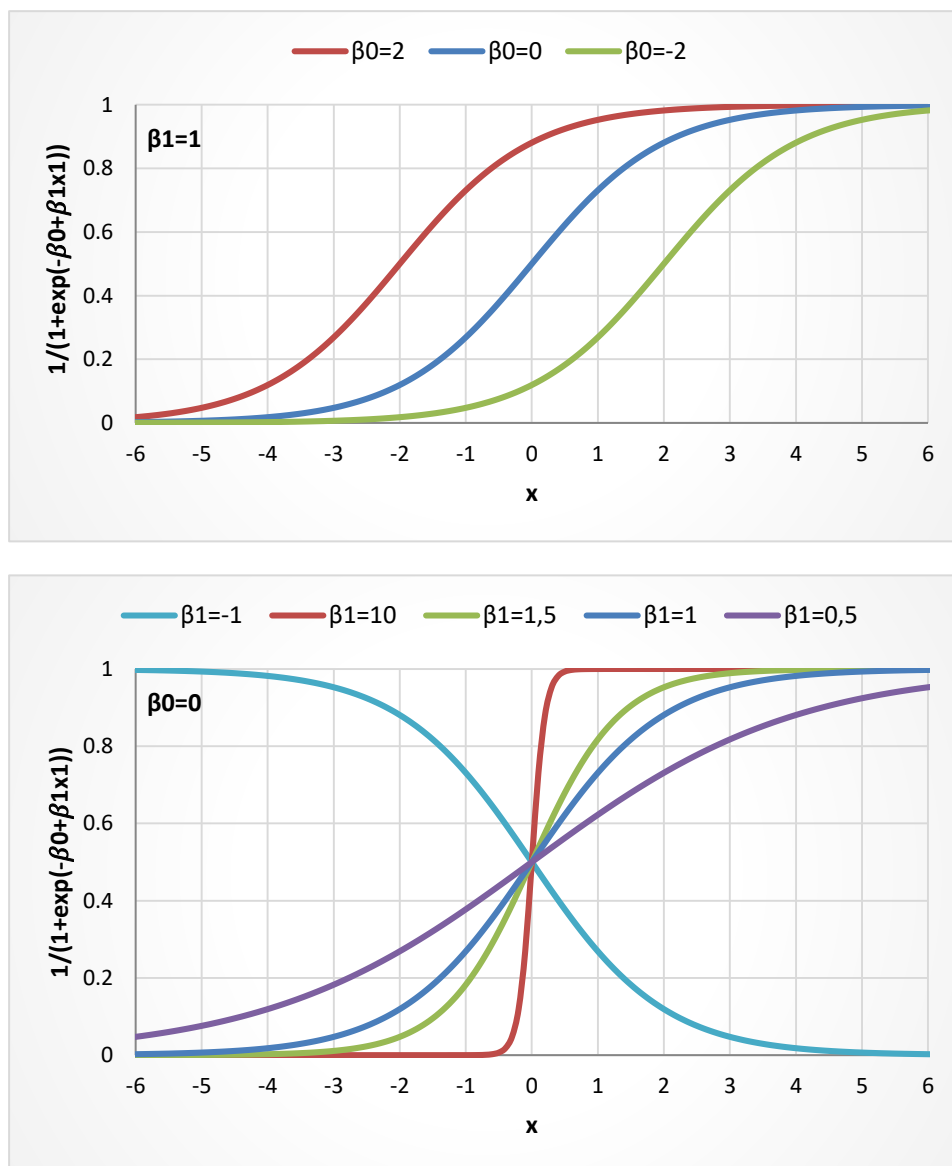
Z modelu logistickej regresie vieme odvodiť aj ďalšie užitočné vzťahy, ktoré sa využívajú pri analýzach. Pomocou využitia inverznej exponenciálnej funkcie môžeme z rovnice pre model logistickej regresie odvodiť vzťah pre šancu:

$$\frac{\pi_i}{1 - \pi_i} = \exp(\beta_0 + \sum_{j=1}^k \beta_j x_{ij})$$

Tento výraz potom vyjadruje, koľkokrát väčšia je pravdepodobnosť, že sledovaný jav nastane, než pravdepodobnosť, že sledovaný jav nenastane. Ďalšími úpravami sa

dostávame k výrazu pre *podmienenu strednú hodnotu* π_i , teda pre pravdepodobnosť nastania sledovaného javu. Tu môžeme vidieť, že vzťah je nelineárny:

$$\pi_i = \frac{\exp(\beta_0 + \sum_{j=1}^k \beta_j x_{ij})}{1 + \exp(\beta_0 + \sum_{j=1}^k \beta_j x_{ij})} = \frac{1}{1 + \exp(-\beta_0 - \sum_{j=1}^k \beta_j x_{ij})}$$



Obr. 3.2.2 – S-krivka pri rôznych hodnotách regresných parametrov

Zdroj: Vlastné spracovanie (MS Excel)

Najdôležitejším vzťahom, ktorý používame na interpretáciu modelu logistickej regresie, je *pomer šancí* (anglicky *odds ratio* alebo skráteno *OR*):

$$OR = \frac{\text{šanca 1}}{\text{šanca 2}} = \exp(\beta_j)$$

Pomer šancí sa interpretuje rôzne vzhľadom na druh a počet vysvetľujúcich premenných. Všeobecne, pre viac vysvetľujúcich premenných môžu nastať dva druhy prípadov:

- a) **Vysvetľujúca premenná, ktorú interpretujeme, je spojitá** – Pomer šancí vyjadruje, koľkonásobne sa zväčší/zmenší šanca, že sledovaný jav nastane v porovnaní so šancou, že sledovaný jav nenastane, ak hodnota vysvetľujúcej premennej, ktorú interpretujeme, vzrastie o 1 meraciu jednotku pri nezmenených hodnotách zvyšných vysvetľujúcich premenných (podmienka *ceteris paribus*).
- b) **Vysvetľujúca premenná, ktorú interpretujeme, je kategóriálna** – Pri takejto interpretácii je najprv potrebné stanoviť, ktorá z kategórií premennej, ktorú interpretujeme, bude referenčná. Pomer šancí potom vyjadruje, koľkonásobne je väčšia/menšia šanca nastania javu v porovnaní so šancou nenastania javu v sledovanej kategórii oproti referenčnej kategórii, pri nezmenených hodnotách zvyšných vysvetľujúcich premenných. (Terek, M. - Horníková, A. - Labudová, V., 2010)

3.2.1 Odhad parametrov pre model logistickej regresie

Na odhad parametrov modelu logistickej regresie sa dá použiť viacero metód, v súčasnosti sa ale najviac používa *metóda maximálnej vierohodnosti*. Zatiaľ čo pri odhade parametrov lineárneho regresného modelu používame metódu najmenších štvorcov, pre logistickú regresiu nám táto metóda neponúka až také dobré výsledky kvôli charakteru vysvetľovanej premennej. Použitiu metódy maximálnej vierohodnosti aj napriek jej dobrým vlastnostiam zase dlho bránila jej výpočtová náročnosť, tento problém ale odstránil rýchly rozvoj informačných technológií.

Metóda maximálnej vierohodnosti je používaná na odhadovanie parametrov v širokom spektre štatistických modelov. Hlavným dôvodom tejto „popularity“ je, že odhady parametrov získané touto metódou majú v súboroch veľkého rozsahu dobré vlastnosti. Odhady získané metódou maximálnej vierohodnosti sú *konzistentné*, *asymptoticky výdatné* a *asymptoticky normálne*. Konzistentnosť znamená, že čím je súbor väčší, tým je väčšia aj pravdepodobnosť, že odhad parametra je blízky jeho skutočnej hodnote. Asymptotická výdatnosť znamená, že pri veľkých súboroch je rozptyl odhadu aspoň tak malý alebo menší, ako rozptyl odhadu pri použití akejkoľvek inej metódy na odhad parametrov. Nakoniec asymptotická normalita hovorí, že odhad pre veľký súbor má

približne normálne rozdelenie. Táto vlastnosť je veľmi dôležitá, nakoľko potom pre výpočet intervalov spoľahlivosti a P - hodnôt môžeme použiť *normálne* a *chí-kvadrát rozdelenie*. (Allison, P. D., 2012)

Základným princípom metódy maximálnej vierohodnosti je odhadnúť parametre modelu tak, aby maximalizoval pravdepodobnosť toho, že v súbore, s ktorým pracujeme, budú práve tie hodnoty premennej y_i , ktoré sme aj v skutočnosti pozorovali (čiže maximalizovať zhodu modelu so skutočnosťou). Prvým krokom v tomto postupe je konštrukcia *funkcie maximálnej vierohodnosti* a druhým jej maximalizácia. Na konštrukciu funkcie maximálnej vierohodnosti potrebujeme poznať pravdepodobnostné rozdelenie vysvetľovanej premennej a jej vzťah s vysvetľujúcimi premennými. Premenná y_i má alternatívne rozdelenie, pre ktoré poznáme pravdepodobnostnú funkciu v tvare $P(y, \pi) = \pi_i^{y_i}(1 - \pi_i)^{1-y_i}$. Pri odhade parametrov sa teda hľadá maximum takejto funkcie vierohodnosti $L(\pi, \mathbf{y})$:

$$L(\pi, \mathbf{y}) = \prod_{i=1}^n \pi_i^{y_i}(1 - \pi_i)^{1-y_i}$$

Funkciu vierohodnosti môžeme zapísať ako súčin pravdepodobnostných funkcií vďaka predpokladu, že pozorovania premennej y_i sú nezávislé. Pri hľadaní maxima sa v praxi využíva logaritmus funkcie vierohodnosti – logaritmická vierohodnostná funkcia. (Rublíková, E. – Labudová, V. – Sandtnerová, S., 2009)

3.2.2 Testovanie štatistickej významnosti modelu ako celku

Na overenie štatistickej významnosti modelu logistickej regresie sa najčastejšie využíva *test pomerom vierohodností (likelihood ratio test)* a *Waldov test*. Oba testy majú rovnaké nulové a alternatívne hypotézy. Pri malých a stredných rozsahoch súborov sa preferuje test pomerom vierohodností. Pri veľkých rozsahoch dávajú oba testy takmer rovnaké výsledky. V mnohých softvéroch sa používa aj *Rauov skórový test*, opäť s rovnakou sadou hypotéz, pri ktorom nie je potrebné na odhad parametrov logistickej regresie použiť metódu maximálnej vierohodnosti. Ďalej rozoberieme test pomerom vierohodností, ktorý je analógiou pre F -test používaný pri overení štatistickej významnosti modelu lineárnej regresie.

Pri tomto teste sa overuje platnosť nulovej hypotézy oproti alternatívnej hypotéze:

H_0 : Model logistickej regresie *nie* je štatisticky významný – $\beta_1, \beta_2, \dots, \beta_k = 0$

H_1 : Model logistickej regresie *je* štatisticky významný – $\exists \beta_j \neq 0, j = 1, 2, \dots, k$

Ako testovaciu štatistiku v teste používame štatistiku G - *pomer vierohodností*:

$G = DEV(\text{model bez premenných}) - DEV(\text{model s premennými})$, kde DEV je deviancia – analógia ku súčtu štvorcov rezíduí v lineárnom regresnom modeli. Vzorec pre DEV je nasledujúci:

$$DEV = -2 \ln \left[\frac{L(\text{testovaného modelu})}{L(\text{saturovaného modelu})} \right]$$

kde L je funkcia vierohodnosti a výraz v hranatej zátvorke sa označuje ako vierohodnostný pomer. Pre vierohodnostnú funkciu saturovaného modelu v prípade kategoriálnej vysvetľovanej premennej platí, že nadobúda hodnotu 1. Vzorec pre devianciu sa v tom prípade zjednoduší – $DEV = -2 \ln L(\text{testovaného modelu})$. Potom môžeme prepísať aj vzťah pre testovaciu štatistiku G do nasledujúcej formy:

$$G = 2 \ln L(\text{model s premennými}) - 2 \ln L(\text{model bez premenných})$$

Tento test teda môžeme použiť aj ako test štatistickej významnosti modelu ako celku, kedy model s premennými predstavuje model so všetkými vysvetľujúcimi premennými a model bez premenných len model s lokujúcou konštantou β_0 , ale aj ako test pre štatistickú významnosť jednej vysvetľujúcej premennej.

Pri platnosti nulovej hypotézy má pomer vierohodností chí-kvadrát rozdelenie s počtom stupňov voľnosti ν , ktorý sa rovná rozdielu počtu parametrov testovaného modelu a saturovaného modelu. Ak testujeme významnosť modelu ako celku, stupne voľnosti určíme ako rozdiel počtu parametrov modelu so všetkými uvažovanými vysvetľujúcimi premennými ($k+1$ parametrov) a počtu parametrov modelu bez vysvetľujúcich premenných (1 parameter – lokujúca konštanta). V takomto špeciálnom prípade tohto testu sa počet stupňov voľnosti rovná k , teda počtu regresných koeficientov. Nulovú hypotézu potom zamietame, ak $G > \chi^2_{1-\alpha}(\nu)$. (Terek, M. - Horníková, A. - Labudová, V., 2010)

3.2.3 Testovanie štatistickej významnosti vplyvu vysvetľujúcich premenných

Na otestovanie štatistickej významnosti vplyvu kategoriálnej vysvetľujúcej premennej sa používa *Waldov test*. Testuje nulovú hypotézu, ktorá hovorí, že vybraná vysvetľujúca premenná nemá štatisticky významný vplyv na pravdepodobnosť výskytu skúmaného javu oproti negujúcej alternatívnej hypotéze. Na overenie tejto hypotézy sa používa *Waldova štatistika*:

$$W = \hat{\beta}^T \cdot S_b^{-1} \cdot \hat{\beta}$$

kde $\hat{\beta}$ je vektor odhadov regresných koeficientov, ktoré stoja pri umelých premenných vytvorených pre vybranú vysvetľujúcu premennú a S_b je výberová kovariačná matica odhadov regresných koeficientov. Waldova štatistika má asymptoticky χ^2 -rozdelenie pravdepodobnosti s počtom stupňov voľnosti rovných počtu vytvorených umelých premenných pre vybranú vysvetľujúcu premennú. (Šoltés, E. – Vojtková, M. – Šoltésová, T., 2018)

Na testovanie štatistickej významnosti regresného koeficienta náležiacoho k vysvetľujúcej premennej – alebo umelej premennej pre túto premennú – sa používa špeciálny prípad predchádzajúceho testu. Navyše môžeme overiť štatistickú významnosť lokujúcej konštanty β_0 . Ide o jednoduchý test s hypotézami:

H_0 : Koeficient β_j nie je štatisticky významný – $\beta_j = \beta_{0j} = 0$, $j = 0, 1, 2, \dots, k$

H_1 : Koeficient β_j je štatisticky významný – $\beta_j \neq 0$, $j = 0, 1, 2, \dots, k$

Waldova premenná má tvar:

$$W = \left(\frac{\hat{\beta}_j - \hat{\beta}_{0j}}{s_{\hat{\beta}_j}} \right)^2 = \left(\frac{\hat{\beta}_j}{s_{\hat{\beta}_j}} \right)^2$$

kde $\hat{\beta}_j$ je bodový odhad parametra β_j a $s_{\hat{\beta}_j}$ je odhad smerodajnej odchýlky $\hat{\beta}_j$. Pri platnosti nulovej hypotézy má testovacia štatistika chí-kvadrát rozdelenie s počtom stupňov voľnosti 1. Nulovú hypotézu potom zamietame pri platnosti vzťahu $W > \chi^2_{1-\alpha}(1)$. V prípade alternatívnej kategoriálnej premennej alebo spojitej premennej sú uvedené testy totožné. (Terek, M. - Horníková, A. - Labudová, V., 2010)

3.2.4 Intervalové odhady pre parametre modelu a pomer šanci

Pri výpočte intervalových odhadov pre parametre logistickej regresie je možné použiť dve metódy. Jednou z nich je metóda, ktorá využíva *Waldove intervaly*. Vychádza z predpokladu, že bodové odhady parametrov modelu sú asymptoticky normálne (čo je pri odhade parametrov metódou maximálnej vierohodnosti zabezpečené). Výhodou teda je, že môžeme použiť kvantily normálneho rozdelenia, avšak pri malých súboroch sú výsledky vychýlené. Waldov $(1-\alpha)$ 100% interval spoľahlivosti pre β_j , $j=0,1,2,\dots,k$ má potom tvar:

$$P\left(\hat{\beta}_j - z_{1-\frac{\alpha}{2}} s_{\hat{\beta}_j} < \beta_j < \hat{\beta}_j + z_{1-\frac{\alpha}{2}} s_{\hat{\beta}_j}\right) = 1 - \alpha$$

kde $z_{1-\frac{\alpha}{2}}$ je kvantil normovaného normálneho rozdelenia.

Vďaka rozvoju informačných technológií sa čoraz viac používa metóda založená na *tvare funkcie vierohodnosti*, ktorá eliminuje vychýlenosť odhadu pri malých súboroch, avšak bez softvérovej podpory ju bolo veľmi zložité vypočítať.

Pri intervalových odhadoch pre pomer šanci musíme rozlišovať, či je nezávislá premenná X_j , ku ktorej prislúcha parameter β_j , $j=1,2,\dots,k$ spojitá alebo kategóriálna. Pre spojité premenné a takisto kategóriálne premenné, ktoré používajú kódovanie 0,1 má tento $(1-\alpha)$ 100% interval spoľahlivosti tvar:

$$P\left(\exp\left(\hat{\beta}_j - z_{1-\frac{\alpha}{2}} s_{\hat{\beta}_j}\right) < OR < \exp\left(\hat{\beta}_j + z_{1-\frac{\alpha}{2}} s_{\hat{\beta}_j}\right)\right) = 1 - \alpha$$

(Rublíková, E. – Labudová, V. – Sandtnerová, S., 2009)

3.3 Logistická regresia s multinomickou vysvetľovanou premennou

Prirodzeným zovšeobecnením binárneho logistického modelu – logitového modelu – je *multinomický logitový model*. Nie vždy totiž môžeme obmedziť počet obmien vysvetľovanej premennej na dve. Zvlášť, ak sú kategórie veľmi odlišné, ich spojením by nedošlo len k strate informácii, ale aj zavádzajúcim výsledkom.

Zovšeobecňujeme teda model s vysvetľovanou premennou s dvoma obmenami, na model s vysvetľovanou premennou, ktorá má viac ako dve obmeny. Označme počet obmien vysvetľovanej premennej s . Vysvetľovaná premenná Y nadobúda jednotlivé

hodnoty s pravdepodobnosťou π_{il} , kde $l = 1, 2, \dots, s-1$. Potom *multinomický logitový model* sa skladá s $s-1$ logitových funkcií vysvetľujúcich premenných X_j , ($j = 1, 2, \dots, k$) s parametrami β_{lj} :

$$\begin{aligned}\ln \frac{\pi_{i1}}{\pi_{i0}} &= \beta_{10} + \beta_{11}x_{i1} + \beta_{12}x_{i2} + \dots + \beta_{1k}x_{ik} \\ \ln \frac{\pi_{i2}}{\pi_{i0}} &= \beta_{20} + \beta_{21}x_{i1} + \beta_{22}x_{i2} + \dots + \beta_{2k}x_{ik} \\ &\vdots \\ \ln \frac{\pi_{i(s-1)}}{\pi_{i0}} &= \beta_{(s-1)0} + \beta_{(s-1)1}x_{i1} + \beta_{(s-1)2}x_{i2} + \dots + \beta_{(s-1)k}x_{ik}\end{aligned}$$

Všeobecne:

$$\ln \frac{\pi_{il}}{\pi_{i0}} = \beta_{l0} + \sum_{j=1}^k \beta_{lj}x_{ij}$$

Model teda spočíva v porovnaní jednotlivých obmien s „nultou“ obmenou, ktorá reprezentuje referenčný jav. Inak, ide o šancu, že nastane niektorý jav oproti šanci, že nastane referenčný jav.

Interpretácia parametrov je teda analogická ako pri binárnom modeli a realizuje sa cez pomer šancí, ktorý je daný vzťahom:

$$OR = \exp(\beta_{lj})$$

Zatiaľ čo pri binárnom modeli pri interpretácii sledujeme, ako sa zväčší/zmenší šanca nastania javu v porovnaní so šancou, že tento jav nenastane, pri multinomickom modeli sledujeme, ako sa zväčší/zmenší šanca, že nastane niektorý jav v porovnaní so šancou, že nastane jav referenčný.

Užitočné je ešte odvodenie bodového odhadu pravdepodobnosti. Pre nultú obmenu platí vzťah:

$$\hat{\pi}_{i0} = \frac{1}{1 + \sum_{l=1}^{s-1} \exp(\hat{\beta}_{l0} + \sum_{j=1}^k \hat{\beta}_{lj}x_{ij})}$$

pre ostatné obmeny platí nasledujúci vzťah:

$$\hat{\pi}_{il} = \frac{\exp(\hat{\beta}_{l0} + \sum_{j=1}^k \hat{\beta}_{lj}x_{ij})}{1 + \sum_{l=1}^{s-1} \exp(\hat{\beta}_{l0} + \sum_{j=1}^k \hat{\beta}_{lj}x_{ij})}$$

(Šoltés, E. – Vojtková, M. – Šoltéssová, T., 2018)

Testy štatistickej významnosti modelu ako celku a jeho parametrov, či intervalové odhady sú konštruované analogicky ako pri modeli s binárnou vysvetľovanou premennou. (Pecáková, I., 2007). Z uvedeného dôvodu vzťahy pre spomenuté indukčné úsudky už nebudeme uvádzať a pre potrebu ilustrácie je možné ich nájsť v predchádzajúcich častiach tejto kapitoly.

3.3.1 Hodnotenie kvality modelu

Pri hodnotení kvality modelu máme k dispozícii množstvo rôznych štatistík. Pri dátach, ktoré nie sú zoskupené do kontingenčnej tabuľky a pri modeli multinomickej logistickej regresie však môžeme použiť len obmedzené množstvo.

Jednou skupinou štatistík, ktorú môžeme použiť na posúdenie kvality modelu sú štatistiky založené na hodnote funkcie maximálnej vierohodnosti L , ktoré opisujú ako dobre sa model zhoduje s výberovými údajmi. Najzákladnejšou z nich je $-2 \log L$, jednoducho povedané maximálna hodnota logaritmu funkcie maximálnej vierohodnosti vynásobená číslom -2. Čím je hodnota tejto štatistiky vyššia, tým je model menej vhodný. Treba si ale uvedomiť, že veľkosť tejto štatistiky vo veľkej miere závisí od počtu pozorovaní. Ďalším problémom je, že modely, ktoré majú viac vysvetľujúcich premenných dosahujú nižšie hodnoty tejto štatistiky aj keď nie sú v skutočnosti „kvalitnejšie“. Tento problém riešia vhodnejšie štatistiky AIC a BIC , ktoré penalizujú modely za počet vysvetľujúcich premenných. Obe charakteristiky môžeme použiť na porovnanie kvality modelov s rôznym počtom vysvetľujúcich premenných. (Allison, P. D., 2012)

Akaikeho informačné kritérium počítame ako $AIC = -2 \log L + 2K$, kde $K = k + 1$. *Schwarzovo-Bayesovo kritérium* vyjadrujeme vzťahom $BIC = -2 \log L + K \log n$, kde n je rozsah výberového súboru. Interpretovať hodnoty týchto štatistík pre jeden model nemá význam, je možné ich použiť na porovnanie kvality viacerých modelov odhadnutých na tej istej dátovej množine. Lepší model je potom ten, ktorý má nižšiu hodnotu týchto štatistík.

Ďalším spôsobom, ako posudzovať kvalitu modelu, je určiť, ako dobre model odhaduje vysvetľovanú premennú na základe hodnôt vysvetľujúcich premenných. Pri logistickej regresii, podobne ako lineárnej sa používa *koeficient determinácie*, upravený pre túto metódu. Existuje veľa rôznych tvarov koeficienta determinácie pre logistickú

regresiu, uvedieme tvar, ktorý bol navrhnutý vedcami Coxom a Snellom a podľa nich má aj pomenovanie *Cox-Snell Pseudo R^2* v tvare:

$$R^2 = 1 - \exp \left[-\frac{2}{n} \ln \frac{L_A}{L_0} \right]$$

Kde L_A je funkcia maximálnej vierohodnosti modelu so zaradenými vysvetľujúcimi premennými a lokujúcou konštantou a L_0 predstavuje model len s lokujúcou konštantou. Keďže výpočet je založený na funkcii vierohodnosti, poskytujú ho prakticky všetky štatistické softvéry. Nevýhodou je, že jeho horná hranica nedosahuje hodnotu 1. Pre nápravu bola vytvorená modifikácia, ktorá tento koeficient determinácie delí jeho hornou hranicou. (Rublíková, E. – Labudová, V. – Sandtnerová, S., 2009)

3.4 Závislosť dvoch kategoriálnych znakov

Základom pri skúmaní závislosti dvoch kategoriálnych znakov sú početnosti, ktoré triedime do tzv. *asociačných tabuliek* (ak oba znaky majú dve obmeny) alebo *kontingenčných tabuliek* (ak niektorý zo znakov má viac ako dve obmeny). Majme k dispozícii n štatistických jednotiek, pričom pre všetky poznáme hodnoty oboch znakov. Označme dva kategoriálne znaky A a B . Nech znak A má s obmien a znak B r obmien. Potom vnútro asociačnej/kontingenčnej tabuľky veľkosti $r \times s$ zachytáva zistené početnosti druhého stupňa n_{ij} , $i = 1, 2, \dots, r$, $j = 1, 2, \dots, s$, t. j. počty štatistických jednotiek, ktoré nadobúdajú jednotlivé kombinácie obmien znakov A a B . Tabuľka okrem toho obsahuje súčtový stĺpec, resp. súčtový riadok, kde sú početnosti prvého stupňa, t. j. počty štatistických jednotiek, ktoré majú jednotlivé obmeny znaku A , resp. znaku B .

Tabuľka 3.4.1 – Kontingenčná tabuľka

Znak	b_1	b_2	...	b_s	Σ
a_1	$n_{11} (a_1 b_1)$	$n_{12} (a_1 b_2)$	...	$n_{1s} (a_1 b_s)$	n_{1*}
a_2	$n_{21} (a_2 b_1)$	$n_{22} (a_2 b_2)$	...	$n_{2s} (a_2 b_s)$	n_{2*}
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
a_r	$n_{r1} (a_r b_1)$	$n_{r2} (a_r b_2)$	...	$n_{rs} (a_r b_s)$	n_{r*}
Σ	n_{*1}	n_{*2}	...	n_{*r}	n

Zdroj: Vlastné spracovanie podľa <http://umv.science.upjs.sk/zezula/stgjax/k14jax.html> (MS Excel)

Početnosti druhého stupňa nazývame aj pozorované početnosti a označujeme ich O_{ij} (z angl. observed; $O_{ij} = n_{ij}$). Pri testovaní ich porovnávame s teoretickými početnosťami,

ktoré označujeme E_{ij} (z angl. *expected*). Ide o teoretické početnosti za predpokladu, že znaky A a B sú nezávislé a vypočítajú sa podľa vzťahu

$$E_{ij} = \frac{n_{i*} \cdot n_{*j}}{n}$$

V prípade, že dva kategoriálne znaky sú úplne nezávislé, ich pozorované a teoretické početnosti sa zhodujú. Na tomto základe je založený *Pearsonov chí-kvadrát test dobrej zhody*:

H_0 : Medzi znakmi A a B nie je štatisticky významná závislosť (asociácia, kontingencia)

H_1 : Medzi znakmi A a B je štatisticky významná závislosť (asociácia, kontingencia)

Testovacia charakteristika má tvar:

$$\chi^2 = \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

s chí-kvadrát rozdelením s počtom stupňov voľnosti $(r - 1) \cdot (s - 1)$. Ak rozdiely medzi pozorovanými a teoretickými početnosťami budú dostatočne malé, môžeme ich pripísať pôsobeniu náhodných činiteľov a teda nemôžeme zamietnuť nulovú hypotézu. Naopak, ak rozdiely budú dostatočne veľké, medzi znakmi existuje štatisticky významná závislosť a nulovú hypotézu zamietame. O štatisticky významnej závislosti môžeme hovoriť pri platnosti vzťahu $\chi^2 > \chi^2_{1-\alpha}[(r - 1) \cdot (s - 1)]$. Na použitie tohto testu musíme splniť podmienky $n > 20$ a $E_{ij} \geq 5$. (Pacáková, V. a kol., 2003)

3.4.2 Sila závislosti dvoch kategoriálnych znakov

Zatiaľ čo test závislosti dvoch kategoriálnych znakov nám povie, či sú dva znaky závislé alebo nie, často je dôležitá aj sila závislosti medzi znakmi. Túto silu nám určuje skupina mier závislosti, ktoré nadobúdajú hodnoty z intervalu $\langle 0; 1 \rangle$ – Čuprovov koeficient a Cramerov koeficient, prípadne $\langle 0; 1 \rangle$ – Pearsonov koeficient. Nulová hodnota niektorej z mier hovorí o úplnej nezávislosti znakov. Hodnota miery rovnajúca sa 1 hovorí o ich úplnej závislosti.

a) Pearsonov koeficient:

$$C = \sqrt{\frac{\Phi^2}{1 + \Phi^2}} = \sqrt{\frac{\chi^2}{n + \chi^2}}, \Phi^2 = \frac{\chi^2}{n}$$

b) Čuprovov koeficient:

$$\tau = \sqrt{\frac{\Phi^2}{\sqrt{(r-1) \cdot (s-1)}}}$$

c) Cramerov koeficient

$$V = \sqrt{\frac{\chi^2}{n \cdot h}}, h = \min[(r-1); (s-1)]$$

(Pacáková, V. a kol., 2003)

4 Výsledky záverečnej práce a diskusia

V tejto časti aplikujeme poznatky z predchádzajúcich kapitol na kvantifikovanie vplyvu relevantných faktorov na konzistentnú chudobu slovenských domácností. Údaje pochádzajú z databázy *EU-SILC 2016*, ktorá je výsledkom pravidelného výberového zisťovania o príjmoch a životných podmienkach domácností, realizovaného vo všetkých krajinách Európskej Únie.

Údaje databázy *EU-SILC 2016* sú rozdelené do štyroch súborov, a to:

- *D_SILC_2016* – Register domácností pre prierezové zisťovanie
- *R_SILC_2016* – Register osôb pre prierezové zisťovanie
- *H_SILC_2016* – Údaje o domácnosti pre prierezové zisťovanie
- *P_SILC_2016* – Osobné údaje pre prierezové zisťovanie

Tieto súbory sú tvorené údajmi z výberového zisťovania za rok 2015, ktoré prebehlo vo viac ako 500 obciach na Slovensku. Celkovo sa zber údajov uskutočnil v 6 646 vybraných domácnostiach. Do konečnej databázy sa zaradilo 5 738 domácností a 14 101 osôb nad 16 rokov (doporučená minimálna efektívna veľkosť súboru pritom bola pre domácnosti 4 250 a pre osoby 11 000). (Vlačuha, R. - Kováčová, Y., 2017)

V analýzach prezentovaných v tejto práci sme použili premenné vzťahujúce sa na celú domácnosť alebo na prednostu domácnosti a abstrahovali sme od premenných charakterizujúcich ostatných členov domácností. Takto sme získali súbor údajov pre 5 715 domácností. Vzhľadom na cieľ diplomovej práce sa zameriavame na výskyt príjmovej chudoby a závažnej materiálnej deprivácie v slovenských domácnostiach. Ide o 2 dimenzie 3-dimenzionálneho konceptu merania chudoby a sociálneho vylúčenia v rámci stratégie Európa 2020, preto v práci vychádzame z ukazovateľa *AROPE* (miera rizika chudoby alebo sociálneho vylúčenia), ktorý zachytáva populáciu, ktorá je v riziku príjmovej chudoby (čiastková miera *ARPT60i*) a/alebo je materiálne deprivovaná (čiastková miera *SEV_DEP*) a/alebo žije v domácnostiach s veľmi nízkou intenzitou práce (čiastková miera *VLWI*), pričom tretia dimenzia charakterizujúca vylúčenie z trhu práce nebude predmetom nášho výskumu

Aby premenná *AROPE* vyhovovala účelom práce, bolo potrebné ju prekódovať. Vznikli štyri obmeny tejto kategoriálnej premennej, vyjadrujúce rôzne kombinácie

príjmovej chudoby a závažnej materiálnej deprivácie, ktorými sú postihnuté slovenské domácnosti. Výsledky sú uvedené v tabuľke.

Tabuľka 4.1 – Kódovanie premennej AROPE

Pôvodné premenné EU-SILC		Nové premenné
AROPE – Miera rizika chudoby a sociálneho vylúčenia		AROPE
0	ARPT60i = 0 / SEV_DEP = 0 / VLWI = 0	0
1	ARPT60i = 0 / SEV_DEP = 0 / VLWI = 1	
100	ARPT60i = 1 / SEV_DEP = 0 / VLWI = 0	1
101	ARPT60i = 1 / SEV_DEP = 0 / VLWI = 1	
10	ARPT60i = 0 / SEV_DEP = 1 / VLWI = 0	2
11	ARPT60i = 0 / SEV_DEP = 1 / VLWI = 1	
110	ARPT60i = 1 / SEV_DEP = 1 / VLWI = 0	3
111	ARPT60i = 1 / SEV_DEP = 1 / VLWI = 1	

Zdroj: Vlastné spracovanie (MS Excel)

Obmena „0“ novej premennej *AROPE* vyjadruje stav, kedy domácnosť nie je ohrozená ani príjmovou chudobou, ani závažnou materiálnou depriváciou. Obmena „1“ nastáva, keď je v domácnosti prítomná príjmová chudoba. Obmena „2“ hovorí o stave, kedy domácnosť je závažne materiálne deprivovaná. Napokon obmena „3“ vyjadruje stav, kedy domácnosť trpí príjmovou chudobou a zároveň závažnou materiálnou depriváciou.

4.1 Vplyv relevantných faktorov na konzistentnú chudobu

V tejto časti verifikujeme závislosť a odhadneme silu závislosti vybraných relevantných faktorov s premennou *AROPE* pomocou analýzy kontingenčných tabuliek. Budeme skúmať izolovaný vplyv nasledujúcich premenných z databázy *EU-SILC 2016* na premennú *AROPE*: *RB210* - Status základnej ekonomickej aktivity - **EAS**, *PE040* - Najvyššia dosiahnutá úroveň vzdelania podľa ISCED - **Education**, *PH010* - Všeobecné zdravie - **Health**, *HT* - Typ domácnosti – **Household Type**, *PB190* - Rodinný stav – **Marital Status**, *KRAJ* – Kraj - **Region** a *DB100* - Stupeň urbanizácie - **Urbanisation**.

Výsledky sa nachádzajú v nasledujúcej tabuľke, kde sa v stĺpci pre každú premennú nachádza v prvom riadku *P* - hodnota pre Pearsonov chí-kvadrát test dobrej zhody (*Chi-Square*) a v nasledujúcich dvoch riadkoch postupne hodnoty koeficientov sily závislosti – Paersonov koeficient (*Contingency Coefficient*) a Cramerov koeficient (*Cramer's V*).

Tabuľka 4.1.1 – Závislosť vybraných faktorov a premennej AROPE

Statistic	EAS	EDU	Health	HT	MS	Region	URB
Chi-Square	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001	<.0001
Contingency Coefficient	0.4206	0.2468	0.1348	0.2442	0.1810	0.1360	0.1154
Cramer's V	0.2676	0.1470	0.0962	0.1454	0.1063	0.0793	0.0821

Zdroj: Vlastné spracovanie (SAS EG, MS Excel)

Na základe P-hodnôt pre Pearsonov chí-kvadrát test dobrej zhody na hladine významnosti 0,05 zamietame nulovú hypotézu a prijímame alternatívnu, že medzi premennou *AROPE* a postupne medzi všetkými faktormi vstupujúcimi do analýzy je štatisticky významná závislosť. Najsilnejšia závislosť je medzi premennými *AROPE* a *EAS*, kde ide o stredne silnú závislosť. O stredne silnej až slabej závislosti s *AROPE* môžeme hovoriť pri premenných *Education* a *Household Type*, zatiaľ čo len o slabej závislosti pri premenných *Health*, *Marital Status*, *Region* a *Urbanisation*. Všetky premenné ale dosahujú významný stupeň závislosti, preto ich ďalej zaradíme do modelu logistickej regresie, kde tieto premenné na rozdiel od kontingencie tabuľky pôsobia na premennú *AROPE* simultánne a môžu nastať rozdiely pri významnosti ich vplyvu.

4.2 Analýza konzistentnej chudoby využitím logistickej regresie

V logistickej regresii budeme skúmať simultánny vplyv vybraných vysvetľujúcich premenných na konzistentnú chudobu pomocou vysvetľovanej premennej *AROPE* využitím softvéru *SAS Enterprise Guide*. Vysvetľujúce premenné vzhľadom na prednostu domácnosti vstupujúce do modelu sú tieto:

a) Spojité premenné:

- Vek na konci príjmového referenčného obdobia

b) Kategoriálne premenné:

- Status základnej ekonomickej aktivity
- Najvyššia dosiahnutá úroveň vzdelania podľa ISCED
- Všeobecné zdravie
- Typ domácnosti
- Rodinný stav
- Kraj
- Stupeň urbanizácie

Premenné sme pre účely práce upravili, pre lepšiu prehľadnosť sme zmenili názvy obmien premenných a keďže početnosti v niektorých obmenách premenných boli pomerne nízke, zlúčili sme viacero podobných obmien do jednej obmeny. Takisto sme určili pre všetky kategoriálne premenné ich referenčnú kategóriu (označená *R*), ku ktorej sa budú vzťahovať interpretácie pomerov šancí.

Tabuľka 4.2.1 – Nezávislé premenné

Pôvodné premenné EU-SILC		Nové premenné	
PX010, RX010 - Vek		Age	
num	Vek na konci príjmového referenčného obdobia		
RB210 - Status základnej ekonomickej aktivity		EAS	
1	pracujúci	at_Work	R
2	nezamestnaný	Unemployed	
3	starobný dôchodca, osoba v predčasnom dôchodku	Retired	
4	iná neaktívna osoba	Inactive_person	
PE040 - Najvyššia dosiahnutá úroveň vzdelania (ISCED)		Education	
000	Nižšie ako primárne vzdelanie	Less_than_Secondary	
100	Primárne vzdelanie		
200	Nižšie sekundárne vzdelanie		
300	Vyššie sekundárne vzdelanie	Upper_Secondary	
400	Post-sekundárne, nie terciárne vzdelanie	Post_Secondary	
500	Krátky cyklus terciárneho vzdelania	Tertiary_1	
600	Bakalárske alebo jeho ekvivalent		
700	Magister alebo jeho ekvivalent	Tertiary_2_3	R
800	Doktorát alebo jeho ekvivalent		
PH010 - Všeobecné zdravie		Health	
1	veľmi dobré	Good	R
2	dobré		
3	ani dobré ani zlé	Fair	
4	zlé	Bad	
5	veľmi zlé		
HT - Typ domácnosti		Household Type	
5	Jednočlenná domácnosť	1Adult	
6	Domácnosť 2 dospelých - obaja pod 65 r.	2Adult_0Ch	
7	Domácnosť 2 dospelých - aspoň jeden vo veku 65+	2A_1R	
8	Ostatné domácnosti bez závislých detí	Other_0Ch	
9	Domácnosť s 1 rodičom a 1 alebo viac závislými deťmi	1A_at_least_1Ch	
10	Domácnosť 2 dospelých s 1 závislým dieťaťom	2A_1Ch	
11	Domácnosť 2 dospelých s 2 závislými deťmi	2A_2Ch	R
12	Domácnosť 2 dospelých s 3 alebo viac závislými deťmi	2A_at_least_3Ch	
13	Ostatné domácnosti so závislými deťmi	Other_with_Ch	

PB190 - Rodinný stav		Marital Status	
1	slobodný / á	Never_married	
2	ženatý / vydatá	Married	R
3	žijúci oddelene	Separated	
4	vdovec / vdova	Widowed	
5	rozvedený / á	Divorced	
KRAJ - Kraj		Region	
1	Bratislavský	Bratislava	R
2	Trnavský	Trnava	
3	Trenčiansky	Trencin	
4	Nitriansky	Nitra	
5	Žilinský	Zilina	
6	Banskobystrický	Banska_Bystrica	
7	Prešovský	Presov	
8	Košický	Kosice	
DB100 - Stupeň urbanizácie		Urbanisation	
1	územie s hustým osídlením	Dense	R
2	územie s priemerne hustým osídlením	Intermediate	
3	územie s riedkym osídlením	Sparse	

Zdroj: Vlastné spracovanie (MS Excel)

Keďže vysvetľovaná premenná *AROPE* má 4 obmeny, ktoré nie je možné usporiadať (nevieme určiť, či je pre domácnosť „lepšie“ keď je ohrozená príjmovou chudobou alebo materiálnou depriváciou), využívame multinomický regresný model. Referenčnou kategóriou pre *AROPE* je obmena „0“ – stav, kedy domácnosť netrpí príjmovou chudobou ani materiálnou depriváciou a všetky ostatné kategórie sú porovnávané vzhľadom na referenčnú.

Kategoriálne vysvetľujúce premenné je potrebné pred vstupom do modelu prekódovať. Keďže každej obmene kategoriálnej vysvetľujúcej premennej je priradený samostatný regresný koeficient, je potrebné určiť kódovanie, ktoré bude charakterizovať vplyv obmeny na vysvetľovanú premennú. V modeli využívame spôsob kódovania {0,1}, čo pre nereferenčné obmeny znamená, že umelá premenná vytvorená pre príslušnú obmenu nadobúda hodnotu 1 a ostatné umelé premenné pre danú premennú majú hodnotu 0. Pri referenčnej obmene nadobúdajú všetky umelé premenné vytvorené pre danú kategoriálnu premennú hodnotu 0. Toto kódovanie môžeme vidieť na výstupe zo SAS EG.

Tabuľka 4.2.2 – Kódovanie kategoriálnych premenných

Class Level Information									
Class	Value	Design Variables							
EAS	Inactive_person	1	0	0					
	Unemployed	0	1	0					
	Retired	0	0	1					
	At_Work	0	0	0					
Education	Less_than_Secondary	1	0	0	0				
	Post_Secondary	0	1	0	0				
	Tertiary_1	0	0	1	0				
	Upper_Secondary	0	0	0	1				
	Tertiary_2_3	0	0	0	0				
Health	Bad	1	0						
	Fair	0	1						
	Good	0	0						
Household Type	1A_at_least_1Ch	1	0	0	0	0	0	0	0
	1Adult	0	1	0	0	0	0	0	0
	2A_1Ch	0	0	1	0	0	0	0	0
	2A_1R	0	0	0	1	0	0	0	0
	2A_at_least_3Ch	0	0	0	0	1	0	0	0
	2Adult_0Ch	0	0	0	0	0	1	0	0
	Other_0Ch	0	0	0	0	0	0	1	0
	Other_with_Ch	0	0	0	0	0	0	0	1
	2A_2Ch	0	0	0	0	0	0	0	0
Marital Status	Divorced	1	0	0					
	Never_married	0	1	0					
	Widowed	0	0	1					
	Married	0	0	0					
Region	Banska_Bystrica	1	0	0	0	0	0	0	
	Kosice	0	1	0	0	0	0	0	
	Nitra	0	0	1	0	0	0	0	
	Presov	0	0	0	1	0	0	0	
	Trencin	0	0	0	0	1	0	0	
	Trnava	0	0	0	0	0	1	0	
	Zilina	0	0	0	0	0	0	1	
	Bratislava	0	0	0	0	0	0	0	
Urbanisation	Intermediate	1	0						
	Sparse	0	1						
	Dense	0	0						

Zdroj: Vlastné spracovanie (SAS EG, MS Excel)

Po potrebných začiatkových úpravách a nastaveniach môžeme pristúpiť k samotnému modelovaniu. Z dostupných údajov pre 5 715 domácností bolo použitých 5 686 pozorovaní z dôvodu chýbajúcich hodnôt pre niektoré vysvetľujúce premenné.

Tabuľka 4.2.3 – Početnosti obmien vysvetľovanej premennej AROPE

Response Profile		
Ordered Value	AROPE	Total Frequency
1	0	4768
2	1	410
3	2	302
4	3	206

Zdroj: Vlastné spracovanie (SAS EG, MS Excel)

Domácností, ktoré nie sú ohrozené príjmovou chudobou ani materiálnou depriváciou, je 4 768. Celkovo je niektorou dimenziou chudoby a sociálneho vylúčenia ohrozených 918 domácností, z toho 410 je v riziku príjmovej chudoby, 302 trpí materiálnou depriváciou a 206 oboma týmito dimenziami.

Tabuľka 4.2.4 – Testovanie štatistickej významnosti modelu a jeho parametrov

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	1365.5014	90	<.0001
Score	1928.3151	90	<.0001
Wald	890.8542	90	<.0001

Type 3 Analysis of Effects			
Effect	DF	Wald Chi-Square	Pr > ChiSq
Age	3	4.7979	0.1872
EAS	9	284.3130	<.0001
Education	12	150.5254	<.0001
Health	6	40.3100	<.0001
Household Type	24	177.5215	<.0001
Marital Status	9	102.2658	<.0001
Region	21	57.6713	<.0001
Urbanisation	6	23.1339	0.0008

Zdroj: Vlastné spracovanie (SAS EG, MS Excel)

Do modelu je potrebné zaradiť len tie faktory (spojité alebo kategóriálne vysvetľujúce premenné), ktoré majú signifikantný vplyv na ohrozenie príjmovou chudobou a/alebo na ohrozenie závažnou materiálnou depriváciou. Z uvedeného dôvodu najskôr overíme štatistickú významnosť modelu ako celku a taktiež jeho parametrov. Na hladine významnosti 0,05 testujeme hypotézy uvedené v kapitole 3.2.2. Model ako celok je na

Ľubovoľnej hladine významnosti štatisticky významný, avšak nie všetky jeho parametre. P-hodnota pri premennej *Age* je väčšia ako zvolená hladina významnosti, preto nemôžeme zamietnuť nulovú hypotézu a prijímame predpoklad, že regresný koeficient pri premennej *Age* je štatisticky nevýznamný. Premennú odstránime z modelu a odhadneme ho nanovo.

Tabuľka 4.2.5 - Testovanie štatistickej významnosti bez premennej Vek

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	1360.6961	87	<.0001
Score	1923.4117	87	<.0001
Wald	888.5343	87	<.0001

Type 3 Analysis of Effects			
Effect	DF	Wald Chi-Square	Pr > ChiSq
EAS	9	314.8343	<.0001
Education	12	148.6977	<.0001
Health	6	36.4264	<.0001
Household Type	24	200.1654	<.0001
Marital Status	9	115.4309	<.0001
Region	21	58.1050	<.0001
Urbanisation	6	22.9547	0.0008

Zdroj: Vlastné spracovanie (SAS EG, MS Excel)

Bez premennej *Age* je na hladine významnosti 0,05 štatisticky významný model a rovnako jeho parametre. Početnosti premennej *AROPE* zostávajú rovnaké. Na základe hodnoty Waldovej premennej (*Wald Chi-Square*) vidíme, že vysvetľujúce premenné majú rôzne veľký vplyv na vysvetľovanú premennú, avšak ich prínos je štatisticky významný. Veľkosť vplyvu sa zhoduje s predpokladmi o sile závislosti z kapitoly 4.1.

Tabuľka 4.2.6 – Hodnotenie kvality modelu

Model Fit Statistics			
Criterion	Intercept Only	Intercept and Covariates (with Age)	Intercept and Covariates (without Age)
AIC	6981.287	5795.786	5794.591
SC	7001.225	6413.842	6392.710
-2 Log L	6975.287	5609.786	5614.591

R-Square	0.2128	Max-rescaled R-Square	0.3011
----------	--------	-----------------------	--------

Zdroj: Vlastné spracovanie (SAS EG, MS Excel)

Na porovnanie máme k dispozícii tri modely – model logistickej regresie len s lokujúcou konštantnou a ďalej modely aj s regresnými koeficientmi, ktoré prislúchajú k daným premenným najprv s premennou *Age* a potom bez nej. Štatistika $-2 \log L$ vzhľadom na rôzne počty vysvetľujúcich premenných nie je vhodná na porovnanie týchto modelov, nakoľko čím je ich viac, tým je jej hodnota nižšia, čo môžeme vidieť aj v porovnaní modelov s premennou *Age* a bez nej. Na porovnanie teda použijeme hodnoty štatistík *AIC* a *BIC* (*SC*). Ako môžeme vidieť, modely s regresnými koeficientmi sú jednoznačne kvalitnejšie ako model len s lokujúcou konštantou, čo potvrdzuje ich významný vplyv na vysvetľovanú premennú.

Navyše po vyradení štatisticky nevýznamnej vysvetľujúcej premennej *Age* nastalo ďalšie mierne zlepšenie kvality modelu. Ďalej teda budeme pracovať s týmto modelom. Zvolený model podľa Cox-Snellého koeficienta determinácie (*R-Square*) a jeho modifikácie (*Max-rescaled R-Square*) vysvetľuje 21,28 %, prípadne 30,11 % variability vysvetľovanej premennej. Môže sa to zdať ako menšie percento, avšak v ekonomike je vždy prítomných veľa ďalších faktorov a takisto vo veľkej miere aj náhodných vplyvov. Keďže však model používame na interpretáciu pomeru šancí a neaplikujeme získané poznatky na iné súbory, či na predikciu do budúcnosti, neznamená to negatívny výsledok. (Nau, R., 2017)

4.2.1 Interpretácia parametrov logistickej regresie

V tejto kapitole odhadujeme parametre modelu logistickej regresie prostredníctvom metódy maximálnej vierohodnosti. Keďže ide o multinomický logistický model so štyrmi obmenami a referenčnou obmenou „0“, pre každú umelú premennú sú odhadnuté tri regresné koeficienty, ktoré sa vzťahujú k obmenám 1, 2 a 3 premennej AROPE. Pre lepšiu prehľadnosť rozdelíme multinomický logistický model na 3 samostatné modely:

- model 1 vs. 0, ktorý porovnáva šancu, že domácnosť bude v riziku príjmovej chudoby, ale nebude závažne materiálne deprivovaná (1), k šanci, že domácnosť nebude v riziku príjmovej chudoby a nebude ani závažne materiálne deprivovaná (0),
- model 2 vs. 0, ktorý porovnáva šancu, že domácnosť bude závažne materiálne deprivovaná, ale nebude v riziku príjmovej chudoby (2) k šanci 0,

- model 3 vs. 0, ktorý porovnáva šancu, že domácnosť bude musieť čeliť konzistentnej chudobe (3), k šanci, že domácnosť nebude v riziku príjmovej chudoby a nebude ani závažne materiálne deprivovaná (0),

Aj keď sú všetky faktory zaradené do modelu štatisticky významné, neznamená to, že aj všetky z nich odvodené umelé premenné štatisticky významne vplyvajú na vysvetľujúcu premennú. Pri testovaní štatistickej významnosti jednotlivých obmien platia analogické hypotézy ako pri testovaní významnosti premenných. Testujeme na hladine významnosti 0,05 a teda všetky obmeny, pri ktorých P-hodnota dosahuje hodnotu menšiu ako je hladina významnosti, determinujú štatisticky významne odlišný pomer šancí pre riziko príjmovej chudoby (model 1 vs. 0), odlišný pomer šancí pre závažnú materiálnu depriváciu (model 2 vs. 0) alebo odlišný pomer šancí pre konzistentnú chudobu (model 3 vs. 0) ako v referenčnej kategórii príslušnej vysvetľujúcej premennej.. Bodové odhady parametrov modelov 1 vs. 0, 2 vs. 0 a 3 vs. 0 sa nachádzajú v tabuľke 4.2.1.1., pričom sú zvýraznené tie P - hodnoty, ktoré potvrdzujú, že príslušné parametre sú štatisticky významné.

Tabuľka 4.2.1.1 – Bodové odhady regresných parametrov

Parameter		1 vs. 0		2 vs. 0		3 vs. 0	
		Est	P-val	Est	P-val	Est	P-val
EAS Inactive_person	vs. at_Work	0.8934	<.0001	0.9271	0.0002	2.1809	<.0001
EAS Retired	vs. at_Work	-0.1316	0.4579	0,0561	0,7637	0,2806	0,3797
EAS Unemployed	vs. at_Work	1,9478	<.0001	1,4988	<.0001	4,2090	<.0001
Education Less_than_Secondary	vs. T_2_3	1,3076	<.0001	1,7817	<.0001	3,6110	<.0001
Education Post_Secondary	vs. T_2_3	0,0698	0,8736	0.8510	0.0879	1.7499	0.0286
Education Tertiary_1	vs. T_2_3	-0.0884	0.8480	0.2558	0.6946	0.2785	0.8364
Education Upper_Secondary	vs. T_2_3	0.5165	0.0033	1.2235	<.0001	1.6106	0.0040
Health Bad	vs. Good	0.4038	0.0241	0.9568	<.0001	0.7614	0.0042
Health Fair	vs. Good	0.3044	0.0287	0.4991	0.0027	0.4542	0.0570
HT 1A_at_least_1Ch	vs. 2A_2Ch	0.0883	0.7774	0.5556	0.1873	2.1850	<.0001
HT 1Adult	vs. 2A_2Ch	-0.6834	0.0047	-0.1556	0.6326	0.7979	0.1001
HT 2A_1Ch	vs. 2A_2Ch	-0.5847	0.0069	-0.4398	0.2526	-0.3473	0.5436
HT 2A_1R	vs. 2A_2Ch	-2.2095	<.0001	0.1935	0.5500	-1.0174	0.1103
HT 2A_at_least_3Ch	vs. 2A_2Ch	0.7873	0.0006	0.1809	0.7310	1.8838	0.0003
HT 2Adult_0Ch	vs. 2A_2Ch	-1.2585	<.0001	-0.2641	0.4197	-0.2210	0.6636
HT Other_0Ch	vs. 2A_2Ch	-1.7599	<.0001	0.0371	0.9025	-0.7714	0.1268

Parameter		1 vs. 0		2 vs. 0		3 vs. 0	
		Est	P-val	Est	P-val	Est	P-val
HT Other_with_Ch	vs. 2A_2Ch	-0.8443	<.0001	0.0172	0.9566	0.3570	0.4265
Marital Status Divorced	vs. Married	0.6035	0.0011	1.0517	<.0001	1.4682	<.0001
Marital Status Never_married	vs. Married	0.5223	0.0078	1.0198	<.0001	1.2538	<.0001
Marital Status Widowed	vs. Married	-0.7574	0.0010	0.2506	0.1995	-0.4352	0.1731
Region Banska_Bystrica	vs. Bratislava	0.3848	0.1123	-0.9531	0.0006	0.7519	0.1413
Region Kosice	vs. Bratislava	0.0128	0.9573	-0.2222	0.3036	1.3800	0.0037
Region Nitra	vs. Bratislava	0.2494	0.3041	-1.1144	0.0001	0.9082	0.0724
Region Presov	vs. Bratislava	0.5387	0.0227	-0.5666	0.0227	1.1003	0.0284
Region Trencin	vs. Bratislava	0.0506	0.8447	-0.1087	0.6288	0.6421	0.2277
Region Trnava	vs. Bratislava	0.1041	0.6883	-0.8751	0.0016	0.6957	0.1890
Region Zilina	vs. Bratislava	0.3144	0.1844	-0.4944	0.0357	0.5808	0.2614
Urbanisation Intermediate	vs. Dense	-0.0534	0.7581	-0.0571	0.7418	0.0921	0.7700
Urbanisation Sparse	vs. Dense	0.4157	0.0128	-0.1879	0.3005	0.5813	0.0568

Zdroj: Vlastné spracovanie (SAS EG, MS Excel)

Interpretácia parametrov modelu sa uskutočňuje prostredníctvom bodového odhadu pomeru šancí – OR. V nasledujúcej tabuľke sa nachádzajú odhady pomerov šancí s ich P - hodnotami, pričom sú opäť označené štatisticky významné položky.

Tabuľka 4.2.1.2 – Bodové odhady pomeru šancí

Parameter		1 vs. 0		2 vs. 0		3 vs. 0	
		OR	P-val	OR	P-val	OR	P-val
EAS Inactive_person	vs. at_Work	2.444	<.0001	2.527	0.0002	8.854	<.0001
EAS Retired	vs. at_Work	0.877	0.4579	1.058	0,7637	1.324	0,3797
EAS Unemployed	vs. at_Work	7.013	<.0001	4.476	<.0001	67.287	<.0001
Education Less_than_Secondary	vs. T_2_3	3.697	<.0001	5.940	<.0001	37.002	<.0001
Education Post_Secondary	vs. T_2_3	1.072	0,8736	2.342	0.0879	5.754	0.0286
Education Tertiary_1	vs. T_2_3	0.915	0.8480	1.291	0.6946	1.321	0.8364
Education Upper_Secondary	vs. T_2_3	1.676	0.0033	3.399	<.0001	5.006	0.0040
Health Bad	vs. Good	1.497	0.0241	2.603	<.0001	2.141	0.0042
Health Fair	vs. Good	1.356	0.0287	1.647	0.0027	1.575	0.0570
HT 1A_at_least_1Ch	vs. 2A_2Ch	1.092	0.7774	1.743	0.1873	8.891	<.0001
HT 1Adult	vs. 2A_2Ch	0.505	0.0047	0.856	0.6326	2.221	0.1001

Parameter		1 vs. 0		2 vs. 0		3 vs. 0	
		OR	P-val	OR	P-val	OR	P-val
HT 2A_1Ch	vs. 2A_2Ch	0.557	0.0069	0.644	0.2526	0.707	0.5436
HT 2A_1R	vs. 2A_2Ch	0.110	<.0001	1.213	0.5500	0.362	0.1103
HT 2A_at_least_3Ch	vs. 2A_2Ch	2.197	0.0006	1.198	0.7310	6.578	0.0003
HT 2Adult_0Ch	vs. 2A_2Ch	0.284	<.0001	0.768	0.4197	0.802	0.6636
HT Other_0Ch	vs. 2A_2Ch	0.172	<.0001	1.038	0.9025	0.462	0.1268
HT Other_with_Ch	vs. 2A_2Ch	0.430	<.0001	1.017	0.9566	1.429	0.4265
Marital Status Divorced	vs. Married	1.829	0.0011	2.862	<.0001	4.342	<.0001
Marital Status Never_married	vs. Married	1.686	0.0078	2.773	<.0001	3.504	<.0001
Marital Status Widowed	vs. Married	0.469	0.0010	1.285	0.1995	0.647	0.1731
Region Banska_Bystrica	vs. Bratislava	1.469	0.1123	0.386	0.0006	2.121	0.1413
Region Kosice	vs. Bratislava	1.013	0.9573	0.801	0.3036	3.975	0.0037
Region Nitra	vs. Bratislava	1.283	0.3041	0.328	0.0001	2.480	0.0724
Region Presov	vs. Bratislava	1.714	0.0227	0.567	0.0227	3.005	0.0284
Region Trencin	vs. Bratislava	1.052	0.8447	0.897	0.6288	1.900	0.2277
Region Trnava	vs. Bratislava	1.110	0.6883	0.417	0.0016	2.005	0.1890
Region Zilina	vs. Bratislava	1.369	0.1844	0.610	0.0357	1.787	0.2614
Urbanisation Intermediate	vs. Dense	0.948	0.7581	0.945	0.7418	1.096	0.7700
Urbanisation Sparse	vs. Dense	1.515	0.0128	0.829	0.3005	1.788	0.0568

Zdroj: Vlastné spracovanie (SAS EG, MS Excel)

V prílohe č.1 tejto práce prikladáme kompletný výstup z bodových odhadov regresných parametrov. V prílohe č.2 a č.3 tiež prikladáme odhady parametrov a takisto pomerov šancí spolu s ich Waldovými intervalmi spoľahlivosti.

Vplyv premennej EAS – Status základnej ekonomickej aktivity

Referenčnou kategóriou pre premennú *EAS* sú domácnosti, v ktorých je *prednosta pracujúci (at_Work)*. Keďže intervaly spoľahlivosti pre odhad pomeru šancí pre domácností, ktoré majú *prednostu na dôchodku (Retired)* vo všetkých troch prípadoch obsahujú 1, so spoľahlivosťou 0,95 môžeme tvrdiť, že neexistuje štatisticky významný rozdiel medzi domácnosťami s pracujúcim prednostom a prednostom na dôchodku v pomere šance, že domácnosť trpí príjmovou chudobou a/alebo materiálnou depriváciou oproti šanci, že domácnosť netrpí žiadnou z týchto dimenzií chudoby a sociálneho vylúčenia.

V prípade *nezamestnaného prednostu (Unemployed)* je pomer šance, že domácnosť bude trpieť príjmovou chudobou a zároveň materiálnou depriváciou oproti šanci, že domácnosť nebude musieť čeliť ani jednému z týchto dvoch negatívnych javov (ďalej aj šanca „0“) až 67,287-násobne vyšší ako v domácnostiach s pracujúcim prednostom. Pomer šance, že domácnosť bude trpieť „len“ príjmovou chudobou oproti šanci „0“ je v domácnostiach s nezamestnaným prednostom 7,013-násobne vyšší ako v referenčnej kategórii a pomer šance, že bude trpieť „len“ materiálnou depriváciou oproti šanci „0“ je 4,476-násobne vyšší. Logicky ide o najviac ohrozené domácnosti, keďže absentuje príjem zo zamestnania.

U *iných neaktívnych osôb v pozícii prednostu (Inactive_person)* domácnosti takisto pozorujeme zvýšené riziko konzistentnej chudoby. V týchto domácnostiach je pomer šance ohrozenia príjmovou chudobou a súčasne materiálnou depriváciou oproti šanci „0“ takmer 9-krát vyšší ako v domácnostiach s pracujúcim prednostom.

Vplyv premennej Education - Najvyššia dosiahnutá úroveň vzdelania

Referenčná kategória je v tomto prípade *Terciárne vzdelanie 2. a 3. stupňa (Tertiary_2_3)* – dokončené magisterské alebo doktorandské vzdelanie. Najväčšie ohrozenie konzistentnou chudobou pozorujeme v domácnostiach, kde *prednosta má nižšie sekundárne alebo nižšie vzdelanie (Less_than_Secondary)*. V týchto domácnostiach je pomer šance, že domácnosť bude trpieť príjmovou chudobou a zároveň materiálnou depriváciou oproti šanci „0“ až 30-násobne vyšší ako v domácnostiach, kde prednosta má dokončené terciárne vzdelanie aspoň 2. stupňa.

V domácnostiach, kde má prednosta ukončené stredoškolské vzdelanie sa pomer šance, že domácnosť trpí konzistentnou chudobou oproti šanci „0“ výrazne znižuje. Stále je tu však vyššie ohrozenie príjmovou chudobou a súčasne materiálnou depriváciou, tento pomer šancí je však pri domácnostiach s *prednostom s ukončenou strednou školou (Upper_Secondary)* nižší až o 86,47% a pri domácnostiach s *prednostom, ktorý má pomaturitné, avšak nie terciárne vzdelanie (Post_Secondary)* nižší o 84,45%.

Vplyv premennej Health - Všeobecné zdravie

Referenčnou kategóriou pre *Health* je *prednosta domácnosti v dobrom zdravotnom stave (Good)*. Najohrozenejšou kategóriou v rámci tejto premennej sú samozrejme domácnosti s *prednostom v zlom zdravotnom stave (Bad)*. Pomer šance, že domácnosť je

ohrozená príjmovou chudobou a zároveň materiálnou depriváciou oproti šanci „0“, pre domácnosť s prednostom v zlom zdravotnom stave oproti domácnosti s prednostom v dobrom zdravotnom stave, je 2,141-násobne vyšší.

Pri domácnostiach s *prednostom*, ktorý nemá dobrý ani zlý zdravotný stav (*Fair*) oproti domácnostiam s prednostom v dobrom zdravotnom stave nie je štatisticky významný rozdiel v pomere šance ohrozenia domácnosti príjmovou chudobou a materiálnou depriváciou. Existujú štatisticky významné rozdiely v pomere šance ohrozenia zvlášť príjmovou chudobou a zvlášť materiálnou depriváciou, nie sú však veľmi výrazné.

Vplyv premennej Household Type - Typ domácnosti

Pre premennú *Household Type* je referenčnou kategóriou domácnosť s 2 dospelými a 2 deťmi (2A_2Ch). V najväčšom riziku konzistentnej chudoby oproti domácnostiam s 2 dospelými a 2 deťmi sú domácnosti, v ktorých sa nachádza 1 dospelý a aspoň 1 dieťa (1A_at_least_1Ch) a domácnosti 2 dospelých s aspoň 3 deťmi (2A_at_least_3ch). Pri domácnostiach s 1 dospelým a aspoň 1 dieťaťom je oproti referenčnej kategórii pomer šance, že budú trpieť príjmovou chudobou a súčasne materiálnou depriváciou oproti šanci „0“ väčší o 8,891 násobok. Pri druhých menovaných domácnostiach je tento pomer šancí 6,578-násobne vyšší.

Naopak, oproti domácnostiam s dvomi dospelými a dvomi deťmi, sú rizikom príjmovej chudoby najmenej ohrozené domácnosti s 2 dospelými a 1 osobou na dôchodku (2A_1R) a domácnosti bez detí (2Adult_0Ch a Other_0Ch) – pomer šance, že domácnosť z prvej spomínanej kategórie nebude trpieť príjmovou chudobou ani materiálnou depriváciou (šanca „0“) oproti šanci, že bude trpieť príjmovou chudobou, je až 9,091-násobne vyšší (1/0,110) ako pri domácnosti s 2 dospelými a 2 deťmi.

Vplyv premennej Marital Status - Rodinný stav

Referenčná kategória pre premennú Marital Status sú domácnosti s vydatým/ženatým prednostom (*Married*), teda domácnosti s manželským párom. Jedinou kategóriou s nižším pomerom šancí rizika príjmovej chudoby sú domácnosti s ovdoveným prednostom, pri ktorých je takýto pomer šancí oproti domácnostiam s manželským párom 0,469-násobne nižší.

Oproti tomu, domácnosti s *rozvedeným* alebo *slobodným prednostom* (*Divorced* a *Never_married*) majú vyšší pomer šanci, že budú trpieť príjmovou chudobou a/alebo materiálnou depriváciou, ako že nebudú trpieť ani jednou z týchto dimenzií. Zvlášť, pre domácnosti s rozvedeným prednostom je pomer šance, že budú trpieť príjmovou chudobou aj materiálnou depriváciou oproti šanci „0“ 4,342-násobne vyšší, zatiaľ čo pre domácnosti so slobodným prednostom je tento pomer šanci vyšší 3,504-násobne.

Vplyv premennej Region – Kraj

Ako referenčnú premennú sme pre túto kategóriu zvolili *Bratislavský kraj*. Domácnosti z ostatných krajov majú oproti domácnostiam z Bratislavského kraja vyšší pomer šance, že budú trpieť príjmovou chudobou a zároveň materiálnou depriváciou oproti šanci „0“. Štatisticky významne vyššie uvedené pomery šanci v porovnaní s Bratislavským krajom mali v roku 2015 domácnosti z *Košického kraja* a *Prešovského kraja*. Pre Prešovský kraj je takýto pomer šance 3-násobne vyšší a pre Košický kraj takmer 4-násobne vyšší.

Naopak, najhoršie je na tom Bratislavský kraj v oblasti ohrozenia materiálnou depriváciou, pričom významne nižší pomer šance, že budú ohrozené touto dimenziou chudoby a sociálneho vylúčenia oproti šanci „0“, majú hlavne domácnosti z krajov *Banská Bystrica*, *Nitra*, *Prešov*, *Trnava* a *Žilina*, pričom pre Nitriansky kraj je to až 0,328-násobne nižší pomer šanci.

Vplyv premennej Urbanisation – Stupeň Urbanizácie

Pri premennej *Urbanisation* sme ako referenčnú kategóriu zvolili domácnosti, ktoré žijú na *území s hustým osídlením* (*Dense*). V ostatných dvoch kategóriách – domácnosti, ktoré žijú na *území s priemerne hustým osídlením* (*Intermediate*) a s *riedkym osídlením* (*Sparse*) – nie sú príliš veľké rozdiely v šanciach ohrozenia dimenziami chudoby a sociálneho vylúčenia. Iba pri domácnostiach, ktoré žijú na území s riedkym osídlením, môžeme pozorovať väčšie ohrozenie príjmovou chudobou – pomer šance, že takéto domácnosti budú ohrozené príjmovou chudobou oproti šanci „0“ je 1,515-násobne vyšší ako pre domácnosti žijúce v oblasti s hustým osídlením.

4.3 Logistická regresia ako nástroj na výpočet pravdepodobnosti

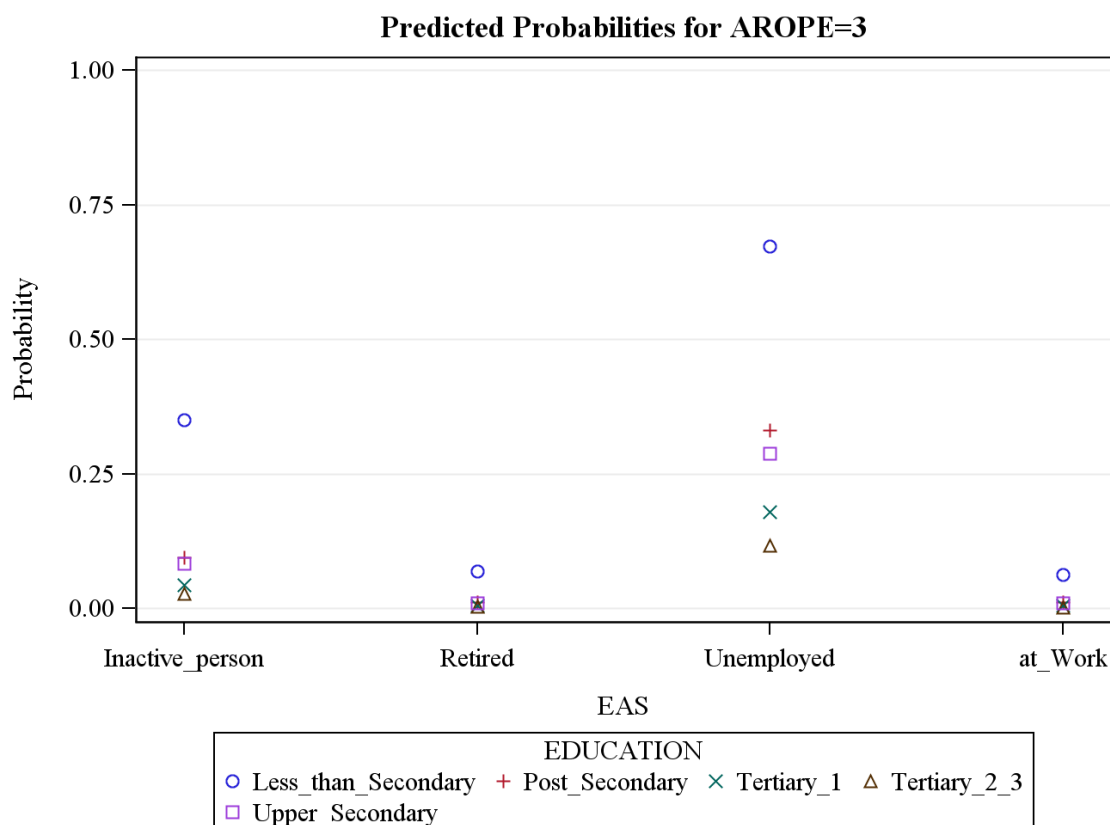
Na základe odhadu multinomického logistického modelu (tab. 4.2.1.1) vieme odhadnúť pravdepodobnosť, že domácnosť bude „len“ v riziku príjmovej chudoby (model 1 vs. 0) alebo bude „len“ závažne materiálne deprivovaná (model 2 vs. 0) alebo bude čeliť konzistentnej chudobe (3 vs. 0) a to v závislosti od všetkých relevantných faktorov uvedených v tab. 4.2.1.1. Ak by sme chceli vypočítať len pravdepodobnosť konzistentnej chudoby pre všetky skupiny domácností, bolo by to 51 840 odhadov pravdepodobností. Išlo by o odhady pre 51 840 skupín domácností, ktoré vznikli v členení podľa relevantných faktorov uvedených v tab. 4.2.5. Aby sme výsledky urobili prehľadnejšími, rozhodli sme sa odhadnúť čiastkové modely, ktoré zahŕňajú 2 relevantné faktory, a to status základnej ekonomickej aktivity prednostu domácnosti a ďalší z ostatných relevantných faktorov. Tieto modely sú odhadnuté v Prílohe č. 5 a ich štatistická významnosť je overená v Prílohe č. 4. Na základe týchto čiastkových modelov môžeme identifikovať najohrozenejšie a najmenej ohrozené skupiny domácností, vypočítať pravdepodobnosti, že skupiny domácností v členení podľa 2 faktorov (EAS a niektorý ďalší relevantný faktor) budú „len“ v riziku príjmovej chudoby alebo budú „len“ závažne materiálne deprivovaná (model 2 vs. 0) alebo budú čeliť konzistentnej chudobe alebo nebudú ohrozené ani príjmovou chudobou ani závažnou materiálnou depriváciou. V prípade takéhoto prístupu získame 20 skupín domácností v členení podľa faktorov EAS a Education, 12 skupín v členení podľa faktorov EAS a Health, atď. Celkovo vznikne 128 skupín domácností, čo je výrazne menej ako pri súčasnom členení podľa všetkých uvažovaných faktorov (51 840 skupín domácností), čo prispeje k väčšej prehľadnosti výsledkov, aj keď na úkor určitej straty informácií.

V tejto časti sa zameriame len na najohrozenejšiu skupinu domácností – domácnosti v konzistentnej chudobe, teda domácnosti, ktoré sú v riziku príjmovej chudoby a súčasne musia čeliť závažnej materiálnej deprivácii. Ukážeme výpočet odhadov konzistentnej chudoby pre jednotlivé skupiny domácností pomocou vzťahu:

$$\hat{\pi}_{i3} = \frac{\exp(\hat{\beta}_{30} + \sum_{j=1}^k \hat{\beta}_{3j} x_{ij})}{1 + \sum_{l=1}^{s-1} \exp(\hat{\beta}_{l0} + \sum_{j=1}^k \hat{\beta}_{lj} x_{ij})}$$

a tieto odhady pre jednoduchšie porovnanie znázorníme v grafoch pravdepodobností predpovedaných modelom (*Predicted Probabilities*), kde na osi *x* sa vždy nachádza premenná *EAS* a na osi *y* predpovedané pravdepodobnosti. Do grafu sú postupne rovnobežne s osou *y* nanášané pravdepodobnosti druhej z dvojice vysvetľujúcich premenných vzhľadom na kategórie premennej *EAS*.

4.3.1 Model s vysvetľujúcimi premennými *EAS* a *Education*



Obrázok 4.3.1.1 – *Pravdepodobnosť predpovedaná modelom pre EAS a Education*
Zdroj: Vlastné spracovanie (SAS EG)

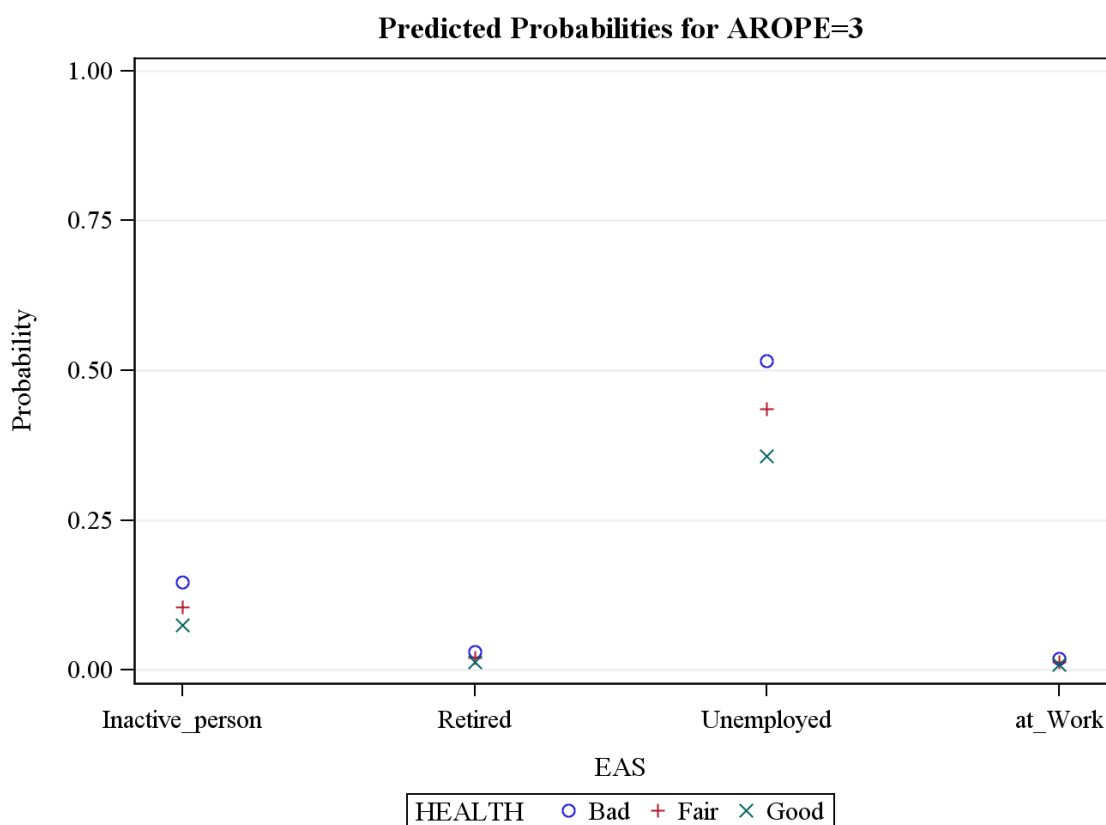
Z grafu môžeme vidieť, že vo všetkých kategóriách ekonomickej aktivity sú najohrozenejšou skupinou domácnosti s prednostom, ktorý nemá dokončené stredoškolské vzdelanie. Pritom najhoršie sú na tom domácnosti s prednostom, ktorý nemá dokončené stredoškolské vzdelanie a je nezamestnaný. Pravdepodobnosť, že takáto domácnosť bude trpieť príjmovou chudobou a zároveň materiálnou depriváciou sa podľa grafu pohybuje až okolo 70%.

Presnejšie sa na túto pravdepodobnosť pozrieme pomocou vzorca pre odhad pravdepodobnosti:

$$\hat{\pi}_{i3} = \frac{\exp(-5,8090 + 4,1785 + 3,3606)}{1 + \exp(-2,8180 + 1,8960 + 1,2428) + \exp(-4,1522 + 1,4820 + 1,6816) + \exp(-5,8090 + 4,1785 + 3,3606)}$$

$$= \frac{5,6718}{1 + 1,3782 + 0,3721 + 5,6718} = 0,7013 = \mathbf{70,13\%}$$

4.3.2 Model s vysvetľujúcimi premennými EAS a Health



Obrázok 4.3.2.1 - Pravdepodobnosť predpovedaná modelom pre EAS a Health

Zdroj: Vlastné spracovanie (SAS EG)

Ako môžeme vidieť, najohrozenejšia skupina domácností vo všetkých kategóriách premennej EAS je tá, ktorá má prednostu v zlom zdravotnom stave. V domácnostiach, kde prednosta je na dôchodku alebo zamestnaný, rozdiely v pravdepodobnosti ohrozenia príjmovou chudobou a súčasne materiálnou depriváciou podľa zdravia nie sú veľmi výrazné. Najohrozenejšie domácnosti sú opäť tie, ktoré majú prednostu nezamestnaného

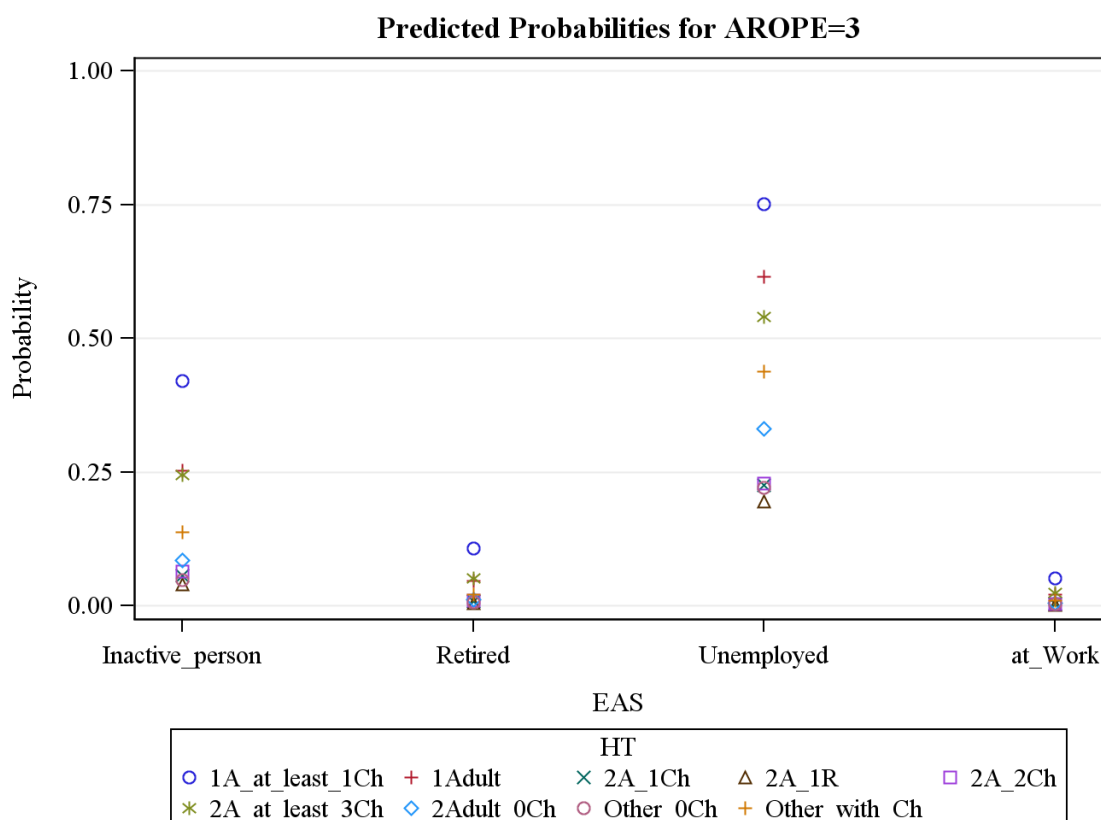
a zároveň v zlom zdravotnom stave. Pravdepodobnosť, že tieto domácnosti trpia konzistentnou chudobou je nad hranicou 50%.

Presnú hodnotu vypočítame využitím vzorca:

$$\hat{\pi}_{i3} = \frac{\exp(-4,6518 + 4,6648 + 0,9204)}{1 + \exp(-2,4618 + 2,0529 + 0,2984) + \exp(-3,3889 + 1,5810 + 1,1086) + \exp(-4,6518 + 4,6648 + 0,9204)}$$

$$= \frac{2,5523}{1 + 0,8954 + 0,4969 + 2,5523} = 0,5162 = \mathbf{51,62\%}$$

4.3.3 Model s vysvetľujúcimi premennými EAS a Household Type



Obrázok 4.3.3.1 - Pravdepodobnosť predpovedaná modelom pre EAS a Household Type

Zdroj: Vlastné spracovanie (SAS EG)

Domácnosti, ktoré pozostávajú z 1 dospelého a aspoň 1 dieťaťa sú najviac rizikovými domácnosťami vo všetkých kategóriách ekonomickej aktivity. Najohrozenejšie sú práve vtedy, ak je prednosta domácnosti zároveň nezamestnaný. Pravdepodobnosť, že

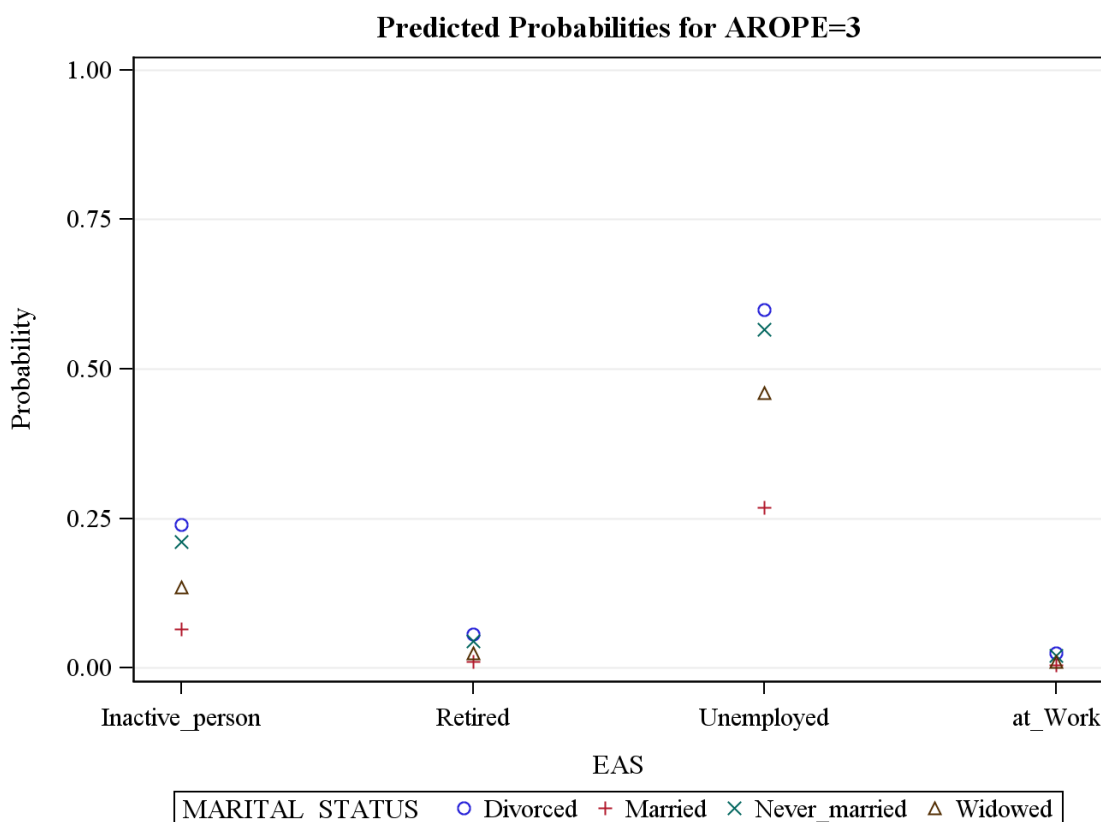
takéto domácnosti sú ohrozené príjmovou chudobou a zároveň materiálnou depriváciou je veľmi vysoká – tesne nad hranicou 75%.

Presnú hodnotu vyrátame nasledovne:

$$\hat{\pi}_{i3} = \frac{\exp(-5,2928 + 4,9151 + 2,6436)}{1 + \exp(-2,0464 + 2,1898 + 0,3469) + \exp(-3,5013 + 1,7015 + 1,2307) + \exp(-5,2928 + 4,9151 + 2,6436)}$$

$$= \frac{9,6400}{1 + 1,6328 + 0,5660 + 9,6400} = 0,7508 = \mathbf{75,08\%}$$

4.3.4 Model s vysvetľujúcimi premennými EAS a Marital Status



Obrázok 4.3.4.1 - Pravdepodobnosť predpovedaná modelom pre EAS a Marital Status

Zdroj: Vlastné spracovanie (SAS EG)

Z hľadiska rodinného stavu osoby na čele domácnosti sú najrizikovejšie domácnosti, kde prednosta je rozvedený a tesne za nimi domácnosti, v ktorých prednosta je slobodný. Pritom ale v kategóriách, kde je prednosta zároveň na dôchodku alebo zamestnaný, sú rozdiely v pravdepodobnosti, že domácnosť bude trpieť príjmovou chudobou a súčasne materiálnou depriváciou minimálne. Najohrozenejšími sú potom

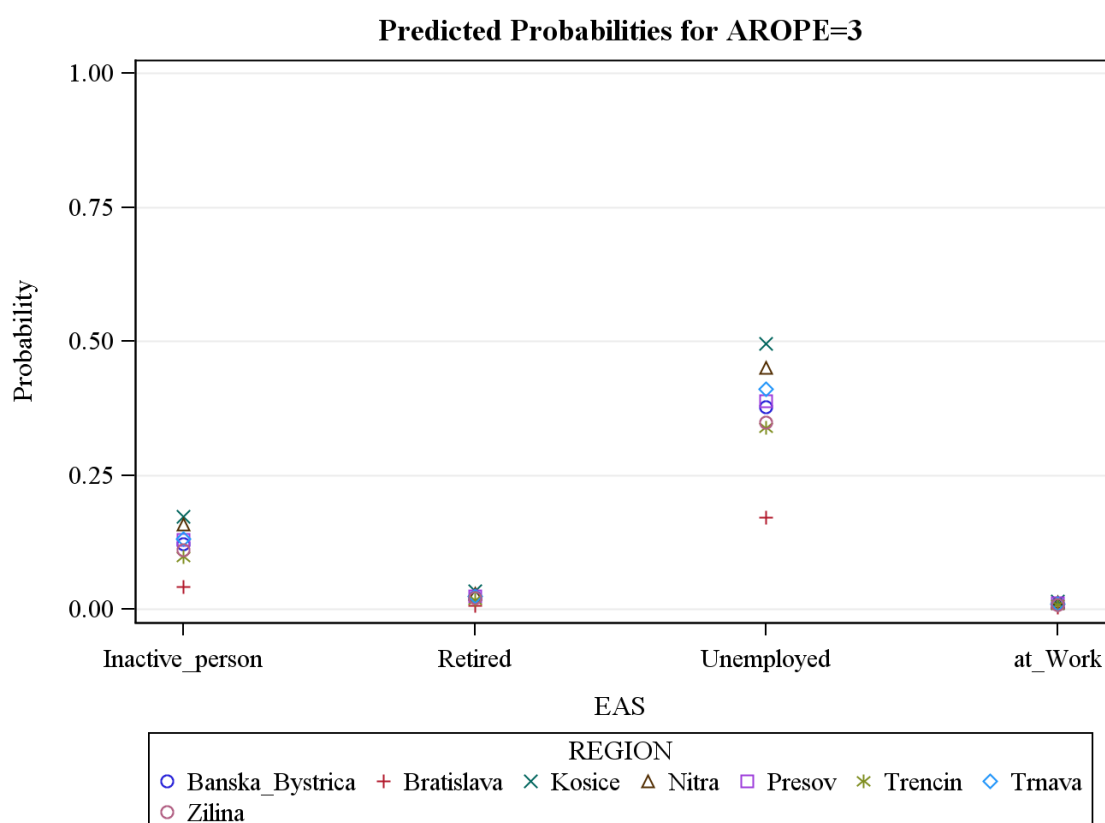
domácnosti, kde prednosta je rozvedený a zároveň nezamestnaný. V tejto skupine domácností je pravdepodobnosť konzistentnej chudoby až okolo 60%.

To si môžeme overiť aj podľa nasledujúceho vzorca:

$$\hat{\pi}_{i3} = \frac{\exp(-5,2579 + 4,8377 + 1,7973)}{1 + \exp(-2,5763 + 2,1126 + 0,5976) + \exp(-3,5587 + 1,7409 + 1,1423) + \exp(-5,2579 + 4,8377 + 1,7973)}$$

$$= \frac{3,9634}{1 + 1,1433 + 0,5089 + 3,9634} = 0,5991 = \mathbf{59,91\%}$$

4.3.5 Model s vysvetľujúcimi premennými EAS a Region



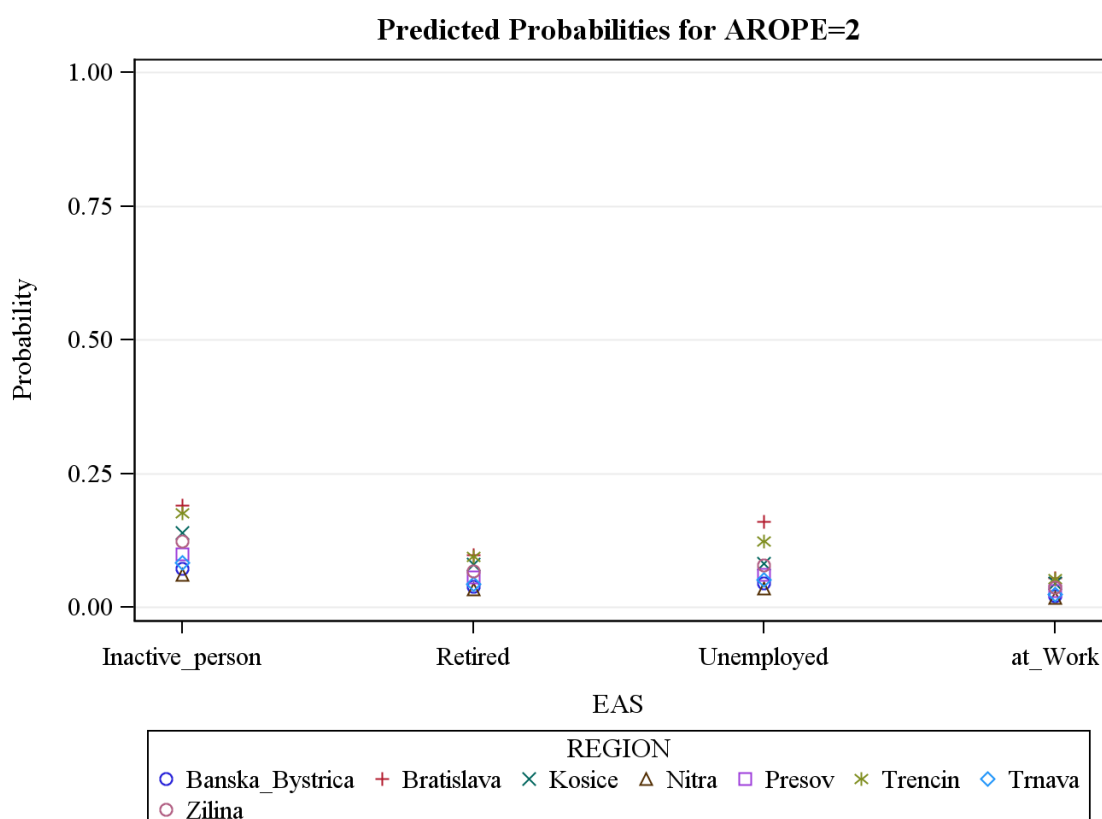
Obrázok 4.3.5.1 - Pravdepodobnosť predpovedaná modelom pre EAS a Region
Zdroj: Vlastné spracovanie (SAS EG)

Najrizikovejší krajom z pohľadu ohrozenia konzistentnou chudobou je Košický kraj, za ktorým nasleduje Nitriansky, Trnavský a Prešovský kraj. Naopak, najmenej rizikovým regiónom je Bratislavský kraj. Najohrozenejšími domácnosťami z hľadiska príjmovej chudoby a súčasne materiálnej deprivácie sú domácnosti z Košického kraja

s nezamestnaným prednostom. Pre takéto domácnosti je pravdepodobnosť, že budú trpieť týmito dimenziami chudoby a sociálneho vylúčenia takmer 50%.

$$\hat{\pi}_{i3} = \frac{\exp(-5,6028 + 4,6440 + 1,5497)}{1 + \exp(-2,7383 + 2,0296 + 0,0955) + \exp(-2,7979 + 1,7686 - 0,1871) + \exp(-5,6028 + 4,6440 + 1,5497)}$$

$$= \frac{1,8056}{1 + 0,5416 + 0,2963 + 1,8056} = 0,4956 = \mathbf{49,56\%}$$

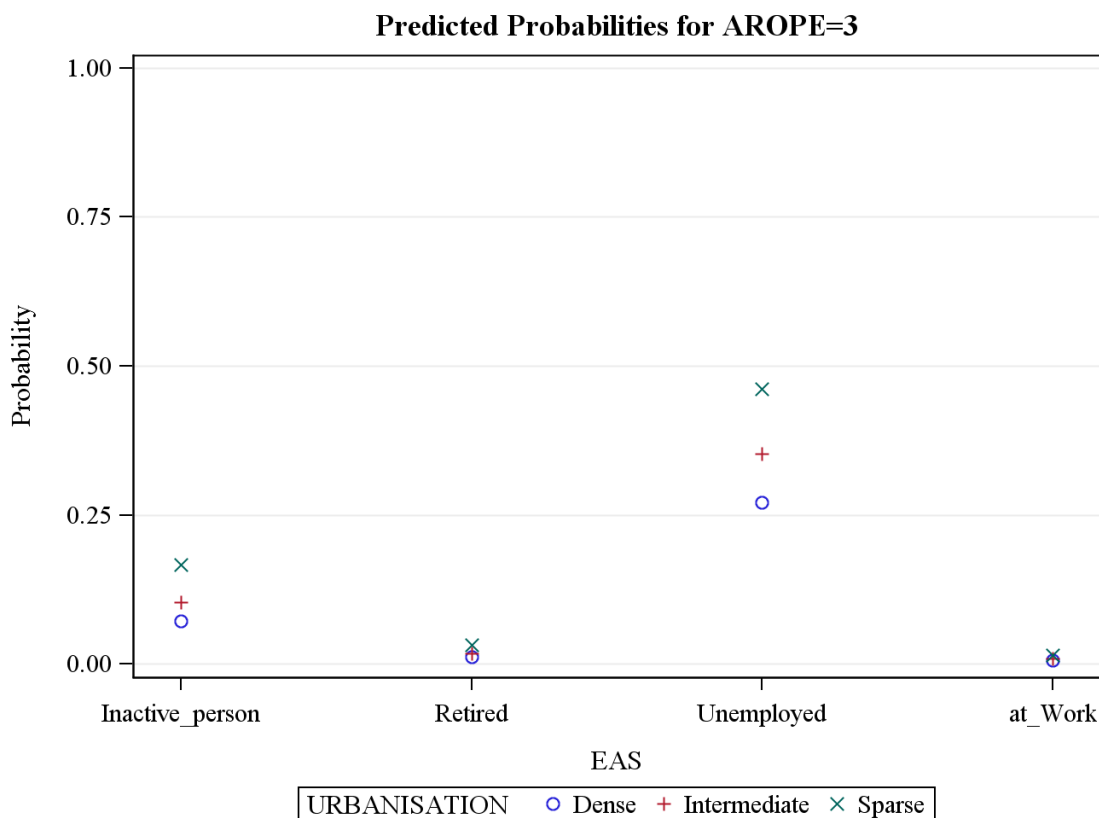


Obrázok 4.3.5.2 - Pravdepodobnosť predpovedaná modelom pre EAS a Region

Zdroj: Vlastné spracovanie (SAS EG)

Zaujímavé ale je, že z pohľadu ohrozenia materiálnou depriváciou je najohrozenejším regiónom práve Bratislavský kraj a domácnosti v ňom žijúce. Zatiaľ čo s približne 20-timi percentami sú v kategórii nezamestnaných prednostov najmenej ohrozené konzistentnou chudobou domácnosti z Bratislavského kraja, tak s približne rovnakým podielom pre domácnosti v závažnej materiálnej deprivácii (obr. 4.3.5.2) sú domácnosti s nezamestnaným prednostom z Bratislavského kraja najrizikovejšou kategóriou.

4.3.6 Model s vysvetľujúcimi premennými EAS a Urbanisation



Obrázok 4.3.6.1 - Pravdepodobnosť predpovedaná modelom pre EAS a Urbanisation

Zdroj: Vlastné spracovanie (SAS EG)

Najohrozenejšími domácnosťami vzhľadom na hustotu osídlenia sú logicky domácnosti sídlia v oblastiach s riedkym osídlením, za nimi nasledujú domácnosti sídlia v oblastiach s priemerne hustým osídlením a najmenej ohrozené sú domácnosti žijúce v oblastiach s hustým osídlením. Súvisí to najmä s ponukou práce – v mestách je najviac pracovných príležitostí. V najväčšom riziku sú domácnosti žijúce v oblasti s riedkym osídlením a zároveň nezamestnaným prednostom. Pravdepodobnosť, že takéto domácnosti trpia príjmovou chudobou aj materiálnou depriváciou je približne 45%.

$$\hat{\pi}_{i3} = \frac{\exp(-5,0599 + 4,6453 + 1,0083)}{1 + \exp(-2,7066 + 2,0129 + 0,6026) + \exp(-3,0142 + 1,7347 - 0,3275) + \exp(-5,0599 + 4,6453 + 1,0083)}$$

$$= \frac{1,8107}{1 + 0,9129 + 0,2005 + 1,8107} = 0,4614 = \mathbf{46,14\%}$$

Záver

Hlavným cieľom tejto záverečnej práce bolo priblíženie problematiky príjmovej chudoby, materiálnej deprivácie a konzistentnej chudoby Slovenských domácností. Tento cieľ sme postupne rozpracovávali na základe stanovených čiastkových cieľov.

V prvých kapitolách sme sa venovali hlavne súčasnému stavu tejto problematiky a vypracovaniu metodológie potrebnej pre uskutočnenie potrebných štatistických analýz. Pomocou metód analýzy kontingencie, resp. asociácie a postupov multinomickej logistickej regresie sme identifikovali relevantné premenné, ktoré v roku 2015 významne vplývali na riziko príjmovej chudoby, hrozbu závažnej materiálnej deprivácie prípadne na ohrozenie slovenských domácností konzistentnou chudobou. Do skupiny relevantných premenných boli zaradené tieto premenné: Status základnej ekonomickej aktivity, Najvyššia dosiahnutá úroveň vzdelania, Všeobecné zdravie, Typ domácnosti, Rodinný stav, Kraj a Stupeň urbanizácie. Multinomická logistická regresia nám navyše umožňuje kvantifikovať vplyv relevantných faktorov na uvedené negatívne sociálne javy a odhadnúť pravdepodobnosť ich výskytu v rôznych skupinách domácností, ktoré vznikli triedeným slovenských domácností podľa relevantných faktorov.

Ako najohrozenejšie konzistentnou chudobou vzhľadom na ekonomickú aktivitu sa ukázali domácnosti s nezamestnaným prednostom, pre ktoré je pomer šance, že domácnosť bude trpieť konzistentnou chudobou oproti šanci, že nebude trpieť príjmovou chudobou a zároveň ani materiálnou depriváciou až 67,287-násobne vyšší ako v domácnostiach s pracujúcim prednostom. V oblasti vzdelania sú najviac ohrozené domácnosti, v ktorých prednosta má nižšie sekundárne alebo nižšie vzdelanie, kde takýto pomer šance oproti domácnostiam s prednostom s dokončeným terciárnym vzdelaním aspoň 2. stupňa je 30-násobne vyšší. V rámci typu domácnosti sú v najväčšom riziku konzistentnej chudoby domácnosti s 1 dospelým a aspoň 1 dieťaťom, oproti domácnostiam s 2 dospelými a dvomi deťmi je pomer šance 8,891-násobne vyšší. Z hľadiska rodinného stavu, zdravotného stavu, ďalej kraja a hustoty osídlenia, kde domácnosť žije, boli v roku 2015 najviac ohrozené domácnosti s rozvedeným prednostom, domácnosti s prednostom v zlom zdravotnom stave, domácnosti z Košického kraja a domácnosti žijúce na území s riedkym osídlením, rozdiely v pomeroch šanci sú tu už ale menšie.

Ďalej sme pomocou modelu logistickej regresie odhadli pravdepodobnosti rizika konzistentnej chudoby pre jednotlivé najohrozenejšie kategórie domácností. Najhoršie z hľadiska ohrozenia konzistentnou chudobou sú na tom domácnosti s nezamestnaným prednostom – v rámci nich najväčšiu pravdepodobnosť, že budú trpieť konzistentnou chudobou, majú domácnosti s 1 dospelým a aspoň 1 dieťaťom, až 75,08%. Nasledujú domácnosti, kde prednosta je nezamestnaný a nemá dokončené stredoškolské vzdelanie, kde je takáto pravdepodobnosť 70,13%. Nasledujú domácnosti s rozvedeným prednostom, ďalej domácnosti s prednostom v zlom zdravotnom stave. Keďže kraj a stupeň urbanizácie mali o niečo menší vplyv na riziko konzistentnej chudoby ako vyššie uvedené faktory, tak aj v najrizikovejších skupinách domácností (domácnosti z Košického kraja s nezamestnaným prednostom a v domácnosti žijúce v riedko osídlených oblastiach taktiež s nezamestnaným prednostom) sme odhadli výrazne nižšie pravdepodobnosti konzistentnej chudoby. V oboch skupinách domácností bola pravdepodobnosť, že domácnosť bude v riziku príjmovej chudoby a súčasne bude závažne materiálne deprivovaná, pod 50 % (presnejšie 49,56 % pre domácnosti s nezamestnaným prednostom z Košického kraja a 46,14 % pre domácnosti s nezamestnaným prednostom žijúce na riedko osídlených územiach).

Téma záverečnej práce je veľmi aktuálna a na základe jej výsledkov je vidno, že problematika chudoby je stále závažná a jej témy neboli vyčerpané. Na tomto základe sme stanovili hlavný cieľ práce, ktorý sa nám podarilo naplniť pomocou jeho čiastkových cieľov a veríme, že práca bude slúžiť ako užitočný zdroj informácií o danej problematike. Jednak sme jednoznačne identifikovali skupiny domácností, ktoré sú v najväčšom riziku konzistentnej chudoby, čo je možné využiť pri tvorbe potrebných opatrení boja proti chudobe či už v rámci Ministerstva práce, sociálnych vecí a rodiny Slovenskej republiky a jeho špeciálnych odborov, napr. Odboru štátnej podpory a stratégie sociálnej a rodinnej politiky. Táto práca môže pomôcť pri lepšom cílení potrebných opatrení a takisto určení potrebného rozsahu týchto opatrení.

Na druhej strane je jasné, že použité metódy nie sú v rámci danej problematiky úplné a vyčerpávajúce a preto je vhodné v rozpracovaných analýzach pokračovať, avšak táto práca poskytuje určitý základ, ktorý je možné v ďalších analýzach využiť, či už v rámci výskumu na Inštitúte pre výskum práce a rodiny alebo pri štatistických hláseniach tvorených na Štatistickom úrade Slovenskej republiky.

Takisto veríme, že práca bude užitočná pre študentov Ekonomickej Univerzity v Bratislave či iných univerzít pri štúdiu problematiky chudoby a sociálneho vylúčenia, alebo pri skúmaní využitia logistickej regresie. Napokon dúfame, že práca zaujme aj širšiu verejnosť, nakoľko poznávanie situácie v oblasti chudoby a sociálneho vylúčenia by malo prirodzene patriť k oblasti ľudského záujmu v snahe zlepšiť nielen našu situáciu v tejto oblasti, ale aj situáciu v celej spoločnosti.

Použitá literatúra

ALLISON, P. D. 2012. *Logistic Regression using SAS: Theory and Application, Second Edition*. Severná Karolína: SAS Institute, 2012. 339 s. ISBN 978-1-59994-641-2

CRAMER, J. S. 2003. *The origins and development of the logit model*. [online]. 2003, [cit. 2018.01.13]. Dostupné na internete: <<http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=D83287C62A2454CB1E781A85C03FBDB3?doi=10.1.1.527.242&rep=rep1&type=pdf>>

KIS, A. – GÁBOS, A. 2016. *Consistent poverty across the EU*. *Corvinus journal of sociology and social policy*. [online]. 2016, [cit. 2018.04.05]. Dostupné na internete: <<http://unipub.lib.uni-corvinus.hu/2837/1/159-758-1-PB.pdf>>

KORMANČÍKOVÁ, B. 2013. *Stratégia Európa 2020*. [online]. 2013, [cit. 2018.04.05]. Dostupné na internete: < http://old.sazp.sk/public/index/open_file.php?file=CTK/GM/Workshop/Kormancikova_EU2020.pdf>

MICHÁLEK, A. 2000. *Chudoba, jej koncepty a geografické dimenzie*. *Geografický časopis* [online]. 2000, [cit. 2018.04.05]. Dostupné na internete: <<https://www.sav.sk/journals/uploads/05131121Michalek.pdf>>

NAU, R. 2017. *What's a good value for R-squared?*. [online]. 2017, [cit. 2018.01.13]. Dostupné na internete: <<https://people.duke.edu/~rnau/rsquared.htm>>

PACÁKOVÁ, V. a kol. 2003. *Štatistika pre ekonómov*. Bratislava: Iura Edition, 2003. 358 s. ISBN 80-89047-74-2

PECÁKOVÁ, I. 2007. *Logistická regrese s vícekategoriální vysvětlovanou proměnnou*. *Acta Oeconomica Pragensia*. [online]. 2007, [cit. 2018.01.13]. Dostupné na internete: <<https://www.vse.cz/polek/download.php?jnl=aop&pdf=42.pdf>>

RUBLÍKOVÁ, E. – LABUDOVÁ, V. – SANDTNEROVÁ, S. 2009. *Analýza kategoriálních údajov*. Bratislava: Vydavateľstvo Ekonóm, 2009. 174 s. ISBN 978-80-225-2710-1

ŘEHÁKOVÁ, B. 2000. *Nebojte se logistické regrese*. *Sociologický Časopis*. [online]. 2000, [cit. 2018.01.13]. Dostupné na internete: <<http://www.jstor.org/stable/41131581>>

ŠOLTÉS, E. – ŠOLTÉSOVÁ, T. 2016. *Diverzita chudoby a sociálneho vylúčenia v krajinách V4 a EÚ-28 v roku 2014. Štatistická analýza chudoby a sociálneho vylúčenia v kontexte stratégie Európa 2020, I. časť*. Bratislava: Vydavateľstvo Ekonóm, 2016. 122 s. ISBN 978-80-225-4304-0

ŠOLTÉS, E. – VOJTKOVÁ, M. – ŠOLTÉSOVÁ, T. 2018. *Work Intensity of Households: Multinomial Logit Analysis and Correspondence Analysis for Slovak Republic*. Statistika. [online]. 2018, [cit. 2018.04.05]. Dostupné na internete: <<https://www.czso.cz/documents/10180/61266317/32019718q1.pdf/aa3a343c-2e16-4a4f-8786-999a54b3579b?version=1.0>> ISSN 1804-8765

TEREK, M. – HORNÍKOVÁ, A. – LABUDOVÁ, V. 2010. *Hĺbková analýza údajov*. Bratislava: Iura Edition, 2010. 265 s. ISBN 978-80-8078-336-5

VLAČUHA, R. – KOVÁČOVÁ, Y. 2017. *EU SILC 2016 Indikátory chudoby a sociálneho vylúčenia*. [online]. 2017, [cit. 2018.03.18]. Dostupné na internete: <http://zbw.eu/econis-archiv/bitstream/handle/11159/826/EU_SILC_2016_Indikatory_chudoby_a_socialneho_vylucenia.pdf?sequence=1&isAllowed=y> ISBN 978-80-8121-557-5

ŽELINSKÝ, T. 2014. *Chudoba a deprivácia na Slovensku: Metodologické aspekty a empiria*. [online]. 2017, [cit. 2018.04.05]. Dostupné na internete: <https://mpra.ub.uni-muenchen.de/76868/1/Zelinsky_Chudoba_online.pdf> ISBN 978-80-8143-134-0

Zoznam príloh

Príloha č. 1: Bodové odhady parametrov modelu so všetkými vysvetľujúcimi premennými

Príloha č. 2: Bodové odhady parametrov modelu so všetkými vysvetľujúcimi premennými a ich Waldove intervaly spoľahlivosti

Príloha č. 3: Bodové odhady OR modelu so všetkými vysvetľujúcimi premennými a ich Waldove intervaly spoľahlivosti

Príloha č. 4: Štatistická významnosť čiastkových modelov a ich parametrov

Príloha č. 5: Bodové odhady parametrov čiastkových modelov